eman ta zabal zazu
UPV EHU
ZIENTZIA ETA TEKNOLOGIA FAKULTATEA
FACULTAD DE CIENCIA Y TECNOLOGÍA
50 URTE AÑOS 1968 - 2018
Biba Zientzia! Ciencia Viva

The General Burnside Problem

Final Degree Dissertation
Degree in Mathematics

Mikel Martínez Puente

Supervisor:
Gustavo A. Fernández Alcober

Leioa, 22 June 2022

Contents

Introduction

Content of the work

The Burnside problems are among the most important problems in group theory in the 20th century. In this project, we will focus on the General Burnside Problem, which asks whether a finitely generated periodic group is necessarily finite, for which the answer is negative. It was proposed by William Burnside in 1902 [1] and it has been subject of study all over the 20th century. In fact, it is considered one of the oldest and most influential questions in group theory. Five years before, in 1897, he wrote the book called "Theory of groups of finite order" [2], which was regarded for several decades as the standard introduction to group theory.

William Burnside (1852-1927) wrote the first dissertation about groups in English and he was the first to develop the theory of groups from a modern abstract point of view. Burnside's contributions to group theory and to the study of group representations are fundamental to the subject. Ironically, the most popular result by which he is often known is an elementary counting lemma erroneously known as Burnside's lemma which is not due to him, although he quotes it in his book attributing it instead to Frobenius.

The main purpose is that someone who is not familiar with group theory learns enough of it in order to understand advanced results. However, it is assumed that the reader is comfortable with basic group theory concepts.

The notes are organized in three chapters. In the first chapter the reader is introduced to commutator theory, which will be useful to define and work with nilpotent and soluble groups, for which the answer to the General Burnside Problem is affirmative. Then, we also study the problem for linear groups, for which the answer is also affirmative.

In the second and third chapters some negative solutions to the General Burnside Problem are introduced. In the second chapter, Golod-Shafarevich groups are constructed using formal power series and polynomials in non-commuting indeterminates. In the third chapter, we introduce Gupta-Sidki and Grigorchuk groups, using graph theory and automorphisms of trees.

Burnside problems

Let us now explain what is the General Burnside Problem about and two similar and less restrictive problems that were proposed after it: the Burnside Problem and the Restricted Burnside Problem.

Observe that every finite group is finitely generated and periodic. What about the other implication? Regarding this question, in 1902 William Burnside introduced what he termed "a still undetermined point" known as the General Burnside Problem, which asks whether a finitely generated periodic (or torsion) group is necessarily finite. This question was answered in the negative in 1964 by Evgeny Golod and Igor Shafarevich, who gave a counterexample of an infinite p-group that is finitely generated, which we will introduce in Chapter 2.

The reason for not having an answer to the problem until 1964 is that the requirements of being finitely generated and periodic give very little information about the possible structure of a group. Due to this difficulty, Burnside immediately suggested a weaker formulation of the General Burnside Problem known as the Burnside Problem, which asks whether a finitely generated group of bounded exponent (Definition 1.9) is necessarily finite.

Let us now introduce the concept of the free Burnside group in order to reformulate the question. The free Burnside group of rank m and exponent n, denoted by $B(m,n)$, is a group with m distinguished generators $x_1, ..., x_m$ for which $x^n = 1$ holds for all elements x of the group, and which is the "largest" group satisfying these requirements. What we mean with the "largest" group is that given any group G with m generators $g_1, ..., g_m$ and of exponent n, there is a unique homomorphism from $B(m,n)$ to G that maps the i-th generator x_i of $B(m,n)$ to the i-th generator g_i of G.

Another way of constructing it is by the quotient $B(m,n) = F_m/F_m^n$ where F is the free group of m generators, that is, the free group of rank m. A group is called a free group if no relation exists between its group generators other than the relationship between an element and its inverse required as one of the defining properties of a group. Elements consist of all words that can be built from these generators.

These definitions then lead to an alternate and more popular formulation of the Burnside Problem: for which positive integers m, n is $B(m,n)$ finite? The full solution to Burnside Problem in this form is not known, since our present state of knowledge of this problem is very incomplete. However, there are some simple cases where the answer is affirmative.

Burnside showed a number of easy results in his 1902 original paper. Among them, we have that $B(1,n)$, which is the cyclic group of order n, and the 2-torsion (abelian) group $B(m,2)$ are both finite. Actually, $B(m,2)$ is the direct product of m copies of C_2.

Moreover, Burnside also showed that $B(m,3)$ and $B(2,4)$ are finite and he gave an upper bound of their orders. It was not until 1940 when Sanov proved that $B(m,4)$ is finite, whose order is known only for $m \leq 5$, which is $2^{12}, 2^{69}, 2^{422}$ and 2^{2728} for $m = 2, 3, 4$ and 5, respectively. An affirmative answer was given for the case $n = 6$ by M. Hall in 1958, whose proof is much harder than the previous ones. At present, no other values of n are known for which $B(m,n)$ is finite and it is still an open question whether $B(2,5)$ is finite or not.

In 1968 Pyotr Novikov and Sergei Adian found a counterexample to the Burnside Problem proving that $B(m,n)$ is infinite for all odd exponents $n \geq 4381$. This bound on the odd exponent was later improved to 665 by Adian himself in 1975, and there have been many improvements since then in terms of even and odd exponents such that $B(m,n)$ is infinite. In 1980 Alexander Yu. Ol'shanskii constructed the so-called Tarski monsters, which are finitely generated infinite groups such that every nontrivial proper subgroup is a finite cyclic group of order a fixed prime number p. He proved that there is a Tarski p-group for every prime $p > 10^{75}$. They form a famous class of counterexamples to the Burnside Problem.

In the early 1930s, the topic was resurrected by the suggestion of a variant on the original problem known as the Restricted Burnside Problem, which asks whether there is a bound for the orders of all m-generated finite groups of exponent n, where this bound depends on m and n. In other words, it asks whether for fixed positive integers m and n there are only finitely many (up to isomorphism) finite groups with m generators and bounded exponent n.

It was not until 1990 when at the age of 34, Efim Zelmanov solved this other problem in the affirmative. He was awarded a Fields Medal for this work in 1994, which in the absence of a Nobel Prize in mathematics, is regarded as the highest professional honour a mathematician can attain. It is given to the most distinguished mathematicians aged 40 or under.

Motivation for doing this work

I came up with the idea of doing my Bachelor's Thesis about group theory once I followed the courses of Commutative Algebra and especially Algebraic Equations with Gustavo Fernández last year. I had always had an special interest in algebra more than in any other branch of mathematics, so I asked him to be my supervisor and he agreed and proposed me this problem. I agreed with him because it tied really well with the idea I initially had, since I was keen on choosing a specific problem related to group theory.

Personal work and acknowledgments

As a general rule, I have studied the material provided from my supervisor's lecture notes of PhD courses, textbooks and also from research papers. Then, I have elaborated my own version which I have complemented with the solution of a set of selected problems.

I want to acknowledge Gustavo Fernández for his many corrections and explanations. Not only did he accept when I asked him whether he could be my supervisor, but he proposed me this problem and trusted in me. He provided me notes about commutator theory, Gupta-Sidki groups and Golod-Shafarevich groups, which have been very useful for me.

I also want to acknowledge the professors from UP Famnit (University of Primorska, Slovenia) for helping me when I was on Erasmus exchange program during the first semester.

Chapter 1

Positive answers to the General Burnside Problem

In this first chapter we provide some positive solutions to the General Burnside Problem. Four of them are addressed: abelian, nilpotent, soluble and linear groups. Before that, we are going to make a brief introduction about commutator theory in order to introduce nilpotent and soluble groups.

All finite groups share some properties such as being finitely generated and periodic. But, are they enough to imply that a group is finite? What makes a group finite? This is exactly what William Burnside wanted to find, which properties are enough to conclude that a group is finite.

In some cases, it suffices to ask for the group to be finitely generated such that the order of the generators is finite, such as for abelian and nilpotent groups. However, for soluble groups this is not enough and we need the orders of all elements to be finite, not just the order of the generators. As a counterexample, the infinite dihedral group can be generated by two elements of order two, but it has an element of infinite order so it does not satisfy the conditions of the General Burnside Problem and it cannot be considered as a negative solution to it.

The whole chapter is mainly based on notes provided by my supervisor [5] and on Derek Robinson's book [13, Chapter 5].

1.1 Commutator theory

In this section the reader is introduced to some basic knowledge about commutator theory, which is an important part of group theory and a very useful tool in order to define and work with nilpotent and soluble groups in Sections 1.2 and 1.3, respectively.

Definition 1.1. Let G be a group and $x, y \in G$. Then, the commutator of x and y is defined as $[x, y] = x^{-1}y^{-1}xy$.

Observe that $[x, y] = x^{-1}x^y$, or equivalently, $x^y = x \cdot [x, y]$.

Definition 1.2. Let G be a group. Then, $[G, G] = \langle [x, y] \mid x, y \in G \rangle$ is known as its commutator or derived subgroup and we denote it by G'.

As we will see in Theorem 1.7, G' is the smallest normal subgroup of G giving abelian quotient, that is, G/G' is the largest abelian quotient of G. Two elements x and y commute if and only if $[x, y] = 1$, hence G is abelian if and only if $G' = \{1\}$.

There is a left-norm convention so that $[x_1, ..., x_i] = [[x_1, ..., x_{i-1}], x_i]$ is recursively defined, which may appear in some properties of commutators.

Proposition 1.3. *Let G be a group, let $x, y, g \in G$ and $n \in \mathbb{N}$. Then:*

(i) $[x, y]^{-1} = [y, x]$;

(ii) $[x, y]^g = [x^g, y^g] = [x, y][x, y, g]$;

(iii) *If y and $[x, y]$ commute, then $[x, y]^n = [x, y^n]$.*

Proof. Let us start proving the first property, which is trivial since we know that $[x, y]^{-1} = (x^{-1}y^{-1}xy)^{-1} = y^{-1}x^{-1}yx = [y, x]$. For the second property, $[x, y]^g = (x^{-1}x^y)^g = (x^{-1})^g x^{yg}$ and since $(x^{-1})^g = (x^g)^{-1}$ and $yg = gy^g$, then $[x, y]^g = (x^g)^{-1}x^{gy^g} = [x^g, y^g]$. On the other hand, we also know that $x^y = x \cdot [x, y]$, hence $[x, y]^g = [x, y][[x, y], g] = [x, y][x, y, g]$.

Let us prove the third property by induction on n. For the base case $n = 1$ it is trivial, so we assume it is true up to $n - 1$ and let us prove it for n. By induction hypothesis we get that $[x, y]^{n-1} = [x, y^{n-1}]$ and since y and $[x, y]$ commute, then:

$$\begin{aligned}[x, y]^n &= [x, y][x, y]^{n+1} = (x^{-1}y^{-1}xy)[x, y]^{n-1} = x^{-1}y^{-1}x[x, y]^{n-1}y \\ &= x^{-1}y^{-1}x(x^{-1}y^{-n+1}xy^{n-1})y = x^{-1}y^{-n}xy^n = [x, y^n].\end{aligned}$$

□

1.2 Abelian and nilpotent groups

The aim of this section is to show that the answer to the General Burnside Problem is positive for abelian and more generally nilpotent groups. In order to define nilpotent groups we need to introduce the lower central series using commutator theory.

It is obvious that abelian groups are positive answers to the General Burnside Problem. Let G be an abelian group with d generators $g_1, ..., g_d$, being d finite and $o(g_j) = n_j < \infty$. Then, we know that

$$G = \langle g_1, ..., g_d \rangle = \{g_1^{i_1} \cdots g_d^{i_d} \mid 0 \leq i_j < n_j\}.$$

Thus, the cardinality of an abelian group with d generators of finite order is bounded by the product of the orders of the generators: $|G| \leq n_1 n_2 \cdots n_d < \infty$, hence G is finite.

Let us now define the lower central series of a group and let us use it in order to define nilpotent groups.

Definition 1.4. Let $G = \gamma_1(G)$ be a group and let $\gamma_{i+1}(G) = [\gamma_i(G), G]$ for $i \geq 1$. Then, $\{\gamma_i(G)\}_{i \in \mathbb{N}}$ is called the lower central series (LCS) of G.

The lower central series of a group is descending, that is, $\gamma_{i+1}(G) \leq \gamma_i(G)$ for all $i \geq 1$. A natural question would be whether the series reaches the trivial subgroup $\{1\}$ or not, which depends on the group.

Definition 1.5. Let G be a group. We say that G is nilpotent if there exists some $n \in \mathbb{N}$ such that $\gamma_n(G) = \{1\}$, that is, its lower central series reaches $\{1\}$.

If G is nilpotent, the lenght of its LCS is called the nilpotency class of G, which is denoted by c. It is indeed the smallest positive integer such that $\gamma_{c+1}(G) = \{1\}$. Now, we are interested in the properties of $\gamma_i(G)$.

Proposition 1.6. *Let G be a group and $H, K \trianglelefteq G$. Then, $[H, K] \trianglelefteq G$.*

Proof. Let $x \in H$, $y \in K$ and $g \in G$. Then, if we conjugate $[x, y] \in [H, K]$ by g we get $[x, y]^g = [x^g, y^g]$. Since $H, K \trianglelefteq G$ it follows that $[x, y]^g \in [H, K]$, hence $[H, K] \trianglelefteq G$. □

Theorem 1.7. *Let G be a group, let G' be its commutator subgroup and let also $N \trianglelefteq G$. Then, G/N is abelian if and only if $G' \leq N$.*

Proof.

$$\begin{aligned} G/N \text{ abelian} &\iff [\overline{x}, \overline{y}] = \overline{1}, \forall \overline{x}, \overline{y} \in G/N \iff \overline{[x, y]} = \overline{1}, \forall x, y \in G \\ &\iff [x, y] \in N, \forall x, y \in G \iff G' \leq N. \end{aligned}$$

□

These two results give us important information about each $\gamma_i(G)$. On the one hand, it is obvious that $G = \gamma_1(G)$ is normal so recursively we deduce that $\gamma_{i+1}(G) = [\gamma_i(G), G] \trianglelefteq G$ for all $i \geq 1$. Moreover, we also know that

$(\gamma_i(G))' = [\gamma_i(G), \gamma_i(G)] \leq [\gamma_i(G), G] = \gamma_{i+1}(G)$, hence $\gamma_i(G)/\gamma_{i+1}(G)$ is abelian for all $i \geq 1$.

The following theorem gives a sufficient condition for $\gamma_i(G)/\gamma_{i+1}(G)$ to be finitely generated.

Theorem 1.8. *Let $G = \langle X \rangle$ and $\gamma_{i-1}(G) = \langle Y, \gamma_i(G) \rangle$. Then:*

$$\gamma_i(G) = \langle [x, y], \gamma_{i+1}(G) \mid x \in X, y \in Y \rangle.$$

Proof. One inclusion is obvious since $\gamma_{i+1}(G) \subseteq \gamma_i(G)$ and $[x, y] \in \gamma_i(G)$ for all $x \in G$, $y \in \gamma_{i-1}(G)$. In order to prove the other inclusion, we set $N = \langle [x, y], \gamma_{i+1}(G) \mid x \in X, y \in Y \rangle$. First of all, let us prove that $N \trianglelefteq G$. We already know that $\gamma_{i+1}(G) \trianglelefteq G$, so it suffices to check that $[x, y]^g \in N$ for all $x \in X$, $y \in Y$ and $g \in G$. By Proposition 1.3 we already know that $[x, y]^g = [x, y][x, y, g]$, hence $[x, y]^g \in N$ since $[x, y, g] \in \gamma_{i+1}(G)$.

Therefore, $N \trianglelefteq G$ and we can factor out N getting the quotient group G/N. We are going to prove now that generators of $\gamma_{i-1}(G)$ commute with generators of G in G/N. Let $x \in X$. Then, if $y \in Y$ it is trivial that $[x, y] \in N$, whereas if $y \in \gamma_i(G)$, then $[x, y] \in \gamma_{i+1}(G) \subseteq N$. In both cases $[\overline{x}, \overline{y}] = \overline{1}$ in G/N, hence $\overline{x}$ and $\overline{y}$ commute and $\gamma_i(G/N) = [\gamma_{i-1}(G/N), G/N] = \{\overline{1}\}$. In particular, $\gamma_i(G/N)$ is a subgroup of G/N, which can be written on the form $\gamma_i(G) \cdot N/N$. Then, $\gamma_i(G) \cdot N/N = \{\overline{1}\} = N/N$, hence $\gamma_i(G) \leq N$ and the proof is completed. □

Equivalently, if $G = \langle X \rangle$ and $\gamma_{i-1}(G)/\gamma_i(G) = \langle \overline{Y} \rangle$, then $\gamma_i(G)/\gamma_{i+1}(G)$ is equal to $\langle \overline{[x, y]} \mid x \in X, y \in Y \rangle$. Therefore, if G and $\gamma_{i-1}(G)/\gamma_i(G)$ are finitely generated, then $\gamma_i(G)$ is not necessarily finitely generated but $\gamma_i(G)/\gamma_{i+1}(G)$ is. In general, for any finitely generated group G and a subgroup H it suffices to ensure $|G : H| < \infty$ in order to conclude H is also finitely generated (see Lemma 1.15).

We have a bound for the number of generators of $\gamma_i(G)/\gamma_{i+1}(G)$ so let us see what happens with the order of its elements. Before that, let us recall what is the exponent of a group.

Definition 1.9. Let G be a group whose elements have finite order (periodic group) in which there is a bound for all these orders. The exponent of G, which is denoted by $\exp G$, is the smallest number n such that $g^n = 1$ for all $g \in G$, i.e., $n = \operatorname{lcm}(o(g) \mid g \in G)$.

The existence of such number n means there is a finite number of different orders for all the elements of G, although these orders could be repeated and G be infinite. In this case, we say that G is a group of finite exponent n or an n-torsion group. This implies that the order of every element in G is finite, i.e., G is periodic.

We want to see if there is a relationship between the exponents of the different quotients $\gamma_i(G)/\gamma_{i+1}(G)$ for all $i \geq 1$.

Theorem 1.10. *Let G be a group and $\{\gamma_i(G)\}_{i\in\mathbb{N}}$ its lower central series. If $\exp(G/G')$ is finite, then:*

$$\exp(\gamma_i(G)/\gamma_{i+1}(G)) \quad | \quad \exp(\gamma_{i-1}(G)/\gamma_i(G)) \quad | \quad \cdots \quad | \quad \exp(G/G').$$

Proof. Without loss of generality we assume that $\gamma_{i+1}(G) = \{1\}$ and since $\gamma_i(G)/\gamma_{i+1}(G)$ is abelian, then $\gamma_i(G)$ is abelian in this particular case. We also know by Theorem 1.8 that generators of $\gamma_i(G)$ are of the form $[x, y]$ where $x \in G$ and $y \in \gamma_{i-1}(G) \subseteq G$ so that $[[x, y], y] \in \gamma_{i+1}(G) = \{1\}$, hence y and $[x, y]$ commute.

Let $n = \exp(\gamma_{i-1}(G)/\gamma_i(G))$, then by Proposition 1.3 we know that $[x, y]^n = [x, y^n]$. Since $y \in \gamma_{i-1}(G)$, then $y^n \in \gamma_i(G)$ and $[x, y]^n = [x, y^n] = 1$. All in all, we have that $\gamma_i(G)$ is abelian and all its generators $[x, y]$ have finite order which divides $n = \exp(\gamma_{i-1}(G)/\gamma_i(G))$, hence the exponent of $\gamma_i(G)$ is finite and also divides n. Therefore, $\gamma_i(G)/\gamma_{i+1}(G)$ is a quotient group of finite exponent, a divisor of n. Applying this result recursively for all i, the proof of the theorem is completed. □

This theorem is very important since it implies that if the exponent of G/G' is finite, then the exponent of each $\gamma_{i-1}(G)/\gamma_i(G)$ is also finite. Let us use this result in order to prove that nilpotent groups are positive solutions to the General Burnside Problem.

Theorem 1.11. *Let G be a finitely generated nilpotent group such that the generators have finite order. Then, G is finite.*

Proof. Let us prove by induction that $\gamma_i(G)/\gamma_{i+1}(G)$ is finite for all $i \geq 1$, hence

$$|G| = |G/\gamma_2(G)| \cdot |\gamma_2(G)/\gamma_3(G)| \cdots |\gamma_c(G)/\gamma_{c+1}(G)| < \infty$$

where c is the nilpotency class of G. Since these quotients are abelian, it suffices to check that each of them is finitely generated with generators of finite order, or equivalently, it is a finitely generated group of finite exponent.

For the base case $i = 1$, the proof is almost trivial. Since G is a finitely generated group such that the generators have finite order, then so is G/G' and since it is abelian G/G' is finite. In addition, the exponent of G/G' is finite. Now, we assume it is true up to $i - 1$ and let us prove it for i.

On the one hand, we know from Theorem 1.10 that $\exp(\gamma_i(G)/\gamma_{i+1}(G))$ divides $\exp(G/G')$, which is finite, hence the exponent of $\gamma_i(G)/\gamma_{i+1}(G)$ also is. On the other hand, we also know from Theorem 1.8 that if G and $\gamma_{i-1}(G)/\gamma_i(G)$ are finitely generated, which is true by induction hypothesis, then $\gamma_i(G)/\gamma_{i+1}(G)$ is also finitely generated. □

Although in this work we deal with general nilpotent groups, there is an important characterization for the finite case, so that a finite group G is nilpotent if and only if it is the direct product of its Sylow subgroups, which are p-groups [13, page 130]. In particular, every finite p-group is nilpotent for a prime p, which is proved in Problem 1.

1.3 Soluble groups

In this section our aim is to prove that the answer to the General Burnside Problem is positive for soluble groups. Moreover, we will also introduce the infinite dihedral group, which is an infinite finitely generated soluble group that can be generated by some generators of finite order, but since it is not periodic it cannot be regarded as a negative solution to the General Burnside Problem.

Now, we need to introduce the derived series of a group in order to characterize soluble groups. Let us first define what a soluble group is as we saw in the third year course of Algebraic Equations.

Definition 1.12. Let G be a group. We say G is soluble or solvable if there exists a series of subgroups $\{1\} = N_k \trianglelefteq N_{k-1} \trianglelefteq \cdots \trianglelefteq N_1 \trianglelefteq N_0 = G$, such that N_i/N_{i+1} is abelian for all $i = 0, ..., k-1$.

Let us now introduce the derived series of a group G, which is the fastest descending series of G such that all successive quotients are abelian.

Definition 1.13. Let G' be the commutator subgroup of G and let us define $G^{(i+1)} = (G^{(i)})' = [G^{(i)}, G^{(i)}]$ for all $i \geq 1$, which are known as derived subgroups. Then, we get a descending series $G = G^{(0)} \trianglerighteq G^{(1)} \trianglerighteq G^{(2)} \trianglerighteq \cdots$, where $G^{(i)}/G^{(i+1)}$ is abelian for all $i \geq 0$. We call it the derived series of G.

Let us now characterize soluble groups once defined the derived series of a group.

Theorem 1.14. *A group G is soluble if and only if its derived series reaches the trivial subgroup $\{1\}$.*

Proof. On the one hand, if there exists some $k \in \mathbb{N}$ such that $G^{(k)} = \{1\}$, then there exists a series of subgroups $\{1\} = N_k \trianglelefteq N_{k-1} \trianglelefteq \cdots \trianglelefteq N_1 \trianglelefteq N_0 = G$, such that N_i/N_{i+1} is abelian for all $i = 0, ..., k-1$, taking $N_i = G^{(i)}$ for all $i = 0, ..., k$. Thus, G is soluble.

On the other hand, let us suppose G is soluble. Since the derived subgroup is the smallest normal subgroup giving abelian quotient, then $G' \leq N_1$ and recursively we get that $G^{(i)} \leq N_i$ for all $i = 0, ..., k$. But since $G^{(k)} \leq N_k = \{1\}$, then $G^{(k)} = \{1\}$. □

If G is soluble, then the first k for which $G^{(k)} = \{1\}$ is called the derived length or solvable index of G. In order to show that the answer to the General Burnside Problem is positive for soluble groups, we need the following auxiliary lemma, whose proof is left as a problem (see Problem 3).

Lemma 1.15. *Let G be a finitely generated group and let H be a subgroup of G of finite index. Then, H is also finitely generated.*

Theorem 1.16. *Let G be a finitely generated soluble periodic group. Then, G is finite.*

Proof. We are going to prove it by induction on n, the derived length of the soluble group G. For $n = 1$ the result is true, since $G^{(1)} = G' = \{1\}$ implies G is abelian, hence G is finite. Now, we assume it true up to length $n-1$ and let us prove it for n.

First of all, G/G' is finitely generated, periodic and abelian, hence G/G' is finite. Moreover, by Lemma 1.15 we know that since $|G : G'| < \infty$, then G' is finitely generated. Now, G' is soluble of derived length $n-1$, periodic (since G is periodic) and finitely generated, then by induction hypothesis G' is finite. Thus, $|G| = |G : G'| \cdot |G'| < \infty$. □

Observe that just like abelian groups are nilpotent, nilpotent groups are particular examples of soluble groups, since all conditions are fulfilled taking the corresponding LCS. Normality of $\gamma_{i+1}(G)$ over G trivially implies normality over smaller subgroups, i.e., $\gamma_{i+1}(G) \trianglelefteq \gamma_i(G)$. Secondly, we had also proved in Section 1.2 that each of these quotient groups $\gamma_i(G)/\gamma_{i+1}(G)$ is abelian.

However, the other implication is not true. Not all soluble groups are nilpotent since stronger conditions are required for nilpotent groups. In fact, we only need N_i/N_{i+1} to be abelian, not necessarily $N_i/N_{i+1} \leq Z(G/N_{i+1})$, or equivalently, $[N_i, G] \leq N_{i+1}$. Moreover, we ask for normality in each step, which is a weaker condition than asking for normality over G.

Let us give some examples of soluble groups that are not nilpotent. The smallest soluble non-nilpotent group is $S_3 \cong D_6$. On the one hand, it is soluble because its commutator subgroup is the abelian alternating group A_3. On the other hand, it cannot be nilpotent since $Z(S_3) = \{1\}$, whereas nilpotent groups have non-trivial center, which can be easily proved. If G is nilpotent with nilpotency class c, then $\gamma_{c+1}(G) = [\gamma_c(G), G] = \{1\}$. Thus, every $x \in \gamma_c(G)$ and $g \in G$ commute, so $\gamma_c(G) \leq Z(G)$.

More generally, any finite dihedral group D_{2n} is soluble for all $n \geq 1$, whereas it is nilpotent if and only if n is a power of 2 (see Problem 2). Thus, we could easily check that D_6 is soluble but not nilpotent since $n = 3$ is not a power of 2.

Let us now introduce the infinite dihedral group, which is constructed as an external semidirect product of C_2 and C_∞. The concept of a semidirect product is a generalization of a direct product and we also have internal and external semidirect product. We are interested in the external one, which is a way to construct a new group from two given groups by using the Cartesian product as a set and a particular multiplication operation.

Definition 1.17. Let H and N be two independent groups and let also $\theta : H \longrightarrow \mathrm{Aut}(N)$ be a group homomorphism. Then, the external semidirect product of H and N with respect to θ, which is denoted by $H \ltimes_\theta N$, is the cartesian product $H \times N$ with multiplication given by the following rule for all $h_1, h_2 \in H$ and all $n_1, n_2 \in N$:

$$(h_1, n_1) \cdot (h_2, n_2) = (h_1 \cdot h_2, \theta(h_2)(n_1) \cdot n_2).$$

The group homomorphism θ is known as the action of H over N and it maps each $h \in H$ to an automorphism of N. The action is essential in order to define the external semidirect product of two groups. We can prove that $H \ltimes_\theta N$ is indeed a group where $e = (1, 1)$ is the neutral or identity element and $(h, n)^{-1} = (h^{-1}, [\theta(h^{-1})(n)]^{-1})$.

Let us now construct the infinite dihedral group as the external semidirect product of C_2 and C_∞ generated by y and x, respectively. Let A be an abelian group and let also $H = \langle y \rangle \cong C_2$, then we could construct a semidirect product $H \ltimes_\theta A$ by choosing $\theta(y)$ as the automorphism of order 2 which maps all elements of A to their inverse, that is, $\theta(y)(a) = a^{-1}$ for all $a \in A$. Let us prove it is indeed an automorphism. It suffices to prove it is a group homomorphism, because each element has a unique inverse, hence it is bijective. Let $a_1, a_2 \in A$, then $\theta(y)(a_1 a_2) = (a_1 a_2)^{-1} = a_2^{-1} a_1^{-1}$ and since A is abelian this is equal to $a_1^{-1} a_2^{-1} = \theta(y)(a_1) \cdot \theta(y)(a_2)$ and we are done.

This group is called the generalized dihedral group associated to A and it is denoted by $\mathrm{Dih}\, A$. If we take $A \cong C_n$, then $\mathrm{Dih}\, A \cong D_{2n}$, whereas for $A = \langle x \rangle \cong C_\infty$ we get the infinite dihedral group, with presentation $D_\infty = \langle x, y \mid y^2 = 1, x^y = x^{-1} \rangle$, where $o(x) = \infty$ and $o(y) = 2$.

However, if we choose the generator yx instead of x, then we get that $D_\infty = \langle x, y \rangle = \langle yx, y \rangle$. Since $(yx)^2 = (yx)(yx) = y^2 \cdot x^y \cdot x = x^{-1} \cdot x = 1$ and $yx \neq 1$, we know that it has order 2. Therefore, for soluble groups it is not enough that the order of the generators is finite, but all elements in the group must have finite order. We have just constructed a finitely generated infinite group whose generators have finite order, but it is not periodic since there exists an element x of infinite order. Thus, it cannot be taken as a negative solution to the General Burnside Problem.

1.4 Linear groups

In this section linear groups are introduced, for which the answer to the General Burnside Problem is also positive. For any field K, the vector space $M(n,K)$ denotes the set of all square matrices of order n over K while $GL(n,K)$ denotes the group of invertible ones among them, with the operation of matrix multiplication. A linear group is a group that is isomorphic to a matrix group, that is, it is isomorphic to a subgroup of $GL(n,K)$ for some $n \in \mathbb{N}$ and some field K. The whole section is mainly based on [14], but there is a mistake in the proof of Theorem 1.29 when it is claimed that "their eigenvalues are in E_{alg}". Thus, for this theorem and the previous three lemmas we have followed [12, pages 149-154].

First of all, we will develop some facts that are closely linked to second year linear algebra when we studied the Jordan normal form.

Definition 1.18. A matrix $g \in GL(n,K)$ is said to be unipotent if all its eigenvalues are 1 over the algebraic closure of K. Equivalently, g is conjugate to an upper triangular matrix with all diagonal entries 1, i.e., there exists $P \in GL(n,K)$ such that

$$P^{-1} \cdot g \cdot P = \begin{pmatrix} 1 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ & 0 & & 1 \end{pmatrix}.$$

Definition 1.19. A matrix $g \in GL(n,K)$ is said to be nilpotent if all its eigenvalues are 0 over the algebraic closure of K, that is, there exists some $n \in \mathbb{N}$ where $g^n = 0$. Equivalently, g is conjugate to an upper triangular matrix with all diagonal entries 0, i.e., there exists $P \in GL(n,K)$ such that

$$P^{-1} \cdot g \cdot P = \begin{pmatrix} 0 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ & 0 & & 0 \end{pmatrix}.$$

The Jordan normal form of unipotent and nilpotent matrices is exactly an upper triangular matrix with all diagonal entries 1 and 0, respectively.

From now on, let us assume that K is algebraically closed without loss of generality, otherwise we take its algebraic closure. This is because we need that the characteristic polynomial of g, $\chi_g(X)$, splits into linear factors over K, that is, all the eigenvalues of g lie in K. This way, the Jordan normal form of g exists, which is very important for the proofs of this section. It is also known as Jordan canonical form and we denote it by JCF.

Lemma 1.20. *A matrix $g \in GL(n,K)$ is unipotent if and only if $g - I$ is nilpotent.*

Proof. The proof is trivial. If $g - I$ is nilpotent, then there exists some $P \in GL(n,K)$ such that

$$P^{-1} \cdot (g - I) \cdot P = P^{-1} \cdot g \cdot P - I = \begin{pmatrix} 0 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ & 0 & & 0 \end{pmatrix}$$

and if we sum the identity matrix in both sides we are done. The other implication is similarly done subtracting the identity matrix. □

Definition 1.21. In linear algebra, the trace of a square matrix A is the sum of the elements on the main diagonal of A and we denote it by $\mathrm{tr}(A)$.

Equivalently, the trace of a square matrix is the sum of its eigenvalues counted with multiplicities and it is invariant with respect to a change of basis. Thus, the trace of a square matrix is equal to the trace of its Jordan normal form.

Remark 1.22. Let K be a field, let $k \in K$ and $A, B \in M(n,K)$. Then, the trace is a linear mapping:

(i) $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$;

(ii) $\mathrm{tr}(k \cdot A) = k \cdot \mathrm{tr}(A)$.

Let us now prove that matrix groups are positive solutions to the General Burnside Problem, and so are linear groups due to isomorphism.

Lemma 1.23 (Burnside's lemma)**.** *Let K be any field and let $G \subset GL(n,K)$ be a subgroup such that the set $\{\mathrm{tr}(g) : g \in G\}$ is finite of cardinality s. Assume also that there is no nontrivial element of G which is unipotent. Then, G must be finite such that $|G| \leq s^{n^2}$.*

Proof. Let $\{g_1, ..., g_d\}$ be elements in G which form a basis for V, the vector subspace of $M(n,K)$ spanned by elements of G. The vector space $M(n,K)$ has dimension n^2, then $d \leq n^2$. In order to "count" elements of G, we associate each $g \in G$ with the ordered d-tuple $(\mathrm{tr}(g_1 g), ..., \mathrm{tr}(g_d g))$. If the same d-tuple is associated with $x, y \in G$, then let us prove that $x = y$ so that there is one-to-one correspondence between matrices and d-tuples.

If the d-tuples are equal, then $\mathrm{tr}(g_i(x - y)) = 0$ for all $i \leq d$. Now, we take $h = I - x^{-1}y$ and since $\{g_1, ..., g_d\}$ form a basis for V, then for each $k \geq 0$ there exist some $\beta_i \in K$ such that $h^k \cdot x^{-1} = (I - x^{-1}y)^k \cdot x^{-1} = \sum_{i=1}^{d} \beta_i g_i$.

Therefore, multiplying the i-th equation $\mathrm{tr}(g_i(x-y)) = 0$ by β_i and adding all of them we get $\mathrm{tr}((I - x^{-1}y)^{k+1}) = 0$ for all $k \geq 0$:

$$\begin{aligned} 0 &= \sum_{i=1}^{d} \beta_i \cdot \mathrm{tr}(g_i(x-y)) = \sum_{i=1}^{d} \mathrm{tr}(\beta_i \cdot g_i(x-y)) = \mathrm{tr}(\sum_{i=1}^{d} \beta_i \cdot g_i(x-y)) \\ &= \mathrm{tr}((\sum_{i=1}^{d} \beta_i \cdot g_i)(x-y)) = \mathrm{tr}((I - x^{-1}y)^k \cdot x^{-1}(x-y)) \\ &= \mathrm{tr}((I - x^{-1}y)^k \cdot (I - x^{-1}y)) = \mathrm{tr}((I - x^{-1}y)^{k+1}). \end{aligned}$$

Thus, we get that $\mathrm{tr}(h^k) = 0$ for all $k \geq 1$, which implies that h is nilpotent, that is, all its eigenvalues are 0. Let us prove this implication. Assume $\mathrm{tr}(h^k) = 0$ for all $k \geq 1$ and suppose h has some non-zero eigenvalues $\lambda_1, ..., \lambda_r$, being n_i the multiplicity of each λ_i, hence h is not nilpotent. Let J be the Jordan normal form of h. Then, the eigenvalues of h^k are exactly the diagonal entries of J^k, that is, the k-th powers of the eigenvalues of h since J is upper triangular. Moreover, each λ_i^k has multiplicity n_i.

Then, for each $k \geq 1$ we know that $\mathrm{tr}(h^k)$ is equal to the sum of all the k-th powers of the eigenvalues of h counting their multiplicities, that is, $\mathrm{tr}(h^k) = \sum_{i=1}^{r} n_i \cdot \lambda_i^k = 0$ and we get the following system of r equations:

$$\begin{cases} n_1\lambda_1 + & \cdots & + n_r\lambda_r = 0 \\ n_1\lambda_1^2 + & \cdots & + n_r\lambda_r^2 = 0 \\ & \vdots & \\ n_1\lambda_1^r + & \cdots & + n_r\lambda_r^r = 0 \end{cases}$$

and if we rewrite the system of equations in matrix form we would obtain

$$\begin{pmatrix} \lambda_1 & \dots & \lambda_r \\ \vdots & & \vdots \\ \lambda_1^r & \dots & \lambda_r^r \end{pmatrix} \begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

If we prove that the determinant of the coefficient matrix is non-zero, then the system has a unique solution $n_1 = \cdots = n_r = 0$, which contradicts the assumption that h has non-zero eigenvalues. Let us compute it:

$$\begin{vmatrix} \lambda_1 & \dots & \lambda_r \\ \vdots & & \vdots \\ \lambda_1^r & \dots & \lambda_r^r \end{vmatrix} = \lambda_1 \cdots \lambda_r \begin{vmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_r \\ \vdots & & \vdots \\ \lambda_1^{r-1} & \dots & \lambda_r^{r-1} \end{vmatrix}.$$

Since all λ_i are non-zero eigenvalues, it suffices to check whether the latter determinant is non-zero or not. In particular, it is the so-called Vandermonde determinant, which is non-zero if and only if all λ_i are distinct, and since these λ_i represent distinct eigenvalues of h we are done.

Hence, zero is the unique eigenvalue of h, with multiplicity n, so that h is nilpotent. Then, by the characterization of Lemma 1.20 we get that $I - h = x^{-1} \cdot y$ is unipotent, but by hypothesis there is no nontrivial unipotent element of G, hence $I = x^{-1} \cdot y$ and $x = y$.

Therefore, the association between g and $(\mathrm{tr}(g_1 g), ..., \mathrm{tr}(g_d g))$ is one-to-one. Since the traces of the elements of G can take at most s values, then the set of d-tuples has cardinality at most $s^d \leq s^{n^2}$. This completes the proof. □

Definition 1.24. Let R be a ring. Then, the characteristic of R is the smallest integer $n \in \mathbb{N}$ such that $na = 0_R$ for all $a \in R$, where 0_R is the additive identity, and we denote it by $\mathrm{Char}\, R$. If no such positive integer exists, then R is said to be of characteristic zero.

If $R = K$ is a field, then the characteristic of K is either 0 or a prime number p.

Theorem 1.25. *Let K be any field and let N be a natural number which is not a multiple of* $\mathrm{Char}\, K$. *If $G \subset GL(n, K)$ is an N-torsion group, i.e., a periodic group of finite exponent N, then G must be finite of cardinality $|G| \leq N^{n^3}$.*

Proof. Without loss of generality, we assume K is algebraically closed. The vector subspace $V \subset M(n, K)$ generated by G has dimension at most n^2 and let $\{g_1, ..., g_d\}$ be elements in G which form a basis for V. We just need to check that all hypotheses from Burnside's lemma are satisfied being N^n an upper bound of the cardinality of the set of traces of elements in G.

If J is the Jordan normal form of $g \in G$, then $J^N = I$ since G is an N-torsion group and $g^N = P^{-1} \cdot J^N \cdot P = I$. The diagonal entries of J^N are precisely the N-th powers of the eigenvalues of g and they are all equal to one, hence the eigenvalues of g are N-th roots of unity. Since K is algebraically closed, then $X^N - 1$ splits into linear factors over K, which has exactly N roots including multiplicities, so there are at most N N-th roots of unity. Thus, we have at most N choices for the eigenvalues of g, and since the trace of g is equal to the sum of all its n eigenvalues, then the trace of g has no more than N^n possibilities.

We only need to check one last condition: there is no nontrivial element of G which is unipotent. Let us prove it by contradiction. Suppose that $I \neq g \in G$ is unipotent and without loss of generality we assume g is upper triangular with all diagonal entries 1, otherwise we conjugate it by a matrix in $GL(n, K)$. Let $g_{ij} \neq 0$ be such that $j - i \geq 1$ and $j - i$ is the least possible. Now, looking at the (i, j)-th entry of $g^N = I$, we have that $0 = N \cdot g_{ij}$, which is a contradiction since $\mathrm{Char}\, K$ does not divide N. Thus, Lemma 1.23 implies the assertion of this corollary since all conditions are satisfied.

Then, it suffices to check that the (i,j)-th entry of g^N is indeed $N \cdot g_{ij}$. We are going to prove by induction on the power r that the (i,j)-th entry of g^r is $g_{ij}^{(r)} = r \cdot g_{ij}$, and thus the proof would be completed taking $r = N$. Taking into account how g_{ij} has been chosen, we know that $g_{ik}^{(1)} = g_{lj}^{(1)} = 0$ for all $k < j$ and $l > i$, except for $g_{ii}^{(1)} = g_{jj}^{(1)} = 1$. Similarly, we could easily check that $g_{ik}^{(m)} = g_{lj}^{(m)} = 0$ and $g_{ii}^{(m)} = g_{jj}^{(m)} = 1$ for all $m \geq 1$.

Let us begin with the induction. For $r = 1$, it trivially holds that the (i,j)-th entry of g is exactly $g_{ij}^{(1)} = g_{ij}$. Now, we assume it is true up to $r-1$ and we need to prove it for r. Since $g^r = g \cdot g^{r-1}$, then $g_{ij}^{(r)} = \sum_{k=1}^{n} g_{ik} \cdot g_{kj}^{(r-1)}$ holds. Moreover, we also know that $g_{ik}^{(r)} = g_{lj}^{(r)} = 0$ for all $k < j$ and $l > i$, except for $g_{ii}^{(r)} = g_{jj}^{(r)} = 1$. Thus:

$$g_{ij}^{(r)} = g_{ii} \cdot g_{ij}^{(r-1)} + g_{ij} \cdot g_{jj}^{(r-1)} = 1 \cdot (r-1)g_{ij} + g_{ij} \cdot 1 = r \cdot g_{ij}.$$

□

We may ask ourselves what happens if $\operatorname{Char} K = p$ divides N. In this case, the last result is not true (see Problem 4). Let us refine the last theorem by dropping the condition of bounded torsion when the group is finitely generated. Before doing that, let us introduce the following lemmas.

Lemma 1.26. *Let $G \subseteq GL(n,K)$ for a field K and let $g \in G$ be a matrix of order m. Then, m is the least possible integer such that the minimal polynomial of g divides $X^m - 1$.*

Proof. Let $\mu_g(X)$ be the minimal polynomial of g. Since $o(g) = m$, then $g^m = I$, which implies that $\mu_g(X) \mid X^m - 1$. On the other hand, if we have that $\mu_g(X) \mid X^k - 1$, then $X^k - 1 = \mu_g(X)f(X)$ for some polynomial $f(X)$ and $g^k - I = \mu_g(g)f(g) = 0$. Thus, $g^k = I$ and $k \geq m$. □

Lemma 1.27. *Let P be a prime field and let E be a finite extension of P. Then, for every $d \in \mathbb{N}$ there is a finite number of monic polynomials of degree d in $E[X]$ with the property that all its roots are roots of unity.*

Proof. Suppose first that $E = P$. If $P \cong \mathbb{F}_p$, then the result is obvious since there are finitely many polynomials of degree d in $\mathbb{F}_p[X]$. Assume now that $P \cong \mathbb{Q}$. Let $f(X) \in \mathbb{Q}[X]$ be a monic polynomial of degree d all of whose roots are roots of unity, say $\lambda_1, ..., \lambda_d$. If λ_i is a primitive n_i-th root of unity, then the minimal polynomial $m_i(X)$ of λ_i over $\mathbb{Q}$ divides $f(X)$, which implies that $\varphi(n_i) \leq d$. Since $\lim_{n\to\infty} \varphi(n) = +\infty$, it follows that there are only finitely many possibilities for the n_i, and consequently also for the λ_i. Since $f(X) = (X - \lambda_1)\cdots(X - \lambda_d)$, we conclude that there are finitely many possibilities for $f(X)$.

Now, we consider the general case when $[E : P] < \infty$. By considering if necessary a normal closure of E over P, we may assume that E/P is a normal extension. Since every extension of $\mathbb{F}_p$ or $\mathbb{Q}$ is separable, then E/P is actually Galois. We define:

$$f^*(X) = \prod_{\sigma \in Gal(E/P)} \sigma(f(X)) \in P[X]$$

where $f(X) \in E[X]$ is a monic polynomial of degree d whose roots are all roots of unity and similarly for $f^*(X)$, which has degree $d \cdot [E : P]$. By the previous case, we know that there are finitely many possibilities for $f^*(X)$ and since $f(X) \mid f^*(X)$, also for $f(X)$. □

Lemma 1.28. *Let L/K be a finitely generated extension and let E be the field of elements in L that are algebraic over K. Then, $[E : K] < \infty$.*

Proof. Write $L = K(\alpha_1, ..., \alpha_r)$. We argue by induction on r. If $r = 1$, then $E = K(\alpha_1)$ or K, according as α_1 is algebraic or trascendental over K. In any case, $[E : K] < \infty$.

Suppose now that $r > 1$. By the induction hypothesis, if F is the field of elements in L that are algebraic over $K(\alpha_1)$, then $[F : K(\alpha_1)] < \infty$. Observe that $E \subseteq F$. If α_1 is algebraic over K, then $[K(\alpha_1) : K] < \infty$, hence $[E : K] \leq [F : K(\alpha_1)] \cdot [K(\alpha_1) : K] < \infty$.

Then, we assume that α_1 is transcendental over K, hence it is also transcendental over E since E/K is algebraic and the property of being algebraic is transitive. Consider a number of elements $\beta_1, ..., \beta_s \in E$ that are K-linearly independent. We claim that they are also $K(\alpha_1)$-linearly independent. This shows that $[E : K] \leq [F : K(\alpha_1)] < \infty$, which completes the proof.

Let us prove the claim. Suppose that $\lambda_1\beta_1 + \cdots + \lambda_s\beta_s = 0$, with $\lambda_i \in K(\alpha_1)$. If we write $\lambda_i = \mu_i/\nu_i$ with $\mu_i, \nu_i \in K[\alpha_1]$ and multiply by $\nu_1 \cdots \nu_s$, then we can further assume that $\lambda_i \in K[\alpha_1]$ for every i, and hence $\lambda_i = f(\alpha_1)$ for some $f_i \in K[X]$. If some $f_i \neq 0$, then we set $d = \max\{\deg f_i \mid i = 1, ..., s\}$ and if we write $\lambda_i = a_{i0} + a_{i1}\alpha_1 + \cdots + a_{id}\alpha_1^d$, with $a_{ij} \in K$, then:

$$\lambda_1\beta_1 + \cdots + \lambda_s\beta_s = \left(\sum_{i=1}^{s} a_{i0}\beta_i\right) + \left(\sum_{i=1}^{s} a_{i1}\beta_i\right)\alpha_1 + \cdots + \left(\sum_{i=1}^{s} a_{id}\beta_i\right)\alpha_1^d = 0.$$

Since α_1 is transcendental over E and all the coefficients in this linear combination lie in E, we deduce that

$$\sum_{i=1}^{s} a_{i0}\beta_i = \sum_{i=1}^{s} a_{i1}\beta_i = \cdots = \sum_{i=1}^{s} a_{id}\beta_i = 0.$$

Since $\beta_1, ..., \beta_s$ are K-linearly independent it follows that $a_{ij} = 0$ for all i, j so $\lambda_i = 0$ for all $i = 1, ..., s$. This proves the claim. □

Theorem 1.29. *Let G be a finitely generated periodic subgroup of $GL(n, K)$. Then:*

(i) *G has finite exponent;*

(ii) *If the orders of all elements of G are not multiples of* $\operatorname{Char} K$*, then G is finite.*

Proof. Let us start proving (i). Let $G = \langle g_1, ..., g_d \rangle$ and let L be the field obtained by adjoining to P all matrix entries of the generators, where P is the prime subfield of K. One can readily check that the entries of every $g \in G$ belong also to L, since g can be written as a finite product of the generators of G, so its matrix entries are obtained by multiplying matrix entries of the generators and adding all these results, hence they lie in L. By Lemma 1.28, if E is the set of elements in L algebraic over P, then $[E : P] < \infty$.

Consider an arbitrary element $g \in G$ of order m. Since the entries of g lie in L, the minimal polynomial $\mu_g(X)$ of g belongs to $L[X]$ and we also know that $\mu_g(X)$ divides $X^m - 1$. Then, all its roots are m-th roots of unity, which are the eigenvalues of g, say $\lambda_1, ..., \lambda_d$. They are algebraic over P and since the product of these algebraic elements is also algebraic over P, then $\mu_g(X) = (X - \lambda_1) \cdots (X - \lambda_d) \in E[X]$ which has degree $d \leq n$. We can now apply Lemma 1.27 to deduce that there are only finitely many possibilities for $\mu_g(X)$, say $\mu_1, ..., \mu_r$. Let m_i be the least possible integer such that μ_i divides $X^{m_i} - 1$. Then, these m_i form a finite set and Lemma 1.26 implies that the order m of g lies in this finite set. This proves that there are only finitely many possibilities for the orders of the elements of g, hence G has finite exponent.

Finally, (ii) follows from (i) and Theorem 1.25. Since the orders of all elements of G are not multiples of $\operatorname{Char} K$, it follows that $\exp(G)$ is not a multiple of $\operatorname{Char} K$. □

Corollary 1.30. *The General Burnside Problem has a positive solution for matrix groups of characteristic* 0*.*

Proof. It trivially holds from Theorem 1.29, since $\operatorname{Char} K = 0$ does not divide the order of any element in G. □

This result was given by I. Schur [15] in 1911, proving that the General Burnside Problem has an affirmative answer for linear groups of characteristic zero. Then, the adaptations needed for the proof in the case of characteristic p were given by I. Kaplansky in 1965 [11], hence the answer to the General Burnside Problem is positive for all linear groups. However, we need more advanced results in ring theory in order to prove it.

Chapter 2

Golod-Shafarevich groups

In this chapter we introduce the first negative solution to the General Burnside Problem, which are Golod-Shafarevich groups. Golod-Shafarevich algebras and groups, which were introduced by Russian mathematicians Evgenii Golod and Igor Shafarevich in 1964, had been used as a powerful tool in ring theory and group theory. They were introduced in relation to the famous class field tower problem, which asks whether the class field tower of any number field must be finite. It was posed by Furtwängler in 1925 and it remained open for almost 40 years until 1964, when Golod and Shafarevich proved that the answer to the problem is negative.

It turns out that their negative solution to the General Burnside Problem goes through the negative solution to a closely connected problem in associative algebras known as Kurosh-Levitzki problem, which asks whether a finitely generated nil algebra is necessarily nilpotent. We are going to construct Golod-Shafarevich algebras, which are negative solutions to this latter problem, in order to construct Golod-Shafarevich groups.

The sources we have used in this chapter are mainly the notes provided by my supervisor [6]. We have also followed [3], [4] and [17] as main references.

2.1 K-algebras and formal power series

Before introducing the Kurosh-Levitzki problem let us make a brief introduction to algebras over a field K and formal power series, which we will use throughout the chapter.

Definition 2.1. Let K be a field and let A be a ring and a vector space over K where multiplication is a K-bilinear map. Then, we say A is an associative algebra over K or a K-algebra.

The multiplication operation in an associative algebra A is not assumed to be commutative, which leads to the concept of commutative and non-commutative algebras. Throughout this chapter we are going to work with non-commutative algebras. Moreover, an algebra does not necessarily have an identity element with respect to multiplication. In case it has an identity element the algebra is called unital or unitary, otherwise it is said to be non-unital.

In the third year course of Commutative Algebra we worked with unital associative commutative algebras, in particular with the polynomial ring $K[X_1, ..., X_n]$ and also with finitely generated K-algebras of the form $A = K[a_1, .., a_n]$ where $a_1, ...a_n \in A$, which are isomorphic to quotients of $K[X_1, ..., X_n]$ by the first isomorphism theorem.

However, most of the results we learned can be generalised for non-commutative and non-unital algebras, which we will use in order to construct Golod-Shafarevich groups. Let us first introduce two important examples of K-algebras: non-commutative polynomial algebras or free algebras and group algebras.

Definition 2.2. A monoid is a set that is closed under an associative binary operation and has an identity element, that is, a semigroup with an identity element. In case all elements have an inverse, which is not necessary, then the monoid is a group.

Definition 2.3. Let X be a fixed set, also called alphabet. Then, the free monoid on X is the set of all finite words (or strings) of zero or more elements of X made into a monoid using string concatenation, denoted by X^*, which is not commutative. It has an identity element ϵ which is the unique word of zero elements known as the empty word.

Let us define non-commutative polynomial algebras, also known as free algebras.

Definition 2.4. Let R be a commutative ring and $X = \{X_1, \ldots, X_d\}$ the set of indeterminates. Then, the free (associative) algebra over R in d variables is the non-commutative analogue of a polynomial ring, where its elements are polynomials in non-commuting variables. It is indeed an R-module with X^* as a basis, for which multiplication is defined such that the product of two basis elements is the concatenation of the corresponding words. Therefore, we also call it the free algebra generated by X and it is denoted by $R\langle X_1, \ldots, X_d\rangle$ or $R\langle X\rangle$.

From now on let $X = \{X_1, \ldots, X_d\}$. If $R = K$ is a field, then X^* is a basis of $K\langle X\rangle$ as a vector space over K, and any element of $K\langle X\rangle$ can be

uniquely written as

$$\sum_{k=0}^{\infty} \sum_{i_1,\ldots,i_k \in \{1,\ldots,d\}} \lambda_{i_1,\ldots,i_k} X_{i_1} X_{i_2} \cdots X_{i_k}$$

where $\lambda_{i_1,\ldots,i_k}$ are elements of K and all but finitely many of these are zero, since elements are polynomials and not series. Unlike in polynomial rings, variables do not commute.

The universal property of non-commutative polynomial algebras is the same as for commutative ones since in both cases polynomials are described uniquely as a sum of monomials. In order to fully determine a K-algebra homomorphism between a polynomial algebra and a K-algebra B, it suffices to give the images of all the free generators, mapping each X_i to any element $b_i \in B$. Thus, $f(X_1, ...X_n)$ is mapped to $f(b_1, ..., b_n)$.

Subalgebras of non-commutative (unital or non-unital) algebras can be generated in a similar way as for the commutative case. Let A be a unital non-commutative algebra and some elements $a_1, ..., a_d \in A$, then these elements generate a subalgebra of A denoted by $K\langle a_1, \ldots, a_d \rangle$. By definition, it is the intersection of all subalgebras containing $a_1, ..., a_d$ and its elements are polynomials in $K\langle X_1, \ldots, X_d \rangle$ evaluated in the generators $a_1, ..., a_d$.

On the other hand, if A is non-unital and generated by $a_1, ..., a_d \in A$, its elements are described as non-commuting polynomials without constant term in these generators. There is a canonical way of constructing a unital algebra B from A. The idea is to construct B as the cartesian product of K and A, identifying elements in B as tuples (λ, a) where $\lambda \in K$ and $a \in A$. Addition is defined by $(\lambda_1, a_1) + (\lambda_2, a_2) = (\lambda_1 + \lambda_2, a_1 + a_2)$, scalar multiplication is defined by $\mu \cdot (\lambda, a) = (\mu\lambda, \mu a)$ and multiplication is defined by $(\lambda_1, a_1) \cdot (\lambda_2, a_2) = (\lambda_1\lambda_2, \lambda_1 a_2 + \lambda_2 a_1 + a_1 a_2)$. Then, $(1, 0)$ is the identity element of B, hence it is unital.

In order to simplify the notation we identify $(\lambda, 0)$ with λ and $(0, a)$ with a, so that $(\lambda, 0) + (0, a) = (\lambda, a)$ corresponds to the element $\lambda + a$. Thus, $B = K \oplus A = \{\lambda + a \mid \lambda \in K, a \in A\}$. In addition, if A is generated by $a_1, ..., a_d$, then B is also generated by the same generators so that $B = K\langle a_1, ..., a_d \rangle$.

Let us now introduce group algebras over a field K, which we will use for the Kurosh-Levitzki problem in Section 2.2.

Definition 2.5. Let K be a field and G a group under multiplication. Then, the group algebra of G over K, which is denoted by $K[G]$, is the set of all (formal) linear combinations of finitely many elements of G with coefficients in K, so that elements are of the form $\sum_{g \in G} \lambda_g g$ where $\lambda_g \in K$.

Group algebras have also a universal property. Suppose that we want to construct a K-algebra homomorphism φ between $K[G]$ and a K-algebra

B. It suffices to give the images of all elements of G, which are not free generators of $K[G]$, so images cannot be arbitrarily chosen. In this case, we need to ensure these two conditions:

(i) $\varphi(g) \in B^\times$, $\forall g \in G$;

(ii) $\varphi(gh) = \varphi(g)\varphi(h)$, $\forall g, h \in G$.

Equivalently, $\varphi|_G : G \longrightarrow B^\times$ must be a group homomorphism.

Finally, let us introduce formal power series and the ring of formal power series, which we will use while working with Hilbert series and computing inverse series in Section 2.3.

Definition 2.6. A formal power series is an infinite sum whose terms are of the form $a_n X^n$ where X^n is the n-th power of a variable X and a_n is called the coefficient of X^n. They can be seen as a generalization of polynomials where the number of terms is allowed to be infinite and with no requirements of convergence or they can also be seen as objects that just record a sequence of coefficients.

Definition 2.7. Let R be a commutative ring. Then, the set of all formal power series in a variable X with coefficients in R is denoted by $R[[X]]$ and it is called the ring of formal power series in X over R. The elements of $R[[X]]$ are infinite expressions of the form $f(X) = a_0 + a_1 X + a_2 X^2 + \cdots$, where $a_n \in R$ for all $n \in \mathbb{N}$.

$R[[X]]$ has indeed a ring structure where addition and multiplication are defined just as for the polynomial ring $R[X]$ and it is commutative since R is. We know polynomials are particular cases of formal power series, hence it is clear that $R[X]$ is a subset of $R[[X]]$ and that the algebraic operations of these two rings agree on this subset.

It would be great to ask for invertible elements in this ring. For instance, $1 + X$ is invertible since the geometric series formula is also valid in $R[[X]]$:

$$(1 + X)^{-1} = \frac{1}{1 + X} = \sum_{n=0}^{\infty} X^n = 1 + X + X^2 + \cdots$$

However, the ring $R[[X]]$ is not a field because, for example, X is not invertible in $R[[X]]$. The following proposition gives a necessary and sufficient condition for a series to be invertible in $R[[X]]$.

Proposition 2.8. *Let R be a commutative ring and let $f(X) = \sum_{n=0}^{\infty} a_n X^n$ be a formal power series in $R[[X]]$. Then, $f(X)$ is invertible if and only if a_0 is invertible in R.*

Proof. We need to determine if there exists $g(X) = \sum_{n=0}^{\infty} b_n X^n \in R[[X]]$ such that $f(X)g(X) = 1$. Expanding the product we have

$$f(X)g(X) = \Big(\sum_{n=0}^{\infty} a_n X^n\Big) \cdot \Big(\sum_{n=0}^{\infty} b_n X^n\Big) = \sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} a_k b_{n-k}\Big) X^n.$$

Comparing the coefficients of X^n on both sides of $f(X)g(X) = 1$ we realise that $g(X)$ satisfies the equation if and only if $a_0 b_0 = 1$ and $\sum_{k=0}^{n} a_k b_{n-k} = 0$ for all $n \geq 1$. Then, $a_0 b_0 = 1$ is a necessary condition, hence a_0 must be invertible in R. Moreover, it is indeed a sufficient condition since recursively defining $b_n = -a_0^{-1} \sum_{k=1}^{n} a_k b_{n-k}$ the second condition is also satisfied. □

Corollary 2.9. *Let $R = K$ be a field. Then, a formal power series in $K[[X]]$ is invertible if and only if the constant term a_0 is non-zero.*

2.2 The Kurosh-Levitzki problem

The Kurosh-Levitzki problem was posed in the early 1940s by Alexander G. Kurosh and Jakob Levitzky and it asks whether a finitely generated nil algebra is necessarily nilpotent, for which the answer is negative. However, if we ask for a bound for the degrees of nilpotency of all elements we get a variant of the problem known as the Ordinary Kurosh-Levitzki problem. In this case, the conclusion that A is nilpotent is true, but we are only interested in the general case.

Definition 2.10. An element $a \in A$ is nilpotent if there exists an integer $n \geq 1$ such that $a^n = 0$, and the minimum of such integers is called degree of nilpotency of a. We say that A is a nil algebra if every $a \in A$ is nilpotent and A is a nilpotent algebra if there exists some integer $n \geq 1$ such that $A^n = \{a_1 \cdots a_n \mid a_i \in A, \forall i \in \{1, ..., n\}\} = \{0\}$.

Observe that nil algebras do not contain identity element 1 since it is not a nilpotent element. Obviously, nilpotent algebras are nil, whereas the converse is not always true. The following theorem gives a sufficient condition for finitely generated nil algebras to be nilpotent, whose proof is left as a problem (see Problem 6).

Theorem 2.11. *If A is a finitely generated nil algebra, then A is nilpotent if and only if A is finite dimensional.*

The Kurosh-Levitzki problem is closely connected to the General Burnside Problem, in that negative answers to one problem lead to negative answers to the other one and vice versa. We are interested in proving just one direction, but we could also construct a negative answer to the Kurosh-Levitzki problem over any field K of characteristic p from a finitely generated infinite p-group (see Problem 7).

Let A be a negative solution to the Kurosh-Levitzki problem over a field K of characteristic p generated by $a_1, ..., a_d$ and let $B = K \oplus A$ be a unital algebra, for which we have the multiplicative group $B^\times$. We want to get a finitely generated infinite p-group $G \leq B^\times$, a negative solution to the GBP. Later on we will construct Golod-Shafarevich algebras, which are negative solutions to the Kurosh-Levitzki problem, and applying this result we get the corresponding negative solutions to the GBP known as Golod-Shafarevich groups.

Since A is a nil algebra of characteristic p, then $1 + A$ is a subgroup of $B^\times$ and it is indeed a p-group. For every $a_1, a_2 \in A$ we get that

$$(1 + a_1)(1 + a_2) = 1 + (a_1a_2 + a_1 + a_2) \in 1 + A.$$

Moreover, every element $a \in A$ is nilpotent, that is, there exists $n \geq 1$ such that $a^n = 0$. Then, there exists also $m \in \mathbb{N}$ such that $p^m \geq n$ and hence $a^{p^m} = 0$. Finally, we get that $(1 + a)^{p^m} = 1 + a^{p^m} = 1$ and hence $(1 + a)^{-1} = (1 + a)^{p^m - 1} \in 1 + A$.

Any subgroup of $1 + A$ is a p-group and we want G to be a finitely generated infinite p-group. Let us take $G = \langle 1 + a_1, \ldots, 1 + a_d \rangle \leq 1 + A$, where $a_1, ..., a_d$ are the generators of A. It remains to prove G is infinite, for which we need to construct an algebra homomorphism φ that goes from $K[G]$ to B, both K-algebras. As we have seen in Section 2.1, we need a group homomorphism $\varphi|_G : G \longrightarrow B^\times$ in order to define φ. Since G belongs to $B^\times$ it suffices to choose the identity map 1_G and φ is surjective because G generates B as a unital algebra.

Since A is not nilpotent, then by Theorem 2.11 it is infinite dimensional. This implies $\dim_K B = \infty$, and by surjectivity of φ, $|G| = \dim_K K[G] = \infty$ as desired.

2.3 Golod-Shafarevich algebras and groups

Our main goal in this section is to construct Golod-Shafarevich algebras, which are negative solutions to the Kurosh-Levitzki problem. As explained in the previous section, this automatically gives Golod-Shafarevich groups. Let us fix the notation we are going to use throughout this section in order to make it clear from scratch.

Definition 2.12. A unital algebra S is said to be graded if it has a direct sum decomposition into K-subspaces: $S = S_0 \oplus S_1 \oplus S_2 \oplus \cdots = \bigoplus_{n=0}^{\infty} S_n$, where $S_0 = K1_S$ and $S_iS_j \subseteq S_{i+j}$ for all $i, j \geq 0$. We say that the elements of S_i are homogeneous elements of degree i.

The free algebra $S = K\langle X_1, \ldots, X_d\rangle$ is a finitely generated, graded and unital algebra. Each S_n is a subspace whose basis is the set of words of X^* of length n, so $\dim_K S_n = d^n$. Let T denote the set of polynomials of S without constant term, $T = \bigoplus_{n=1}^{\infty} S_n$. Similarly as we have done before, T is non-unital and $S = K \bigoplus T$.

In order to get a negative solution to the Kurosh-Levitzki problem, our aim is to find a suitable ideal N of S and consider the quotient algebra $Q = S/N = K1_Q \bigoplus A$, where $A = T/N$ is the negative answer we are looking for and $K1_Q = K + N/N \cong K/K \cap N \cong K$, since $K \cap N = \{0\}$. Let us simplify the notation and write $Q = K \bigoplus A$. Since T is finitely generated as an algebra, then A also is by taking the cosets as its generators. Therefore, we need A to be a nil algebra and infinite dimensional, which is equivalent to satisfying these two conditions:

(i) For every $f \in T$, some power of f is in N, i.e., every element in A is nilpotent.

(ii) $\dim_K Q = \infty$. In fact, since $\dim_K Q = 1 + \dim_K A$, then A is infinite-dimensional if and only if Q is.

We search for an ideal N satisfying both properties, which is generated by a sequence of homogeneous polynomials $\{f_1, f_2, \ldots\}$ with finitely many polynomials of every degree, say k_n polynomials of degree n. We assume that polynomials have at least degree 2, hence $k_0 = k_1 = 0$. An arbitrary element of the homogeneous two-sided ideal $N = (f_1, f_2, \ldots)$ is of the form $\sum_{i \in \mathbb{N}} g_i f_i h_i$ where $g_i, h_i \in S$ and only finitely many summands are non-zero.

Similarly as we have done for S, the homogeneous ideal N can also be decomposed as a direct sum of K-subspaces: $N = \bigoplus_{n=0}^{\infty} N_n$, where $N_n \subseteq S_n$ is the set of homogeneous polynomials of degree n in N. Since N_n is a vector subspace of S_n, then we can take a complement B_n such that $S_n = N_n \bigoplus B_n$, hence $B_n \cong S_n/N_n$. Let us define the vector subspace $B = \bigoplus_{n=0}^{\infty} B_n \leq S$ so that we get the following isomorphism as vector spaces:

$$Q = S/N \cong \bigoplus_{n=0}^{\infty} S_n/N_n \cong \bigoplus_{n=0}^{\infty} B_n = B.$$

Therefore, $S = N \bigoplus B$ and observe that Q and B are isomorphic as vector spaces, so that Q is infinite-dimensional if and only if B is, which is easier to work with. Our goal is to get infinitely many B_n different from zero so that B is infinite-dimensional, and thus Q is infinite-dimensional. Let $b_n = \dim_K B_n$, then equivalently we want to find a suitable sufficient condition such that $b_n > 0$ for infinitely many n.

Lemma 2.13 (The fundamental inequality of Golod-Shafarevich)**.** *Let S be the free algebra $K\langle X_1, \ldots, X_d\rangle$ and let N be the ideal generated by homogeneous $f_i \in S$ of degree d_i with $2 \leq d_1 \leq d_2 \leq \cdots$, where d_i tends to infinity. Let k_i be the number of generators f_i of degree i and let $b_i = \dim_K B_i$. Then:*

$$b_n \geq d \cdot b_{n-1} - \sum_{i=2}^{n} k_i b_{n-i} \quad \textit{for} \quad n \geq 1.$$

Proof. Let $R = \langle f_1, f_2, \ldots\rangle$ be a linear graded (since all f_i are homogeneous) subspace over K. Every element of $N = (f_1, f_2, \ldots)$ is of the form $\sum_{i=1}^{\infty} g_i f_i h_i$, where $g_i, h_i \in S$ and only finitely many summands are non-zero. Then, it follows that $N = SRS$.

Moreover, we also need to realise that $T = SS_1$, hence $S = SS_1 + K$ and since R is a linear subspace over K, then $RK = R$. Finally, let us remark that $R \subseteq T$ and $NS = N$, hence $NR \subseteq NSS_1 = NS_1$. All in all, we get the following chain of equalities and inclusions:

$$\begin{aligned} N = SRS = SR(SS_1 + K) = NS_1 + SRK = NS_1 + SR \\ = NS_1 + (N + B)R \subseteq NS_1 + BR. \end{aligned}$$

For a fixed $n \geq 2$ we get that $N_n \subseteq N_{n-1}S_1 + \sum_{i=2}^{n} R_i B_{n-i}$, hence

$$\dim N_n \leq (\dim N_{n-1})(\dim S_1) + \sum_{i=2}^{n} (\dim R_i)(\dim B_{n-i}).$$

Observe that the sum starts from $i = 2$ since all f_i are homogeneous of degree at least 2. Let us recall that $\dim N_n = \dim S_n - \dim B_n = d^n - b_n$ and $\dim R_i = k_i$. Then,

$$d^n - b_n \leq (d^{n-1} - b_{n-1}) \cdot d + \sum_{i=2}^{n} k_i b_{n-i}$$

hence

$$b_n \geq d \cdot b_{n-1} - \sum_{i=2}^{n} k_i b_{n-i}.$$

□

Definition 2.14. Let $B = \bigoplus_{n=0}^{\infty} B_n$ be a graded vector space and let also $b_i = \dim_K B_i < \infty$ for all i. Then, the associated Hilbert series is the formal power series $\sum_{n=0}^{\infty} b_n t^n$.

Let $\sum_{n=0}^{\infty} a_n t^n$ and $\sum_{n=0}^{\infty} b_n t^n$ be two series. We write $\sum_{n=0}^{\infty} a_n t^n \leq \sum_{n=0}^{\infty} b_n t^n$ provided that $a_n \leq b_n$ for all $i \geq 0$. Observe that we do not want the Hilbert series associated to B to be a polynomial, hence we need $b_n > 0$ for infinitely many n.

Theorem 2.15 (Golod-Shafarevich Theorem)**.** *With the above notation we have that*

$$\big(1 - dt + \sum_{n=2}^{\infty} k_n t^n\big)\big(\sum_{n=0}^{\infty} b_n t^n\big) \geq 1.$$

Proof. In the fundamental inequality of Golod-Shafarevich, if we multiply t^n and we sum over n we get that

$$\sum_{n=1}^{\infty} b_n t^n \geq \sum_{n=1}^{\infty} d \cdot b_{n-1} t^n - \sum_{n=1}^{\infty} \big(\sum_{i=2}^{n} k_i b_{n-i}\big) t^n.$$

Rewriting our inequality we have

$$\big(\sum_{n=0}^{\infty} b_n t^n\big) - 1 \geq dt\big(\sum_{n=0}^{\infty} b_n t^n\big) - \big(\sum_{n=2}^{\infty} k_n t^n\big)\big(\sum_{n=0}^{\infty} b_n t^n\big)$$

and taking the Hilbert series associated to B as common factor we finally get

$$\big(1 - dt + \sum_{n=2}^{\infty} k_n t^n\big)\big(\sum_{n=0}^{\infty} b_n t^n\big) \geq 1.$$

□

From this inequality we can get a negative answer to the Kurosh-Levitzki problem. Let us now look for sufficient conditions for $Q = K\langle X_1, \ldots, X_d\rangle/N$ to be infinite-dimensional.

Theorem 2.16. *If all coefficients of* $(1 - dt + \sum_{n=2}^{\infty} k_n t^n)^{-1} = \sum_{n=0}^{\infty} c_n t^n$ *are non-negative, then* Q *is infinite-dimensional.*

Proof. If the inverse series has non-negative coefficients, we can multiply it on both sides of the inequality in Theorem 2.15 and we get the following:

$$\sum_{n=0}^{\infty} b_n t^n \geq \big(1 - dt + \sum_{n=2}^{\infty} k_n t^n\big)^{-1} = \sum_{n=0}^{\infty} c_n t^n \geq 0.$$

Thus, $b_n \geq c_n \geq 0$ for all n. It suffices to prove that the inverse series cannot be a polynomial in order to get $c_n > 0$ for infinitely many n, and thereby similarly for b_n. Let us prove it by contradiction assuming that the inverse is a polynomial of degree s, $f(t) = \sum_{n=0}^{s} c_n t^n$, such that all c_i are non-negative coefficients. Then,

$$\big(1 + \sum_{n=2}^{\infty} k_n t^n\big) \cdot f(t) = 1 + dt \cdot f(t).$$

On the right-hand side we get a polynomial of degree $s + 1$. The coefficient for t^{s+1} in both sides must coincide, that is, $k_2 c_{s-1} + \cdots + k_{s+1} c_0 = d c_s \neq 0$.

Therefore, there is at least one k_i among $k_2, ..., k_{s+1}$ which is non-zero and comparing coefficients at degree $s+i$ we get $c_s \cdot k_i + \sum_{n=1}^{s} k_{i+n} c_{s-n} = 0$.

This is a contradiction since $c_s \cdot k_i > 0$ and $\sum_{n=1}^{s} k_{i+n} c_{s-n} \geq 0$, which completes the proof. □

Theorem 2.17. *Suppose that with the previous notation we have $k_n \leq s_n$ for all n. If all coefficients of the inverse series $(1 - dt + \sum_{n=2}^{\infty} s_n t^n)^{-1}$ are non-negative, then Q is infinite-dimensional.*

Proof. If $k_n \leq s_n$ for all n, then $1 - dt + \sum_{n=2}^{\infty} k_n t^n \leq 1 - dt + \sum_{n=2}^{\infty} s_n t^n$. We get that $(1 - dt + \sum_{n=2}^{\infty} s_n t^n)(\sum_{n=0}^{\infty} b_n t^n) \geq 1$ and we are in the same case as in Theorem 2.16. □

Let us now think about bounding k_n with the same constant s for all n. This will provide us a very interesting result, although it is not the final sufficient condition we are looking for.

Corollary 2.18. *If $k_n \leq (d-1)^2/4$ for all n, then Q is infinite-dimensional.*

Proof. Assume $k_n \leq s$ for all n. We are going to compute the inverse series of $1 - dt + \sum_{n=2}^{\infty} st^n$ explicitly, for which we need to recall from Calculus I the power series representations for some particular functions, such as

$$\frac{1}{1-at} = \sum_{n=0}^{\infty} (at)^n \quad \text{and} \quad \frac{1}{(1-at)^2} = \sum_{n=0}^{\infty} n(at)^{n-1}. \tag{2.1}$$

Since we are working with formal power series we do not worry about convergence. Then:

$$\begin{aligned} 1 - dt + \sum_{n=2}^{\infty} st^n &= (1 + t + t^2 + \cdots) - (d+1)(t + t^2 + \cdots) + (s+d)(t^2 + t^3 + \cdots) \\ &= \frac{1}{1-t} - (d+1)\frac{t}{1-t} + (s+d)\frac{t^2}{1-t} = \frac{1 - (d+1)t + (s+d)t^2}{1-t}. \end{aligned}$$

Therefore, if we take $s = (d-1)^2/4$ we get $(1 - \frac{d+1}{2}t)^2$ in the numerator, hence the inverse series is $(1-t)(1-\frac{d+1}{2}t)^{-2}$. Using (2.1) and making all the calculations (see Problem 8) we obtain the following series:

$$\begin{aligned} \left(1 - dt + \sum_{n=2}^{\infty} st^n\right)^{-1} &= \sum_{n=0}^{\infty} \left(\frac{d+1}{2}\right)^{n-1} \left[\frac{(d+1) + n(d-1)}{2}\right] t^n \\ &= 1 + dt + \frac{3d^2 + 2d - 1}{4} t^2 + \cdots \end{aligned}$$

All the coefficients of the inverse series are non-negative if $k_n \leq s = (d-1)^2/4$ for all n. Thus, Q is infinite-dimensional by Theorem 2.17. □

We want to go just one step further and in greater generality our goal is to bound $k_n \leq s_n$ in such a way that $1 - dt + \sum_{n=2}^{\infty} s_n t^n = (1-\lambda t)^2/(1-\mu t)$. If we multiply both sides by $1-\mu t$ we get

$$1 - (\mu + d)t + (s_2 + \mu d t^2) + \sum_{n=3}^{\infty} (s_n - \mu s^{n-1})t^n = 1 - 2\lambda t + \lambda^2 t^2$$

and if we equal the coefficients of t^n on both sides we get the following system of equations:

$$\begin{cases} 2\lambda = d + \mu \\ s_2 + \mu d = \lambda^2 \\ s_n = \mu s_{n-1}, \text{ for } n \geq 3 \end{cases} \iff \begin{cases} s_2 = (d - \mu/2)^2 \\ s_n = \mu^{n-2}(d-\mu/2)^2, \text{ for } n \geq 2 \end{cases} .$$

Thus, we get that $s_n = \epsilon^2(d-2\epsilon)^{n-2}$ for all $n \geq 2$, where $\epsilon = d - \mu/2$. Now, if we fix $0 < \epsilon < d$ so that $d - \epsilon > 0$ and again using (2.1) we get the expansion of the inverse series as a power series with non-negative coefficients (see Problem 8):

$$\left(1 - dt + \sum_{n=2}^{\infty} s_n t^n\right)^{-1} = \frac{1-(d-2\epsilon)t}{(1-(d-\epsilon)t)^2} = \sum_{n=0}^{\infty} (d-\epsilon)^{n-1}[d + (n-1)\epsilon]t^n.$$

Corollary 2.19. *If $k_n \leq \epsilon^2(d-2\epsilon)^{n-2}$ for all n with a fixed $0 < \epsilon < d$, then Q is infinite-dimensional.*

Thus, we have found a suitable sufficient condition for Q to be infinite-dimensional. Now, we want $A = T/N$ to be a nil algebra in order to get a negative solution to the Kurosh-Levitzki problem. In other words, for every polynomial $f \in T$ there must be an exponent m such that $f^m \in N$. In order to get that, we fix ϵ such that $0 < \epsilon < (d-1)/2$, hence $d - 2\epsilon > 1$ and we construct our ideal N recursively defining $N_{(k)} = (f_1, \ldots, f_{n_k})$, such that the following properties are satisfied:

(i) For every $f \in T$ of degree at most k, some power of f is in $N_{(k)}$.

(ii) In the sequence of homogeneous polynomials $f_1, \ldots, f_{n_k}$, the number of polynomials of degree n is less or equal to $\epsilon^2(d-2\epsilon)^{n-2}$ for all n.

We are going to proceed by induction on k. We first fix ϵ and we construct the sequence $\{f_1, f_2, \ldots\}$ step by step. For the base case $k = 0$, we do not have any non-zero constant polynomial $f \in T$, then it suffices to set $N_{(0)} = \{0\}$.

Now, by induction hypothesis let $N_{(k-1)} = (f_1, \ldots, f_{n_{k-1}})$ such that for every polynomial $f \in T$ with degree less than k, we have that $f^m \in N_{(k-1)}$ for some m. We need to prove that adding some new homogeneous polynomials of higher degree than the previous ones, let us denote them by

$f_{n_{k-1}+1}, \ldots, f_{n_k}$, for every $f \in T$ of degree k we can find a suitable M such that $f^M \in N_{(k)} = (f_1, \ldots, f_{n_k})$. Moreover, we have to take into account that the number of new polynomials of degree n we add must be less than $\epsilon^2(d-2\epsilon)^{n-2}$ for every n.

Let $f = c_1 X_1 + \cdots + c_d X_d + c_{1,2} X_1 X_2 + \cdots + c_{d,\ldots,d} X_d^k$ be the general expression for a polynomial of degree k belonging to T, and thus $f^M = c_1^M X_1^M + \cdots + c_{d,\ldots,d}^M X_d^{kM}$. We have to see it from another point of view, considering the coefficients as commuting indeterminates and f^M as a homogeneous polynomial of degree M in $c_1, \ldots, c_{d,\ldots,d}$ whose coefficients are homogeneous polynomials in $X_1, \ldots, X_d$. These homogeneous polynomials are indeed the ones we are going to add to the sequence.

Let us understand it with an example. For $f = c_1 X_1 + c_2 X_2 + c_3 X_3$, we would have the following expression for the square:

$$\begin{aligned} f^2 = {} & c_1^2 X_1^2 + c_2^2 X_2^2 + c_3^2 X_3^2 + c_1 c_2 (X_1 X_2 + X_2 X_1) \\ & + c_1 c_3 (X_1 X_3 + X_3 X_1) + c_2 c_3 (X_2 X_3 + X_3 X_2). \end{aligned} \tag{2.2}$$

This particular case for $d = 3$, $k = 1$ and $M = 2$ is a second degree polynomial in commuting variables c_1, c_2 and c_3. Thus, the set of new homogeneous polynomials we would add to the sequence is

$$\{X_1^2,\ X_2^2,\ X_3^2,\ X_1X_2 + X_2X_1,\ X_1X_3 + X_3X_1,\ X_2X_3 + X_3X_2\}$$

which correspond to the homogeneous polynomials in X_1, X_2 and X_3 which are regarded as coefficients.

Observe that if $M > \max\{\deg f_i \mid i = 1, \ldots, n_{k-1}\}$, the degrees of the new homogeneous polynomials we add to the sequence are between M and kM, hence we only have to care about the number of new polynomials of degree n to be bounded by $\epsilon^2(d-2\epsilon)^{n-2}$ for all n between M and kM. But which is the total number of new homogeneous polynomials we are adding to the sequence? The answer to this problem is easy.

The total number of homogeneous polynomials we are adding is exactly the same as the number of monomials of degree M in the commuting indeterminates $c_1, \ldots, c_{d,\ldots,d}$. This equality holds since there is a one-to-one correspondence between these monomials and their respective coefficients in f^M, which are homogeneous polynomials in $X_1, \ldots, X_d$ of degree between M and kM. Let us consider the previous example $f = c_1 X_1 + c_2 X_2 + c_3 X_3$. Then, looking at (2.2) we observe that the homogeneous polynomial corresponding to c_1^2 is X_1^2 and the one corresponding to $c_1 c_2$ is $X_1 X_2 + X_2 X_1$, among others.

Let $c_1, \ldots, c_{d,\ldots,d}$ be the commuting indeterminates and $q = d + d^2 + \cdots + d^k$ be the total number of them. In order to compute the number of monomials of degree M in these q commuting indeterminates, we can draw a parallel

with a well-known combinatorial problem which consists in computing the total number of ways in which M identical balls can be distributed into q different boxes. In this case, the M balls would be the degrees we have to distribute among $c_1, \ldots, c_{d,\ldots,d}$ since we look for monomials of degree M and the q boxes correspond to these commuting indeterminates. We know from the second year course of Discrete Mathematics that the solution to this problem is $\binom{M+q-1}{q-1}$.

At the beginning we have fixed ϵ in such a way that $d - 2\epsilon > 1$ and q is also fixed for each k-th step of the induction. Then, there exists an M big enough such that $(M+q-1)^{q-1} \leq \epsilon^2(d-2\epsilon)^{M-2}$. This happens because $d - 2\epsilon > 1$ and M is in the exponent whereas in $(M+q-1)^{q-1}$ it is in the base. Therefore, when M tends to infinity $\epsilon^2(d-2\epsilon)^{M-2}$ increases faster than $(M+q-1)^{q-1}$ and we have the following chain of inequalities:

$$\binom{M+q-1}{q-1} \leq (M+q-1)^{q-1} \leq \epsilon^2(d-2\epsilon)^{M-2} \leq \epsilon^2(d-2\epsilon)^{n-2}.$$

This way, the number of new homogeneous polynomials of degree n we add is bounded by $\epsilon^2(d-2\epsilon)^{n-2}$ for all n between M and kM, hence we have completed our construction at the k-th step.

Let $\{f_{n_{k-1}+1}, \ldots, f_{n_k}\}$ be the set of new homogeneous polynomials we add at the k-th step and $N_{(k)} = (f_1, \ldots, f_{n_k})$. Then,

$$N_{(0)} \subset N_{(1)} \subset N_{(2)} \subset \cdots \subset N_{(k)} \subset \cdots$$

is an ascending chain of recursively defined ideals. Let us take $N = \bigcup_{k=0}^{\infty} N_{(k)}$. Then, $A = T/N$ is a nilpotent finitely generated nil algebra, hence a negative solution to the Kurosh-Levitzki problem. These algebras are known as Golod-Shafarevich algebras.

Now, let A be a Golod-Shafarevich algebra over a field K of characteristic p generated by $\overline{X_1}, \ldots, \overline{X_d}$ as a K-algebra and let $Q = K \oplus A$. Then, we have previously seen at Section 2.2 that $G = \langle \overline{1+X_1}, \ldots, \overline{1+X_d} \rangle \leq Q^\times$ is a finitely generated infinite p-group, hence a negative solution to the General Burnside Problem. These groups are called Golod-Shafarevich groups.

Chapter 3

Gupta-Sidki and Grigorchuk groups

In the last chapter we introduce some other negative solutions to the General Burnside Problem which are known as Gupta-Sidki and Grigorchuk groups. The construction of these particular groups is done by using graph theory and groups of automorphisms of p-adic rooted trees acting on the set of vertices of the trees. These constructions show the existence of finitely generated infinite p-groups for some fixed prime p.

In the case of the Gupta-Sidki group p is an odd prime, whereas for Grigorchuk groups $p = 2$. The orders of the elements in these groups are unbounded, hence we cannot consider them as negative answers to the Burnside problem where the torsion is bounded.

The main references we have followed in the first two sections of this chapter are the notes provided by my supervisor [6, 7]. In Section 3.3 we have also followed [10] and [16].

3.1 Groups of automorphisms of p-adic rooted trees

Let us first define p-adic rooted trees, the groups of automorphisms of these trees and some basic notions in order to introduce Gupta-Sidki and Grigorchuk groups. All the concepts of this section can be generalised for any integer $p \geq 2$, in fact, the particular cases $p = 2$ and $p = 4$ correspond to Grigorchuk groups, which are introduced later in Section 3.3.

However, in order to construct the Gupta-Sidki group associated to p we want p to be an odd prime, so from now on we assume p is an odd prime. Let us first define some basic concepts about p-adic rooted trees.

Definition 3.1. A tree $\mathcal{T}$ is a connected graph with no cycles and we say that it is rooted if it has a special vertex labelled as the root, which is denoted by ϕ.

The root serves as a point of reference for other vertices in the tree, and we usually keep it at the top in order to list other vertices below it.

Definition 3.2. A vertex v is a descendant of u if u and v are connected and vertex v is listed below u, so that it belongs to the next level. Equivalently, we could say that u is an ancestor of v.

Definition 3.3. A rooted tree $\mathcal{T}$ is p-adic if every vertex has exactly p descendants.

From now on, we assume $\mathcal{T}$ is a p-adic rooted tree.

Definition 3.4. A path is a sequence $\{u_1, \ldots, u_k\}$ of vertices of $\mathcal{T}$ starting at the root such that u_{i+1} is a descendant of u_i for $i = 1, ..., k-1$.

Regarding the notation, in order to label the vertices of the p-adic tree we use the alphabet X which is customary to take as $X = \mathbb{Z}/p\mathbb{Z} = \{0, 1, \ldots, p-1\}$. Then, we form words in X and using string concatenation the set of all finite words is the free monoid X^*, as we have seen in Definition 2.3. These words represent vertices of the tree, in fact, the root is labelled as the empty word ϕ and its p descendants form the first level which corresponds to X. Then, each of these p vertices has p descendants and so on, hence in total there are p^n vertices on the n-th level.

We denote by $X^n = \{x_1 x_2 \ldots x_n \mid x_i \in X\}$ the set of all words of length n, which represent vertices on the n-th level.

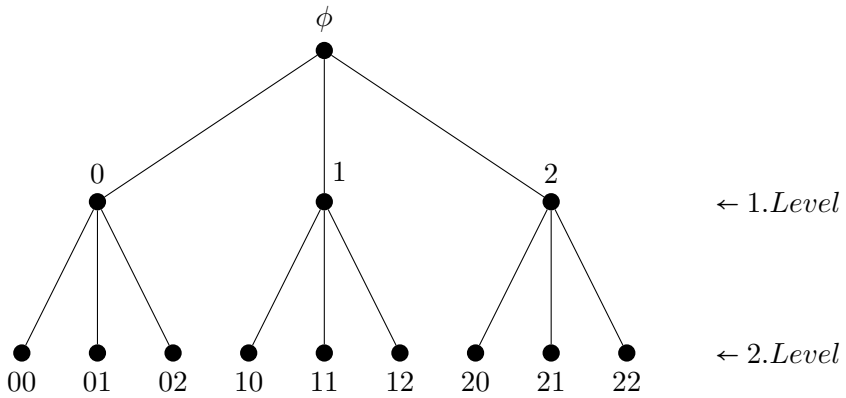

Figure 3.1: First two levels of a 3-adic rooted tree.

Definition 3.5. An automorphism f of $\mathcal{T}$ is a bijection of the set of vertices $V(\mathcal{T})$ that preserves incidence, that is, if u and v are connected, then $f(u)$ and $f(v)$ also are.

The set of all automorphisms, $\mathrm{Aut}\,\mathcal{T}$, is a group with respect to composition.

Example 3.6. These are two examples of automorphisms of a p-adic tree:

(i) The identity map is an automorphism.

(ii) *Rooted automorphisms*: the automorphism f_σ permutes rigidly the main subtrees according to a permutation σ of S_p, such that $f_\sigma(\phi) = \phi$ and $f_\sigma(xv) = \sigma(x)v$, for all $x \in X, v \in X^*$. Notice that the order of f_σ is equal to the order of $\sigma \in S_p$.

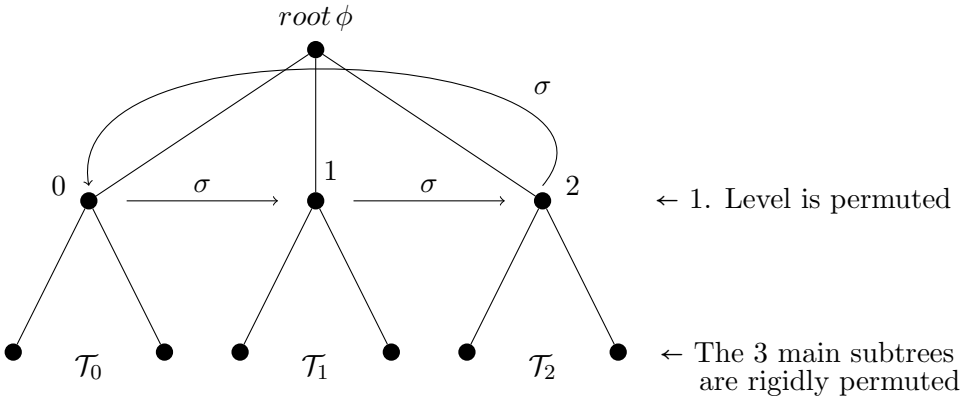

Figure 3.2: Rooted automorphism of the 3-adic tree corresponding to $\sigma = (0, 1, 2) \in S_3$.

Proposition 3.7 (General properties of automorphisms)**.** *If $f \in \mathrm{Aut}\,\mathcal{T}$, then:*

(i) $f(\phi) = \phi$*;*

(ii) $f(X^n) = X^n \ , \ \forall n \geq 0$ *(the n-th level is fixed);*

(iii) *f sends a descendant of u to a descendant of $f(u)$, so it sends paths to paths;*

(iv) *If we know the image of a vertex u, then we know the images of all vertices in the path from ϕ to u.*

Proof. Observe that every vertex is connected to $p+1$ vertices (its ancestor and its p descendants), except for ϕ and that is why $f(\phi) = \phi$.

Let us prove $f(X^n) = X^n$ by induction on n. For the base case $n = 0$ we are in (i) so we assume it is true up to $n-1$ and we have to prove it for n. Since X^n is finite and f is bijective, then it suffices to prove that $f(X^n) \subseteq X^n$.

Let $v \in X^n$ and $u \in X^{n-1}$ be connected, so that v is a descendant of u. Since f preserves incidence, then $f(v)$ is either the ancestor or a descendant of $f(u) \in X^{n-1}$. By induction hypothesis we know that $f(X^{n-2}) = X^{n-2}$ and since f is bijective, if $f(v) \in X^{n-2}$, i.e., $f(v)$ is the ancestor of $f(u)$, then we have that also $v \in X^{n-2}$ which is a contradiction. Thus, $f(v) \in X^n$ is a descendant of $f(u)$ and we have also proved the third property.

Finally, the fourth property follows from the third one, since we are sending paths to paths. If we know $f(u)$, it follows that its ancestor is the image of the ancestor of u and recursively we get the images of all vertices in the path from ϕ to u. □

Once we have defined automorphisms of $\mathcal{T}$ and their properties, the idea is to describe them without giving explicitly all images of all vertices. In order to do that, we have to introduce the concepts of the label and the portrait of an automorphism.

Let $u \in X^*$ be an arbitrary vertex and its descendants are of the form ux where $x \in X = \{0, 1, \ldots, p-1\}$ and let also $f \in \operatorname{Aut}\mathcal{T}$. Then, there exists a permutation $\alpha \in S_p$ such that $u0$ is mapped to $f(u)\,\alpha(0)$, $u1$ is mapped to $f(u)\,\alpha(1)$ and so on.

Definition 3.8. We call such α the label of f at the vertex u, and we denote it by $\alpha = f_{(u)}$. The collection of the labels of f at all vertices is called the portrait of f.

The main formula to use labels is $f(ux) = f(u)f_{(u)}(x)$, for all $u \in X^*$ and $x \in X$. This leads to a general expression for every vertex $u = x_1x_2\cdots x_n$ in the tree:

$$f(u) = f_{(\phi)}(x_1)\, f_{(x_1)}(x_2)\, f_{(x_1x_2)}(x_3) \ldots f_{(x_1x_2\cdots x_{n-1})}(x_n). \tag{3.1}$$

Conversely, any portrait on the tree defines an automorphism by using (3.1) to define $f(u)$ for every vertex u. Thus, there is a one-to-one correspondence between automorphisms of the tree and portraits.

Definition 3.9. The stabilizer of the n-th level, which is denoted by $\operatorname{st}(n)$, is the set of all automorphisms in $\operatorname{Aut}\mathcal{T}$ such that all vertices up to the n-th level (included) are fixed.

In addition, $\operatorname{st}(n)$ is a normal subgroup of $\operatorname{Aut}\mathcal{T}$ of finite index, which is proved in Problem 10.

Definition 3.10. Let $f \in \operatorname{Aut}\mathcal{T}$ and $u \in X^*$ be a vertex of $\mathcal{T}$. Then, the section of f at u, denoted by f_u, is the automorphism of $\mathcal{T}$ defined by $f(uv) = f(u)f_u(v)$ for all $v \in X^*$.

Example 3.11. Let us show two examples of portraits where sections can be easily identified. On the left-hand side (see Figure 3.3) we have the portrait of the rooted automorphism corresponding to $\sigma = (0, 1, \ldots, p-1)$, whereas on the right-hand side (see Figure 3.4) the one of an automorphism belonging to $\mathrm{st}(n)$, which clearly fixes the first n levels.

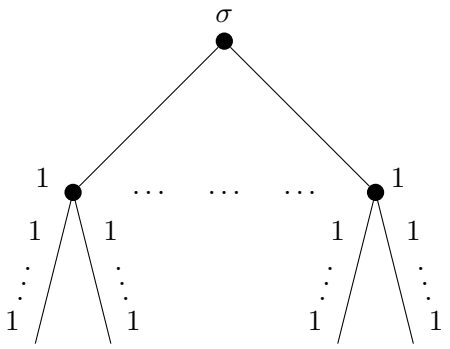

Figure 3.3: Portrait of the rooted automorphism corresponding to $\sigma = (0, 1, \ldots, p-1)$.

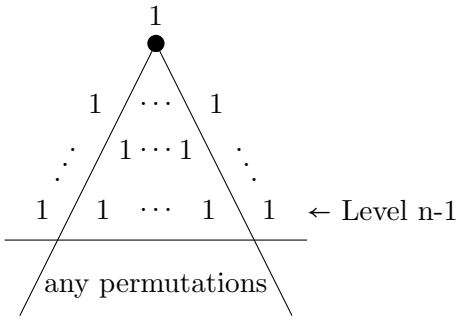

Figure 3.4: Portrait of an arbitrary automorphism belonging to $\mathrm{st}(n)$.

Proposition 3.12. *Let $f \in \mathrm{st}(n)$ for any $n \geq 1$. Then, the following map is a group isomorphism:*

$$\begin{aligned} \psi_n : \mathrm{st}(n) &\longrightarrow \mathrm{Aut}\,\mathcal{T} \times \overset{p^n}{\cdots} \times \mathrm{Aut}\,\mathcal{T} \\ f &\longmapsto (f_u)_{u \in X^n} \end{aligned}$$

Proof. Let us first prove ψ_n is a group homomorphism. We have to check whether $\psi_n(fg) = \psi_n(f)\psi_n(g)$, or equivalently, $(fg)_u = f_u g_u$. We know from Problem 9 that $(fg)_u = f_u\, g_{f(u)}$ and since in this case $f(u) = u$ for every $u \in X^n$, then we are done.

Let us now prove it is a bijection. It is obvious that $\ker \psi_n = \{1_{\mathrm{Aut}\,\mathcal{T}}\}$, since it consists of the set of automorphisms in $\mathrm{st}(n)$ such that the section at every vertex on the n-th level is $1_{\mathrm{Aut}\,\mathcal{T}}$. By definition of $\mathrm{st}(n)$, only $1_{\mathrm{Aut}\,\mathcal{T}}$ satisfies this property (see Figure 3.4), hence ψ_n is injective. Let us prove ψ_n is also surjective. For each vertex u on the n-th level, we choose an arbitrary automorphism of $\mathcal{T}$ as f_u. We have seen that portraits fully determine automorphisms, so we are done since for levels up to n we have identity elements as labels in the portrait (see again Figure 3.4), and consequently $f \in \mathrm{st}(n)$. □

In order to construct the Gupta-Sidki group for each odd prime p in Section 3.2, we need to introduce one more thing: recursively defined automorphisms.

Let $b \in \mathrm{st}(1)$ be defined by $\psi_1(b) = (a, a^{-1}, 1, \ldots, 1, b)$, where a is the rooted automorphism corresponding to $\sigma = (0, 1, \ldots, p-1)$. This recursive definition makes sense and it defines an automorphism since we can draw its portrait (see Figure 3.5).

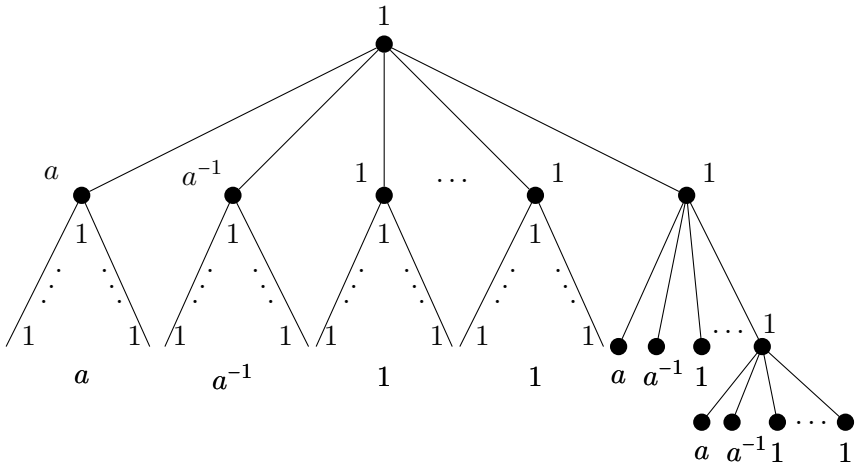

Figure 3.5: Portrait of $b \in \mathrm{st}(1)$, which is recursively defined by $\psi_1(b) = (a, a^{-1}, 1, \ldots, 1, b)$.

As we can observe in the picture above, the first level is fixed and hence $b \in \mathrm{st}(1)$. In addition, the descendants of the rightmost vertex are fixed in every level and the same pattern is repeated recursively.

3.2 Gupta-Sidki group

The principal aim of this section is to introduce the associated Gupta-Sidki group for each odd prime p and prove that it is a finitely generated infinite p-group, hence a negative solution to the General Burnside Problem. The construction of the Gupta-Sidki groups was due to Narain Gupta and Said Sidki [9] in 1983.

The Gupta-Sidki groups have some remarkable properties such as being just-infinite and residually finite, that is, they are infinite groups with all proper quotients finite and the intersection of all their normal subgroups of finite index is trivial.

Let p be an odd prime and let us define the associated Gupta-Sidki group, for which we are going to use the notions of the previous Section 3.1.

Definition 3.13. Let $\mathcal{T}$ be a p-adic rooted tree. The Gupta-Sidki group is the finitely generated group $G = \langle a, b \rangle \leq \mathrm{Aut}\,\mathcal{T}$, where a is a rooted automor-

phism corresponding to $\sigma = (0, 1, \dots, p-1) \in S_p$ and $b \in \mathrm{st}(1)$ is recursively defined by $\psi_1(b) = (a, a^{-1}, 1, \dots, 1, b)$.

On the one hand, the order of a is equal to the order of $\sigma \in S_p$ since it is rooted, so $o(a) = o(\sigma) = p$. On the other hand, since ψ_1 is a group homomorphism we know that $\psi_1(b^p) = (1, 1, \dots, b^p)$ and looking to its portrait this defines the identity map. Thus, $o(b)$ divides p, but since b is not the identity map and p is a prime, then $o(b) = p$.

Let us now introduce some basic concepts we will need for the next proposition.

Definition 3.14. Let G be a group and let H and K be two subgroups of G. Let also $N_G(H) = \{x \in G \mid H^x = H\}$ be the normalizer of H in G. Then, we say that K normalizes H if $K \leq N_G(H)$.

In particular, H is a normal subgroup of G if and only if G normalizes H, that is, $N_G(H) = G$. This is trivial since by definition $H \trianglelefteq G$ if $H^g = H$ for all $g \in G$, which is equivalent to saying that $g \in N_G(H)$.

Definition 3.15. Let $G = \langle X \rangle$ be a group and $H \leq G$. We define the normal closure of H as the smallest normal subgroup of G containing H, which we denote by H^G.

It is easy to show that $H^G = \langle h^g \mid h \in H, g \in G \rangle$, which we are going to leave as a problem (see Problem 13).

Let us now fix some notation. We will denote $b^{a^i} = b_i$ for every $i \in \mathbb{Z}$. Since a has order p, if $i \equiv j \pmod{p}$, then $b_i = b_j$.

Proposition 3.16. *Let* $\mathrm{st}_G(1) = G \cap \mathrm{st}(1) \trianglelefteq G$. *Then,* $G = \langle a \rangle \ltimes \mathrm{st}_G(1)$ *where* $\mathrm{st}_G(1) = \langle b \rangle^G = \langle b_0, b_1, \dots, b_{p-1} \rangle$.

Proof. First of all, let us prove that $\mathrm{st}_G(1) = \langle b \rangle^G$. Since $b \in \mathrm{st}_G(1)$, then it is obvious that $\langle b \rangle^G \leq \mathrm{st}_G(1) \trianglelefteq G$, so it suffices to prove that the indices of $\langle b \rangle^G$ and $\mathrm{st}_G(1)$ in G are equal. We know that $G = \langle a, b \rangle$ and $b \in \mathrm{st}_G(1)$, hence $G/\mathrm{st}_G(1) = \overline{\langle a, b \rangle} = \langle \overline{a} \rangle$ and $|G : \mathrm{st}_G(1)| = o(\overline{a})$ divides $o(a) = p$. Thus, the index is either 1 or p, but since $\mathrm{st}_G(1) \neq G$ we get $|G : \mathrm{st}_G(1)| = p$. Similarly, we get that $|G : \langle b \rangle^G| = p$ and equality holds.

Secondly, we need to check that $\langle b \rangle^G = \langle b_0, b_1, \dots, b_{p-1} \rangle$. We know that $\langle b \rangle^G \supseteq \langle b_0, b_1, \dots, b_{p-1} \rangle$ so let us prove the other inclusion. It suffices to check that $N = \langle b_i \mid i \in \mathbb{Z} \rangle = \langle b_0, b_1, \dots, b_{p-1} \rangle$ is a normal subgroup of G, which is equivalent to proving that $G = \langle a, b \rangle = N_G(N)$. This is satisfied if $a, b \in N_G(N)$. On the one hand, $b \in N \subseteq N_G(N)$ and on the other hand, we have that $b_i^a = (b^{a^i})^a = b^{a^{i+1}} = b_{i+1} \in N$, hence $a \in N_G(N)$.

Finally, we need to prove that $G = \langle a \rangle \ltimes \mathrm{st}_G(1)$, that is, G can be decomposed as the internal semidirect product of $\langle a \rangle$ and $\mathrm{st}_G(1)$. These three conditions must be satisfied:

(i) $\langle a \rangle \leq G$ and $\mathrm{st}_G(1) \trianglelefteq G$;

(ii) $G = \langle a \rangle \cdot \mathrm{st}_G(1)$;

(iii) $\langle a \rangle \cap \mathrm{st}_G(1) = \{1\}$.

We already know that (i) is satisfied. The second condition is also fulfilled since $\langle b \rangle \leq \mathrm{st}_G(1) \trianglelefteq G$, so that $G = \langle a, b \rangle = \langle a, \mathrm{st}_G(1) \rangle = \langle a \rangle \cdot \mathrm{st}_G(1)$. Finally, the third condition also holds since $o(a) = p$ and $a \notin \mathrm{st}_G(1)$. □

The main goal of the last proposition is to prove that any $g \in G$ can be written as $g = a^i b_{j_1} b_{j_2} \cdots b_{j_r}$, with $i, j_1, ..., j_r \in \{0, 1, ..., p-1\}$. Let us call length of g to the smallest r for which such an expression exists and let us denote it by $l(g)$. Since $G = \langle a, b \rangle$, then g can be written as product of these generators (we do not need inverses since they have finite order), so in order to get the previous expression we move all a to the left hand side. Whenever we move a^i to the left through b we get $ba^i = a^i b^{a^i} = a^i b_i$. Therefore, for any $g = a^{i_1} b \cdots a^{i_d} b a^{i_{d+1}} \in G$ where all i_k lie in $\{0, ..., p-1\}$, the length of g is at most d.

Theorem 3.17. *The Gupta-Sidki group $G = \langle a, b \rangle$ is an infinite p-group.*

Proof. We are going to split the proof in two parts. Firstly, we are going to prove that G is infinite and in the second part that G is indeed a p-group.

Let $\psi_1 : \mathrm{st}(1) \longrightarrow \mathrm{Aut}\,\mathcal{T} \times \overset{p}{\cdots} \times \mathrm{Aut}\,\mathcal{T}$, where $\psi_1(b) = (a, a^{-1}, 1, \ldots, 1, b)$ and let ψ be its restriction to $\mathrm{st}_G(1)$. Then, it suffices to give the images of the generators of $\mathrm{st}_G(1)$ in order to define ψ completely, which are given by $(b^a)_x = b_{\sigma^{-1}(x)}$ as we have proved in Problem 9. Observe that all the components of the images, which are p-tuples, are in G for each generator. Then, this holds true for every $g \in \mathrm{st}_G(1)$.

$$\begin{aligned} \psi : \mathrm{st}_G(1) &\longrightarrow G \times G \times \overset{p}{\cdots} \times G \\ b_0 &\longmapsto (a, a^{-1}, 1, \ldots, 1, b) \\ b_1 &\longmapsto (b, a, a^{-1}, 1, \ldots, 1) \\ &\vdots \\ b_{p-1} &\longmapsto (a^{-1}, 1, \ldots, 1, b, a) \end{aligned}$$

Let $\pi_1 : G \times \overset{p}{\cdots} \times G \longrightarrow G$ be the projection on the first component and consider the composition $\pi_1 \circ \psi$. Our claim is that this composition is a

surjective map. It suffices to prove that a and b have preimages, which is trivial since $(\pi_1 \circ \psi)(b_0) = a$ and $(\pi_1 \circ \psi)(b_1) = b$. Thus, G must be infinite since $|G : \mathrm{st}_G(1)| = p$ and $\pi_1 \circ \psi : \mathrm{st}_G(1) \longrightarrow G$ is surjective. Otherwise, both groups would be finite and $|\mathrm{st}_G(1)| \geq |G|$, which is a contradiction.

In the second part of the proof we are going to show that G is a p-group. Let us prove that $o(g)$ is a power of p by induction on $l(g)$, the length of g. For the base case $l(g) = 0$, we have that $g = a^i$ so $o(g)$ is either 1 or p. Now, we assume it is true up to $l(g) < r$ and we have to prove it for $l(g) = r$. We consider two different cases for $g \in G$ of length r.

The first case is that $g \in \mathrm{st}_G(1)$. Then, g can be represented as a word in $b_0, ..., b_{p-1}$, that is, $g = w(b_0, ..., b_{p-1})$ and $\psi(g) = (w(a, b, 1, ..., 1, a^{-1}), ...)$. For each of these components, the length is at most the number of b's that appear in the word representing it. If the length of the first component is less than r, by induction hypothesis it has p-power order. Otherwise, all the b_i arising in the word that represents g are equal to b_1, since $(\pi_1 \circ \psi)(b_1) = b$. In that case $g = b_1^r$ and since $o(b_1) = p$, then $o(g)$ is either 1 or p. We use the same argument for the rest of the components.

The order of $\psi(g)$ is the least common multiple of the orders of all its components, in this case the maximum among them. Thus, the order of $\psi(g) = \psi_1(g)$ is a p-power. Moreover, since ψ_1 is a group isomorphism, then $o(g) = o(\psi_1(g))$ and we are done.

The second case is that $g \notin \mathrm{st}_G(1)$. Let $g = a^i b_{j_1} b_{j_2} \cdots b_{j_r}$, such that $i, j_1, ..., j_r \in \{0, 1, ..., p-1\}$ and $i \neq 0$. Since $\mathrm{st}_G(1) \trianglelefteq G$ and $|G : \mathrm{st}_G(1)| = p$, then $g^p \in st_G(1)$. Let us compute g^p:

$$g^p = (a^i b_{j_1} b_{j_2} \cdots b_{j_r}) \overset{p}{\cdots} (a^i b_{j_1} b_{j_2} \cdots b_{j_r}) = b_{j_1 + i(p-1)} \cdots b_{j_r + i(p-1)} \cdots b_{j_1} \cdots b_{j_r}$$

where for each j_k have that $\{j_k, j_k + i, j_k + 2i, ..., j_k + i(p-1)\} = \{0, 1, ..., p-1\}$ in $\mathbb{Z}/p\mathbb{Z}$. This holds because multiplication by $i \not\equiv 0 \pmod{p}$ and addition of j_k are indeed bijections in $\mathbb{Z}/p\mathbb{Z}$. Therefore, we deduce that $\{b_{j_k}, b_{j_k+i}, b_{j_k+2i}, ..., b_{j_k+i(p-1)}\} = \{b_0, b_1, ..., b_{p-1}\}$ for every $k \in \{1, 2, ..., r\}$, and hence g^p is a product of $b_0, ..., b_{p-1}$ in some order, where each b_i appears exactly r times, once per each j_k. Thus, every component of $\psi_1(g^p) = \psi(g^p)$ is a product of r times a, r times a^{-1} and r times b, in some order. This implies that each of these components can be written in the form $b_{i_1} b_{i_2} \cdots b_{i_r}$ so that they belong to $\mathrm{st}_G(1)$, with length at most r. By the first case, all components have order a power of p, hence $o(g^p)$ is a power of p and so is $o(g)$. □

More generally, one can consider $G = \langle a, b \rangle$ where a is a rooted automorphism which corresponds to $\sigma = (0, 1, \ldots, p-1)$ as before and $b \in \mathrm{st}(1)$ is recursively defined by $\psi_1(b) = (a^{e_1}, a^{e_2}, \ldots, a^{e_{p-1}}, b)$ with $e_1, e_2, ..., e_{p-1} \in \mathbb{Z}/p\mathbb{Z}$, not all zero. Previous proofs for the Gupta-Sidki group can be adapted in

order to show G is always infinite, and that G is periodic if and only if $e_1 + e_2 + \cdots + e_{p-1} = 0$ in $\mathbb{Z}/p\mathbb{Z}$. These are called GGS-groups, i.e., Grigorchuk-Gupta-Sidki groups. Observe that for $p = 2$, the condition $e_1 = 0$ cannot happen, so we do not obtain a periodic group in this case.

3.3 Grigorchuk groups

In the last section we introduce Grigorchuk groups, in particular the first Grigorchuk group (also known simply as Grigorchuk group), which was first constructed in a 1980 paper [8] by the mathematician Rostislav Grigorchuk providing another counterexample to the General Burnside Problem. In 1984 he proved that this group has intermediate growth, that is, faster than polynomial but slower than exponential. In fact, this was the first finitely generated group proven to show such growth, answering the open problem posed by John Milnor in 1968 of whether such a group existed.

In order to find a finitely generated periodic infinite group, the first idea can be to think whether the same construction as for the Gupta-Sidki group is possible for $p = 2$. However, if $G = \langle a, b\rangle \leq \operatorname{Aut}\mathcal{T}$ such that $o(a) = o(b) = 2$, we obtain a dihedral group of order $2n$ where $n = o(xy)$ is possibly ∞, so either it is periodic and finite or infinite but not a periodic group. Thus, we have to think about other possibilities such as taking a third generator of G or increasing the order of the two generators. Based on this idea we introduce the Grigorchuk groups.

Definition 3.18. Let $\mathcal{T}$ be a 2-adic rooted tree. The first Grigorchuk group, also known simply as the Grigorchuk group, is the finitely generated group $\Gamma = \langle a, b, c, d\rangle \leq \operatorname{Aut}\mathcal{T}$, where a is a rooted automorphism corresponding to $\sigma = (0, 1) \in S_2$ and $b, c, d \in \operatorname{st}(1)$ are recursively defined in such a way that $\psi_1(b) = (a, c)$, $\psi_1(c) = (a, d)$ and $\psi_1(d) = (1, b)$.

Definition 3.19. Let $\mathcal{T}$ be a 4-adic rooted tree. The second Grigorchuk group is the finitely generated group $G = \langle a, b\rangle \leq \operatorname{Aut}\mathcal{T}$, where a is a rooted automorphism which corresponds to $\sigma = (0, 1, 2, 3) \in S_4$ and $b \in \operatorname{st}(1)$ is recursively defined by $\psi_1(b) = (a, 1, a, b)$.

We are not going focus on second Grigorchuk group, but it is indeed an infinite 2-group generated by two elements of order 4.

Let us go deeper into the first Grigorchuk group $\Gamma = \langle a, b, c, d\rangle$, where a, b, c and d are defined as above. These four generators are automorphisms of order 2 since $o(a) = o(\sigma) = 2$ and in case of b, c and d it suffices to look at portraits (see Figure 3.7 for b^2). In fact, $\psi_1(b^2) = (a^2, c^2) = (1, c^2)$, $\psi_1(c^2) = (1, d^2)$ and $\psi_1(d^2) = (1, b^2)$, hence $b^2 = c^2 = d^2 = 1$.

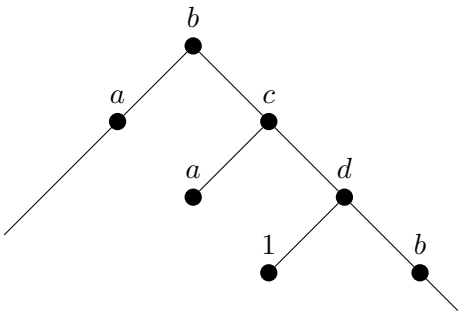

Figure 3.6: Portrait of $b \in \mathrm{st}(1)$ recursively defined as $\psi_1(b) = (a, c)$.

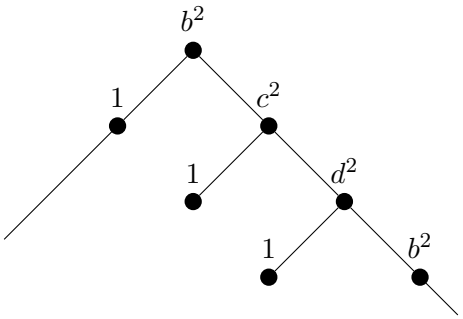

Figure 3.7: Portrait of $b^2 = 1$.

Portraits are very useful while working with recursively defined automorphisms since they provide us a lot information and they completely determine automorphisms. For instance, we can use them to check that b, c and d commute with each other and $bc = d = cb$, hence $\langle b, c, d\rangle = \{1, b, c, d = bc\}$ is an abelian group of order 4 isomorphic to $C_2 \times C_2$. From $bc = d$ we deduce that $dc = b$ and $bd = c$. Since b, c and d depend from each other, any two of them generate Γ along with a, hence $\Gamma = \langle a, b, c\rangle$ without loss of generality.

For the first Grigorchuk group we have similar results as for Gupta-Sidki groups. The first one is the analogue of Proposition 3.16, whose proof is given in Problem 14.

Proposition 3.20. *The first Grigorchuk group* Γ *can be decomposed as* $\Gamma = \langle a\rangle \ltimes \mathrm{st}_\Gamma(1)$ *where* $\mathrm{st}_\Gamma(1) = \langle b, c\rangle^\Gamma = \langle b, b^a, c, c^a\rangle$.

Theorem 3.21. *The first Grigorchuk group* Γ *is an infinite 2-group.*

Proof. First of all, let us prove that Γ is infinite. It suffices to see that $\pi_1 \circ \psi : \mathrm{st}_\Gamma(1) \longrightarrow \Gamma$ is surjective, where π_1 is the projection on the first component and $\psi : \mathrm{st}_\Gamma(1) \longrightarrow \Gamma \times \Gamma$ is given by $\psi(b) = (a, c)$, $\psi(b^a) = (c, a)$, $\psi(c) = (a, d)$ and $\psi(c^a) = (d, a)$.

In order to prove that $\pi_1 \circ \psi$ is surjective, it suffices to prove that a, c and d have preimages, since they generate Γ. This is trivial since $(\pi_1 \circ \psi)(b) = a$, $(\pi_1 \circ \psi)(b^a) = c$ and $(\pi_1 \circ \psi)(c^a) = d$. Thus, Γ is infinite.

In the second part of the proof we have to show that Γ is a 2-group. Let us prove it by induction on $l(g) = \min\{k \in \mathbb{N} \mid g = w_1 \cdots w_k,\ w_i \in \{a, b, c, d\}\}$. For $l(g) = 0$ it is trivial and for $l(g) = 1$ we know that all a, b, c and d are of order 2. Let us recall that $bc = d, dc = b$ and $bd = c$, hence every element in $\Gamma = \langle a, b, c, d\rangle$ can be written as a product of the generators, alternating a with either b, c or d over and over. Let us assume it is true for $l(g) < k$ and

we have to prove it for $l(g) = k$. We are going to prove it in different ways for k odd and for k even.

Let k be odd. Then, g is either of the form $g = aua = u^a$ or $g = uxv$ where $u, v \in \{b, c, d\}$. In the first case, $l(u) = k - 2$ and by induction hypothesis $o(u^a) = o(u)$ is a power of 2. In the second case, we have that $l(x) = k - 2$ and $g^u = ugu = xvu = xu'$ where $u' \in \{b, c, d\}$. By induction hypothesis, $o(g^u) = o(g)$ is a power of 2.

Let $k = 2l$ be even. Without loss of generality $g = aw_1aw_2 \cdots aw_l$, where $w_i \in \{b, c, d\}$. Otherwise, if $g = w_1aw_2 \cdots aw_la$ we take the conjugate by a and we are in the first case. We have to consider again two different cases.

If $l = 2m$ is even, then $g \in \mathrm{st}_\Gamma(1)$ since the number of a's is even and $g = w_1^a w_2 \cdots w_{2m-1}^a w_{2m}$. Thus, $\psi(g) = \psi(w_1^a)\psi(w_2) \cdots \psi(w_{2m}) = (g_0, g_1)$ where $l(g_0), l(g_1) < k$. By induction hypothesis both have order a power of 2, hence $o(g) = o(\psi(g)) = \mathrm{lcm}(o(g_0), o(g_1))$ is also a power of 2.

If $l = 2m - 1$ is odd, then $g \notin \mathrm{st}_\Gamma(1)$ since the number of a's is odd, but $g^2 \in \mathrm{st}_\Gamma(1)$, where $g^2 = w_1^a w_2 \cdots w_{2m-1}^a w_1 w_2^a \cdots w_{2m-1}$. Thus:

$$\psi(g^2) = \psi(w_1^a)\psi(w_2) \cdots \psi(w_{2m-1}^a)\psi(w_1) \cdots \psi(w_{2m-1}) = (g_0, g_1).$$

In total there are $k = 2(2m - 1)$ factors, which are either 1 or one of the generators, so that g_0 and g_1 have length at most k. Suppose that $w_j = d$ for some j, then $\psi(d) = (1, b)$ and $\psi(d^a) = (b, 1)$, which implies that in both elements g_0 and g_1 one of the k factors is the identity element and hence g_0 and g_1 have length less than k. Thus, the proof is completed by induction hypothesis since $o(g) = 2 \cdot o(g^2) = 2 \cdot o(\psi(g^2)) = 2 \cdot \mathrm{lcm}(o(g_0), o(g_1))$ is also a power of 2.

If $w_j \neq d$ for all j, then all w_j are either b or c. Looking at $\psi(b)$, $\psi(b^a)$, $\psi(c)$ and $\psi(c^a)$ we deduce that all the factors of g_0 and g_1 alternate between a and either c or d, depending on whether w_j is equal to b or c, respectively. Thus, g_0 and g_1 have length k and we cannot apply induction, but let us prove that their orders are a power of 2 and hence $o(g)$ also is. Suppose that $w_j = c$ for some j. Then, d is one of the factors of g_0 and g_1 so we are in the previous case for some factor being equal to d, hence we are done. In the worst case, that is, if $w_j = b$ for all j, then g_0 and g_1 are products of alternating factors a and c, in fact, $g_1 = g_0^a$. We are in the previous case for some factor being equal to c, and hence the proof is completed. □

Appendix A

Solved problems

A.1 Problems of Chapter 1

Problem 1. Prove that every finite p-group is nilpotent, given a prime number p.

Solution. Let G be a p-group, that is, the order of G is p^n for some $n \geq 0$. Then, let us prove by induction on n that G is nilpotent.

For the base case $n = 0$, $G = \{1\}$ is trivially nilpotent. Let us now assume it is true up to $n-1$ and we have to prove it for n. We are going to split the proof in two parts. Firstly, let us show that if $G \neq \{1\}$, then $Z(G) \neq \{1\}$ and secondly, we are going to prove that G is indeed nilpotent.

Let $n \geq 1$, hence $G \neq \{1\}$ is a finite non-trivial p-group. Let $Z(G)$ be the center of G and let $C_G(x) = \{g \in G \mid x^g = x\}$ be the centralizer of x in G. Let also $C_1, .., C_r$ be all the conjugacy classes of G of cardinality greater than 1 and let $x_1, ..., x_r$ be their representatives, respectively. Then, the class equation of G is:

$$|G| = |Z(G)| + \sum_{i=1}^{r} |G : C_G(x_i)|$$

where $|G : C_G(x_i)| = |C_i|$ is a power of p greater than 1, since $\{1\} \neq C_i \subseteq G$ for all $i = 1, ..., r$. Since $|G|$ is a power of p and the sum of $|C_i|$ is a multiple of p, the same happens with $|Z(G)|$. In particular, $|Z(G)| > 1$.

Now, let us prove that G is nilpotent. Since $\{1\} \neq Z(G) \trianglelefteq G$, then $G/Z(G)$ is a p-group of order p^r where $r < n$, hence by induction hypothesis it is nilpotent and let $c-1$ be its nilpotency class. Then, $\gamma_c(G/Z(G)) = \{\overline{1}\}$

and similarly as we have done in the proof of Theorem 1.8 we get that $\gamma_c(G) \leq Z(G)$. Then, $\gamma_{c+1}(G) = [\gamma_c(G), G] = \{1\}$ so that G is a nilpotent group, with nilpotency class c. □

Problem 2. Prove that any finite dihedral group D_{2n} is soluble for all $n \geq 1$, whereas it is nilpotent if and only if n is a power of 2.

Solution. Firstly, let us prove that a finite dihedral group D_{2n} is soluble for all $n \geq 1$. We have seen in Section 1.3 that $D_{2n} \cong \operatorname{Dih} A$, which can be constructed as an external semidirect product of cyclic groups $H = \langle y \rangle \cong C_2$ and $A = \langle x \rangle \cong C_n$. In this case, D_{2n} is an internal semidirect product of cyclic subgroups $\langle y \rangle$ and $\langle x \rangle$, where y is a reflection of order 2 and x is a rotation of order n. Its presentation is $D_{2n} = \langle x, y \mid x^n = y^2 = 1, x^y = x^{-1} \rangle$.

Let us take $\langle x \rangle \trianglelefteq D_{2n}$, so that the quotient $D_{2n}/\langle x \rangle = \langle \overline{y} \rangle$ is cyclic and hence abelian, just like $\langle x \rangle$. Therefore, D_{2n} is soluble taking the series of subgroups $\{1\} \trianglelefteq \langle x \rangle \trianglelefteq D_{2n}$.

Secondly, we have to prove that D_{2n} is nilpotent if and only if n is a power of 2. If n is a power of 2, then D_{2n} is a 2-group and we know from Problem 1 that it is nilpotent. On the other hand, let us prove that if n is not a power of 2, then D_{2n} is not nilpotent.

Let $n = 2^k m$ for some $k \geq 0$ and an odd m. Then, let us take $h = x^{2^k}$ and let $H = \langle h \rangle$. We know from the presentation of D_{2n} that $x^y = x^{-1}$, hence $h^y = h^{-1}$ and it is obvious that $h^x = h$. Thus, h^g is equal to h or h^{-1} for all $g \in D_{2n}$, which implies that $[h, g] = h^{-1}h^g$ is either 1 or h^{-2} for all $g \in D_{2n}$. Therefore, $[H, D_{2n}] = \langle h^2 \rangle \leq H$. Moreover, since x is of order $n = 2^k m$, then h has order m and H is a subgroup of order m. Observe that h^2 has also order m since $\gcd(m, 2) = 1$, hence $H = \langle h^2 \rangle$ and $[H, D_{2n}] = H$. Recursively, we get that $[H, D_{2n}, \overset{i}{\ldots}, D_{2n}] = H$ for all $i \geq 1$.

If D_{2n} is a nilpotent group with nilpotency class c, then we get that $[H, D_{2n}, \overset{c}{\ldots}, D_{2n}] \leq \gamma_{c+1}(D_{2n}) = \{1\}$, but we have just seen that it is H, hence D_{2n} cannot be nilpotent. □

Problem 3. Let G be a finitely generated group and let H be a subgroup of G of finite index. Prove that H is also finitely generated.

Solution. Let X be a finite generating system of G so that $G = \langle X \rangle$ and we consider $Y = X \cup X^{-1}$ and T a left transversal of H in G, i.e., every left coset of H contains exactly one element of T. Thus, $|T| = |G : H| < \infty$. Without loss of generality $1 \in T$.

Firstly, we prove that for every $y \in Y$, $t \in T$ there exist elements $t' \in T$ and $h_{y,t} \in H$ such that $y \cdot t = t' \cdot h_{y,t}$. Since left cosets of the subgroup H form a partition of G, then $y \cdot t \in G$ lies in exactly one left coset and since

T is a left transversal that coset is of the form $t'H$ for some $t' \in T$. Then, $y \cdot t \in t'H$ or equivalently $y \cdot t = t' \cdot h_{y,t}$ for some $h_{y,t} \in H$.

Secondly, we prove $\{h_{y,t} \mid y \in Y, t \in T\}$ is a generating set of H. Since Y is a generating set of G and $Y = Y^{-1}$, we can write any $h \in H$ as $h = y_1 \cdots y_r$ where $y_i \in Y$ for $i = 1, ..., r$. Let us now take y_r and since $1 \in T$ we could apply the previous property such that $y_r = y_r \cdot 1 = t_r \cdot h_{y_r,1}$ for some $t_r \in T$ and $h_{y_r,1} \in H$. Similarly, we get that $y_{r-1} \cdot t_r = t_{r-1} \cdot h_{y_{r-1},t_r}$ and recursively we continue with the procedure until we get $y_1 \cdot t_2 = t_1 \cdot h_{y_1,t_2}$.

Finally, we get that for an arbitrary element $h \in H$, $h = t_1 \cdot h_{y_1,t_2} \cdots h_{y_r,1}$, where $t_1 \in T$ and $h_{y_1,t_2}, ..., h_{y_r,1} \in H$. Since H is a subgroup of G, then $t_1 \in H$. In addition, since H itself is a left coset containing $1 \in T$, then $T \cap H = \{1\}$ and $t_1 = 1$. Thus, $\{h_{y,t} \mid y \in Y, t \in T\}$ is a generating set of H and $|\{h_{y,t} \mid y \in Y, t \in T\}| \leq |Y| \cdot |T| < \infty$, i.e., H is finitely generated. □

Problem 4. Let K be a field an let $G \subseteq GL(n, K)$ be a periodic group of finite exponent N. Prove that if $\operatorname{Char} K = p$ divides N, then G is not necessarily finite.

Solution. It suffices to find a counterexample. Let us take the unitriangular matrix group for $n = 2$, $UT(2, K) \leq GL(2, K)$:

$$UT(2,K) = \left\{ \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \middle| \lambda \in K \right\}, \text{ where } \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \lambda + \mu \\ 0 & 1 \end{pmatrix}.$$

Looking at how multiplication works in $UT(2, K)$, it is easy to check that $UT(2, K) \cong (K, +)$. Suppose K is infinite with $\operatorname{Char} K = p$, $K = \mathbb{F}_p(X)$ for instance. Then, $p \cdot a = 0$ for all $a \in K$. In fact, since p is prime every non-zero element $a \in K$ has order p and $\exp(K, +) = p$. Due to the isomorphism between $(K, +)$ and $UT(2, K)$ we deduce that $\exp(UT(2, K)) = p$, hence $UT(2, K) \leq GL(2, K)$ is a p-torsion group, but it is an infinite group. □

Problem 5. Let $G = \langle X \rangle$ be a group. Prove that $G' = \langle [x, y] \mid x, y \in X \rangle^G$.

Solution. Let $N = \langle [x, y] \mid x, y \in X \rangle^G \trianglelefteq G$ and let us prove that $N = G'$. Since $[x, y] \in G'$ and $G' \trianglelefteq G$, it is trivial that $N \subseteq G'$ so we have to prove the other inclusion. Let us take G/N and generators $\overline{x}, \overline{y} \in G/N$, with $x, y \in X$. Since $[x, y] \in N$, then $[\overline{x}, \overline{y}] = \overline{1}$, hence all generators commute with each other and G/N is abelian. Thus, by Theorem 1.7 we know that $G' \leq N$ and the proof is completed. □

A.2 Problems of Chapter 2

Problem 6. Prove that a finitely generated nil-algebra A is nilpotent if and only if A is finite-dimensional.

Solution. On the one hand, let A be nilpotent. Then, we have the following descending chain of ideals $A \supseteq A^2 \supseteq A^3 \supseteq \cdots \supseteq A^n = \{0\}$, for some $n \in \mathbb{N}$. If $a_1, ..., a_d$ generate A as an algebra, then the monomials of degree i in the generators $a_1, ..., a_d$ generate A^i/A^{i+1}, hence the dimension is finite for every $i = 1, ..., n-1$. The sum of all these dimensions is indeed the dimension of A, hence we are done.

On the other hand, if A is finite-dimensional, it suffices to see that it is impossible to get $A^i = A^{i+1} \neq \{0\}$, hence $A^i = A^j$ for all $j \geq i$. In particular, $A^i = A^{2i} = A^i \cdot A^i \neq \{0\}$. Then, take a left ideal $L \neq \{0\}$ of A^i, which is minimal with the property $A^i L = L$. Its existence is obvious since A^i is a left ideal satisfying this property. Take now $l \in L$ such that $A^i l \neq \{0\}$. Then, we get that $A^i(A^i l) = A^i l \subseteq L$, and by minimality of L equality holds, hence $L = A^i l$.

Therefore, $l = xl$ for some $x \in A^i$ and hence $(1-x)l = 0$. Observe that $(1-x)$ is invertible since x is nilpotent and $(1-x)(1+x+\cdots+x^{n-1}) = 1-x^n = 1$ for some $n \in \mathbb{N}$. This implies $l = 0$, which is a contradiction since $A^i l \neq \{0\}$ and we are done. □

Problem 7. Prove that every finitely generated infinite p-group G can be used to construct a negative solution to the Kurosh-Levitzki problem over any field K of characteristic p.

Solution. Let $G = \langle g_1, ..., g_d \rangle$. We are going to prove that the negative solution to the Kurosh-Levitzki problem is in particular the augmentation ideal of the group algebra $K[G]$, denoted by Δ, which is the set of elements $\sum_{g \in G} \lambda_g g \in K[G]$ such that $\sum_{g \in G} \lambda_g = 0$.

Then, since $K[G] = K \oplus \Delta$ and $\dim_k K[G] = |G| = \infty$ we get that $\dim_K \Delta = \infty$ and by the previous problem it is not nilpotent. Observe that $\Delta = \langle g - 1 \mid g \in G \rangle$ as a vector space over K. This can be easily proved taking into account the definition of the augmentation ideal and the general expression for elements in $\Delta \subseteq K[G]$, which is

$$\sum_{g \in G} \lambda_g g = \sum_{g \in G} \lambda_g (g-1) + \sum_{g \in G} \lambda_g = \sum_{g \in G} \lambda_g (g-1).$$

Since $G = \langle g_1, ..., g_d \rangle$ and $xy - 1 = (x-1)(y-1) + (x-1) + (y-1)$ for all $x, y \in G$, we deduce that $\Delta = K[g_1 - 1, ..., g_d - 1]$, that is, Δ is a finitely generated K-algebra.

It remains to prove that Δ is a nil algebra. An arbitrary element $x \in \Delta$ is of the form $\sum\limits_{g\in S} \lambda_g(g-1)$ where S is a finite subset of G.

We know that K has characteristic p and G is a p-group, so let us fix $p^m = \max\{o(g) \mid g \in S\}$. Then, we get the following:

$$x^{p^m} = \sum_{g\in S} \lambda_g^{p^m}(g^{p^m} - 1) = 0$$

and hence every element in Δ is nilpotent. Thus, the augmentation ideal Δ is a negative solution to the Kurosh-Levitzki Problem. □

Problem 8. Compute explicitly the inverse series of Corollary 2.18 and 2.19 in order to show that all coefficients are non-negative.

Solution. We will use the power series representations (2.1) from Chapter 2. Let us start with $(1 - dt + \sum_{n=2}^{\infty} st^n)^{-1} = (1-t)(1-\frac{d+1}{2}t)^{-2}$, where $s = (\frac{d+1}{2})^2$.

$$\begin{aligned}
\left(1 - dt + \sum_{n=2}^{\infty} st^n\right)^{-1} &= (1-t)\left(1 - \frac{d+1}{2}t\right)^{-2} = (1-t)\sum_{n=0}^{\infty} n\left(\frac{d+1}{2}\right)^{n-1} t^{n-1} \\
&= \sum_{n=0}^{\infty}(n+1)\left(\frac{d+1}{2}\right)^{n} t^n - \sum_{n=0}^{\infty} n\left(\frac{d+1}{2}\right)^{n-1} t^n \\
&= \sum_{n=0}^{\infty}\left(\frac{d+1}{2}\right)^{n-1}\left[(n+1)\left(\frac{d+1}{2}\right) - n\right]t^n \\
&= \sum_{n=0}^{\infty}\left(\frac{d+1}{2}\right)^{n-1}\left[\frac{(d+1)+n(d-1)}{2}\right]t^n.
\end{aligned}$$

Secondly, let us compute $(1 - dt + \sum_{n=2}^{\infty} s_n t^n)^{-1}$ where $s_n = \epsilon^2(d-2\epsilon)^{n-2}$ for all $n \geq 2$.

$$\begin{aligned}
\left(1 - dt + \sum_{n=2}^{\infty} s_n t^n\right)^{-1} &= \frac{1-(d-2\epsilon)t}{(1-(d-\epsilon)t)^2} = (1-(d-2\epsilon)t)\sum_{n=0}^{\infty} n(d-\epsilon)^{n-1}t^{n-1} \\
&= \sum_{n=0}^{\infty}(n+1)(d-\epsilon)^n t^n - (d-2\epsilon)\sum_{n=0}^{\infty} n(d-\epsilon)^{n-1}t^n \\
&= \sum_{n=0}^{\infty}(d-\epsilon)^{n-1}[(n+1)(d-\epsilon) - n(d-2\epsilon)]t^n \\
&= \sum_{n=0}^{\infty}(d-\epsilon)^{n-1}[d + (n-1)\epsilon]t^n.
\end{aligned}$$

□

A.3 Problems of Chapter 3

Problem 9. Let $f, g \in \operatorname{Aut}\mathcal{T}$ and $h \in \operatorname{st}(1)$. Let also $u \in X^*$, $x \in X$ and let a be a rooted automorphism corresponding to the permutation σ. Prove the following properties of sections:

(i) $(fg)_u = f_u\, g_{f(u)}$;

(ii) $(h^a)_x = h_{\sigma^{-1}(x)}$.

Solution. Let $v \in X^*$ and let us start with the first property. On the one hand, we have that $(fg)(uv) = (fg)(u)(fg)_u(v)$. On the other hand, we get the following expression which implies that $(fg)_u = f_u\, g_{f(u)}$:

$$\begin{aligned}(fg)(uv) &= g(f(uv)) = g(f(u)f_u(v)) = g(f(u))g_{f(u)}(f_u(v))\\ &= (fg)(u)(f_u g_{f(u)})(v).\end{aligned}$$

Let us now prove the second property. We need to realise that since $h \in \operatorname{st}(1)$, then $h^a \in \operatorname{st}(1)$. On the one hand, $(h^a)(xv) = (h^a)(x)(h^a)_x(v) = x(h^a)_x(v)$. On the other hand, we get the following expression:

$$\begin{aligned}(h^a)(xv) &= a(h(a^{-1}(xv))) = a(h(\sigma^{-1}(x)v))\\ &= a(\sigma^{-1}(x)h_{\sigma^{-1}(x)}(v)) = xh_{\sigma^{-1}(x)}(v).\end{aligned}$$

This completes the proof, since looking at both expressions we deduce that $(h^a)_x = h_{\sigma^{-1}(x)}$. □

Problem 10. Prove that $\operatorname{st}(n)$ is a normal subgroup of $\operatorname{Aut}\mathcal{T}$ of finite index for $n \geq 0$ and hence $\operatorname{Aut}\mathcal{T}$ is residually finite.

Solution. Let $\mathcal{T}_n$ be a finite truncated tree, which has only n levels. Let us now consider the natural restriction homomorphism:

$$\begin{aligned}\phi : \operatorname{Aut}\mathcal{T} &\longrightarrow \operatorname{Aut}\mathcal{T}_n\\ f &\longmapsto f|_{\mathcal{T}_n}\end{aligned}$$

which is clearly surjective since in order to get $f_1 \in \operatorname{Aut}\mathcal{T}_n$, it suffices to choose an arbitrary automorphism in $\operatorname{Aut}\mathcal{T}$ with the same labels at all the vertices on the first $n-1$ levels. Moreover, the kernel of ϕ is the set of all automorphisms in $\operatorname{Aut}\mathcal{T}$ such that the label at all the vertices on the first $n-1$ levels is the identity. This is exactly $\operatorname{st}(n)$, as we can observe in Figure 3.4, hence by the first isomorphism theorem $\ker\phi = \operatorname{st}(n) \trianglelefteq \operatorname{Aut}\mathcal{T}$ and $\operatorname{Aut}\mathcal{T}/\operatorname{st}(n) \cong \operatorname{Aut}\mathcal{T}_n$, which is finite. Thus, $|\operatorname{Aut}\mathcal{T} : \operatorname{st}(n)| < \infty$. In addition, since the intersection of all $\operatorname{st}(n)$ for $n \geq 0$ is trivial, then $\operatorname{Aut}\mathcal{T}$ is residually finite. □

Problem 11. Use Zelmanov's positive solution to the Restricted Burnside Problem to show that the Burnside Problem has positive solution in the class of residually finite groups: if G is a finitely generated group of finite exponent and G is residually finite, then G is finite. As a consequence, if $\mathcal{T}$ is a p-adic rooted tree, then $\operatorname{Aut}\mathcal{T}$ does not contain any subgroups providing a negative solution to the Burnside Problem.

Solution. Assume that G can be generated with m elements and also that $\exp(G) = n < \infty$. By way of contradiction, suppose G is infinite. In particular, if K is the bound provided by Zelmanov's result for m generators and exponent n, then there exist $K+1$ different elements $g_1, ..., g_{K+1} \in G$. Now, since G is residually finite, there exists $N \trianglelefteq G$ of finite index not containing all products $g_i g_j^{-1} \neq 1$, with $1 \leq i \neq j \leq K+1$. Then, G/N is finite and $|G/N| \geq K+1$. Moreover, G/N is also finitely generated with at most m generators and of finite exponent a divisor of n, so by Zelmanov's result the bound for the order of G/N is at most K and we have reached a contradicition.

As a consequence, if $\mathcal{T}$ is a p-adic rooted tree, then $\operatorname{Aut}\mathcal{T}$ does not contain any subgroup G providing a negative solution to the Burnside Problem, since G is residually finite. This can be proved similarly as for $\operatorname{Aut}\mathcal{T}$, with $\mathrm{st}_G(n) \trianglelefteq G$. □

Problem 12. Let G be the Gupta-Sidki p-group, for p an odd prime. By the previous problem, we know that $\exp(G) = \infty$, that is, that G has elements of arbitrarily high order. In this problem, we prove this result without relying on Zelmanov's positive solution to the Restricted Burnside Problem. For simplicity, we assume that $p \geq 5$.

(i) Prove that $G' \times \{1\} \times \cdots \times \{1\} \subseteq \psi(G')$.

(ii) Prove that the projection of $\psi(G')$ on the first component, that is $\pi_1 \circ \psi(G')$ is the whole G.

(iii) Prove by induction on $k \geq 1$ that G has an element ag_k of order greater or equal to p^k, with $g_k \in G'$.

Solution. Let us start proving (i), for which we need to recall that ψ is

defined in the following way:

$$\begin{aligned}\psi : \mathrm{st}_G(1) &\longrightarrow G \times G \times \overset{p}{\cdots} \times G \\ b_0 &\longmapsto (a, a^{-1}, 1, \ldots, 1, b) \\ b_1 &\longmapsto (b, a, a^{-1}, 1, \ldots, 1) \\ &\vdots \\ b_{p-1} &\longmapsto (a^{-1}, 1, \ldots, 1, b, a)\end{aligned}$$

Let $h = [a, b]$. We have seen in Problem 5 that $G' = \langle h \rangle^G = \langle h^g \mid g \in G \rangle$. Our goal is to find some $x \in G'$ such that $\psi(x) = (h, 1, ..., 1)$. Looking at the images of every b_i for $p \geq 5$, we know that $\psi([b_0, b_1]) = \psi([b, b^a]) = (h, 1, ..., 1)$ where $[b, b^a] \in G'$. We also know from the proof of Theorem 3.17 that $\pi_1 \circ \psi : \mathrm{st}_G(1) \longrightarrow G$ is surjective, that is, for every $g \in G$ there exists some $y \in \mathrm{st}_G(1)$ such that $\psi(y) = (g, *, ..., *)$. Thus, since ψ is a group homomorphism we get that $\psi([b, b^a]^y) = \psi([b, b^a])^{\psi(y)} = (h^g, 1, ..., 1)$ where $[b, b^a]^y \in G'$, which implies that $G' \times \{1\} \times \cdots \times \{1\} \subseteq \psi(G')$.

Let us now prove (ii), that is, $\pi_1 \circ \psi(G') = G$. On the one hand, we know that $\psi([b, a]) = \psi(b^{-1}b^a) = \psi(b^{-1})\psi(b^a) = (a^{-1}b, a^2, a^{-1}, 1, ..., b^{-1})$. On the other hand, if we take the conjugates by powers of a and using property (ii) from Problem 9, we get that $\psi([b, a]^a) = (b^{-1}, a^{-1}b, a^2, a^{-1}, 1, ..., 1)$ and so on. In fact, if we conjugate $[b, a]$ by a^i the image is obtained by shifting the components of $\psi([b, a])$ to the right i times. Therefore, $\pi_1(\psi([b, a])) = a^{-1}b$ and $\pi_1(\psi([b, a]^a)) = b^{-1}$, where $[b, a], [b, a]^a \in G'$. Since $G = \langle a^{-1}b, b^{-1} \rangle$, it follows that $\pi_1 \circ \psi(G') = G$.

Finally, we are going to prove (iii) by induction on $k \geq 1$. For the base case $k = 1$ it is trivial if we take $g_k = 1$. Then, we assume it is true up to k and let us prove it for $k + 1$. By (i) we know that there exists some $h_k \in G'$ such that $\psi(h_k) = (g_k, 1, ..., 1)$, and by (ii) we also know that there exists some $x \in G'$ such that $\psi(x) = (a, *, ..., *)$. Thus, $\psi(xh_k) = (ag_k, *, ..., *)$ and $o(xh_k) \geq o(ag_k) \geq p^k$. By (i) again, there exists some $g_{k+1} \in G'$ such that $\psi(g_{k+1}) = (xh_k, 1, ..., 1)$. Then, we know that $(ag_{k+1})^p = g_{k+1}^{a^{p-1}} g_{k+1}^{a^{p-2}} \cdots g_{k+1}^{a} g_{k+1}$ and if we apply ψ using property (ii) from Problem 9, we get the following:

$$\begin{aligned}\psi((ag_{k+1})^p) &= \psi(g_{k+1}^{a^{p-1}}) \cdots \psi(g_{k+1}) = (1, ..., 1, xh_k) \cdots (xh_k, 1, ..., 1) \\ &= (xh_k, xh_k, ..., xh_k).\end{aligned}$$

Hence, $o(ag_{k+1}) = p \cdot o(xh_k) \geq p^{k+1}$, as desired. □

Problem 13. Prove that $H^G = \langle h^g \mid h \in H, g \in G \rangle$.

Solution. We know that the normal closure of H is the smallest normal subgroup of G containing H, let us call it K. Each $h \in H$ lies in K and since

K is normal, $h^g \in K$ for all $g \in G$ and thus $\langle h^g \mid h \in H, g \in G\rangle \leq K$. On the other hand, it suffices to prove that $\langle h^g \mid h \in H, g \in G\rangle$ is a normal subgroup of G, which completes the proof. It is obvious that the set $\{h^g \mid h \in H, g \in G\}$ is stable under all conjugations of G and hence the group it generates also is. Therefore, $K = \langle h^g \mid h \in H, g \in G\rangle$. □

Problem 14. Prove that the first Grigorchuk group can be decomposed as $\Gamma = \langle a\rangle \ltimes \mathrm{st}_\Gamma(1)$ where $\mathrm{st}_\Gamma(1) = \langle b, c\rangle^\Gamma = \langle b, b^a, c, c^a\rangle$.

Solution. First of all, let us prove that $\mathrm{st}_\Gamma(1) = \langle b, c\rangle^\Gamma$. Since $b, c \in \mathrm{st}_\Gamma(1)$ we know that the normal closure of $\langle b, c\rangle$, denoted by $\langle b, c\rangle^\Gamma$, must be a subgroup of $\mathrm{st}_\Gamma(1) \trianglelefteq \Gamma$. Then, it suffices to prove that the indexes of $\langle b, c\rangle^\Gamma$ and $\mathrm{st}_\Gamma(1)$ in Γ are equal. We know that $\Gamma = \langle a, b, c\rangle$, hence $\Gamma/\mathrm{st}_\Gamma(1) = \overline{\langle a, b, c\rangle} = \langle \overline{a}\rangle$ and $|\Gamma : \mathrm{st}_\Gamma(1)| = o(\overline{a})$ divides $o(a) = 2$, but since $a \notin \mathrm{st}(1)$ we get $|\Gamma : \mathrm{st}_\Gamma(1)| = 2$. Similarly, we also get that $|\Gamma : \langle b, c\rangle^\Gamma| = 2$, hence $\mathrm{st}_\Gamma(1) = \langle b, c\rangle^\Gamma$.

Secondly, we need to check that $\langle b, c\rangle^\Gamma = \langle b, b^a, c, c^a\rangle$. It is obvious that $\langle b, c\rangle^\Gamma \supseteq \langle b, b^a, c, c^a\rangle$ so we have to prove the other inclusion. Since $\langle b, c\rangle^\Gamma$ is the normal closure of $\langle b, c\rangle \leq \Gamma$, it suffices to check that $N = \langle b, b^a, c, c^a\rangle \trianglelefteq \Gamma$, which is equivalent to proving that $\Gamma = \langle a, b, c\rangle = N_\Gamma(N)$. This is satisfied if $a, b, c \in N_\Gamma(N)$, which is true since $b, c \in N \subseteq N_\Gamma(N)$ and $b, c, b^a, c^a \in N$ where $b = (b^a)^a$ and $c = (c^a)^a$, hence $a \in N_\Gamma(N)$.

Finally, we need to prove that $\Gamma = \langle a\rangle \ltimes \mathrm{st}_\Gamma(1)$, that is, Γ can be decomposed as the internal semidirect product of $\langle a\rangle$ and $\mathrm{st}_\Gamma(1)$. These three conditions must be satisfied:

(i) $\langle a\rangle \leq \Gamma$ and $\mathrm{st}_\Gamma(1) \trianglelefteq \Gamma$;

(ii) $\Gamma = \langle a\rangle \cdot \mathrm{st}_\Gamma(1)$;

(iii) $\langle a\rangle \cap \mathrm{st}_\Gamma(1) = \{1\}$.

We already know that (i) is satisfied. The second condition is also fulfilled since $\langle b, c\rangle \leq \mathrm{st}_\Gamma(1) \trianglelefteq \Gamma$, and hence $\Gamma = \langle a, b, c\rangle = \langle a, \mathrm{st}_\Gamma(1)\rangle = \langle a\rangle \cdot \mathrm{st}_\Gamma(1)$. Finally, the third condition also holds since $o(a) = 2$ and $a \notin \mathrm{st}_\Gamma(1)$. □

Bibliography

[1] W. Burnside, On an unsettled question in the theory of discontinuous groups, *Quart. J. Pure Appl. Math.*, **33** (1902), 230-238.

[2] W. Burnside, *Theory of Groups of Finite Order*, Cambridge University Press, 1897.

[3] T. Ceccherini-Silberstein, M. D'Adderio, and E. Zelmanov, *Topics in Groups and Geometry: Growth, Amenability, and Random Walks*, Springer Monographs in Mathematics, Springer International Publishing, 2021.

[4] M. Ershov, Golod-Shafarevich groups: a survey, *International Journal of Algebra and Computation*, 2012.

[5] Gustavo Fernández-Alcober, *A Course in Nilpotent and Soluble Groups*, lecture notes of a PhD course delivered at the University of the Basque Country, 2019.

[6] Gustavo Fernández-Alcober, *Groups of Automorphisms of p-adic Rooted Trees*, lecture notes of a PhD course delivered at the University of Padova, 2009.

[7] Gustavo Fernández-Alcober, *Groups of Automorphisms of p-adic Rooted Trees*, lecture notes of a PhD course delivered at the University of Salerno, 2017.

[8] R. I. Grigorchuk, Burnside Problem on Periodic Groups, *Funktsional. Anal. i Prilozhen.*, **14**:1 (1980), 53–54.

[9] N. Gupta and S. Sidki, On the Burnside Problem for Periodic Groups, *Mathematische Zeitschrift*, **182** (1983), 385-388.

[10] Jake Huryn, *What is the Grigorchuk Group?*, Ohio State University, 2018.

[11] I. Kaplansky, *Notes on Ring Theory*, Mathematics lecture notes, University of Chicago, 1965.

[12] T. Y. Lam, *A First Course in Noncommutative Rings*, Graduate Texts in Mathematics, Springer New York, NY, 2001.

[13] Derek J.S. Robinson, *A Course in the Theory of Groups*, Springer, second edition, 1995.

[14] B. K. Sahoo and B. Sury, What is the Burnside Problem?, *Resonance*, pp. 1-5, 2005.

[15] I. Schur, *Ueber Gruppen periodischer linearer Substitutionen*, Sitzungsberichte der Königlich-Preussischen Akademie der Wissenschaften zu Berlin, pp. 619-627, 1911.

[16] Katie Waddle, *The Grigorchuk Group*, 2008. It is available at `http://www.math.uchicago.edu/ may/VIGRE/VIGRE2008/ REUPapers/Waddle.pdf.`

[17] Naomi Watson, *Formal Power Series*, 2021. It is available at `http://www.math. uwaterloo.ca/ dgwagner/co430I.pdf.`