

On the Burnside Problem

Drew A. Hudec

University of Chicago REU 2006

Abstract

This paper will investigate the classical Burnside problem as it was originally proposed in 1902 and will focus on developing a counterexample to this conjecture involving automorphism groups on binary rooted trees. This example is more widely known as the first Grigorchuk group. I will introduce all the relevant ideas about binary rooted trees and their connection to the Grigorchuk group, so that this survey will be more self-contained. Although we will see that Burnside's original conjecture is false in its full generality, I will also present a short historical overview of some related results that have since been attained.

1 Introduction to Burnside's Problem

Although having studied under the great algebraist Arthur Cayley while at Cambridge, William Burnside did not begin his mathematical career in this field. In fact, his university studies were focused in applied mathematics and later hydrodynamics. It was years later that his research shifted from applied mathematics to finite group theory, the subject that would later bring him fame. Today Burnside can be regarded as one of the most influential figures in the history of this subject, after having drawn the interest of many mathematicians throughout the twentieth century with his proposed problems. Before stating his question, we will need a few preliminary definitions.

Definition 1.1: Let G be a group and $S \subset G$. Let $\mathfrak{J} = \{H \leq G \mid S \subset H\}$. Then $\langle S \rangle = \bigcap_{H \in \mathfrak{J}} (H)$ is the subgroup of G generated by S . A group G is said to be finitely generated if it is generated by some finite subset $S \subset G$, ie. $G = \langle S \rangle$.

Definition 1.2: Let G be a group. An element $g \in G$ has finite order if $\exists n \in \mathbb{Z}$ such that $g^n = 1$.

Definition 1.3: A group G is periodic if $\forall g \in G, \exists n \in \mathbb{Z}$ such that $g^n = 1$, ie. Every element in G has finite order.

Now we are ready to state Burnside's question:

Is it the case that all finitely generated, periodic groups are finite?

It may seem at first that this question should be answered in the affirmative. For there are only finitely many generators for the group, each of which has finite order. So this amounts to at most a finite number of distinct elements. Then all the finite products of the generators are also elements in the group, and thus these elements will each have finite order as well. It may seem as though we could somehow make the leap from this information to conclude that the group must then be finite. But this intuitive simplicity is only apparent. In the following sections we will define the first Grigorchuk group - one of the best known counterexamples to the Burnside problem - which will show that the answer to this question is *no*.

2 Graphs and Trees

The definition of the Grigorchuk group involves permutations that act on certain types of graphs. Formally, a graph G is a pair of sets $G = (V, E)$ where V is a set of vertices and E is a set of edges. Edges are defined by pairs of vertices, and visualized by a line segment connecting them. Two vertices in a graph are said to be adjacent if the edge that they form is contained in E , and adjacent vertices are called neighbors. In a graph, a walk is a sequence of vertices v_0, v_1, \dots, v_k such that v_i and v_{i+1} are adjacent for all $i \in \{0, \dots, k-1\}$. A cycle is a walk in which $v_0 = v_k$ and there are no repeated vertices in the sequence beside $v_0 = v_k$. A graph is connected if there is a path between any two vertices in V . The following definition is necessary for our study of the Grigorchuk group.

Definition 2.1: A tree is a connected graph with no cycles.

2.1 d -ary Rooted Trees

A type of tree that will be of special interest to us will be the d -ary rooted tree $T^{(d)}$, ($d > 1$). These trees can be defined as follows: The set of vertices of $T^{(d)}$ consists of finite sequences of the integers $\{0, \dots, d-1\}$. The edges of $T^{(d)}$ connect pairs of vertices whose sequences differ in length by exactly one, and the shorter sequence can be obtained from the longer by deleting the last term. For example, $v_i = (1, 0, 0, 3,)$ and $v_j = (1, 0, 0, 3, 2)$ would form an edge, whereas $v_l = (1, 0, 0, 2, 2)$ and v_i would not. The root of the tree is defined to be the empty sequence, and we shall denote it by (\emptyset) . Every vertex $(x_1, \dots, x_k) \neq (\emptyset)$ has exactly one parent, which is defined to be (x_1, \dots, x_{k-1}) (Note that the root of $T^{(d)}$ has no parent). In the d -ary tree $T^{(d)}$, every vertex (x_1, \dots, x_k) has d children, which are the vertices of the form $(x_1, \dots, x_k, x_{k+1})$ where $x_{k+1} \in \{0, \dots, d-1\}$. In the d -ary tree, we have a notion of levels, where

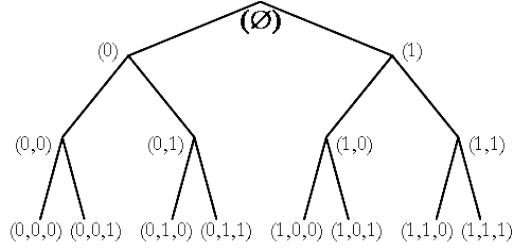


Figure 1: A binary rooted tree

the k^{th} level of $T^{(d)}$ is defined to be the subset of the vertex set V consisting of all the sequences with length k , denoted by $L^{(d)}(k)$. Thus, we can visualize the tree as having the root at the top and $L^{(d)}(k-1)$ as being above $L^{(d)}(k)$. The Giorchuk group is defined using binary rooted trees (trees in which every vertex has exactly two children, as in Figure 1), so as results are presented in the general setting of d -ary rooted trees, keep in mind that the binary case will be most important to us.

2.2 Automorphisms of $T^{(d)}$

Now that we have established the basic structure of $T^{(d)}$, we can consider functions acting on these trees. The first concept will be that of graph isomorphism.

Definition 2.2: Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$. A bijective function $f : V_1 \rightarrow V_2$ is an isomorphism of graphs G_1 and G_2 if the function f preserves the structure of adjacency, ie. for any two adjacent vertices $v_i, v_j \in G_1$, $f(v_i)$ and $f(v_j)$ are adjacent in G_2 . In the case that $G_1 = G_2$, we say that f is an automorphism.

Note that an isomorphism of a graph maps vertices to vertices and edges to edges due to the preservation of adjacency. For any vertex $x \in T^{(d)}$, there exists a subtree $T_x^{(d)}$ of $T^{(d)}$ which contains all the vertices whose sequences begin with x . By concatenation of x with a vertex $v \in T^{(d)}$, we obtain a new vertex xv in $T_x^{(d)}$. This defines an isomorphism:

$$\delta_x : T^{(d)} \rightarrow T_x^{(d)} \text{ where } v \mapsto xv$$

The automorphism group of all automorphisms on $T^{(d)}$ is denoted by $G^{(d)}$. Given any $g \in G^{(d)}$, g must fix the root. This is because the root has d neighbors, whereas every other vertex of $T^{(d)}$ has $d+1$ neighbors. So automorphisms of $T^{(d)}$ only permute the subtrees $T_0^{(d)}, \dots, T_{d-1}^{(d)}$ and fix (\emptyset) .

Definition 2.3: $St_{G^{(d)}}(k) = \{g \in G^{(d)} \mid \forall x \in L^{(d)}(k), g(x) = x\}$ is the stabilizer subgroup of $G^{(d)}$ fixing $L^{(d)}(k)$.

Suppose we are given d automorphisms $g_0, g_1, \dots, g_{d-1} \in G^{(d)}$. We can then define a new automorphism $g \in G^{(d)}$ that is in $St_{G^{(d)}}(1)$ (i.e. g fixes the first level of the tree), such that each subtree $T_j^{(d)}$ is acted upon by g_j , $j \in \{0, \dots, d-1\}$. We may formalize this notion by noticing that the action of g_j on the j^{th} subtree is precisely the automorphism $\delta_j g_j \delta_j^{-1}$, where δ_j is as defined above.

We can reverse this procedure to show that every element in the stabilizer of the first level, $St_{G^{(d)}}(1)$, is of this form. Consider an automorphism $g \in St_{G^{(d)}}(1)$. Then, by definition, g fixes the first level of $T^{(d)}$. Furthermore, g acts on each of the subtrees $T_0^{(d)}, \dots, T_{d-1}^{(d)}$ as a collection of d automorphisms $g_0, \dots, g_{d-1} \in G^{(d)}$ where each $\delta_j g_j \delta_j^{-1}$ is the restriction of g to the subtree $T_j^{(d)}$. Now we can define the following function, which is an isomorphism by the preceding discussion.

$$\begin{aligned} \psi = (\phi_0, \dots, \phi_{d-1}) : St_{G^{(d)}}(1) &\rightarrow G^{(d)} \times \dots \times G^{(d)} \\ \text{where } g &\mapsto (g_0, \dots, g_{d-1}) \end{aligned}$$

Although it is correct to write $\psi(g) = (g_0, \dots, g_{d-1})$, for $g \in St_{G^{(d)}}(1)$ we will agree to write $g = (g_0, \dots, g_{d-1})$ instead. Now that a general theory of rooted d -ary trees has been established, we can focus our attention on the binary case and on the Burnside problem.

3 The First Grigorchuk Group

It took over fifty years for a conclusive answer to Burnside's question to be discovered. This was first shown in 1964 by Golod and Shafarevich who found a finitely generated infinite p -group. Years later in 1980, Rostislav Grigorchuk published a paper containing the construction of the first Grigorchuk group. This group is of special interest to mathematicians for a number of reasons including growth rates of words, the various interesting properties exhibited by its subgroups, and the close relationship it shares with the Burnside problem. This final aspect of the Grigorchuk group is what leads us to investigate its properties.

The first Grigorchuk group is constructed out of elements in $G^{(2)}$, which are automorphisms of the binary rooted tree, $T^{(2)}$. For $j \in \{0, 1\}$, let \bar{j} be defined by $\bar{0} = 1$ and $\bar{1} = 0$. Our first automorphism on the binary rooted tree will be denoted by a and is the automorphism that simply flips the first two subtrees of $T^{(2)}$ by interchanging the vertices (0) and (1). This function can be written more precisely using the sequence notation: for some vertex (i_1, \dots, i_k) in $T^{(2)}$, $i_j \in \{0, 1\}$,

$$a(i_1, \dots, i_k) = (\bar{i}_1, i_2, \dots, i_k)$$

Note that $a^2 = 1$ since $a(a(i_1, \dots, i_k)) = a(\bar{i}_1, i_2, \dots, i_k) = (\bar{\bar{i}}_1, i_2, \dots, i_k)$ and $\bar{\bar{i}}_1 = i_1$. Also note that 1 is being used in this context to represent the identity

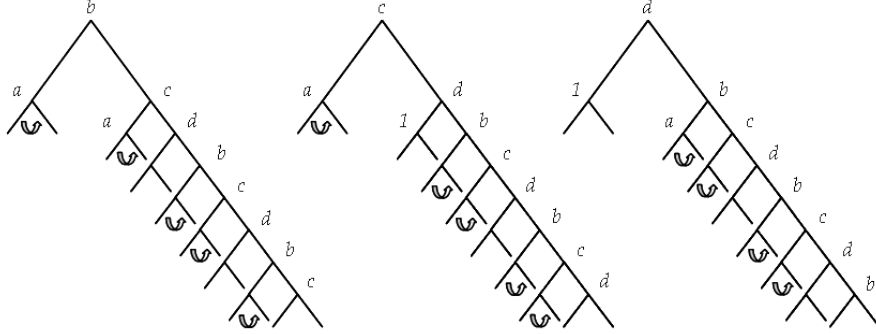


Figure 2: The automorphisms b , c , and d .

automorphism of $T^{(2)}$. Three automorphisms in $St_{G^{(2)}}(1)$ can now be defined recursively from the definition of a in the following way. Let $b = (a, c)$, $c = (a, d)$, and $d = (1, b)$. This manner of expressing b , c , and d is implicitly using the function ψ as we defined earlier. These definitions are actually formulated as $\psi(b) = (a, c)$, $\psi(c) = (a, d)$, and $\psi(d) = (1, b)$, but we will usually refer to them without reference to ψ . One can think of these automorphisms which are of the form (γ_0, γ_1) as γ_0 acting on $T_0^{(2)}$ and γ_1 acting on $T_1^{(2)}$. For example, the function b applies a to the left subtree $T_0^{(2)}$ of $T^{(2)}$ and c to the right subtree $T_1^{(2)}$. The following examples will illustrate the actions of these automorphisms on vertices of the tree.

Example: Consider the vertex $v = (1, 0, 0, 1, 1)$.

- (i.) $a(v) = a(1, 0, 0, 1, 1) = (0, 0, 0, 1, 1)$.
- (ii.) $b(v) = b(1, 0, 0, 1, 1) = (1, c(0, 0, 1, 1)) = (1, 0, a(0, 1, 1)) = (1, 0, 1, 1, 1)$.
- (iii.) $c(v) = c(1, 0, 0, 1, 1) = (1, d(0, 0, 1, 1)) = (1, 0, 1(0, 1, 1)) = (1, 0, 0, 1, 1)$.
- (iv.) $d(v) = d(1, 0, 0, 1, 1) = (1, b(0, 0, 1, 1)) = (1, 0, a(0, 1, 1)) = (1, 0, 1, 1, 1)$.

Definition 3.1: The first Grigorchuk group Γ is defined to be the group generated by $\{a, b, c, d\}$, ie., $\Gamma = \langle a, b, c, d \rangle$.

Clearly, the group Γ is finitely generated, for it has only four generators. These four automorphisms of Γ are shown in Figure 2. We can see that they only permute the subtrees directly next to the rightmost subtree on each level, and they act in a pattern of two nontrivial permutations followed by the identity. Notice that b , c , and d only differ in where they begin this pattern. The following section will provide an interesting perspective on a , b , c and d that connects the Grigorchuk group to finite state automata and may provide some insight into the nature of these recursively defined automorphisms.



Figure 3: The adding machine automaton.

3.1 Finite State Automata

The notion of an automaton comes from computer science, and roughly, it is just a type of computing machine. I will not present a detailed introduction to finite state automata, but instead this section is only designed to acquaint the reader with the main ideas that will be relevant to our discussion of the Grigorchuk group. A finite state automaton is a device that takes a string from a finite alphabet as an input into one of its finitely many states, called an initial state. Each of the states in the automaton may perform some predefined operation on the string. The operations that will be of interest to us are permutations on a term in the string that permute it with the other letters in the alphabet. The automaton accepts the input string in the initial state and looks to the next letter in the sequence. It then matches this letter to a transition defined by the same letter, which sends the string to another state. In each state, the automaton acts in the same manner, performing a permutation on the term of the string and then matching the next letter in the string to the appropriate transition until the string enters one of the final states. We picture a finite state automaton as a directed graph where each of the states is a vertex, and the transitions between states are the edges. Take, for example, the simple adding machine that can be defined using a finite state automaton. In this example, we wish to add 1 to a given number. Our alphabet will be $\{0, 1\}$ and we will express the given number in binary form. The string will be the binary representation in reverse order (eg. 6 is 110 in binary so the string will be $(0, 1, 1)$). Figure 3 shows the desired “adding machine” automaton. The permutation ϵ is defined by $\epsilon(0) = 1, \epsilon(1) = 0$. Say we input the string $(0, 1, 1)$ for the number 6. Then the automaton changes the 0 to 1 and sends the string $(1, 1, 1)$ to the final state, which represents the number $7 = 6 + 1$. Further inspection of this automaton reveals that it does indeed add one to any input value.

We can also use an automaton to model the action of the automorphisms a, b, c, d on vertices of $T^{(2)}$. This will provide another perspective from which one can understand these recursively defined functions. Again, we let ϵ flip a 0 to a 1 and vice versa. This will serve the same purpose as the flips that a performs that were described above by $a(i_1, \dots, i_k) = (\bar{i}_1, i_2, \dots, i_k)$. Figure 4 shows the corresponding automaton that emulates the actions of these four automorphisms on the vertex sequences. If we would like to apply $\gamma \in \{a, b, c, d\}$ to a given vertex v , we simply input v into the initial state labelled by γ and the output string will be the vertex $\gamma(v)$. In the given automaton, the permutations to be performed in each state are shown in the circles depicting each state. The

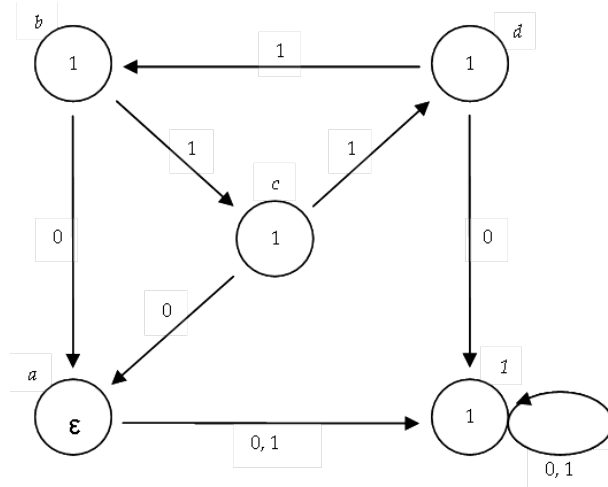


Figure 4: A finite state automaton for b , c , and d .

permutation 1 is just the identity permutation, and ϵ is as defined above.

3.2 Properties of the Grigorchuk Group

So far we have seen a great deal about how the generators of the Grigorchuk group act on the tree $T^{(2)}$, but the group structure of Γ has not been mentioned. How do the elements in the Grigorchuk group interact under the group operation? What are some of the relations between elements in the group? We have seen one such relation so far, namely $a^2 = 1$, but this is by no means the only relation. There exist similar identities for each of the generators as we will soon see in Theorem 3.1. But before this proof, a few notational customs must be stated. Recall that there exists an isomorphism $\psi : St_{G^{(2)}}(1) \rightarrow G^{(2)} \times G^{(2)}$ and that we have agreed to write, for example, $b = (a, c)$ instead of $\psi(b) = (a, c)$. The restriction of ψ to $St_\Gamma(1)$ is still a homomorphism (further properties of this restricted function will be given in Theorem 3.2). Although they are different functions, we will continue to write ψ for the restriction $\psi|_{St_\Gamma(1)}$. Now how could we write $\psi(b^2)$? Well, $\psi(b^2) = \psi(b)^2 = (a, c)^2$. Which shows that applying b twice applies (a, c) twice, and a acts on the left subtree while c acts on the right subtree. So overall, this is the same as applying a^2 to the left subtree and c^2 to the right subtree. This shows that we should define multiplication of elements in $\Gamma \times \Gamma$ coordinatewise. Thus $\psi(b^2) = (a^2, c^2)$, or with our notational convention, $b^2 = (a^2, c^2)$.

Theorem 3.1: $b^2 = c^2 = d^2 = 1$.

Proof: We know that the root of the tree is fixed under any automorphism, so it will be fixed by b^2 , c^2 , and d^2 . Since $b, c, d \in St_\Gamma(1)$, b^2 , c^2 , and d^2 will leave

the first level fixed as well. Now we will proceed by induction on the level of $T^{(2)}$. Suppose b^2 , c^2 , and d^2 all fix the first $k-1$ levels of $T^{(2)}$. We have the identities: $b^2 = (a, c)(a, c) = (a^2, c^2) = (1, c^2)$, $c^2 = (a, d)(a, d) = (a^2, d^2) = (1, d^2)$, and $d^2 = (1, b)(1, b) = (1, b^2)$. Now, any vertex on the k^{th} level containing a 0 in its sequence will be fixed since in the recursive definitions of b^2 , c^2 , and d^2 , the left subtree is always acted upon by the identity automorphism, thus fixing all the descendants thereafter, and any vertex in a left subtree must have a zero somewhere in its representative sequence. Now consider a vertex x on the k^{th} level which contains no zeros in its sequence. Since b^2 , c^2 , and d^2 all fix the first $k-1$ levels by the induction hypothesis, the only way for a vertex in $L^{(2)}(k)$ to be moved is to switch with its sibling (where siblings are two vertices that are children of the same vertex). But since the sibling of x must contain a zero in its sequence it will remain fixed. Thus, x must also remain fixed under b^2 , c^2 , and d^2 because it has nowhere to move. Therefore, the vertices on every level will remain fixed, and the theorem follows. ■

In addition to the results of Theorem 3.1, one can also establish the following relations between the different generators of Γ , which will be stated without proof:

$$bc = cb = d, bd = db = c, cd = dc = b$$

These relations show that we could have generated Γ with only three of a , b , c , and d . In fact, a and any two of b , c , and d would suffice.

Due to the relations provided by Theorem 3.1, we can see that the conjugate of $\gamma \in \{b, c, d\}$ by a is $a^{-1}\gamma a$, and $a^2 = 1$ implies that $a^{-1} = a$. Thus the conjugate of γ by a is $a\gamma a$. Also, if we had $\gamma = (\gamma_0, \gamma_1)$, and we applied $a\gamma a$ to the tree $T^{(2)}$, this would be the same as interchanging the two subtrees $T_0^{(2)}$ and $T_1^{(2)}$, then applying the automorphism γ , and then switching the subtrees back to their initial position. Thus, γ_0 acted on the right subtree, and γ_1 acted on the left subtree. So the conjugate $a\gamma a = (\gamma_1, \gamma_0)$.

We will now return to the subgroup $St_\Gamma(1) \leq \Gamma$ and characterize all of its elements. Consider some $g \in St_\Gamma(1)$. We have already stated that $b, c, d \in St_\Gamma(1)$, so we must have $\langle b, c, d \rangle \leq St_\Gamma(1)$. Now we must ask if there may be any words in $St_\Gamma(1)$ containing an a , since those are the only words not contained in $\langle b, c, d \rangle$. Consider a word containing exactly one a . Then every term in the word will fix the first level except the a , in which case it would permute the vertices (0) and (1). So any word containing one a will not be in $St_\Gamma(1)$. What if we had two a 's in the word? Then one a would switch (0) and (1) while the other a would switch them back to the original position. All of the other terms in the word would fix the first level, so the word would appear in $St_\Gamma(1)$. We can generalize this idea to classify all the elements in the stabilizer of the first level. Any word containing an odd number of a 's will permute the two vertices on the first level, and therefore not be in $St_\Gamma(1)$, whereas a word containing an even number of a 's will switch the vertices on the first level an even number of times and ultimately place (0) and (1) back into their initial positions, thus fixing the first level and placing the word in $St_\Gamma(1)$. This shows that for a word

$g \in \Gamma$, $g \in St_\Gamma(1)$ if and only if the word g contains an even number of a 's. It is now clear that we can generate $St_\Gamma(1)$ by $b = (a, c)$, $c = (a, d)$, $d = (1, b)$, $aba = (c, a)$, $aca = (d, a)$, and $ada = (b, 1)$.

Theorem 3.2: The restricted homomorphism $\psi = (\phi_0, \phi_1) : St_\Gamma(1) \rightarrow \Gamma \times \Gamma$ is injective and both of $\phi_0, \phi_1 : St_\Gamma(1) \rightarrow \Gamma$ are surjective.

Proof: The injectivity of the restricted ψ follows from the fact that the unrestricted ψ is injective. From the paragraph above, we see that $\psi(b) = (a, c)$, $\psi(c) = (a, d)$, $\psi(d) = (1, b)$, $\psi(aba) = (c, a)$, $\psi(aca) = (d, a)$, and $\psi(ada) = (b, 1)$ so that:

$$\begin{array}{llll} \phi_0(b) = a & \phi_1(b) = c & \phi_0(aba) = c & \phi_1(aba) = a \\ \phi_0(c) = a & \phi_1(c) = d & \phi_0(aca) = d & \phi_1(aca) = a \\ \phi_0(d) = 1 & \phi_1(d) = b & \phi_0(ada) = b & \phi_1(ada) = 1 \end{array}$$

Notice that a, b, c , and d are in the image of both ϕ_0 and ϕ_1 , so that, for any word $g \in \Gamma$, we can find words $g'_0, g'_1 \in St_\Gamma(1)$ such that $\phi_0(g'_0) = g = \phi_1(g'_1)$ since ϕ_0 and ϕ_1 are homomorphisms. So we see that ϕ_0 and ϕ_1 surject onto Γ . ■

A direct consequence of Theorem 3.2 is that the first Grigorchuk group, Γ , must be infinite. This is because we have found homomorphisms ϕ_0, ϕ_1 that map the proper subgroup $St_\Gamma(1) \subset \Gamma$ onto Γ itself, and the only way for this to happen is for Γ to be infinite. Now we have shown that the first Grigorchuk group possesses two of the three properties it must have to serve as a counterexample to the Burnside problem (it is finitely generated and infinite). We will now enter into the investigation of the third and final property: that Γ is periodic.

3.3 Subgroups of Γ and the Dihedral Groups

As was mentioned earlier, the Grigorchuk group is an interesting mathematical object for a number of reasons other than its relationship to the Burnside problem. One of these reasons concerns the various subgroups of the Grigorchuk group - many of which have been studied in great detail. This section will be devoted to showing the structure of a few of the finite subgroups of Γ for later use in proving the periodicity of Γ . We will show that there are three subgroups of Γ isomorphic to certain dihedral groups. To review, the dihedral group of order $2n$, denoted hereafter by D_n , is informally recognized as the group of symmetries of a regular n sided polygon. For our purposes, it will be useful to be familiar with some presentations of D_n .

$$D_n = \langle r, f \mid r^n = f^2 = 1, frf = r^{-1} \rangle = \langle x, y \mid x^2 = y^2 = (xy)^n = 1 \rangle$$

The first presentation may seem more intuitive geometrically if r is taken to be a symmetry by rotation and f is taken to be a symmetry by reflection, or a flip. Applying the rotation r n times rotates the figure completely and restores it

to its initial position, just as performing the flip twice would yield the identity. The second presentation, although not being as intuitive as the first, will prove itself to be quite useful in the following theorem.

Theorem 3.3: (i.) $\langle a, d \rangle \cong D_4$. (ii.) $\langle a, c \rangle \cong D_8$. (iii.) $\langle a, b \rangle \cong D_{16}$.

Proof: We already know that $a^2 = b^2 = c^2 = d^2 = 1$, so all that remains to be proven in each case is the relation $(a\gamma)^n = 1$ for each $\gamma \in \{b, c, d\}$ with the corresponding n . (i.) $(ad)^2 = (ad)(ad) = (ada)d = (b, 1)(1, b) = (b, b)$, and since $b^2 = 1$, $(ad)^4 = 1$. Since $(ad)^2 = (ad)(ad) = (ada)d = (b, 1)(1, b) = (b, b) = (1, b)(b, 1) = d(ada) = (da)^2$, $(ad)^4 = (da)^4 = 1$. (ii.) $(ac)^2 = (ac)(ac) = (aca)c = (d, a)(a, d) = (da, ad)$. Also, $(ca)^2 = (ca)(ca) = c(aca) = (a, d)(d, a) = (ad, da)$. With the result from (i.), we have $(ac)^8 = (ca)^8 = 1$. (iii.) $(ab)^2 = (ab)(ab) = (aba)b = (c, a)(a, c) = (ca, ac)$, so by (ii.), $(ab)^{16} = 1$. ■

3.4 Γ as a 2-Group

The previous two properties that we have shown (that Γ is finitely generated and has infinite order) were relatively easy to prove. In order to complete the verification that it is in fact a counterexample to Burnside's problem, we must now show that Γ is periodic. In fact, we will see that a stronger statement holds, that is, Γ is a 2-group. The proof of this fact is not as simple as was the proof of the infinitude of Γ . These two properties did not rely too heavily upon our particular choice of Γ and are rather common properties of groups. But the real reason why the Grigorchuk group is so special is that it is periodic while also being infinite and finitely generated. Thus, it may not come as a surprise that this will be our most technical proof yet. Without further ado, here is our final observation about the Grigorchuk group, thus completing the proof that it is finitely generated, periodic, and infinite, and that it is a counterexample to the Burside problem.

Theorem 3.4: Γ is a 2-group. That is, $\forall g \in \Gamma, \exists n \in \mathbb{Z}, n \geq 0$ such that $g^{2^n} = 1$.

Proof: Given some $g \in \Gamma$, we will let k be the length of the reduced word representing g . If the length of g is $k = 0$, then g is the empty word, so $g = 1$ is the identity automorphism. So, in this case, $g^{2^0} = g = 1$. If the length of g is $k = 1$, then $g \in \{a, b, c, d\}$, so $g^{2^1} = g^2 = 1$ by Theorem 3.1. In the case that $k = 2$, there exist some $g_1, g_2 \in \{a, b, c, d\}$ such that $g = g_1 g_2$. We may assume that $g_1 \neq g_2$, for if they were then $g = g_1^2 = 1$, and g would then be a word of length 0. We may also assume that one of g_1 and g_2 is a , then by Theorem 3.2, $g^{16} = g^{2^4} = 1$ and the order of g must divide $16 = 2^4$, ie., $ord(g)$ must be a power of 2. This assumption is justified because if both $g_1 \neq a$ and $g_2 \neq a$, then using the reductions $bc = cb = d$, $cd = dc = b$, and $bd = db = c$ from Theorem 3.3, we see that the word representing g would be of length 1 and $g = g_1 g_2$ would not have been reduced, contrary to our initial assumption. So all words of length $k \leq 2$ have order a power of two. Now we will proceed by induction on k , the length of g . Assume that $k \geq 3$ and that the theorem holds

for all elements of Γ represented by reduced words of length $k - 1$ and shorter. The proof now breaks into a number of cases. First, suppose that k is odd. We know that g can contain no two consecutive terms both of which are not equal to a since we have the relations $bc = cb = d$, $cd = dc = b$, and $bd = db = c$. Thus g must consist of an alternating sequence of a 's and elements from $\{b, c, d\}$. Since k is odd, g must either begin and end with a or both the first and last terms in the word are not a . In the first case, $g = au_1au_2a \cdots au_na$, where the $u_i \in \{b, c, d\}$. Then g is conjugate to aga which is a word of length $k - 2$. In the second case, $g = v_1av_2a \cdots av_ma$, where the $v_i \in \{b, c, d\}$. Then g is conjugate to v_1gv_1 , and this is a word of length $k - 1$ or $k - 2$. Since any element in a group has the same order as its conjugate, it follows that if k is odd, g has the same order as a reduced word in Γ of length $k - 1$ or smaller, and by the induction hypothesis we are done.

Now, suppose that k is even, so that $\exists l \in \mathbb{Z}$ such that $k = 2l$. I claim that g has the same order as a word of the form $au_1 \cdots au_l$ with $u_i \in \{b, c, d\}$. This is because we have two possibilities: 1. $g = au_1 \cdots au_l$ in which case the claim follows automatically, and 2. $g = u_1a \cdots u_la$ in which case we may obtain the desired word by replacing g by its conjugate u_1gu_1 . Suppose first that l is even and set $w = au_1 \cdots au_l$. Then by the claim we see that w has the same order as g . Then we can reassociate the terms of w as follows $w = au_1 \cdots au_l = (au_1a)u_2 \cdots (au_{l-1}a)u_l$. We see that we must have an even number of a 's appearing in w , and from our classification of the elements of $St_\Gamma(1)$, we see that $w \in St_\Gamma(1)$. Then if we apply the isomorphism ψ to w we obtain:

$$\psi(w) = \psi(au_1a)\psi(u_2) \cdots \psi(au_{l-1}a)\psi(u_l) = (w_0, w_1)$$

for some $w_0, w_1 \in \Gamma$. Notice that w_0 and w_1 can be represented by reduced words of at most $l = \frac{k}{2}$ since in the above form they are of length l and we may be able to make further reductions. So by the induction hypothesis, there exist integers $m, n \geq 0$ such that the $ord(w_0) = 2^m$ and $ord(w_1) = 2^n$. I claim that the order of w must be the least common multiple of the orders of w_0 and w_1 . Let $lcm(ord(w_0), ord(w_1)) = x$. Then $\psi(w^x) = \psi(w)^x = (w_0, w_1)^x = (w_0^x, w_1^x) = (1, 1)$, and since ψ is injective, $w^x = 1$. Conversely, suppose that $ord(w) = y$. Then $w^y = 1$ and $(1, 1) = \psi(1) = \psi(w^y) = \psi(w)^y = (w_0, w_1)^y = (w_0^y, w_1^y)$. Then $w_0^y = 1$ and $w_1^y = 1$. So, $ord(w_0) \mid y$ and $ord(w_1) \mid y$. Since x is the smallest number that both $ord(w_0)$ and $ord(w_1)$ divide, it must be the case that $x \leq y$. From the argument above, $w^x = 1$, so we must have $y = x = lcm(ord(w_0), ord(w_1))$ since $y = ord(w)$ is the smallest number such that $w^y = 1$. Thus, $ord(g) = ord(w) = lcm(ord(w_0), ord(w_1)) = lcm(2^m, 2^n)$ which is a power of 2.

Now, suppose $l = \frac{k}{2}$ is odd. Again, we may assume that g has the same order as the word $w = au_1 \cdots au_l$, but $w \notin St_\Gamma(1)$ because there are a total of l a 's appearing in w , which is an odd number. Although we do not have $w \in St_\Gamma(1)$, we do have $w^2 = (au_1 \cdots au_l)(au_1 \cdots au_l) \in St_\Gamma(1)$. Now, after reassociation, we obtain $w^2 = (au_1a)u_2 \cdots u_{l-1}(au_la)u_1(au_2a) \cdots (au_{l-1}a)u_l$ and so,

$$\psi(w^2) = \psi(au_1a)\psi(u_2) \cdots \psi(u_{l-1})\psi(au_la)$$

$$\begin{aligned} & \psi(u_1)\psi(au_2a)\cdots\psi(au_{l-1}a)\psi(u_l) \\ &= (\alpha, \beta) \end{aligned}$$

where $\alpha, \beta \in \Gamma$ and both are words of length at most k . There are now three possible subcases:

(i.) There is some $j \in \{1, \dots, l\}$ such that $u_j = d$. Since both $u_j = d$ and its conjugate $au_ja = ada$ appear in w^2 , we know that both $\psi(u_j) = \psi(d) = (1, b)$ and $\psi(au_ja) = \psi(ada) = (b, 1)$ appear in $\psi(w^2)$ so that the reduced words α and β will have length at most $k - 1$. Then, by the induction hypothesis, $\exists m, n \in \mathbb{Z}$ such that $\text{ord}(\alpha) = 2^m$ and $\text{ord}(\beta) = 2^n$. Thus, by the argument given above, $\text{ord}(w^2)$ is the least common multiple of 2^m and 2^n , and so it is a power of two. This shows that $(w^2)^{\text{lcm}(2^m, 2^n)} = w^{2 \cdot \text{lcm}(2^m, 2^n)}$ and so $\text{ord}(w) \mid 2 \cdot \text{lcm}(2^m, 2^n)$ and is a power of two as well.

(ii.) There is some $j \in \{1, \dots, l\}$ such that $u_j = c$. Since both u_j and its conjugate au_ja appear in w^2 , we have both $\psi(u_j) = \psi(c) = (a, d)$ and $\psi(au_ja) = \psi(aca) = (d, a)$ in $\psi(w^2)$. Then, either both α and β are words containing d of length k and we may apply the result from (i.), or α and β have reduced in length and we may apply the induction hypothesis directly to them and repeat the least common multiple argument to show that $\text{ord}(w)$ is a power of two.

(iii.) There is some $j \in \{1, \dots, l\}$ such that $u_j = b$. In this case, we have both $u_j = b$ and $au_ja = aba$ so that $\psi(u_j) = \psi(b) = (a, c)$ and $\psi(au_ja) = \psi(aba) = (c, a)$ both appear in $\psi(w^2)$. Then, we either 1. apply the result from (ii.) if α or β still have length k (since now they will both have c 's) or 2. α and β have reduced in length and we may apply the induction hypothesis.

Thus, it has been shown in every case that the order of w is power of two, and since $\text{ord}(w) = \text{ord}(g)$, the theorem follows. ■

4 A Brief History of the Burnside Problem

In the time that it took to find a counterexample to Burnside's question, which asks if all finitely generated, periodic groups are finite, a great interest in a number of related problems arose. Upon failing to answer his own question, Burnside turned to the simpler one: *Is it the case that all finitely generated groups of bounded exponent are finite?* Here, a group G has finite exponent if $\exists n \in \mathbb{Z}$ such that $\forall g \in G, g^n = 1$. Having bounded exponent differs from being periodic in that periodic means each element in the group has its own integer power equal to the identity element, whereas bounded exponent means that there exists a single integer such that every element raised to its power is the identity element. This second question raised a great interest among mathematicians, but the world had to wait until 1968 for Adian and Novikov to provide the first major result, which was only partial. They showed that for a bounded exponent of $n \geq 4381$ there always exist counterexamples. Later, in 1975, Adian proved a tighter bound, showing that there existed counterexamples for all $n \geq 665$. The final restriction of this bound was recently proven by Sergei V. Ivanov in 1992, which showed that the conjecture was false for $n \geq 13$.

Meanwhile, there was another related problem being investigated by Burnside that evaded a definitive answer for more than 90 years. This problem asks if you fix $m, n \in \mathbb{N}$ are there are only finitely many groups generated by m elements of bounded exponent n ? Many of the results that were found regarding the previous question on finitely generated groups of bounded exponent came from the investigation of this new problem, which is referred to as the Restricted Burnside Problem. A number of small results concerning the Restricted Burnside Problem surfaced throughout the twentieth century, but it wasn't until 1994 that a definitive answer was found. Efim Zelmanov was awarded the Fields medal for his affirmative answer to the Restricted Burnside Problem, showing that there are only finitely many groups generated by m elements and of bounded exponent n . It may seem that all of Burnside's questions have now been resolved, thanks to counterexamples like the first Grigorchuk group and Zelmanov's solution of the restricted problem, but these few results are by no means exhaustive. There are a number of other problems that have since been investigated, most of which apply various finiteness restrictions to the Burnside problems. In fact, there are still some related problems that remain unsolved - for example: *Do there exist infinite groups on 2 generators of bounded exponent 5?* - but we will again have to wait for future researchers of the Burnside problem to provide the answers.

References

de la Harpe, Pierre. *Topics Geometric Group Theory*, The University of Chicago Press, (2000)

O'Connor, J. J., E. F. Robertson, and A. Isherwood, *A history of the Burnside problem*, http://www-history.mcs.st-andrews.ac.uk/HistTopics/Burnside_problem.html, (2002)

O'Connor, J. J. and E. F. Robertson, *William Burnside*, <http://www-history.mcs.st-andrews.ac.uk/Biographies/Burnside.html>, University of St. Andrews, (2005)