PROOF OF THE NEGATIVE OF THE BURNSIDE PROBLEM

JASON CHEN

Abstract. The Burnside Problem asks whether a finitely generated group in which every element has finite order must necessarily be a finite group. We utilize important properties of profinite p-groups to help prove the Golod-Shafarevich Theorem, an infinite group that satisfies the negative of the Burnside Problem.

1. Introduction

In 1902, William Burnside posed the Burnside problem, which asks whether every finitely generated group in which all elements have finite order must itself be finite [Bur02]. This question, which is deceptively simple in formulation, gave rise to several major subproblems in group theory, including the general Burnside problem, the bounded Burnside problem, and the restricted Burnside problem.

The Bounded Burnside problem asks:

Question 1. If G is a finitely generated group with exponent n, is G necessarily finite? More generally, if we define the free Burnside group of rank m and exponent n, to be denoted $B(m,n)$, to be a group with m distinguished generators and $x^n = e$ for every $x \in B(m,n)$, for which m and n is the free Burnside group $B(m,n)$ finite?

For instance, $B(1,n)$ is merely a cyclic group of order n, while $B(m,2)$ is the direct product of m copies of the cyclic group of order 2. In Burnside's paper, he proved that $B(m,3)$, $B(m,4)$, and $B(m,6)$ are finite for all m. In 1968, Pyotr Novikov and Sergei Adian made a breakthrough. Using a complicated combinatorial method, they were able to prove that for every odd number $n > 4381$, there exist infinite, finitely generated groups of exponent n. Adian later lowered it to 665 [NA68].

Formulated in the 1930s, the Restricted Burnside problem asks:

Question 2. If it is known that a group G with m generators and exponent n is finite, can one conclude that the order of G is bounded by some constant depending only on m and n?

The case of arbitrary exponent has been completely settled in the affirmative by Efim Zelmanov [Zel91], who was awarded the Fields Medal in 1994 for his work.

In 1964, over 60 years after the Burnside problem was posed, Evgeny Golod and Igor Shafarevich proved that there did in fact exist finitely generated groups where every element had finite order, with infinitely many elements [GS64]. This theorem, the Golod-Shafarevich Theorem, is the main subject of this paper.

In this paper, we look to [Zho17] and [Cas22] for the proof of the Golod-Shafarevich Theorem. Due to the complex nature of both this proof and the problem, complex topics like cohomologies may pop up.

Date: July 2025.

2. Preliminaries

We first begin with some important preliminaries. This is to help the reader generally understand what the Burnside Problem even asks, while also supporting future proofs. Then, we introduce some more complex definitions, which may not be as easily understandable to a lay audience.

2.1. Burnside Problem.

We now begin with three important definitions to help understand the Burnside Problem. We define **Group, Order, and Generators.**

Definition 1. A **Group** (G,+) is a nonempty set defined under a binary operation, with the following properties:

- **Identity:** There exists some $e \in (G, +)$ such that for all $g \in (G, +)$, $e+g = g+e = g$.
- **Inverse:** For every element $g \in (G, +)$, there exists some inverse $g^{-1} \in (G, +)$ such that $g + g^{-1} = g^{-1} + g = e$.
- **Associativity:** For arbitrary elements $a, b, c \in (G, +)$, $(a + b) + c = a + (b + c)$.
- **Closure:** For arbitrary elements $a, b \in (G, +), a + b \in G$.

One intuitive example of a group is a 12-hour clock. The elements of our group would be the hours, while the operation would be addition, albeit a little altered from the typical interpretation. For example, 3 hours after 11 isn't 14, since it's not on our clock. Rather, it'd be 2, the remainder after 14 is divided by 12.

Our identity would simply be 12 (typically the identity would be denoted as 0, but clocks don't have a 0 on them), since 12 hours after 1 is still 1, 12 hours after 2 is still 2, and so on. To obtain the inverse of some hour g, simply add $12 - g$ to get the identity. For example, 3 hours after 9 would be 12, 4 hours after 8 would be 12, and so on.

Definition 2. The **order** of an element $g \in G$ is the smallest positive integer m such that g combined m times results in the identity. Under multiplicative notation, $g^m = e$. If there exists no such integer, then we say that the order of g is infinite.

In our clock example, some elements have different order. For instance, 1 has order 12, while 2 has order 6. However, no element has an infinite order.

Definition 3. The **generators** S of a group G are a subset such that every element of G can be written under group operation of finitely many elements in S and their inverses.

In our clock example, a generator could be 1, since you can get to every single hour on the hand by adding enough 1's. However, this example is rather simple, as it needs only one generator, while many other groups are much more complex.

We now define some other important terms.

Definition 4. For a group G, a **normal subgroup** $N \lhd G$ is a subgroup that is invariant under conjugation. That is, N is normal if and only if $gng^{-1} \in N$ for all $g \in G$ and $n \in N$.

Definition 5. Let N be a normal subgroup of G. Then, the **quotient group** G/N is the set of left cosets, $\{gN : g \in G.\}$ Keep in mind that $gN = \{gn : n \in N.\}$

We now have enough background to understand what the Burnside Problem asks. In the clock example, it is equivalent (isomorphic) to $\mathbb{Z}/12\mathbb{Z}$, the integers modulo 12, and is an example of a finitely generated group (generated by 1), where every element has finite order, with a finite number of elements.

We will now begin defining important information which will be utilized for the Golod-Shafarevich Theorem.

Definition 6. For a prime p, the group G is a **finite p-group** if $|G| = p^n$ for some positive integer n.

We'll see this definition pop up a lot more, so make sure to keep this in mind.

Definition 7. For a group G, the **commutator** of two elements $a, b \in G$ is defined as $[a, b] = a^{-1}b^{-1}ab$. The **commutator subgroup** of G, denoted $[G, G]$ is the set of all commutators of G, defined as $\{[a, b] : a, b \in G\}$.

The commutator serves as a measure for how commutative two elements are. In fact, if we take the quotient of G by the closure of our commutator subgroup $[G, G]$, we obtain an abelian version of G, called the **abelianization** of G, often denoted as G^{ab}. We now turn from our group theory reminders over to algebras and rings, which are defined under two operations, instead of just one.

2.2. **Rings and Group Algebras.**

Definition 8. A ring R is a set defined under two binary operations, $+, \cdot$ that satisfy the following properties:

- It is an **abelian group under addition**, meaning that it is commutative under addition ($a + b = b + a$) and forms a group under addition.
- It is a **monoid under multiplication**, meaning that under multiplication, it has associativity and an identity, but not necessarily an inverse. A **monoid** is a weaker version of a group.
- Multiplication is **distributive**, meaning that $(a+b)\cdot c = (a\cdot c)+(b\cdot c)$, and $c\cdot(a+b) = (c \cdot a) + (c \cdot b)$.

We now are going to define an ***R*-module** to help define a **group algebra**, which we'll be using in future proofs.

Definition 9. For a ring R, and 1 being its multiplicative identity, the **left *R*-module** M is an abelian group under addition and has multiplicative operation $\cdot : R \times M \to M$ such that for all elements $r, s \in R$ and $x, y \in M$, the following properties are satisfied:

- $r \cdot (x + y) = r \cdot x + r \cdot y$
- $(r + s) \cdot x = r \cdot x + s \cdot x$
- $(rs) \cdot x = r \cdot (s \cdot x)$
- $1 \cdot x = x$.

A **right *R*-module** is one defined similarly in terms of the operation $\cdot : M \times R \to M$. If R is commutative, the left and right ***R*-modules** are simply called ***R*-modules**.

Modules can be thought of as generalizations of a vector space, though in this case, our scalars, usually taken from a field, are taken from a ring.

Definition 10. Let G be a multiplication group, and Λ a commutative ring (under multiplication). The **group algebra** of G over Λ, also commonly known as the **group ring**, denoted $\Lambda[G]$, is defined to be the free $\Lambda-$module on the basis G. That is, we can think of it as all linear combinations

$$\alpha = \sum_{g \in G} a_g g$$

with $a_\Lambda \in R$ and $a_g = 0$ for all but finitely many g. We define the sum of two elements to be:

$$\left(\sum_{g\in G} a_g g\right) + \left(\sum_{g\in G} b_g g\right) = \left(\sum_{g\in G} (a_g + b_g) g\right)$$

We also define the product to be:

$$\left(\sum_{g\in G} a_g g\right)\left(\sum_{g\in G} b_g g\right) = \sum_{g\in G} c_g g$$

where $c_g = \sum_{x\in G} a_x b_{x^{-1}g}$

One intuitive way of thinking about the group algebra is as a polynomial space (a vector space if Λ is a field). The group ring $\Lambda[G]$ can be thought of as the ring of polynomials with coefficients in Λ, and variables coming from G, except multiplication is defined with groups rather than by "adding" the order of groups, which is what we're used to for polynomials.

For example, consider the group ring $\mathbb{Z}[C_3]$, where C_3 denotes the cyclic group of order 3, generated by element a. Then, for integers z_0, z_1, z_2, $\mathbb{Z}[C_3] = \{z_0 e + z_1 a + z_2 a^2\}$, which is isomorphic to $\mathbb{Z}[a]/(a^3 - 1)$.

Definition 11. The group ring $\Lambda[G]$ has an **augmentation map** $\epsilon : \Lambda[G] \to \Lambda$ defined as

$$\sum_{g\in G} a_g g \to \sum_{g\in G} a_g.$$

The kernel $I(G)$ of this map (elements sent to the identity) is the **augmentation ideal** of $\Lambda[G]$.

Proposition 1. The set $\{g - 1 : g \neq 1, g \in G\}$ is a Λ-basis for $I(G)$.

Proof: We follow [Zho17]'s proof, on page 6 of their paper.

From above, we know that $\epsilon(g) = 1$ for every $g \in G$. Thus, $\epsilon(g - 1) = 1 - 1 = 0$, so we conclude that $\{g - 1 : g \neq 1, g \in G\} \subseteq \ker(\epsilon) = I(G)$. For any element $\sum_{g\in G} a_g g \in I(G), \sum_{g\in G} a_g = 0$. We can combine this information to see:

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

Since every element of $I(G)$ has a corresponding preimage in $\{g - 1 : g \neq 1, g \in G\}$, we conclude that this set spans $I(G)$. We now show linear independence. Suppose

$$\sum_{g\in G} a_g (g - 1) = 0.$$

Then, we see that

$$\sum_{g\in G} a_g g = \sum_{g\in G} a_g.$$

Since this right side is a constant, this can only be true if all a_g terms are equal to 0, demonstrating linear independence. Since $\{g - 1 : g \neq 1, g \in G\}$ spans and is linearly independent, we conclude that this set is a basis for $I(G)$.

We now turn to our final section, which is on group cohomologies.

2.3. **Group Cohomology.** Group cohomology is a set of mathematical tools used in group theory to help study groups. We look at the group actions (special mappings from a group onto itself) of a group in an associated ***G*-module** (similar to the Λ-module but defined over a group). There is a category theory definition of something called a **cohomology group**, but we can define it through an alternate method, utilizing **cochains**.

Definition 12. A **chain complex** $(A_\bullet, d_\bullet)$ is a sequence of abelian groups of modules $\cdots, A_0, A_1, A_2, \cdots$ linked by homomorphisms $d_n : A_n \to A_{n-1}$, where $d_n \circ d_{n+1} = 0$ for all n. That is, the composition of two consecutive maps nets the zero map. The chain complex can be written out like so:

$$\cdots \xleftarrow{d_0} A_0 \xleftarrow{d_1} A_1 \xleftarrow{d_2} A_2 \xleftarrow{d_3} A_3 \xleftarrow{d_4} A_4 \xleftarrow{d_5} \cdots .$$

A **cochain complex**, $(A^\bullet, d^\bullet)$ is the dual notion to a chain complex. It is a sequence of abelian groups or modules linked by homomorphisms $d_n : A_n \to A_{n+1}$, where $d_{n+1} \circ d_n = 0$ for all n. It can be written as:

$$\cdots \xrightarrow{d^{-1}} A_0 \xrightarrow{d^0} A_1 \xrightarrow{d^1} A_2 \xrightarrow{d^2} A_3 \xrightarrow{d^3} A_4 \xrightarrow{d^4} \cdots$$

Notably, we can think of a cochain complex as a chain complex that's simply going in the opposite direction, hence why it serves to be the dual of the chain complex.

While these chain and cochain complexes will be useful in defining cohomology groups, it is also important to note that they are able to provide insight onto the algebraic structure, such as the image and kernels of maps.

Definition 13. Let G be a group and M a G-module, and an integer $n \geq 0$. The **group of *n*-cochains** is defined as $C^n(G, M) = \{f : G^n \to M\}$, the continuous maps from G^n to M. These result in a cochain complex

$$\cdots \to C^n \xrightarrow{d^n} C^{n+1} \xrightarrow{d^{n+1}} C^{n+2} \to \cdots$$

with our d^n being the **coboundary homomorphisms** defined as

$$(d^n f)(g_1, \ldots, g_{n+1}) = g_1 \cdot f(g_2, \ldots, g_{n+1}) + \sum_{i=1}^{n} (-1)^i f(g_1, \ldots, g_{i-1}, g_i g_{i+1}, g_{i+2}, \ldots, g_{n+1}).$$

Under this definition, our **coboundary homomorphisms** satisfy the property $d^{n+1} \circ d^n = 0$, so we know these are still cochains.

A little intuition behind what this definition of the differential d^n does comes from understanding what happens to each of the specific g's you put into the differential. Each d^n alternates the signs, and "drops" some of the values. For instance,

$$(d^1 f)(g, h) = gf(h) - f(gh) + f(g),$$

while

$$(d^2 f)(g, h, k) = gf(h, k) - f(gh, k) + f(g, hk) - f(g, h).$$

Note how going from d^2 to d^1 drops some terms (we have 3 instead of 2), and the signs are alternated.

Definition 14. Let us define G, M, n to be the same as in Definition 13. Then, the set of ***n*-cocyles** of G with coefficients in M, is defined as

$$Z^n(G, M) = \ker(d^n).$$

The group of ***n*-coboundaries** is defined as

$$B^n(G,M) = \begin{cases} 0 & n = 0 \\ \text{im}(d^n) & n \geq 1 \end{cases}$$

Now that we have enough information, we can define the ***n*th cohomology group** $H^n(G,M)$ of G to be

$$H^n(G,M) = Z^n(G,M)/B^n(G,M).$$

The kernel of d^n ($Z^n(G,M)$) represents the cochains that are sent to 0, while the image of $d^n(B^n(G,M))$ represents the cochains that are equal to d^n composed of something. So, when we divide by $B^n(G,M)$, we essentially say that when these cocyles differ by a co-boundary, they are equivalent.

For example's sake, let us consider the 0th cohomology group, $H^0(G,M)$. The 0-cochains are going to simply represent all functions $c : G \to M$. The 0-cocycles are cochains where $(d^0c) = 0$. Thus, by utilizing the formula above, we see that this implies that

$$(d^0c)(g) = gc - c.$$

In order for c to be a 0-cocycle, we need to have $(d^0c)(g) = 0$ for every $g \in G$, which means that

$$gc = c.$$

Thus, we can conclude that 0-cocyles will merely send every element to itself, so they are invariant. The 0-coboundaries is $B^0(G,M) = \text{im}(d^0)$, but since there exist no -1-cochains, we know that $B^0(G,M) = \{0\}$, the zero map. Thus, $H(G,M) = \{m \in M : mg = m \forall g \in G\}$.

3. Pro-p Groups

Before we get to the proof of the Golod-Shafarevich Theorem (it's the next section, so it's coming quick!), we must first introduce pro-p groups, a type of profinite group. We will first define them through topology, which utilizes some important terminology we must first define.

3.1. Relevant Definitions for Pro-p groups.

Definition 15. A set X with collection of subsets T is said to be a **Topology** if it falls under the following criteria:

- The trivial subsets X and $\{\emptyset\}$ are in T.
- For $A, B \in T, A \cap B \in T$.
- For two or more sets in T, their union is also in T.

We say that **Topological space** is a pair (X,τ) where X is a set and τ is a topology on X.

Definition 16. We say that a function is **continuous** over G if for any open set $U \subseteq G, f^{-1}(U)$ is open in the domain of f.

Definition 17. A **Topological Group** G is a topological space under a group operation for which the group operation $(\cdot : G \times G \to G, \cdot(x,y) = xy)$ and inverse maps $({}^{-1} :^{-1} (x) = x^{-1})$ are **continuous**.

Definition 18. A topological space X is **compact** if every open cover of X has a finite subcover. That is, if $\{U_\alpha\}_{\alpha\in A}$ is a collection of open sets with $X \subseteq \bigcup_{\alpha\in A} U_\alpha$, then there exists a finite subcollection $U_{\alpha_1,\dots,\alpha_n}$ where

$$X \subseteq \bigcup_{i=1}^{n} U_{\alpha_i}.$$

A more intuitive definition can come from the Heine-Borel Property, which states that a subset is compact if it is closed and bounded. However, this only works in euclidean spaces, which we aren't working with, so it's not technically the correct definition.

Definition 19. A topological space X is a **Hausdorff space** if any two points in X are separated by a neighborhood. That is, for any two points x, y, there exists neighborhoods U about x and V about y such that U and V are disjoint.

A compact space is closed in a Hausdorff space, but boundedness isn't a required property for the Hausdorff space (it isn't necessarily a metric space).

Definition 20. A topological space T is **totally disconnected** if all the **connected components** are one-point sets (essentially the point itself). The **connected components** of a point $x \in T$ is the union of all connected subsets that contain x.

The **connected component** of a point x can be thought of as the largest connected subset that still contains x, so if a topological space is **totally disconnected**, we can imagine that it means that every single point is "by itself".

Definition 21. A **profinite group** is a **compact**, **Hausdorff**, **totally disconnected topological group**.

We will introduce an example shortly. Note that open subgroups of profinite groups are closed due to compactness, and every closed subgroup of a profinite group is also profinite. This definition is topological, and we will shortly define this in algebraic notation, which will be more useful for this specific context. However, before doing so, we must include the notion of **inverse limits**.

Definition 22. A **directed partially ordered set** is a poset $(I, \leq)$ such that for every $i, j \in I$, there exists $k \in I$ such that $k \geq i, j$. An **inverse system** over I, sometimes also called a **projective system**, is a family of groups $(A_i)_{i\in I}$ with a family of homomorphisms $f_{i,j} : A_i \to A_j$ for all $i \geq j$, where $f_{i,i}$ is the identity on A_i and $f_{ij} \circ f_{jk} = f_{ik}$ for all $i \geq j \geq k$. The **inverse limit** of the inverse system $((A_i)_{i\in I}, (f_{ij})_{i\geq j\in I}$, sometimes called the **projective limit**, is a subgroup of the direct product of the A_i's, defined as follows:

$$\varprojlim_{i\in I} A_i = \left\{ \vec{a} \in \prod_{i\in I} A_i \,\middle|\, a_j = f_{ij}(a_i), \forall : i \geq j \in I \right\}.$$

We can think of the **inverse limit** as the infinitely long chain of elements as they slowly get projected. For example, with a sequence

$$X_0 \leftarrow X_1 \leftarrow X_2 \leftarrow \cdots,$$

we find the space which compiles the sequences in a manner that shows us what a term is mapped to: $(x_0, x_1, x_2, \dots)$, where $x_0 \in X_0, x_1 \in X_1, \dots$

One example of a **directed partially ordered set** is the natural numbers, $\mathbb{N}$. Every element precedes the next by increasing (directed), and for every pair, it's also ordered (totally ordered, which is also partially ordered).

Now, we can put a profinite group into algebraic notation, which will be quite helpful.

Proposition 2. If G is a profinite group, then G is isomorphic to $\varprojlim(G/N)$, where N ranges over the open normal subgroups of G. Furthermore, the inverse limit of an inverse system of discrete finite groups is profinite. Therefore, the topological definition and inverse limit definition are equivalent.

We omit this proof for sake of brevity, but it can be found in Chapter 1 of [DSSD99].

An example of a profinite group is the p-adic integers, $\mathbb{Z}_p$. A p-adic integer can be represented as the sequence

$$x = \{x \pmod p, x \pmod{p^2}, x \pmod{p^3}, \dots\}$$

In this instance, N approaches $p^n\mathbb{Z}$, denoted by the inverse system

$$\cdots \to \mathbb{Z}/p^3\mathbb{Z} \to \mathbb{Z}/p^2\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}.$$

We now define another example of profinite groups, called the **profinite completion** of a group.

Definition 23. Let G be an arbitrary group, and $\mathcal{N}$ the family of normal subgroups $N \lhd G$ of finite index in G, ordered by reverse inclusion (left is largest, right is smallest). The **profinite completion** of G is the inverse limit $\varprojlim(G/N)_{N\in\mathcal{N}}$.

Now, we can define what a **pro-p group** is.

Definition 24. For a prime p, a **pro-p group** is a profinite group that is the inverse limit of finite p-groups. A finite group is pro-p if and only if its order is a power of p.

A common example of the pro-p groups are the p-adic integers $\mathbb{Z}_p$ under the addition operation, since they're defined as $\varprojlim \mathbb{Z}/p^n\mathbb{Z}$. In fact, this is the pro-p completion of the ring Z, since every $p^n\mathbb{Z}$ is a normal subgroup of $\mathbb{Z}$.

3.2. **Presentations.** The Golod-Shafarevich theorem is deeply related to the minimal presentation of p-groups, which is a method of representing groups in terms of their relations and generators. Because of this, it is important to look at how we define the finitely generated pro-p groups and their presentations.

Definition 25. Let G be a profinite group, and $N \subset G$ a closed normal subgroup of G. Let I be an index set and $S = \{s_i : i \in I\}$ be a convergent subset of N. That is, every open subgroup of N contains all but finitely many elements of S. Then, we say that the s_i **generate N** if N is the smallest closed normal subgroup of G containing S. Equivalently, we can say that these s_i **generate N** if the subgroup generated by the conjugates of s_i is dense in N.

This definition doesn't vary too much from the definition of generators for a standard group. One key point to note is the fact that E is slowly converging to a subset of N, which is relevant since we are working over a topology, not just a group. The subgroup containing S must contain all the elements of S in combination with each other, mirroring the group

theoretic definition above. However, something to note is that this subgroup needs to be normal.

In the context of this problem, we can take G to be a normal subgroup of itself (under the property of closure), so thus we now have a definition for the generators of a pro-p group.

Definition 26. Let G be a profinite group, and let $S = \{s_i : i \in I\}$ be a convergent subset of G. Then S is a **system of generators** for G if G is the smallest closed subgroup containing S. We say that S is **minimal** if no proper subset of S generates G. The cardinality of a **minimal system of generators** is the **generator rank** of G, denoted by $d(G)$ or d.

The Golod-Shafarevich Theorem will utilize this generator rank $d(g)$, so keep this definition in mind (or reference it later).

Given the similarities between the generators in standard group theory and that of the profinite group, it makes sense that there also exist free pro-p groups, constructed through profinite completions (more specifically, pro-p completions).

Definition 27. Let I be an index set and F_I be the free group on the generators $\{s_i | i \in I\}$. Let $\mathfrak{U}$ be the set of normal subgroups $N \lhd F_I$ of p-power index (the index of the group is a power of p) containing all but finitely many generators. Then, the inverse limit

$$F(I) = \varprojlim (F_I/N)_{N \in \mathfrak{U}}$$

is called the **free pro-p group** with system of generators $\{s_i | u \in I\}$.

Since F_I is defined to be a free group based on some generators, we see that $F(I)$ is the inverse limit of the free group as we quotient it by normal subgroups. Thus, the free group structure should still be preserved, though we lose some of the relations that are defined within the normal subgroups. Thus, we can see that the free group in standard group theory is still similar enough to that of the pro-p groups.

Now, we move to define the presentation of these pro-p groups, but we need one more definition.

Definition 28. A sequence

$$G_0 \xrightarrow{f_1} G_1 \xrightarrow{f_2} G_2 \xrightarrow{f_3} \cdots \xrightarrow{f_n} G_n$$

of groups G_i and group homomorphisms f_i is said to be **exact** at G_n if $\mathrm{im}(f_n) = \ker(f_{n+1})$. The sequence is an **exact sequence** if it is exact at every G_i for $1 \leq i \leq n$.

The exact sequence is useful since it lets us know how one group is influencing another. In the definition below, we will see how this is useful.

Now, we can move onto the definition for the presentation of a pro-p group.

Definition 29. Suppose E is a system of generators of the pro-p group G. Let F be the **free pro-p group** on the system of generators E. Then, similarly to the presentation of ordinary groups, we have the **exact sequence**

$$1 \to R \to F \to G \to 1,$$

which is called a **presentation** of G by F, denoted by $\langle E; R \rangle$.

Notably, 1 denotes the trivial group, $\{e\}$, so our first homomorphism (from R to F) is injective, while the last homomorphism (from F to G) is surjective. This means that the image of R onto F is exactly the elements which get mapped from F to G's identity. In other words, F is like an extension of G by A.

Definition 30. Let R, F, G, E be defined like Definition 29. We let a set $S \subseteq R$ to be a **system of relations** with respect to E if S is a system of generators for R, a normal subgroup of F. A system of relations S is minimal if there is no proper subset of S that can generate R. The cardinality of S is the **relation rank** of G, denoted $r(G)$, or r.

Essentially, $G = F/R$ since we've done restrictions upon the free group through R, and now we want to see what the cardinality of the system of generators for this specific restriction is.

The **relation rank** is another key definition which will be utilized in the proof of the Golod-Shafarevich Theorem, so keep it in mind.

3.3. **Cohomological Interpretations of Generators and Relations.** In this section, we will begin actually computing $d(G)$ and $r(G)$, which is extremely important. We'll do this by putting relations and generators in terms of the cohomology groups of pro-p groups, while utilizing tools such as exact sequences to simplify our work.

The main motivation for this comes from Pontryagin duality, which states the category of discrete torsion abelian groups is dual to the category of profinite abelian groups.

A topological group is **discrete** if there is no limit point for it. Equivalently, it means that the identity is isolated. Such a group is a torsion group if every element has finite order, a type of group that the Burnside Problem is about.

The Pontryagin dual of a group G is defined as $G^* = \mathrm{Hom}(G, \mathbb{Q}/\mathbb{Z})$, the set of continuous maps from G to $\mathbb{Q}/\mathbb{Z}$.

3.3.1. *Computing $d(G)$.*

Definition 31. The **Frattini subgroup** $\Phi(G)$ of a group G is the intersection of all maximal closed subgroups of G.

A maximal subgroup H is a group for which no other proper subgroup contains H strictly. They will always share elements (at a minimum, e), but H cannot be contained within any other subgroup.

Remark 1. If G is a finite p-group, then $\Phi(G) = \overline{G^p[G, G]}$.

This is a subgroup that is generated by products of pth powers (elements that are raised to the pth power)**and** the commutators in G (elements of the form $a^{-1}b^{-1}ab$).

This subgroup is useful because this subgroup becomes abelian when quotienting G by $\Phi(G)$, while also killing all pth powers. If we quotient by $\Phi(G)$, we set all commutators to be equivalent to the identity, and do the same with all pth powers, which means that we're left with a commutative group that doesn't have any pth powers. This means that every element raised to the pth power will get sent to the identity, meaning this group has exponent p. This means that $G/\Phi(G)$ is the largest profinite abelian quotient of exponent p of G.

We can also interpret $G/\Phi(G)$ as a vector space over $\mathbb{F}_p$, a finite field of order p (essentially just $\mathbb{Z}/p\mathbb{Z}$). If we take the Pontryagin dual of $G/\Phi(G)$, we get

$$(G/\Phi(G))^* = \mathrm{Hom}(G/\Phi(G), \mathbb{Q}/\mathbb{Z}) = \mathrm{Hom}(G, \mathbb{F}_p) = H^1(G, \mathbb{F}_p).$$

We can combine this result with Burnside's Basis Theorem to give us a new interpretation of generator and relation rank.

Theorem 1. *(Burnside's Basis Theorem) Let G be a pro-p group and let $S = \{s_i : i \in I\}$ be a convergent subset of G. Then S is a system of generators of G if and only if the subset $\overline{S}$ of residue classes modulo $\Phi(G)$ generates $G/\Phi(G)$.*

To note, the residue classes are just the possible values of $\overline{S}$ after modding by $\Phi(G)/$ A proof can be found in [Lin10], but we will not prove this theorem.

Taking the dual doesn't actually change the dimension of the group, which is extremely useful in this case. By Burnside's Basis Theorem, we see that G and $G/\Phi(G)$ have the same dimension, and thus this is the same for $(G/\Phi(G))^* = H^1(G, \mathbb{F}_p)$. Therefore, we can draw equalities to the generator rank like so:

$$d(G) = \dim_{\mathbb{F}_p}(G/\Phi(G)) = \dim_{\mathbb{F}_p}(G/\Phi(G))^* = \dim_{\mathbb{F}_p} H^1(G).$$

3.3.2. *Computing $r(G)$.* We first note that through the same cohomological interpretation from above, we can conclude that the rank of the relation of G is $\dim_{\mathbb{F}_p} H^1(R)$ (this is in terms of R, not G like above). Furthermore, the presentation of G by the free pro-p group F from Definition 29 yields the isomorphism $G \cong F/R$. This allows us to apply a special type of sequence called the Hochschild-Serre spectral sequence. That is, given our sequence

$$1 \to R \to F \to G \to 1,$$

for a G-module A, the Hochschild-Serre spectral sequence is $H^p(G/R, H^q(R, A)) \implies H^{p+q}(G, A)$. This inner cohomology $(H^q(R, A))$ describes how our module A acts on the restriction R, while the outer cohomology, $(H^p(G/N, -))$ tells us how the quotient group acts on the other cohomology groups. From this, we obtain the sequence:

$$0 \to H^1(G, \mathbb{F}_p^R) \to H^1(R, \mathbb{F}_p)^G \to H^2(G, \mathbb{F}_p^R) \to H^2(F, \mathbb{F}_p),$$

given that $\mathbb{F}_p^R$ is the set of functions from R to $\mathbb{F}_p$, and $H^1(R)^G$ is the set of invariants under the group action of G on $H^1(R)$. We now can start talking about dimensions, but we first need to define **cohomological dimension**.

Definition 32. The **cohomological dimension** of G, denoted $\mathrm{cd}(G)$, is the least integer n such that $H^k(G) = 0$ for all $k > n$.

In this case, our free group F has cohomological dimension $\mathrm{cd}(F) \leq 1$. For a proof of this, reference [Ser01]. Because of this, we know that $H^2(F) = 0$, and we can utilize sum dimensions, yielding

$$\dim_{\mathbb{F}_p} H^1(G) - \dim_{\mathbb{F}_p} H^1(F) + \dim_{\mathbb{F}_p} H^1(R)^G - \dim_{\mathbb{F}_p} H^2(G) = 0.$$

Since $\dim_{\mathbb{F}_p} H^1(G) = \dim_{\mathbb{F}_p} H^1(F)$, we can conclude that

$$\dim_{\mathbb{F}_p} H^1(R)^G = \dim_{\mathbb{F}_p} H^2(G).$$

Note that the action of any element $\overline{g} \in G$ on any $f \in H^1(R)$ is defined as $(g \cdot f)(r) = g \cdot f(g^{-1}rg)$, where $g \in F$ is a representative of the residue class $\overline{g}$ (think of g as being the element that represents the entire class $\overline{g}$). We can thus conclude the set of G-invariants of $H^1(R)$ is $\{f : R \to \mathbb{F}_p \mid f(r) = f(g^{-1}rg) \forall r \in R\}$. So, the dimension of $H^1(R)^G$ (the set of invariants under the group action of G on $H^1(R)^G$) is the number of conjugacy classes of generators on R. However, R is the smallest normal subgroup containing the generators of R, which lets us know that $\dim_{\mathbb{F}_p} H^1(R) = \dim_{\mathbb{F}_p} H^1(R)^G$. Therefore, we come to the following conclusion:

Theorem 2. *For a pro-p group G,*

$$d(G) = \dim_{\mathbb{F}_p} H^1(G, \mathbb{F}_p) \text{ and } r(G) = \dim_{\mathbb{F}_p} H^2(G, \mathbb{F}_p).$$

3.4. **Completed Group Algebras.** We now want to provide a little more structure to our pro-p groups, and we can do this by defining an object with structure similar to a polynomial ring, allowing us to treat it like a graded algebra.

Definition 33. A **graded algebra** A over a ring is a direct sum of groups or modules such that

$$A = \bigoplus_{n=0}^{\infty} A_n$$

where A_n is a degree n component, and $A_m \cdot A_n \subseteq A_{m+n}$.

The grading in a graded algebra is analogous to the degree of a polynomial in polynomial space.

Definition 34. The **completed group algebra** $\Lambda[[G]]$ of the pro-p group G is

$$\varprojlim (\Lambda[G/N])_{N \in \mathfrak{R}},$$

where $\mathfrak{R}$ is the set of all open normal subgroups of G.

The completed group algebra is the profinite completion of group rings of the quotients of open normal subgroups. One key property of group algebras is the following:

Theorem 3. *If $\phi : G \to G'$ is a morphism of profinite groups with $N = \ker \phi$, we have that the kernel of the induced morphism $\phi' : \Lambda[[G]] \to \Lambda[[G']]$ is the closed ideal $I(N)$ generated by all $h - 1$ where $h \in N$.*

While we will not fully prove this (the full proof can be found in [Koc02], Theorem 7.3 (iii)), we provide an outline for the proof. We know that $I(N) \subseteq \ker \phi$ since $N = \ker \phi$, so we can induce a morphism $\hat{\phi} : \Lambda[[G]]/I(N) \to \Lambda[[G']]$. To prove inclusion in the other direction, we restrict the morphism $\hat{\phi}$ to G, yielding the isomorphism $\{G + I(N)\}/I(N) \to G'$. The inverse map $G' \to \{G + I(N)\}/I(N)$ gives us an inverse of $\hat{\phi}$, letting us know that $I(N) = \ker \phi$.

This theorem is useful for the Golod-Shafarevich Theorem as the presentation of a pro-p group G gives us morphisms $F \to G$ of pro-p groups with kernel R. Thus, we are able to apply what we know about free pro-p groups to general pro-p groups. Free pro-p groups are easy to work with because their group algebras have structures similar to polynomial rings, shown below.

Definition 35. Let Λ be a ring with identity and let m be a positive integer. The **Magnus Algebra $\Lambda(\mathbf{m})$** in the variables $x_1 \ldots, x_m$ over Λ is the algebra of formal noncommutative power series in the x_i with coefficients in Λ.

We can now easily relate rings to Magnus algebras, through the following theorem:

Theorem 4. *If F is a free pro-p group with system of generators $\{s_1, \ldots, s_m\}$, then the completed group ring $\Lambda[[F]]$ is isomorphic to the ring $\Lambda(m)$ by linearly extending the homomorphism $\psi(s_1) = 1 + x_i$*

The details of this proof will be omitted, as they are contained within Theorem 7.16 of [Koc02].

We can now relate $\Lambda[[F]]$ to $\Lambda(m)$, and in the proof of the Golod-Shafarevich theorem, we will consider $\Lambda = \mathbb{F}_p$.

3.5. **Filtrations.**

Definition 36. Let G be a finitely generate pro-p group, and $\mathbb{F}_p[G]$ the group ring over $\mathbb{F}_p$ with augmentation ideal $I(G) = (g-1)\mathbb{F}_p[G]$. For positive integer n, let the ideal $I^n(G)$ in $\mathbb{F}_p[[G]]$ denote the closure of the nth power of $I(G)$ in $\mathbb{F}_p[[G]]$. Then, define the ***n*th modular dimension subgroup of G**, G_n, as

$$G_n = \{g : g - 1 \in I^n(G)\}.$$

The chain of dimension subgroups

$$G = G_1 \supseteq G_2 \supseteq \dots$$

forms the **Zassenhaus Filtration of G.**

Recall the first part of the definition requires Proposition 1. This filtration allows us to divide the relations of a pro-p group into **levels**.

Definition 37. Let G be a pro-p group with minimal presentation $1 \to R \to F \to G \to 1$. The **level** of any $r \in R$ is given my the largest integer m such that $r \in F_m \setminus F_{m+1}$, where $\{F_m : m \in \mathbb{N}\}$ is the Zassenhaus Filtration of F.

If F is a free pro-p group with d generators, each element of $\mathbb{F}_p[[F]] = \mathbb{F}_p(d)$ can be uniquely represented as a linear combination $\sum_K \lambda_K M_K$, where $\lambda_K \in \mathbb{F}_p$ and M_k is a monomial of the variables $\{x_1, \dots, x_d.\}$ This allows us to define something that acts like a degree function on $\mathbb{F}_p(d)$.

Definition 38. Let F be a d-generated free pro-p group, and let $\tau_1, \dots, \tau_d$ be positive integers. The **Lazard valuation of type $(\tau_1, \dots, \tau_d)$** on $\mathbb{F}_p(x_1, \dots, x_d)$ is an additive function $v : \mathbb{F}_p(x_1, \dots, x_d) \to \mathbb{Z} \cup \{\infty\}$ where $v(x_i) = \tau_i$, with the valuation determined additively:

$$v(x_{i_1} x_{i_2} \dots x_{i_d}) = \tau_{i_1} + \tau_{i_2} + \dots + \tau_{i_d},$$
$$v(1) = 0, v(0) = \infty.$$

The valuation of an element $\sum_K \lambda_K M_K$ is defined as

$$v\left(\sum_K \lambda_K M_K\right) = \min\{v(M_K) : \lambda_K \neq 0\}.$$

Remark 2. Extending the definition, for all $a, b \in \mathbb{F}_p(x_1, \dots, x_d)$, we also hold the following properties:

$$v(ab) = v(a) + v(b)$$
$$v(a+b) \geq \min\{v(a), v(b)\}.$$

We can intuitively think of the Lazard Valuation as similar to the $\log(x)$ function.

Definition 39. $\sum_K \lambda_K M_K$ is **homogenous of degree *m*** if $\lambda_K = 0$ for all $v(M_K) \neq m$.

If G is a pro-p group with minimal presentation $1 \to R \to F \xrightarrow{\phi} G \to 1$, then the Lazard valuation on F will also result in a Lazard valuation on G. For any $\beta \in \mathbb{F}_p[[G]]$, we define

$$v(\beta) = \max\{v(\alpha \mid \alpha \in \mathbb{F}_p[[F]], \phi(\alpha) = \beta\}.$$

We can thus define a filtration

$$G_n^v = \{g \in G : v(g-1) \geq n\}$$

defining the **level** of any element in G to be the greats n where $g \in G_n^v$.

In this context, the Zassenhaus filtration is the induced filtration given by the Lazard valuation of type $(1, \dots, 1)$, where valuation is equivalent to taking the degree of a power series.

4. Golod-Shafarevich Theorem

Now that we have the preliminary background for understanding the Golod-Shafarevich theorem, we march forward and begin proving it.

4.1. **Setup.** Like previously, let G be a finitely generated pro-p group with minimal presentation

$$1 \to R \to F \xrightarrow{\phi} G \to 1.$$

The map ϕ induces a map $\mathbb{F}_p[[F]] \to \mathbb{F}_p[[G]]$, which will also be denoted ϕ. For simplicity's sake, let

$$A = \mathbb{F}_p[[F]] \text{ and } \mathbb{F}_p[[G]].$$

Let $\{s_1, \dots, s_d\}$ be a lift of the generators of G to F (what we send the generators to), and let $\{\rho_1, \dots, \rho_r\}$ be a system of relations for G (and thus a system of generators for R). Using the results from Magnus Algebras (which relate rings to specific polynomials), we see that

$$A \cong \mathbb{F}_p(x_1, \dots, x_d)$$

under the isomorphism given by $s_i \mapsto x_i + 1$. Then, since the kernel of $\phi : A \to B$ is given by $I(R)$ (generated by $\{\rho_1 - 1, \dots, \rho_r - 1\}$), we have $B \cong A/I(R)$. Denote the generators of B by $y_i = \phi(s_i) - 1 = \phi(x_i)$.

Let v be a Lazard valuation of type $(\tau_1, \dots, \tau_d)$ on A. Without loss of generality, we assume that $\tau_i \leq \tau_{i+1}$ for $1 \leq i \leq d-1$(if not, we can reorder x_i to fit this inequality). Furthermore, we can suppose that the relations are ordered so the levels are monotonically increasing. The valuation v (on A) induces a valuation on B (similarly denoted v), giving us the filtration

$$I_n = \{b \in B \mid v(b) \geq n\}$$

for $n \in \mathbb{Z}$. If $n \leq 0$, we let $I_n = B$. We also introduce the sequence

$$c_n = \dim_{\mathbb{F}_p} B/I_N.$$

With this filtration, we count relations and generators by level, defining

$$r_n = |\{\rho_i \mid v(\rho_i - 1) = n\} \text{ and } d_n = |\{x)i \mid \tau_i = n\}|.$$

Note that $r_0 = 1$, since the valuation of a constant is 0.

Our proof will rely on the sequence

$$B^r \xrightarrow{\phi_1} B^d \xrightarrow{\phi_0} B \xrightarrow{\epsilon} \mathbb{F}_p \to 0,$$

with its restriction to the sequence

$$\bigoplus_{i=1}^{r} I_{n-v(\rho_i-1)} \xrightarrow{\psi_1} \bigoplus_{i=1}^{d} I_{n-\tau_i} \xrightarrow{\psi_0} I_n \to 0.$$

We first prove that these sequences are exact. Then, after a little work with dimensions, we will produce an inequality in terms of r_i and d_i, yielding the Golod-Shafarevich Theorem.

4.2. **Proof.** Consider the sequence

$$B^r \xrightarrow{\phi_1} B^d \xrightarrow{\phi_0} B \xrightarrow{\epsilon} \mathbb{F}_p \to 0, \tag{1}$$

where ϵ is the augmentation map of $B \to \mathbb{F}_p$. One way to think about ϵ is to consider B as $A/I(R)$ (shown in Equation 4.1) so ϵ needs to only be $(0, \dots, 0)$. We define

$$\phi_0(b_1, \dots, b_d) = \sum_{i=1}^{d} b_i y_i.$$

To define ϕ_1, we first note that each $\rho_i - 1$ has unique representation in the ring of formal power series $\mathbb{F}_p(x_1, \dots, x_d)$:

$$\rho_i - 1 = \sum_{j=1}^{d} z_{ij} x_j,$$

since we can reference monomials based on the last free variable. Then, we can define

$$\phi_1(b_1, \dots, b_r) = \left(\sum_{i=1}^{r} b_i \phi(z_{i1}), \dots, \sum_{i=1}^{r} b_i \phi(z_{id}) \right).$$

Proposition 3. The sequence given in Equation 1 is exact.

As a little refresher, a sequence is exact if the image of a morphism is the kernel of the next ($\mathrm{im}(f_n) = \ker(f_{n+1})$).

Proof: Since ϵ is surjective, the sequence is exact at $\mathbb{F}_p$. For B, we note the augmentation idea of B is generated by the elements y_i, so $\mathrm{im}(\phi_0) = \ker(\epsilon)$.

Exactness at B^d is found from the fact that $\ker(\phi)$ is generated by the $\rho_i - 1$. Specifically, let $(b_1, \dots, b_r) \in B^r$. Then

$$\begin{aligned}
\phi_0(\phi_1(b_1, \dots, b_r)) &= \sum_{j=1}^{d} \sum_{i=1}^{r} b_i \phi(z_{ij}) y_j \\
&= \sum_{j=1}^{d} \sum_{i=1}^{r} b_i \phi(z_{ij} x_j) \\
&= \sum_{i=1}^{r} b_i \sum_{j=1}^{d} \phi(z_{ij} x_j) \\
&= \sum_{i=1}^{r} b_i \phi(\rho_i - 1) \\
&= 0.
\end{aligned}$$

This last step follows from the fact that $\ker(\phi)$ is generated by the $\rho_i - 1$, so thus $\phi(\rho_i - 1) = 0$. From these equations, we know that $\mathrm{im}(\phi_1) \subseteq \ker(\phi_0)$. To show equality, we do double-containment, so we show inclusion in the other direction. Let $b_1, \dots b_d \in B$ such taht $\sum_{j=1}^{d} b_j y_j = 0$. We lift the b_j to $a_j \in A$, so that $\sum_{j=1}^{d} a_j x_j \in \ker(\phi)$. Since $\ker(\phi)$ is generated by the $\rho_i - 1 = \sum_{j=1}^{d} z_{ij} x_j$, we have

$$\sum_{j=1}^{d} a_j x_j = \sum_{i=1}^{r} a_i' \sum_{j=1}^{d} z_{ij} x_j = \sum_{j=1}^{d} \sum_{i=1}^{r} a_i' z_{ij} x_j,$$

which gives us $a_j = \sum_{i=1}^{r} a_i' z_{ij}$. Thus, we have

$$\begin{aligned}\phi_1(\phi(a_1'), \ldots, \phi(a_r')) &= \left(\sum_{i=1}^{r} \phi(a_i' z_{i1}), \ldots, \sum_{i=1}^{r} \phi(a_i' z_{id}) \right) \\ &= (\phi(a_1), \ldots, \phi(a_d)) \\ &= (b_1, \ldots, b_d),\end{aligned}$$

so thus we know that $(b_1, \ldots, b_d) \in \mathrm{im}(\phi_1)$. Therefore, Equation 1 is exact. □

For a fixed integer n, the restriction of Equation 1 induces the sequence

$$\bigoplus_{i=1}^{r} I_{n-v(\rho_i-1)} \xrightarrow{\psi_1} \bigoplus_{i=1}^{d} I_{n-\tau_i} \xrightarrow{\psi_0} I_n \to 0, \tag{2}$$

where ψ_1 and ψ_0 denote the restriction of ϕ_1 and ϕ_0, respectively.

First, we verify that $\oplus_{i=1}^{r} I_{n-v(\rho_i-1)}$ is sent to $\oplus_{i=1}^{d} I_{n-\tau_i}$ under ϕ_1. Let $(h_1, \ldots, h_r) \in \oplus_{i=1}^{r} I_{n-v(\rho_i-1)}$, which gives us

$$v(h_i) \geq n - v(\rho_i - 1). \tag{3}$$

Furthermore, since $\rho_i - 1 = \sum_{j=1}^{d} z_{ij} x_j$, we have that

$$v(\rho_i - 1) = \min\{v(z_{ij} + \tau_j : 1 \leq j \leq d\}.$$

This gives us

$$v(z_{ij}) = v(\phi(z_{ij})) \geq v(\rho_i - 1) - \tau_j. \tag{4}$$

Combining Equation 3 and Equation 4 together, we have

$$v(h_j) + v(\phi(z_{ij})) \geq n - \tau_j.$$

We also have that the jth term $\phi(1)(h_1, \ldots, h_r)$ is

$$v\left(\sum_{i=1}^{r} h_j \phi(z_{ij}) \right) = \min\{v(h_j) + v(\phi(z_{ij}))\} \geq n - \tau_j,$$

so $\phi(h_1, \ldots, h_r) \in \oplus_{i=1}^{d} I_{n-\tau_i}$, as desired.

Similarly, ϕ_0 sends elements of $\oplus_{i=1}^{d} I_{n-\tau_i}$ to I_n. Now that we've verified the restriction under the maps is correct, we show that Equation 2 is exact at I_n.

Proposition 4. The map ψ_0 is surjective.

Proof: Let $h \in I_n$. Select a $g \in A$ such that $\phi(g) = h$ and $v(h) = v(g)$. Such a g exists, since ϕ is surjective and $v(h) = \max\{v(g) : \phi(g) = h\}$. This g has a unique representation $\sum_{i=1}^{d} g_i x_i$. Denoting the homogenous components of g with degree m by $g^{(m)}$, we have

$$g^{(m)} = \sum_{i=1}^{d} g_i^{(m-\tau_i)} x_i.$$

Since g cannot have homogenous components of degree less than n ($\phi(g) \in I_n$), we have $v(g_i) \geq n - \tau_i$, which means that $v(h_i) \geq n - \tau_i$, where $h_i = \phi(g_i)$. Thus, $(h_1, \ldots, h_d \in \oplus_{i=1}^{d} I_{n-v(x_i)}$, and

$$\psi_0(h_1, \ldots, h_d) = \sum_{i=1}^{d} h_i y_i = \sum_{i=1}^{d} \phi(g_i x_i) = \phi(g) = h.$$

□

Since Equation 2 is exact in I_n and Equation 1 is exact, the factor sequence

$$\bigoplus_{i=1}^{r} B/I_{n-v(\rho_i-1)} \to \bigoplus_{i=1}^{d} B/I_{n-\tau_i} \to B/I_n \to \mathbb{F}_p \to 0$$

is exact, which gives us

$$\sum_{i=1}^{n} r_i c_{n-i} - \sum_{i=1}^{n} d_i c_{n-i} + c_n \geq 1.$$

Noting that $r_0 = 1, d_0 = 0, c_0 = 1$, we can rewrite this as

$$\sum_{i=0}^{n} (r_i - d_i) c_{n-i} \geq 1. \tag{5}$$

With this, we make the following claim:

Proposition 5. Let G be a finite pro-p group, with d_i and r_i defined as above, Then

$$\phi_v(t) = 1 + \sum_{n=1}^{\infty} (r_n - d_n) t^n$$

converges, and is greater than 0 for $0 < t < 1$.

Proof: Multiplying each side of Equation 5 by r^n and summing for all n, we have

$$\sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} (r_i - d_i) c_{n-i} \right) t^n \geq \sum_{n=0}^{\infty} t_n = \frac{1}{1-t}.$$

The Cauchy product of the two series gives us

$$\left(\sum_{n=0}^{\infty} (r_n - d_n) t^n \right) \left(\sum_{n=0}^{\infty} c_n t^n \right) = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} (r_i - d_i) c_{n-i} \right) t^n.$$

Thus, we have that

$$\left(\sum_{n=0}^{\infty} (r_n - d_n) t^n \right) (c_n t_6 n) \geq \frac{1}{1-t}.$$

We have that $\sum_{n=0}^{\infty} c_n t^n$ is convergent for $0 < t < 1$, since c_n is bounded above if G is finite. See lemma 7.9 in [Koc02] for a proof that $I^n(G) = 0$ for sufficiently large n.

Furthermore, since r is finite, we can also assume without loss of generality that almost all of the $r_n = 0$. Because almost all the d_n and r_n are 0 and $\sum c_t^n$ converges, the lefthand side is a polynomial in t and also converges.

Dividing both sides by $\sum c_n t^n > 0$, and noting that $\frac{1}{1-t} > 0$, we get

$$\infty > \sum_{n=0}^{\infty} (r_n - d_n) t^n > 0.$$

Rewriting gives us

$$\phi_v(t) = 1 + \sum_{n=1}^{\infty} (r_n - d_n) t^n > 0,$$

as desired. □

Corollary 1. *Let G be a finite pro-p group with $d(G)$ d and $r(G) = r$, and let $1 \to R \to F \to G \to 1$ be a minimal presentation. of G with $R \subset F_m$ where $\{F_n\}$ is the Zassenhaus filtration of G. Then,*

$$r > \frac{d^m}{m^m}(m-1)^{m-1}.$$

Proof: Since $R \subset F_m$, we have $\phi_v(t) = 1 - dt + rt^m$. Suppose for the sake of contradiction that

$$r \leq \frac{d^m}{m^m}(m-1)^{m-1}.$$

This gives us

$$\sqrt[m-1]{\frac{d}{mr}} \geq \frac{m}{d(m-1)}.$$

Setting $t = \sqrt[m-1]{\frac{d}{mr}}$ gives

$$\begin{aligned}\phi_v\left(\sqrt[m-1]{\frac{d}{mr}}\right) &= 1 - d \cdot \sqrt[m-1]{\frac{d}{mr}} + r \cdot \frac{d}{mr} \cdot \sqrt[m-1]{\frac{d}{mr}} \\ &\leq 1 - \frac{m}{m-1} + \frac{1}{m-1} = 0,\end{aligned}$$

which contradicts Proposition 5. □

For $m = 2$, Corollary 1 gives us the Golod-Shafarevich Theorem in the form given of Gaschütz and Vinberg's refinement [Koc69].

Theorem 5 (Golod-Shafarevich Theorem)**.** *Let G be a nontrivial finite pro-p group with $d(G) = d$ denoting the generator rank, and $r(G) = r$ denoting the relation rank. Then, $r > \frac{d^2}{4}$.*

By the Golod-Shafarevich Theorem, if you have $r \leq \frac{d^2}{4}$, then this group G must be an infinite pro-p group. However, d and r can still be finite integers, which provides a negative to the general Burnside Problem.

Acknowledgements

I am deeply grateful for everyone who contributed to the completion of my paper.

I would like to first express my sincere gratitude to Simon Rubinstein-Salzedo, whose guidance, support, and insightful feedback were invaluable throughout the development of this work. I am constantly amazed by not just the sheer amount of knowledge he has, but how well he conveys it to high schoolers like me, and how hard of a worker he is. I am also thankful to the Euler Circle community for providing the academic environment that made this paper-writing experience so enjoyable. While I may eventually forget some of the content of the talks I listened to, I won't forget the conversations and feelings I had when listening to people excitedly talk about their papers.

I would like to also extend my thanks to my TA, Zarif Ahsan. Every time that we met, he was able to give me some good intuition behind the concepts that I was confused about, and he was always there to help me. I won't forget the conversations we had about math, college, and his calm and inviting personality.

Finally, I am grateful to my family and friends for their unwavering support and patience.

References

[Bur02] William Burnside. *On an unsettled question in the theory of discontinuous groups*, volume 33. Quarterly Journal of Mathematics, 1902.

[Cas22] Jordi Casadevall. The golod-shafarevich inequality and the class field tower problem, 2022.

[DSSD99] Marcus Du Sautoy, Dan Segal, and JD Dixon. *Analytic pro-p groups*. Cambridge University Press, 1999.

[GS64] Evgeniy Golod and Igor Rostislavovich Shafarevich. On the class field tower. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 28(2):261–272, 1964.

[Koc69] Helmut Koch. Zum satz von golod-schafarewitsch. *Mathematische Nachrichten*, 42(4-6):321–333, 1969.

[Koc02] Helmut Koch. *Galois theory of p-extensions*. Springer Science & Business Media, 2002.

[Lin10] Julie Linman. *Burnside's Theorem*. Oregon State University, 2010.

[NA68] Petr S Novikov and Sergei I Adjan. Infinite periodic groups. ii. *Mathematics of the USSR-Izvestiya*, 2(2):241, 1968.

[Ser01] Jean-Pierre Serre. *Galois Cohomology*. Springer, 1 edition, 2001.

[Zel91] Efim Isaakovich Zel'manov. Solution of the restricted burnside problem for groups of odd exponent. *Mathematics of the USSR-Izvestiya*, 36(1):41, 1991.

[Zho17] Kathleen Zhou. The golod-shafarevich theorem and the class field tower problem, 2017.