Topics in Infinite Groups

Lectures by Jack Button

Notes by Tony Feng

Lent 2014
These are lecture notes for a course taught in Cambridge during Lent 2014 by Jack Button, on a topics in infinite group theory. There are likely to be errors, which are solely the fault of the scribe. If you discover any, please contact me tonyfeng009@gmail.com.
# Contents

**Preface**  

| 1 Preliminaries | 1  
|----------------|---  
| 1.1 Subgroups | 1  
| 1.2 Normal subgroups | 3  
| 1.3 Homomorphisms | 4  
| 1.4 Group Properties | 4  
| 1.5 Actions | 5  
| 1.6 Automorphisms | 6  
| 1.7 Direct Products | 6  
| 1.8 Semidirect products | 7  

| 2 Abelian Groups | 8  
|----------------|---  
| 2.1 Finitely generated abelian groups | 8  
| 2.2 Infinitely generated abelian groups | 9  
| 2.3 Abelianization | 9  

| 3 Free groups | 11  
|---------------|---  
| 3.1 Construction of free groups | 11  
| 3.2 General properties of free groups | 12  
| 3.3 Free Products | 15  
| 3.4 Nielson-Schreier Theorem | 16  

| 4 Presentations of Groups | 20  
|--------------------------|---  
| 4.1 Construction and basic properties | 20  
| 4.2 Free products with amalgamation and HNN extensions | 23  

| 5 Soluble, Polycyclic, and Nilpotent Groups | 27  
|---------------------------------------------|---  
| 5.1 Soluble groups | 27  
| 5.2 Polycyclic Groups | 28  
| 5.3 Nilpotent groups | 30  


# Contents

6 Finite index subgroups and virtual properties 34  
6.1 Finite index subgroups ........................................... 34  
6.2 The Regular Representation ...................................... 35  
6.3 Virtual properties ................................................... 38  
6.4 Virtually polycyclic Groups ........................................ 39  
6.5 Virtually soluble groups ............................................ 40  

7 Maximal (normal) subgroups 42  
7.1 Maximal subgroups .................................................. 42  
7.2 Infinite Simple Groups .............................................. 43  

8 Residually Finite, Hopfian, Baumslag-Solitar 44  
8.1 Residually Finite Groups ........................................... 44  
8.2 Hopfian Groups ...................................................... 48  
8.3 Baumslag-Solitar Groups ........................................... 48  

9 The Generalized Burnside Problem 51  

10 Example Sheet 1 54  

11 Example Sheet 2 58
Chapter 1

Preliminaries

This is a course on infinite groups. We first establish some basic technical definitions and properties of groups.

1.1 Subgroups

Let $G$ be a group. If $H$ is a subgroup of $G$, then we write $H \leq G$. If $L \leq H$ and $H \leq G$, then $L \leq G$. An arbitrary intersection of subgroups of $G$ is again a subgroup. In contrast, a union of subgroups is not in general a subgroup. However, it is true in the following case.

**Proposition 1.1.** If $H_1 \leq H_2 \leq \ldots \leq G$ is an ascending chain of subgroups, then $\bigcup_{n=1}^{\infty} H_n = G$.

**Example 1.2.** Let $G = S^1 \subset \mathbb{C}$. Let $G_n$ be the $2^n$ roots of unity. Then $G_n \subset G_{n+1} \subset \ldots$, and the union consists of all roots $2^n$-torsion roots of unity.

If $G$ is a subgroup and $X \subset G$ a subset, then the group generated by $X$, denoted $\langle X \rangle$, is the intersection of all subgroups of $G$ containing $X$. This intersection is nonempty, as $G$ is itself such a group.

**Definition 1.3.** $G$ is **finitely generated** if there exists a finite set $\{g_1, \ldots, g_k\}$ such that $\langle g_1, \ldots, g_k \rangle = G$.

**Example 1.4.** $G$ is **cyclic** if $\langle g \rangle = G$ for some $g \in G$. Either $G = \mathbb{Z}$ or $G$ has order $n$, in which case we write $G = C_n$.

**Proposition 1.5.** If $X \subset G$, then the elements of $\langle X \rangle$ are

$$x_1^{n_1}x_2^{n_2} \ldots x_k^{n_k}$$

where the $x_i$ are not necessarily distinct.

**Corollary 1.6.** Any finitely generated group is countable.
Any ascending sequence of subgroups

\[ H_1 \leq H_2 \leq \ldots \leq G \]

either terminates, i.e. \( H_N = H_{N+n} \) for sufficiently large \( N \), or is infinitely strictly ascending after throwing away duplicates.

**Definition 1.7.** A group \( G \) has the \textit{max} property if every ascending sequence of subgroups terminates.

**Theorem 1.8.** \( G \) has \textit{max} if and only if all subgroups \( H \leq G \) are finitely generated.

**Proof.** Consider the forward direction first. If \( H \) is a subgroup, let \( h_1 \) be any element of \( H \). If \( H = \langle h_1 \rangle \), then \( H \) is finitely generated. Otherwise, we can pick some \( h_2 \in H \) such that \( h_2 \notin \langle h_1 \rangle \), giving the strict inclusion

\[ \langle h_1 \rangle \leq \langle h_2 \rangle \leq \ldots \]

This sequence must terminate at a finite point, meaning that \( H \) is finitely generated.

Conversely, if any subgroup is finitely generated, and

\[ H_1 \ldots H_2 < \ldots \]

is an ascending chain, then \( \bigcup H_i \) is a subgroup, hence finitely generated. Its generators must all appear in the sequence at some finite point, after which the sequence must terminate.

**Example 1.9.** It is possible for \( G \) to be finitely generated, but have subgroups that are not finitely generated. For an example, consider the permutation group of \( \mathbb{Z} \), and let \( S_0(\mathbb{Z}) \) be the group of \textit{finitely supported} permutations, i.e. permutations that fix everything outside \([-N, N]\) for some \( N \) (depending on the group element). This is generated by the transposition \( (n, n+1) \) since these generate \( S_N \) for each finite \( N \). Obviously, \( S_0(\mathbb{Z}) \) cannot be finitely generated. Now let \( f \) be the permutation sending \( n \mapsto n+1 \) for all \( n \). Then the group \( \langle f, S_0(\mathbb{Z}) \rangle \) is generated by \( f \) and \( (0,1) \), since \( f^n(01)f^{-n} = (n, n+1) \). So here we have a group generated by two elements, with a subgroup that is not finitely generated.

If \( H \subset G \), then the \textit{left cosets} of \( H \) in \( G \) are the sets of the form \( gH = \{ gh : h \in H \} \).

**Proposition 1.10 (Lagrange).** The \textit{left cosets} of \( H \) in \( G \) form a partition of \( G \), and any left coset is in bijection with \( H \).

The same holds for right cosets.
Definition 1.11. 1. The \textit{index} of $H$ in $G$ is the cardinality of the set of left cosets. It is written $[G: H]$ if it is finite.

2. A left (or right) \textit{transversal} is a choice of left (or right) coset representatives, with exactly one for each coset.

1.2 Normal subgroups

A subgroup $N \leq G$ is normal if any of the following equivalent conditions holds:

1. $gN = Ng$ for all $g \in G$,
2. $gNg^{-1} = N$ for all $g \in G$,
3. $gNg^{-1} \leq N$ for all $g \in G$,
4. $N$ is a union of conjugacy classes of $G$.

Example 1.12. If $G$ is abelian, then all its subgroups are normal. If $[G : H] = 2$, then $H$ is normal in $G$.

Proposition 1.13. If $N \triangleleft G$ and $H \leq G$, then $N \cap H \triangleleft H$. Also, an arbitrary intersection of normal subgroups is normal.

Example 1.14. If $N \triangleleft H$ and $H \triangleleft G$, then it is not necessarily true that $N \triangleleft G$.

Proposition 1.15. If $N_1 \leq N_2 \leq \ldots \leq G$, and each $N_n \triangleleft G$, then

$$\bigcup_{n=1}^{\infty} N_n \triangleleft G.$$  

If $X \subset G$ is a subset, then we denote by $\langle \langle X \rangle \rangle$ the \textit{normal closure} of $X$ in $G$, which is the intersection of all normal subgroups containing $X$. Intuitively, this is the smallest normal subgroup containing $X$.

Proposition 1.16. If $X \subset G$, then the elements of $\langle \langle X \rangle \rangle$ are

$$g_1x_1^{n_1}g_1^{-1}g_2x_2^{n_2}g_2^{-1}\ldots g_kx_k^{n_k}g_k^{-1}$$

for $n_1, \ldots, n_k \in \mathbb{Z}$, $x_1, \ldots, x_k \in X$ and $g_1, \ldots, g_k \in G$ (not necessarily distinct).

The set product of two subgroups $A, B \leq G$ is

$$AB = \{ab : a \in A, b \in B\}.$$  

This is not in general a subgroup, but it is if one of $A, B$ is normal.
Proposition 1.17. 1. \( AB \leq G \iff AB = BA \).
2. If \( AB \leq G \), then \( AB = \langle A, B \rangle \).

1.3 Homomorphisms

A homomorphism between groups \( G \) and \( H \) is a set map \( \theta : G \to H \) satisfying \( \theta(gg') = \theta(g)\theta(g') \) for all \( g, g' \in G \). It is an isomorphism if it is bijective, which is equivalent to the existence of an inverse homomorphism.

Theorem 1.18. If \( \theta : G \to H \) is a homomorphism and \( A \leq G, B \leq H \), then \( \theta(A) \leq H \) and \( \theta^{-1}(B) \leq G \).

If \( B \triangleleft H \) then \( \theta^{-1}(B) \triangleleft G \), and if \( \theta \) is surjective then \( A \triangleleft G \implies \theta(A) \triangleleft H \).

In particular, \( \theta(G) \) is a subgroup of \( H \) called the image, and \( \theta^{-1}(e) \) is a subgroup of \( G \) called the kernel.

If \( N \triangleleft G \) then \( G/N \) is a group under the obvious multiplication.

Theorem 1.19 (Homomorphism theorem). If \( \theta : G \to H \) is a homomorphism, then \( G/\ker \theta \cong \theta(G) \).

Theorem 1.20 (Correspondence theorem). If \( N \triangleleft G \), then the subgroups of \( G/N \) are exactly \( H/N \) for \( N \leq H \leq G \), and the normal subgroups are exactly \( L/N \) for \( N \leq L \triangleleft G \).

Theorem 1.21 (Product isomorphism Theorem). If \( H \leq G \) and \( N \triangleleft G \), then \( H/H \cap N \cong HN/N \).

Theorem 1.22 (Quotient Isomorphism Theorem). Let \( N, L \triangleleft G \) with \( N \leq L \). Then
\[
\frac{G/N}{L/N} \cong G/L.
\]

If \( G/N \cong Q \), we say that \( G \) is an extension of \( N \) by \( Q \), or \( G \) is \( N \)-by-\( Q \) for short.

Lemma 1.23. If \( G \) is \( A \)-by-\((B\text{-by-}C) \), then \( G \) is \( (A\text{-by-}B)\text{-by-}C \).

Proof. If \( G/A \cong Q \), where \( Q/B \cong C \), then we have a map \( A \to C \) by composing. The kernel of this map to \( Q \), with kernel \( A \).

1.4 Group Properties

When we have a group property, it is worth asking whether it is preserved by (i) subgroups, (ii) quotients, and (iii) extensions (if \( N, Q \) have the property and \( G/N \cong Q \), does \( G \) have the property?)
Theorem 1.24. The properties finitely generated and max are preserved by extensions.

Proof. Suppose $G$ is $N$-by-$Q$. For finite generation, take a generating set for $N$ and the pre-image of a generating set for $Q$; this generates $G$.

For max, take any ascending sequence $H_1 \leq H_2 \leq \ldots \subset G$. The image in $Q$ is an ascending sequence of subgroups, so it stabilizes. The intersections $H_1 \cap N \leq H_2 \cap N \leq \ldots$ also stabilize, so the original sequence does too. 

1.5 Actions

A group $G$ acts on a set $X$ if there is a function $\psi : G \times X \to X$ such that

$$\psi(e, x) = x \quad \text{and} \quad \psi(g_1, \psi(g_2, x)) = \psi(g_1 g_2, x) \quad \forall g_1, g_2 \in G, x \in X.$$ 

Observe that $\psi(g, -)$ is a permutation of $X$ for each $g$.

Equivalently, an action is a homomorphism $\rho : G \to S(X)$, where $S(X)$ is the symmetric group on $X$.

We say that $G$ acts

- faithfully (or effectively) if $\rho$ is injective.
- free if $\rho(g)(x) = x \implies g = e$.

The orbit of $x \in X$ is

$$\{\rho(g)(x) : g \in G\}$$

and the stabilizer is

$$G_x = \{g \in G : \rho(g)(x) = x\}.$$ 

We say that the action is transitive if there is only one orbit.

Observe that if $y = \rho(g)(x)$, then $G_y = gG_x g^{-1}$.

Theorem 1.25. If $G$ acts on $X$, then for all $x \in X$ the sets $\text{Orb}(x)$ and $G/G_x$ are in bijection.

Example 1.26. 1. $G$ acts on itself by $\rho(g)(x) = gx$. This is transitive and free, giving an injection $G \leq S(G)$.

2. $G$ acts on itself by conjugation. For this action, $\text{Orb}(x)$ is called the conjugacy class of $x$ and the stabilizer is called the centralizer of $x$, and denoted $C_G(x)$. 

For $H \leq G$, we define

$$C_G(H) = \bigcap_{h \in H} C_G(h).$$

3. $G$ acts on its set of subgroups by conjugation. The stabilizer of $H$ is called the normalizer of $H$, and denoted $N(H)$. If $H$ is normal then $N(H) = G$.

### 1.6 Automorphisms

An isomorphism $G \to G$ is called an automorphism of $G$.

**Example 1.27.** For any $g \in G$, there is an isomorphism $\alpha_g(x) = gxg^{-1}$ is an automorphism. Such automorphisms are called “inner.”

Note that $\alpha_g = e \iff g \in Z(G)$, so $G/Z(G) \cong \text{Inn}(G)$. Moreover, all automorphisms form a group $\text{Aut}(G)$, and $\text{Inn}(G) \triangleleft \text{Inn}(G)$. The quotient is called the outer automorphism group, $\text{Out}(G)$.

**Definition 1.28.** A subgroup $C \leq G$ is *characteristic* in $G$ if $\alpha(C) = C$ for all $\alpha \in \text{Aut}(G)$.

In particular, a characteristic subgroup is normal because normal subgroups are those which are preserved by inner automorphisms. We saw that normality is a property that is not well-behaved under inclusion; characteristic is stronger condition that can be used to preserve normality.

**Proposition 1.29.**

1. If $A$ is characteristic in $B$ and $B$ is characteristic in $C$, then $A$ is characteristic in $C$.

2. If $A$ is characteristic in $B$ and $B \triangleleft C$, then $A \triangleleft C$.

### 1.7 Direct Products

If $G_1, G_2$ are two groups, then the direct product is defined to be $G_1 \times G_2$ as a set, with multiplication

$$(x_1, x_2) \cdot (y_1, y_2) = (x_1y_1, x_2y_2).$$

This is called an *external direct product* (without reference to inclusion in a common enveloping group).

**Proposition 1.30.** If $M, N \triangleleft G$ with $MN = G$ and $M \cap N = \{e\}$, then $\theta : M \times N \to G$ given by $\theta(m, n) = mn$ is an isomorphism.
Definition 1.31. If $G_n$ is a family of groups indexed by $n \in \mathbb{N}$, the Cartesian product $\prod_n G_n$ is the group whose underlying set is the set Cartesian product, with pointwise multiplication.

The direct product $\times_n G_n$ is the subgroup of sequences which are eventually $e$.

1.8 Semidirect products

Definition 1.32. Given groups $G_1, G_2$ and a homomorphism $\phi : G_2 \to \text{Aut}(G_1)$ the semidirect product $G_1 \rtimes_\phi G_2$ is the group whose underlying set is $G_1 \times G_2$, and with multiplication

$$(x_1, x_2) \cdot (y_1, y_2) = (x_1(\phi(x_2) : y_1), x_2 y_2).$$

Example 1.33. If $\phi$ is the trivial homomorphism, then we recover the direct product.

Example 1.34. Let $Z = \langle z \rangle, C_2 = \{e, c\}$, and $\phi(c)(z) = -z$. Then $Z \rtimes_\phi C_2$ is the “infinite dihedral group.” With $Z$ replaced by $Z/n$, we obtain the usual dihedral groups.

If

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$$

is a short exact sequence of abelian groups, then a splitting $s : C \rightarrow B$ gives a semidirect product structure $B \cong A \rtimes C$. Indeed, for any $c \in C$ we have $s(c)as(c)^{-1} \in A$, so this defines a homomorphism $C \rightarrow \text{Aut}(A)$. Then any $b \in B$ can be written uniquely as $ac$, and if $b_1 = a_1s(c_1), b_2 = a_2s(c_2)$, then

$$b_1b_2 = a_1s(c_1)a_2s(c_2) = a_1(s(c_1)a_2s(c_1)^{-1})s(c_1c_2)$$

which is the semidirect product group structure. This proves:

Proposition 1.35. If $H \leq G$ and $N \triangleleft G$ with $NH = G$ and $N \cap H = I$, then $\theta : N \rtimes H \rightarrow G$ given by $\theta(n, h) = nh$ and $\phi(h)(n) = hnh^{-1} \in N$ is an isomorphism.

Example 1.36. Consider $\text{SL}(2, \mathbb{C}) \triangleleft \text{GL}(2, \mathbb{C})$ with quotient $\mathbb{C}^\times$. The quotient map $\text{SL}(2, \mathbb{C}) \rightarrow \text{PSL}(2, \mathbb{C})$ does not split because there are elements of order four in $\text{SL}(2, \mathbb{C})$ having image of order two in $\text{PSL}(2, \mathbb{C})$, but the only element of order two in $\text{SL}(2, \mathbb{C}) = -I$. 
Chapter 2
Abelian Groups

2.1 Finitely generated abelian groups

The fundamental theorem of abelian groups says that a finitely generated abelian group is isomorphic to a product of cyclic groups, which are either $\mathbb{Z}$ or $\mathbb{Z}/n$. There are two normal forms for such a presentation.

Theorem 2.1 (Rational form). Any finitely generated abelian group $G$ is isomorphic to a finite product of cyclic groups

$$G \cong \mathbb{Z}^r \times C_{d_1} \times \ldots \times C_{d_s}$$

where $d_i \mid d_{i+1}$.

Theorem 2.2 ($p$-primary). Any finitely generated abelian group $G$ is isomorphic to a finite product group

$$G \cong \mathbb{Z}^r \times C_{p_1} \times \ldots \times C_{p_m}$$

where each $p_i$ is prime, and $C_{p_i}$ is a product of simple $p_i$-groups.

Definition 2.3. Let $G$ be any group. Then $G$ is torsion if all elements have finite order. $G$ is torsion-free if $e$ is the only element with finite order.

Definition 2.4. If $G$ is a finitely generated group, we let $d(G)$ be the minimum size of a generating set.

Proposition 2.5. For a finitely generated abelian group $G \cong \mathbb{Z}^r \times C_{d_1} \times \ldots \times C_{d_s}$, we have $d(G) = r + s$.

Proof. It’s obvious that $G$ can be generated by $r + s$ elements, so we have to show that it cannot be generated by fewer elements. To see this, take a prime $p \mid d_s$. Then there exists a map $\theta : G \to (C_p)^{r+s}$. The image is a vector space over $\mathbb{F}_p$, and a generating set for $G$ must be sent to a spanning set over the vector space $\mathbb{F}_p^{r+s}$. By linear algebra, the spanning set must have size at least $r + s$. \qed
2.2 Infinitely generated abelian groups

What about abelian groups that are not finitely generated? (Examples include \( \mathbb{Q}, \mathbb{Q}/\mathbb{Z}, \mathbb{R}, \mathbb{R}/\mathbb{Z} \), etc.) The theory has quite a different flavor, and we won’t spend much time on it. However, here is a brief discussion.

If \( \mathcal{P} \) is a property preserved by subgroups, we say that \( G \) is locally \( \mathcal{P} \) if \( H \) has \( \mathcal{P} \) for all finitely generated subgroups \( H \leq G \).

- Consider \( G = \mathbb{Q} \). Any finitely subgroup \( H \) is of the form \( \langle p_1/q_1, \ldots, p_k/q_k \rangle \). Then \( H \leq \langle 1/q_1 \ldots q_k \rangle \) and is therefore cyclic. So \( \mathbb{Q} \) is locally cyclic, and torsion free.

- By extending the previous analysis, \( \mathbb{Q}/\mathbb{Z} \) is locally cyclic and locally finite.

**Example 2.6.** If \( A \times Z \cong B \times Z \), then it is not necessarily the case that \( A \cong B \). But if \( A \) and \( B \) are (arbitrary) abelian groups satisfying this property, then it is true that \( A \cong B \).

**Proof.** Suppose that \( \Gamma = A \times Y = B \times Z \), where \( Y, Z \cong Z \). We consider two cases.

- If \( A \leq B \), then \( \Gamma/A \cong Z \rightarrow \Gamma/B \cong Z \). This map must be an isomorphism, so we get \( A = B \) (if \( b \in B \setminus A \), then \( b \) is in the kernel).

- If we have no containments between \( A \) and \( B \), then the second isomorphism theorem tells us that

\[
A/A \cap B \cong AB/B \leq \Gamma/B \cong Z,
\]

and we have a proper containment \( B < AB \). So then \( A/A \cap B \) is isomorphic to a nontrivial subgroup of \( Z \), all of which are isomorphic to \( Z \). We haven’t yet used the abelian assumption, but we do now. Since \( Z \) is free in the category of abelian groups, this gives a splitting \( A \cong (A \cap B) \times Z \). Now swap the roles of \( A \) and \( B \).

\[ \square \]

2.3 Abelianization

Let \( G \) be any group. If \( x, y \in G \), then we define the commutator

\[
[x, y] = xyx^{-1}y^{-1}.
\]

**Definition 2.7.** The commutator (derived) subgroup \( G' \) of \( G \) is the group generated by \( [x,y] \) for all \( x, y \in G \).
Chapter 3. Free groups

Proposition 2.8. $G' \triangleleft G$ with $G/G'$ abelian.

Proof. For $\alpha$ any automorphism of $G$, we have

$$\alpha([x,y]) = [\alpha(x), \alpha(y)] \in G'. $$

So in fact, $G'$ is characteristic in $G$, and that immediately implies that it is a normal subgroup.

In $G/G'$, we have $xyG' = yx[x^{-1}, y^{-1}]G' = yxG'$. \hfill $\square$

Corollary 2.9. $G/G'$ is the largest abelian quotient of $G$, i.e. if $G/N$ is abelian, then $G' \subset N$, so the map $G \to G/N$ factors through $G \to G/G'$.

Definition 2.10. For any group $G$, the abelianization $\overline{G}$ of $G$ is this quotient $G/G'$.

Note that if $G$ is finitely generated, then $\overline{G}$ is a finitely generated abelian group. Then we can apply the classification theorem and try to deduce information about $G$. For instance, it can be hard to tell whether an arbitrary group $G$ is infinite, e.g. from a presentation. It can be easier to tell if $G$ or a subgroup has an infinite abelianization.
Chapter 3

Free groups

3.1 Construction of free groups

We want to define the notion of a free group on a set.

Let $X = \{a, b, c, \ldots\}$ be a set of symbols, and let $X^{-1} = \{a^{-1}, b^{-1}, \ldots\}$ be a set disjoint from $X$, but in bijection with $X$. We write $X^\pm = X \cup X^{-1}$.

A word in $X^\pm$ is a finite sequence of elements of $X^\pm$ (letters), including $\emptyset$. Let $W$ be the set of all words, and $W_0 \subset W$ be the set of “reduced” words, i.e. words containing no subword $xx^{-1}$ or $x^{-1}x$.

Intuitively, we want to define the free group to be the set of reduced words, with multiplication being concatenation plus simplification. It turns out that this approach is kind of messy. For instance, associativity is nasty to check, because one has to keep track of all the cancellation of inverses. (Everything is obvious, but writing down rigorous arguments is awkward.) So we will instead use a trick.

The key idea is that groups should be thought of as acting on a set, so we instead think of the free group as acting on the reduced words.

Definition 3.1. The free group $F(X)$ on the set $X = \{x_i\}$ is the subgroup of $S(W_0)$ generated by the elements $\chi_i(w) = x_iw$ if $w$ doesn’t start with $x_i^{-1}$, and $w'$ if $w = x_i^{-1}w'$.

Note that this is well-defined because $\chi_i(w)$ is in $W_0$, and this permutation does have an inverse.

Proposition 3.2. The map $M : W_0 \rightarrow F(X)$, given by replacing by replacing $x_i^\pm$ by $\chi_i$ and multiplying in the group law, is an injection.

Proof. (Don’t blink for this proof.) Suppose that $M(w_1) = M(w_2)$. Then $M(w_1)(\emptyset) = w_1$, and $M(w_2)(\emptyset) = w_2$ so $w_1 = w_2$. \(\Box\)

Corollary 3.3. The map $M : W_0 \rightarrow F(X)$ is surjective. Furthermore, given a word $w \in W$, if we delete all subwords $xx^{-1}, x^{-1}x$ in any order, then we reach a unique $w_0 \in W_0$. 

11
Proof. Extend the map to $\tilde{M}: W \to F(X)$ in the obvious way. This is surjective because the generators of $F(X)$ are all hit, and $W$ is closed under concatenation (corresponding to multiplication in the group law of $F(X)$). The surjectivity of $M$ then follows from the second assertion. To see that, observe that each deletion reduces the length of $w$, and does not change $M(w)$, so we reach a word $w_0 \in W_0$ in finite time satisfying $M(w) = M(w_0)$, so $w_0$ is unique.

This establishes that the underlying set of $F(X)$ is $W_0$, as intuition says it should be.

Theorem 3.4 (Universal property of free groups). Any set map $f : X \to G$, where $G$ is any group, extends uniquely to a group homomorphism $f^* : F(X) \to G$ such that $f^*(\chi_i) = f(x_i)$.

Proof. First extend $f$ to $X^\pm$ by setting $f(x_i^{-1}) = f(x_i)^{-1} \in G$. Then let

$$f^*(t_1 \ldots t_k) = f(t_1) \ldots f(t_k)$$

where $t_j \in \{\chi_i^\pm 1\}$. Then $f^*$ is well-defined by 3.2, 3.3, and is a homomorphism by definition. This is clearly unique.

3.2 General properties of free groups

Proposition 3.5. If $F(X)$ and $F(Y)$ are free groups on sets $X$ and $Y$, then $F(X) \cong F(Y) \iff |X| = |Y|$.

Proof. $\iff$. If $f : X \to Y$ is a bijection with inverse $g : Y \to X$, then we get homomorphisms

$$f^* : F(X) \to F(Y) \quad \text{and} \quad g^* : F(Y) \to F(X)$$

such that $g^*f^*$ fixes $X$, hence is the identity homomorphism (by uniqueness), and similarly for $f^*g^*$.

$\implies$ For any group $G$, let $S_G \triangleleft G$ be the group $\langle g^2 : g \in G \rangle$. Then $G/S_G$ is abelian since every non-identity element has order 2, and is thus a vector space over $\mathbb{F}_2$. The image of $X$ in $F(X)/S_{F(X)}$ is linearly independent and spans, so $|X|$ is the dimension of this vector space.

By this proposition, we can unambiguously define the free group of rank $n$, $F_n$.

- $F_0 = I$,
- $F_1 = \mathbb{Z}$,
• If $|X| \geq 2$, then $F(X)$ is non-abelian and contains $F_2$.

**Corollary 3.6.** Every finitely generated group is a quotient of a finitely generated free group.

*Proof.* Straightforward from the universal property. Take a generating set \( \{g_i\} \) for \( G \), and take the free group on a set in bijection with \( \{g_i\} \). This set bijection induces a surjective group homomorphism. □

**Corollary 3.7.** If \( \theta : G \to F_n \) is a surjective group homomorphism, then it splits, i.e. there is a section \( F_n \to G \) and \( G = \ker \theta \rtimes F_n \).

*Proof.* If \( F_n = F(X = \{x_i\}) \), then take \( g_0, \ldots, g_n \in G \) with \( \theta(g_i) = x_i \). By the universal property, the set map \( x_i \mapsto g_i \) extends to a group homomorphism \( s : F_n \to G \) such that \( \theta \circ s \) is the identity on \( F_n \).

Now set \( K = \ker \theta \). The split short exact sequence

\[
0 \to K \to G \to F_n \to 0
\]

exhibits \( G \) as the semidirect product. □

We have defined a free group on a set. This is like defining a vector space in terms of its basis, and we would like a more invariant way of discussing it.

**Definition 3.8.** We say that the set \( S = \{s_i\} \subset F(X) \) is a *free basis* for \( F(X) \) if \( \langle S \rangle = F(X) \) (generates) and for any reduced word \( w_0 = \emptyset \) on \( S^{\pm 1} \), we have \( w_0 \neq \text{id} \) when evaluated in \( F(X) \) (freeness).

This is analogous to the notion of basis for vector spaces, but not quite as nice. However, we do have some nice properties.

**Proposition 3.9.** *Free bases for \( F(X) \) have cardinality \( X \).*

*Proof.* Given a free basis \( S \) for \( F(X) \), we will define a homomorphism

\[
f^* : F(S) \to F(X)
\]

by extending the set map \( s_i \in S \mapsto s_i \in F(X) \). This is surjective because \( S \) generates, and injective by the freeness condition. Therefore, \( f^* \) is an isomorphism, so \( |S| = |X| \) by Proposition 3.5 □

**Proposition 3.10.** *The automorphisms of \( F(X) \) are exactly (extensions of) bijective functions of free bases \( f : X \to S \).*

*Proof.* An automorphism \( \alpha \) must send a free basis bijectively to a free basis as \( \alpha(w(x))0 = w(\alpha(x)) \), and it is determined by this bijection.

Moreover, given such a bijection, we can use the universal property to extend to a group automorphism. □
We now discuss conjugacy in free groups.

**Definition 3.11.** A word \( w \) in \( X^{\pm 1} \) is **cyclically reduced** if it is reduced and the first and last letters are not inverses.

A reduced word may not be cyclically reduced, e.g. \( aba^{-1} \in F(\{a, b\}) \).

We can write any reduced word \( w_0 \) as

\[
w_0 = u|w_1|u^{-1}
\]

where \( w_1 \) is cyclically reduced, and \(|\ldots|\) means no cancellation between two words on the boundary.

**Lemma 3.12.** If \( w, w' \) are cyclically reduced words in a free group, then they are conjugate if and only if \( w' \) is a cyclic permutation of \( w \).

**Proof.** This is saying that if \( w = L_1|L_2|\ldots|L_n \) for \( L_i \in X^{\pm 1} \), then \( w' \) must be \( L_k|L_{k+1}|\ldots|L_{k-1} \) for some \( k \). That two such words are conjugate is obvious, so we prove the converse.

If \( w' = cwc^{-1} \), where \( c \) is reduced and non-empty, then as \( w' \) is cyclically reduced, there exists cancellation in \( cw \) or \( wc^{-1} \) (but not both, since \( w \) is cyclically reduced). Without loss of generality, we suppose that the cancellation is between \( c \) and \( w \). Without loss of generality, say \( c = d|L_1^{-1} \) and \( w = L_1|v \). Then \( w' = dvL_1d^{-1} \), but \( vL_1 \) is a cyclic permutation of \( w \). If \( d = \emptyset \), we are done. Otherwise, we can continue the process, decreasing the length of the conjugating element at each step until it is the empty set.

**Corollary 3.13.** A free group is torsion-free.

**Proof.** Any reduced word \( w_0 \) can be written as \( w_0 = u|w|u^{-1} \) for some cyclically reduced \( w \). But for \( n > 0 \), we get

\[
w_0^n = u|w^n|u^{-1}.
\]

There can be no cancellation since \( w \) is cyclically reduced, hence so is \( w^n \). \( \square \)

**Example 3.14.** Take \( F_2 \) to be free on \( a, b \). Let \( H_n = \langle a, bab^{-1}, \ldots, b^nab^{-n} \rangle \leq F_2 \). Then any element of \( H_n \) is of the form

\[
h = b^{i_1}a^{k_1}b^{j_1}a^{-i_1}a^{k_2}b^{j_2}a^{-i_2} \ldots b^{-i_m} \quad 0 \leq i_j \leq n.
\]

Hence \( b^{n+1}ab^{-n+1} \notin H_n \). So the ascending chain \( H_1 \leq H_2 \leq \ldots \) does not stabilize, implying that \( H = \bigcup H_n \) is not finitely generated.

This shows that \( F_2 \) does not have the max property. Since any \( F(X) \) with \( |X| > 1 \) contains \( F_2 \), we get the same for all such free groups.

**Corollary 3.15.** A finitely generated group \( G \) containing a non-abelian free group does not have max.
3.3 Free Products

Let \( \{G_\lambda : \lambda \in \Lambda\} \) be an indexed family of groups. We want to construct the “free product” of the \( G_\lambda \). What should this look like?

The usual notion of “direct product” in terms of tuples is a highly abelian notion. The embeddings of the factors in the direct product commute with each other, so it is a highly abelian construction.

A reduced sequence in \( \{G_\lambda\} \) is a finitely sequence \( g_1 \ldots g_n \) where \( g_j \in \coprod_{\lambda \in \Lambda} G_\lambda \) (the disjoint union) such that \( g_j \neq e \) and no successive \( g_j, g_{j+1} \) are from the same \( G_\lambda \). Let \( \mathcal{R} \) be the set of all reduced sequences and \( \mathcal{A} \) the set of all sequences.

**Definition 3.16.** The free product \( \ast_{\lambda \in \Lambda} G_\lambda \) is the subgroup of \( S(\mathcal{R}) \) generated by elements \( \gamma(g, \lambda) \) for \( \lambda \in \Lambda \) and \((g, \lambda) \in G_\lambda \setminus I\) acting as

\[
\gamma(g, \lambda)(g_1 \ldots g_n) = \begin{cases} 
(g, \lambda)g_1 \ldots g_n & g_1 \notin G_\lambda \\
[(g, \lambda)g_1]g_2 \ldots, g_n & g_1 \in G_\lambda
\end{cases}
\]

where \([ (g, \lambda)g_1 \] is understood to be deleted if \((g, \lambda)g_1 = e_\lambda\).

**Proposition 3.17.** The function \( f : \mathcal{A} \to \ast_{\lambda \in \Lambda} G_\lambda \) given by

\[
f(g_1 \ldots g_n) = \gamma(g_1) \cdot \ldots \cdot \gamma(g_n)
\]

restricts to a bijection on \( \mathcal{R} \).

**Proof.** The surjectivity is obvious. The injectivity follows as before by consider the action on the empty set. \( \square \)

**Theorem 3.18** (Universal property for free products). For any group \( H \) and any collection of homomorphisms \( \theta_\lambda : G_\lambda \to H \), there exists a unique \( \theta : \ast_{\lambda \in \Lambda} G_\lambda \to H \) such that

\[
G_\lambda \to \ast_{\lambda \in \Lambda} G_\lambda \xrightarrow{\theta} H = \theta_\lambda.
\]

**Proof.** Define \( \theta(g_1 \ldots g_n) = \theta_{\lambda_1}(g_1) \ldots \theta_{\lambda_n}(g_n) \). \( \square \)

**Example 3.19.** If \( F(X) \) is free on \( X = \{x_i : i \in I\} \), then it is \( \ast_{i \in I} G_i \) where \( G_i = F(\{x_i\}) \).

Note that \( G_1 \ast G_2 \) is always infinite if \( G_1 \) and \( G_2 \) are nontrivial, as \( g_1g_2 \) is torsion-free since \((g_1g_2)^n\) are distinct reduced elements. Furthermore, it is non-abelian since \( g_1g_2 \neq g_2g_1 \).

How do free groups and free products come up in practice? In fact, they come up a lot in topology. Suppose \( X \) is a topological space and \( G \) is a group of homeomorphisms \( X \to X \). We say that a subspace \( S \subset X \) is \( G \)-packing if \( g(S) \cap S = \emptyset \) whenever \( g \neq e \).
Chapter 3. Free groups 16

Theorem 3.20 (Klein). $KFG_1, G_2 \leq \text{Homeo}(X)$ and $S_1 \subset X$ is $G_1$-packing such that $S_1 \cup S_2 = X$, and $S_1 \cap S_2 \neq \emptyset$, then $G := \langle G_1, G_2 \rangle = G_1 * G_2$.

Proof. Note that for $x \in S_1$, $g(x) \notin S_1$ for any $g \in G_1 \setminus I$. Therefore, $g(x) \in S_2$. Similar observations apply when swapping 1 and 2.

Take $s \in S_1 \cap S_2$ and a reduced sequence $g_1, \ldots, g_n$ with (without loss of generality) $g_n \in G_1$. Then $g_n(s) \in S_2 \setminus S_1$, $g_n-1g_n(s) \in S_1 \setminus S_2$, etc. so $g_1 \ldots g_n(s) \neq s$. This shows that the homomorphism $\theta : G_1 * G_2 \to \langle G_1, G_2 \rangle$ is injective. It is obviously surjective, so an isomorphism.

Example 3.21. Let $X = \mathbb{R}^2$ with reflections $a$ in the $y$-axis and $b$ in the line $x = \frac{1}{2}$. A packing for $a$ is the half of the plane to the right of the $y$ axis, and a packing for $b$ is the half of the plane to the left of the line $x = \frac{1}{2}$. They satisfy the conditions of the theorem, so $\langle a, b \rangle = C_2 * C_2$.

This is the infinite dihedral group we encountered before, as $ba(x) = x + 1$ and $a(ba)a^{-1} = (ba)^{-1}$.

Möbius transformations

We now discuss an example from hyperbolic geometry. Recall the group of Möbius transformations, which is $\text{PSL}_2(\mathbb{C})$ acting on $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ as

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}.$$  

Consider the elements $f(z) = z + 2$ and $g(z) = \frac{z}{z + 1}$. Both preserve the (open) upper half-plane. Let’s think about their geometric effects. A packing for $f$ is the region $\text{Re} z \in (-1, 1)$. The automorphism $g$ sends 0 to 0 and $-1$ to 1, so it swaps the semicircles (hyperbolic geodesics) between $(-1, 0)$ and $(0, 1)$. So a packing for $g$ is the region outside these semicircles. These two packings satisfy the conditions for the theorem, so we get:

Proposition 3.22. $F = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $G = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$ freely generate $F_2$.

Proof. It is clear that the corresponding elements $f, g \in \text{PSL}_2(\mathbb{Z})$ have infinite order, hence generate $F_2$. Now send $f, g \mapsto F, G$ and extend by the universal property to some section $F_2 = \langle f, g \rangle \to \langle F, G \rangle$. 

3.4 Nielson-Schreier Theorem

For us, a graph $\Gamma$ consists of a set $V$ of vertices with discrete topology, with a set $E$ of edges $I_\alpha \cong [0, 1]$ with interior $e_\alpha$. The edge endpoints are
attached to points in $V$ by maps $f_0, f_1 : E \to V$, and $\Gamma$ is the topological space

$$\Gamma = \prod_\alpha I_{e/\alpha} \sim f_0(I_{e/\alpha})$$

This has the quotient topology, i.e. a subset $S \subset \Gamma$ is open if its intersection with each $\overline{e}_\alpha$ is open. A basic open neighborhood of some vertex $v \in \Gamma$ consists of an open about each edge glued to $v$. So $\Gamma$ is definitely locally path-connected and locally contractible. In particular, $\Gamma$ is connected if and only if it is path-connected.

If you’ve done some algebraic topology, you will recognize that this is just the notion of a 1-dimensional CW complex.

A subgroup of $\Gamma$ is a union $\Delta$ of edges and vertices such that $e_\alpha \in \Delta \implies \overline{e}_\alpha \in \Delta$.

**Lemma 3.23.** If $\tilde{X}$ covers the (connected) graph $\Gamma$ then $\tilde{X}$ is a graph, with vertices and edges the being the lifts (inverse images) of those in $\Gamma$.

**Proof.** If $p : \tilde{X} \to X$ denotes the covering map, define $V(\tilde{X}) = p^{-1}(V)$ and for the each edge $I_{e/\alpha} \in \Gamma$ and $\tilde{v}$ above $f_0(I_{e/\alpha})$, you get a unique lift of $\tilde{I}_{e/\alpha}$ with $f_0(I_{e/\alpha}) = \tilde{v}$. Thus $\tilde{X}$ is a graph and the graph topology agrees with its given one. 

If $S$ is a simple graph, an edge path in $S$ is a finite sequence of vertices such that any two consecutive vertices span an edge in $S$. A cycle $v_0 \ldots v_n$ is a closed edge path, i.e. $v_0 = v_n$. A tree is a connected simple graph with no reduced cycles.

**Proposition 3.24.** Given any connected graph $\Gamma$ and any vertex $v_0 \in \Gamma$, there exists a subgraph $\Delta \simeq \{v_0\}$ such that $\Delta$ contains all of $V(I')$.

This is basically obvious: imagine “breadth-first search” starting at $v_0$. However, we do have to write down a formal proof.

**Proof.** Let $\Gamma_0 = \{v_0\} \subset \Gamma_1 \subset \Gamma_2 \cup \ldots$ be a sequence of subgraphs where $\Gamma_{i+1}$ is $\Gamma_i$ unioned with all edges $\overline{e}_\alpha$ for all $e_\alpha \in \Gamma \setminus \Gamma_i$ with an endpoint in $\Gamma_i$.

Then $\bigcup \Gamma_i$ is open (since a neighborhood of a point in $\Gamma_i$ is in $\Gamma_{i+1}$) and closed in $\Gamma$ (since it is a union of closed edges), hence is all of $\Gamma$.

Now we begin to define $\Delta$. Let $\Delta_0 = \Gamma_0 \subset \Delta_1 \subset \Delta_2 \ldots$ where $\Delta_i$ has the same vertices as $\Gamma_i$, but for each $v \in \Gamma_{i+1} \setminus \Gamma_i$ we choose only one edge connecting $v$ to $\Delta_i$ to form $\Delta_{i+1}$. (We are using the axiom of choice here.) Then $\Delta_{i+1}$ deformation retracts onto $\Delta_i$. Setting $\Delta = \bigcup \Delta_i$, we can deformation retract to $\Delta_0$ by performing the $i$th deformation retract in the interval $[\frac{1}{2^{i+1}}, \frac{1}{2^i}]$.

**Corollary 3.25.** A tree $T$ is contractible.
Proof. Let $\Delta \subset T$ be a subgraph guaranteed by Proposition 3.24. We claim that $\Delta = T$. If an edge $e$ spanning $v, w$ is not in $\Delta$, then take the reduced edge paths $v_0, \ldots, u, v$ and $w, x, \ldots, v_0$ in $\Delta$. By assumption, $u \neq w, v \neq x$ since the edge between $v, w$ is not in $\Delta$. But then concatenating this gives a reduced cycle in $T$, and no such thing exists.

**Theorem 3.26.** The fundamental group of a connected graph $\Gamma$ is a free group.

Proof. Let $T \subset \Gamma$ (without loss of generality, assume it to be simple) be $\Delta$ as guaranteed by Proposition 3.24. This is a tree, and let $\{e_\alpha : \alpha \in A\}$ be the set of edges in $\Gamma \setminus T$. It will turn out that the fundamental group is the free group on this set.

First assume that there is only one $\alpha$. Let $w_\alpha, w'_\alpha$ be the endpoints of $e_\alpha$. For each $\alpha$, take a reduced cycle $c_\alpha$ from $v_0$ including $w_\alpha w'_\alpha$, so

$$c_\alpha = v_0 u_1 u_2 \ldots w_\alpha w'_\alpha \ldots v_1 v_0.$$ Deleting repeats in $c_\alpha$, we may assume that it is a reduced cycle with every edge except $e_\alpha \in T$. The corresponding subgraph is topologically $L \simeq S^1$.

We claim that $\Gamma \simeq L$. Indeed, the components of $\Gamma \setminus L$ are in $T$, and so are trees. If such a component $K$ intersected $L$ in two distinct points $a, b$ then you could form a path from $a$ to $b$ in $K$, and then back via $L$ not including the forbidden edge, which would give a reduced cycle in $T$. So we can deformation retract each component to its unique intersection point $L$.

Now, for general $\Gamma$ let $m_\alpha$ be the “midpoint” of $e_\alpha$. Let

$$A_\alpha = \left( \Gamma \setminus \bigcup_{\beta \in A} m_\beta \right) \cup m_\alpha$$

be the subspace obtained by cutting each edge in $A$ except $\alpha$. These are open in $\Gamma$, path-connected, and

$$A_\alpha \cap A_\beta = \Gamma \setminus \bigcup_\alpha m_\alpha.$$ This deformation retracts onto $T$, hence is simply-connected. Also, $A_\alpha$ retracts onto $T \cup e_\alpha$, which has fundamental group $\mathbb{Z}$ by the first case. So, by the Seifert-Van Kampen theorem, $\pi_1(\Gamma, v) \simeq *_{\alpha \in A} \mathbb{Z}$.

**Corollary 3.27.** For every free group $F(X)$, there exists a simple connected graph which has $\pi_1(\Gamma) \simeq F(X)$.

Proof. Basically just use Theorem 3.26. Take a loop $i_\alpha$ for each $x_\alpha \in X$, all joined at $v$. To make the graph simple, turn each self loop into a triangle by adding a vertex.
Theorem 3.28 (Nielsen-Schreier). A subgroup of a free group is free.

Proof. For \( H \leq F(X) \), Corollary 3.29 guarantees the existence of a connected graph \( \Gamma \) with \( \pi_1(\Gamma) = F(X) \). By the theory of covering spaces, there is a covering space \( \tilde{X} \to X \) with fundamental group \( H \). By Lemma 3.23, \( \tilde{X} \) can be given a graph structure, and then Theorem 3.26 implies that \( H \) is free.

\[ \square \]

Theorem 3.29 (Nielsen-Schreier Index Formula). If \( H \) has index \( i \) in the free group \( F_n \), then \( H \) is free of rank \( i(n - 1) + 1 \).

Proof. One just has to keep track of the number of vertices and edges for a covering space. For a finite graph \( \Gamma \), define the Euler characteristic \( \chi(\Gamma) = |V| - |E| - 1 \). Note that \( \pi_1(\Gamma) \) is free on \( -\chi(\Gamma) \) generators. If \( p: \tilde{\Gamma} \to \Gamma \) has degree \( i \), then \( \chi(\tilde{\Gamma}) = i\chi(\Gamma) \).

If \( \Gamma \) has \( \pi_1(\Gamma) = F_n \), then \( \chi(\Gamma) = 1 - n \). So \( \chi(\tilde{\Gamma}) = i(1 - n) \), hence \( \tilde{\Gamma} \) is free on \( i(n - 1) + 1 \) generators.

\[ \square \]

This shows that one can have \( F_2 \leq F_3 \) and \( F_3 \leq F_2 \).
Chapter 4

Presentations of Groups

4.1 Construction and basic properties

By Theorem 3.4, any group $G$ can be written as $F(X)/N$ where the image of $X$ is a generating set for $G$.

**Definition 4.1.** A presentation $\langle X \mid R \rangle$ for a group $G$ is a set $X$ and a subset $R$ of $F(X)$ such that

$$G \cong F(X)/\langle\langle R \rangle\rangle.$$ 

We say that elements of $X$ are *generators* and elements of $R$ are *relators*.

The following theorem says that quotients are can be viewed as just adding more relators, which is kind of obvious.

**Theorem 4.2** (von Dyck). If $G = \langle X \mid R \rangle$ then $Q$ is a quotient of $G$ if and only if $Q \cong \langle X \mid R \cup S \rangle$.

**Proof.** We certainly do have a map $F(X) \rightarrow \langle x \mid R \cup S \rangle$ which factors through $G$ as $\langle\langle R \rangle\rangle \leq \langle\langle R \cup S \rangle\rangle$.

Conversely, if $Q$ is a quotient of $G$, we can write $Q = G/L$ for $G = F(X)/M$, where $M = \langle\langle R \rangle\rangle$. Then there exists a normal subgroup $N \triangleleft F(X)$ such that $L = NM/M$ (we can take $N$ to be the pre-image of $L$ in $G$), so $Q \cong F(X)/NM$. Take any $S$ whose normal closure is $N$. Then $\langle\langle R \cup S \rangle\rangle = \langle\langle N \cup M \rangle\rangle = MN$. Therefore, $Q \cong \langle X \mid R \cup S \rangle$.

**Definition 4.3.** A group $G$ is *finitely presented* if $G$ has a presentation with finitely many generators and relators:

$$G = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_m \rangle$$

**Remark 4.4.** If $N \neq I$ is a normal subgroup with infinite index in $F(X)$, then it is not finitely generated (although it may very well be finitely normally generated, e.g. $F_1 \subset F_2$).
Now, you may wonder: if a group is finitely presented, is it possible to give it a nontrivial infinite presentation? (We won’t allow something stupid like giving infinite copies of the same relator/generator.) Fortunately, the answer turns out to be pretty nice.

**Proposition 4.5.** If \( G = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_m \rangle = \langle y_1, \ldots, y_k \mid s_1, \ldots \rangle \), then there exists some finite \( L \) such that \( G = \langle y_1, \ldots, y_k \mid s_1, \ldots, s_L \rangle \).

**Proof.** The idea for this and many subsequent proofs is to write down a finite presentation for a group that we think is isomorphic to \( G \). We can verify this by checking that the relevant maps are well-defined.

Write the \( x \)'s and \( y \)'s as words in terms of each other: \( y_i = v_i(x_1, \ldots, x_n) \) and \( x_i = w_i(y_1, \ldots, y_k) \). This means that we can write \( y_i = v_i(w_1(y), \ldots, w_n(y)) \) and \( r_i(w_1(y), \ldots, w_n(y)) = e \). Let \( N = \langle \langle s_1, s_2, \ldots \rangle \rangle \triangleleft F(\{y_i : i\}) \) and form

\[
\overline{G} = \langle \overline{y}_1, \ldots, \overline{y}_k \mid \overline{y}_i = v_i(\overline{w}_1(\overline{y}), \ldots, \overline{w}_n(\overline{y})), r_i(\overline{w}_1(\overline{y}), \ldots, \overline{w}_n(\overline{y})) \rangle.
\]

This is an abstract presentation for a group that we think is \( G \). So there is certainly a surjective homomorphism \( \theta : \overline{G} \to G \) sending \( \overline{y}_i \mapsto y_i \), since the relations above are satisfied in \( G \). On the other hand, there is a map \( \varphi : G \to \overline{G} \) sending \( x_i \mapsto \overline{x}_i \), since the relations for \( G \) are satisfied in \( \overline{G} \). The composition \( \theta \varphi(x_i) = x_i \) and \( \varphi \) is onto, so \( \theta \) is an isomorphism. This shows that \( N \) is generated by a finite set.

The usual argument then shows that any set of generators of \( N \) must terminate. \( \blacksquare \)

**Proposition 4.6.** The group \( G = \langle a, b \mid [a^{2n+1}ba^{-2n+1}, b] : n \in \mathbb{N} \rangle \) is not finitely presented.

**Proof.** In \( A_j \) (the alternating group), let \( \alpha = (1 2 \ldots j) \) and \( \beta = (1 2 3) \) then \( \alpha^k \beta \alpha^{-k} \) commutes with \( \beta \) if \( 3 \leq k \leq j-3 \) but not if \( k = j-2 \).

There is a homomorphism \( F_2 \to A_{2l+3} \) sending \( a \mapsto \alpha \) and \( b \mapsto \beta \). Letting \( c_\ell = [a^{2n+1}ba^{-2n+1}, b] \), we see that in \( A_{2l+3} \) the images of \( c_1, \ldots, c_{\ell-1} \) are trivial but the image of \( c_\ell \) is not. That means that \( c_\ell \notin \langle \langle c_1, \ldots, c_{\ell-1} \rangle \rangle \leq F_2 \).

Then by Proposition 4.5, \( G \) is not finitely presented. \( \blacksquare \)

**Example 4.7.** Finite groups are finitely presented. Finitely generated free groups are finitely presented.

**Proposition 4.8.** If \( N = \langle n_i \mid r_j \rangle \) and \( H = \langle h_k \mid s_l \rangle \) then \( G = N \rtimes_\varphi H \) has presentation

\[
P = \langle n_i, h_k \mid r_j, s_l, h_k n_i h_k^{-1} = \varphi(h_k) n_i \rangle.
\]

**Proof.** By the relations, any element \( p \in P \) can be written as a product of words in \( n_i \) and \( h_k \), \( p = v(n)w(h) \). Define \( \theta : P \to G \) as the quotient of
the obvious map $F(\{n_i\} \cup \{h_k\}) \to G$. If $\theta(p) = e$, then $\theta(v(n))\theta(w(h)) = e$.

But since $N \cap H = I$, we must have $\theta(v(n)) = e$ and $\theta(w(h)) = e$, and a little thought shows that this forces $v(n) = w(h) = e$ (if it’s $e$ in $N$, it’s in the normal closure of the relators, which are also in $P$).

The problem with presentations is that many different presentations can be given for the same group. It can even be difficult to tell from a presentation of $G$ that it is not trivial. One trick is to map to finite fields.

If $G = \langle x_1, \ldots, x_n \mid r_1, r_2, \ldots \rangle$ and $p$ is prime, let $\overline{r_j} \in (\mathbb{F}_p)^n$ be the exponential sum vector modulo $p$ of $r_j$, i.e. taking $r_i$ to the sum of the number of $x_1$’s, $x_2$’s, etc modulo $p$. This gives a homomorphism $F_n \to (\mathbb{C}_p)^n$.

**Proposition 4.9.** If $S = \text{span}\{\overline{r_j}\} \neq \mathbb{F}_p^n$, then $G \neq I$.

**Proof.** There exists a linear function $f : \mathbb{F}_p^n \to \mathbb{F}_p$ with $S \leq \ker f$. Then composing, we get $F_n \to \mathbb{F}_p^n \to \mathbb{F}_p$ a surjective homomorphism factoring through $G$.

**Example 4.10.** It’s difficult to tell if a presentation yields an infinite group. But the above argument shows that if the number of relations is less than the number of generators in a finite presentation, then the group surjects onto $\mathbb{F}_p$ for all $p$, hence is infinite.

**Proposition 4.11.** If $G = \langle x_i \mid r_k \rangle$ and $H = \langle y_j \mid s_l \rangle$ then

$$G * H = \langle x_i, y_j \mid r_k, s_l \rangle.$$

In fact, this can be useful in identifying a presentation as a free product of simpler groups.

**Proof.** Let $G \mid H = \langle x_i, y_j \mid r_k, s_l \rangle$. There is a map to $G * H$ sending $x_i \mapsto x_i, h_j \mapsto h_j$. There is also a homomorphism $G * H \to G \mid H$ fixing each $x_i$. We can extend this to $\varphi : G * H \to G \mid H$. By uniqueness, $\theta \varphi$ is the identity, and $\varphi$ is surjective so $\theta$ is an isomorphism.

Recall that whenever we encounter a group property, we want to know if it’s preserved under the basic operations of subgroups, quotients, or extensions.

<table>
<thead>
<tr>
<th></th>
<th>Free</th>
<th>Finitely Presented</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subgroups</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Quotients</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Extensions</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Theorem 4.12 (Hall).** If $N \triangleleft G$ and $N, G/N$ are finitely presented, then $G$ is.
Proof. Let $N = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_m \rangle$ and $G/N = \langle y_1, \ldots, y_l \mid s_1, \ldots, s_k \rangle$. Take $g_1, \ldots, g_l$ such that the image of $g_i$ in $G/N$ is $y_i$. It is easy to see that $G$ is generated by $x_1, \ldots, x_n, g_1, \ldots, g_l$ (this is how you prove that an extension of finitely generated groups is finitely generated). For relations, we take $r_i = e,$ $s_i(g_1, \ldots, g_l) = t_i(x_1, \ldots, x_n),$ $g_jx_ig_j^{-1} = u_{ij}(x_1, \ldots, x_n),$ $g_j^{-1}x_i = v_{ij}(x_1, \ldots, x_n).$

Take $G = \langle x_1, \ldots, x_n, g_1, \ldots, g_l \mid \text{above relations in the new variables} \rangle$. Then there is an obvious (surjective) homomorphism $\theta: G \to G$. Let $K = \ker \theta$ and $\overline{N} = \langle \overline{x_1}, \ldots, \overline{x_n}, \overline{g_1}, \ldots, \overline{g_l} \mid \text{above relations in the new variables} \rangle$. Then there is an obvious (surjective) homomorphism $\theta: \overline{G} \to G$.

Let $K = \ker \theta$ and $\overline{N} = \langle \overline{x_1}, \ldots, \overline{x_n} \rangle \triangleleft \overline{G}$. This is normal by the relations. Restricting $\theta$ to $\overline{N}$, it’s an isomorphism, so $K \cap \overline{N} = I$. Since $\theta(\overline{N}) = N$, we get a map $\theta_0: \overline{G}/\overline{N} \to G/N$ which is also injective. Its kernel (which is $I$) is $K\overline{N}/\overline{N} \cong K/K \cap \overline{N}$ so $K = I$. 

### 4.2 Free products with amalgamation and HNN extensions

In this section we define a few constructions on groups.

First, the free product with amalgamation is a group formed by “gluing” two isomorphic subgroups of a given one.

**Definition 4.13.** Let $G, H$ be groups with $A \leq G, B \leq H$ and an isomorphism $\varphi: A \to B$. The **free product with amalgamation** is the group $G \ast H/(\langle a = \varphi(a) \; \forall a \in A \rangle)$.

**Remark 4.14.** In fact, it suffices to add in these relations for a generating set for $a$. In particular, if $G, H$ are finitely presented and $A$ is finitely generated, then the free product with amalgamation is also finitely presented.

The **HNN extension** is the universal group in which two isomorphic subgroups of $G$ are conjugate.

**Definition 4.15.** If $G$ is a group with $A, B \leq G$ and $\varphi: A \to B$ an isomorphism, then the **HNN extension** $G*_{\varphi} \langle t \rangle/\langle tat^{-1} = \varphi(a) \rangle$.

It’s not obvious that $G$ embeds into the HNN extension or the free product with amalgamation, but we wish to show this.

In the HNN extension, note that $ta = \varphi(a)t, t^{-1}b = \varphi^{-1}(b)t^{-1}$, so we can move $a$’s and $b$’s to the left of the $t$’s.

Choose right transversals $T_A$ and $T_B$ for $A, B \in G$, both including $e$. A **normal form** is a sequence $g_0t^{\epsilon_1}g_1t^{\epsilon_2} \ldots t^{\epsilon_n}g_n$ for $n \geq 0, \epsilon = \pm 1$ such
that \( \epsilon_i = +1 \implies g_i \in T_A \) and \( \epsilon_i = -1 \implies g_i \in T_B \) and there is no subsequence \( t^\epsilon t^{-\epsilon} \).

Every element of \( G* \varphi \) can be put into normal form. Indeed, imagine a generic element of \( G * \langle t \rangle \). If it ends with \( g \in G \), we can write \( g = ar \) for some \( r \in T_A \) and \( a \in A \), and then we can move the \( a \) past \( t \) at the cost of twisting by \( \varphi \), or past \( t^{-1} \) at the cost of twisting by \( \varphi^{-1} \). From this algorithmic description, you might intuit the following.

**Theorem 4.16.** Every element in \( G* \varphi \) has a unique normal form.

**Proof.** The idea is, again, to find an action for \( G* \varphi \) in which any two reduced elements visibly act differently.

Let \( N \) be the set of normal forms. Define \( \rho : G* \varphi \to S(N) \) by the obvious action, namely

1. \( \rho(g)(g_0t^{\epsilon_1} \ldots g_n) = gg_0t^{\epsilon_1} \ldots g_n \) for \( g \in G \). Note that for \( g, h \in G \),
   \[ \rho(g)\rho(h) = \rho(gh). \]

2. If \( \epsilon_1 = -1 \) and \( g_0 \in A \) then
   \[ \rho(t)(g_0t^{-1} \ldots g_n) = (\varphi(g_0)g_1)t^{\epsilon_2} \ldots g_n. \]
   Otherwise,
   \[ \rho(t)(g_0t^{-1} \ldots g_n) = \varphi(a)t\gamma_0t^{\epsilon_1} \ldots g_n \]
   where \( g_0 = a\gamma_0 \) for \( \gamma_0 \in T_A \).

This has inverse
   \[ \rho(t^{-1})(g_0tg_1 \ldots g_n) = (\varphi^{-1}(g_0)g_1)t^{\epsilon_2} \ldots g_n \text{ if } g_0 \in B \]
   or
   \[ g_0t^{\epsilon_1} \ldots g_n \mapsto \varphi^{-1}(b)t^{-1}g_0t^{\epsilon_1} \ldots g_n. \]

Also, \( \rho(a) = \rho(t^{-1})\rho(\varphi(a))\rho(t) \) so \( \rho \) is a well-defined homomorphism on \( G* \varphi \) and
   \[ \rho(g_0t^{\epsilon_1} \ldots g_n)(e) = g_0t^{\epsilon_1} \ldots g_n \text{ for normal forms}. \]

We say that a sequence \( g_0t^{\epsilon_1} \ldots g_n \) is reduced if there is no subsequence \( tg_it^{-i} \) for \( g_i \in A \) for \( g_i \in A \) or \( t^{-1}g_it \) for \( g_i \in B \) (called a "pinch"). Note that such a sequence could be simplified by replacing with \( \varphi^{\pm 1}(g_i) \), by construction.

**Corollary 4.17** (Britton’s Lemma). The group \( G \) embeds in \( G* \varphi \) by the map \( g \mapsto g \). If for \( n \geq 1 \), the element \( g_0t^{\epsilon_1} \ldots g_n \) is reduced then it is not the identity in \( G* \varphi \).
**Proof.** If \( g \neq e \in G \), then \( g, e \in G_\phi \) are both normal forms so \( g \neq e \) by Theorem 4.16. On changing a reduced sequence into a normal one, no \( t^{\pm 1} \) cancel so the normal looks like \( g_0 t^{\epsilon_1} \ldots g_n' \), and in particular is not the identity. \( \square \)

**Corollary 4.18** (Torsion in HNNs). If \( \gamma \in G_\phi \) has finite order, then \( \gamma \) is conjugate to some \( g \in G \) (which has the same order).

**Proof.** If \( \gamma = g_0 t^{\epsilon_1} \ldots g_n \) (we assume \( n \geq 1 \) since the case \( n = 0 \) is trivial) is a reduced sequence, and \( t^{\epsilon_1} g_0 t^{\epsilon_1} \) is not a pinch (this is the middle part of \( \gamma \gamma \)), then \( \gamma^k \) is reduced for all \( k > 0 \), hence not \( e \). If it is a pinch, then we “smooth the pinch” by replacing \( \gamma \) with

\[
t^{\epsilon_n} g_n \gamma (t^{\epsilon_n} g_n)^{-1} = gt^{\epsilon_2} \ldots t^{\epsilon_{n-1}} g_{n-1}
\]

which is either reduced or in \( G \). Now repeat. \( \square \)

**Example 4.19.** There exists an infinite group where every non-identity element is conjugate. This is hard to achieve for finite groups by orbit-stabilizer! Take \( G \) to be countably infinite and torsion-free (e.g. \( \mathbb{Z} \)). Let \( \{g_0, g_1, \ldots\} \) be an enumeration of the non-identity elements. Form the following HNN extensions:

1. \( G_1 = \langle G, t_1 \mid t_1 g_0 t_1^{-1} = g_1 \rangle \). (Since \( G \) is torsion-free, \( \langle g_0 \rangle \cong \langle g_1 \rangle \cong \mathbb{Z} \).)
2. Let \( G_2 = \langle G_1, t_2 \mid t_2 g_0 t_2^{-1} = g_2 \rangle \) (\( G \) embeds in \( G_1 \), so \( \langle g_0 \rangle \cong \langle g_2 \rangle \cong \mathbb{Z} \).)
3. \( G_3 = \langle G_3, t_3 \mid t_3 g_0 t_3^{-1} = g_3 \rangle \), \( \Gamma(G) = \bigcup_n G_n \), etc. This is an ascending union of countably infinite groups and torsion-free. Note that any two elements of \( G \) are conjugate in \( \Gamma(G) = \Gamma_1 \).
4. Let \( \Gamma_2 = \Gamma(\Gamma_1) \), etc. \( \Gamma_n = \Gamma(\Gamma_{n-1}) \).
5. Now if \( \gamma, \delta \in \Delta \) then they are both in some \( \Gamma_M \), so they are conjugate in \( \Gamma_{M+1} \) and hence in \( \Delta \).

This is a crazy construction, and you might think that it is impossible to produce a finitely generated example. Actually, D. Osin showed in 2010 that there are finitely generated examples. On the other hand, the analogous question for finitely presented groups is completely open.

For free product with amalgamation \( G \ast_\phi H \), we say that an element \( c_1 \ldots c_n \in G \ast H \) is “a-reduced” if for \( i > 1 \), no \( c_i \) is in \( A \) or \( B \). We can turn any \( c \in G \ast H \) into an a-reduced one by “absorbing” extra elements of \( A \) or \( B \), since these are identified with a subgroup of \( G \) and \( H \).

**Corollary 4.20.** If \( c_1, \ldots, c_n \) is a-reduced, then it’s not the identity element.
Proof. The idea is to map $G *_{\varphi} H$ into an HNN extension, where the image of an a-reduced word will be reduced.

Let $F = (G * H) * t/\langle\langle tat^{-1} = \varphi(a)\rangle\rangle$ be the HNN extension. Let $\psi : G *_{\varphi} H \to F$ be defined by $\psi(g) = tgt^{-1}$ and $\psi(h) = h$ (this is well-defined by the universal property). For an a-reduced element, if $c_1 \in A$ then $c_1 \mapsto \varphi(a) \neq 1$ (this is a normal form). Otherwise, it goes to a reduced sequence in $F$ so we are done by 4.17. □
Chapter 5

Soluble, Polycyclic, and Nilpotent Groups

5.1 Soluble groups

Definition 5.1. The group $G$ is soluble (solvable) if there exists a sequence

$$ I = G_n \triangleleft G_{n-1} \triangleleft \ldots \triangleleft G_0 = G $$

such that each $G_i/G_{i+1}$ is abelian.

Theorem 5.2. Solubility is preserved by subgroups, quotients, and extensions.

Proof. 1. (Subgroups) If $H \leq G$, and $G$ is solvable then let

$$ I = G_n \triangleleft G_{n-1} \triangleleft \ldots \triangleleft G_0 = G $$

be as above. Then

$$ I = G_n \cap H \triangleleft G_{n-1} \cap H \triangleleft \ldots \triangleleft G_0 \cap H = H. $$

Furthermore, $G_i \cap H/G_{i+1} \cap H \hookrightarrow G_i/G_{i+1}$.

2. (Quotients) If $N \triangleleft G$ then $G_{i+1}N/N \triangleleft G_iN/N$ with quotient

$$ G_iN/G_{i+1}N \cong G_i/G_i \cap G_{i+1}N $$

is a quotient of $G_i/G_{i+1}$.

3. (Extensions) If $G/N$ is solvable, then we get a sequence in $G/N$ which lifts (by the correspondence theorem) to

$$ N \leq G_{i+1} \leq G_i \ldots \leq G $$
such that $G_{i+1}/N \leq G_i/N$ and $G_i/N \cong G_{i+1}/N$ abelian. This implies that $G_{i+1} \triangleleft G_i$ and $G_i/G_{i+1}$ is abelian. String this together with a sequence

$$I = N_j \leq N_{j-1} \ldots \leq N_0 = N$$

with $N_i/N_{i+1}$ abelian.

**Definition 5.3.** For any group $G$, the derived series is the sequence of subgroups $G = G^{(0)} \geq G^{(1)} \geq \ldots$ where $G^{(i+1)} = (G^{(i)})'$ (the commutator subgroup).

Note that $G'$ is characteristic in $G$. Since the characteristic property is transitive with respect to inclusions, this is a sequence of characteristic subgroups of $G$.

**Proposition 5.4.** $G$ is soluble if and only if the derived series terminates at $I$. If so, then it has the smallest length among any derived series.

**Proof.** By definition of the commutator subgroup, $G^{(i)}/G^{(i+1)}$ is abelian. Therefore, if the derived series terminates then $G$ is soluble by definition.

Conversely, if $I = G_n \triangleleft G_{n-1} \triangleleft \ldots \triangleleft G_0 = G$ is any finite series with $G_i/G_{i+1}$ abelian, then we claim that $G^{(i)} \geq G_i$, essentially because $G_i$ is the minimal subgroup we can quotient by to get an abelian group. Inductively, since $H \leq G \implies H' \leq G'$, we get $G^{(i+1)} \geq (G^{(i)})' \geq G_{i+1}$.

**Definition 5.5.** We say a group $G$ is perfect if $G = G'$.

A nontrivial perfect group is not soluble (in fact, it’s “as far as possible” from being soluble). Alternatively, a perfect group cannot surject onto a nontrivial abelian group, so this property is preserved by quotients.

**Example 5.6.** If $G$ is simple and non-abelian then it is perfect, since it cannot have any nontrivial abelian quotients.

**Corollary 5.7.** If $G$ contains a non-abelian free subgroup then $G$ is not soluble.

**Proof.** Since solubility is preserved by subgroups and quotients, if $G$ is soluble and $F_2 \leq G$, then $F_2/N \cong A_5$ would be soluble.

### 5.2 Polycyclic Groups

**Definition 5.8.** The group $G$ is polycyclic if there exists a sequence

$$I = G_n \triangleleft G_{n-1} \triangleleft \ldots \triangleleft G_0 = G$$

such that each $G_i/G_{i+1}$ is cyclic.
Chapter 5. Soluble, Polycyclic, and Nilpotent Groups.

Obviously, polycyclic implies solvable. The converse is not true in general. However, it is true for finite groups, so these definitions are sometimes used interchangeably in the finite group world.

**Theorem 5.9.** Polycyclicity is preserved by subgroups, quotients, and extensions.

*Proof.* Everything is completely analogous to the proof of Theorem 5.2. □

In fact, this proof shows that if property P is closed under subgroups and quotients but not extensions, then we can define a property “poly-P” (there exists a filtration with successive quotients having P) which is closed under all three.

**Corollary 5.10.** Finitely generated abelian groups A are polycyclic.

*Proof.* This follows from Theorem 2.1. □

**Theorem 5.11.** A group G is polycyclic if and only if G is soluble and has max.

*Proof.* Suppose G is polycyclic. A cyclic group is soluble and has max, and both of these properties are preserved by extensions, so G does as well.

If G is soluble and has max, then in a derived series each $G_i/G_{i+1}$ is polycyclic by Corollary 5.10, so G is a (repeated) extension of polycyclic groups. □

**Corollary 5.12.** If G polycyclic and $H \leq G$, then H is finitely presented.

*Proof.* A cyclic group is finitely presented and this is preserved by extensions (Theorem 4.12), so polycyclic groups are finitely presented. If $H \leq G$ and G is polycyclic, then H is polycyclic as well (Theorem 5.9). □

**Example 5.13.** We will produce a finitely generated soluble group that is not polycyclic. Let $D \leq \mathbb{Q}$ be the dyadics, i.e. $\mathbb{Z}_2 = \{ \frac{n}{2^i} : a \in \mathbb{Z} \}$. This is an infinitely generated abelian group (in particular, it does not have max and is therefore not polycyclic). Let $B = D \rtimes \mathbb{Z}(t)$ where $\varphi(t)$ is the automorphism $d \mapsto 2d$. Then B is “metabelian” (abelian by abelian). Moreover, we claim that it is finitely generated, by $1 \in D$ and $t \in \mathbb{Z}$. To see this, note that all elements of B are of the form $(\frac{n}{2^i}, t)$. Note that

$$(0, t^k) \left( \frac{n}{2^i}, 1 \right) (0, t^{-k}) = \left( \frac{2^k n}{2^i}, 1 \right).$$

What about a finitely presented example? Consider

$G = \langle a, b \mid bab^{-1} = a^2 \rangle.$
Then sending $a \mapsto (1, 1)$ and $b \mapsto (0, t)$ gives a surjective homomorphism $G \to B$. Now $G/G' \simeq \mathbb{Z}$ generated by $bG'$. As the relation implies $ba = a^2b$ (since $ba^{-1} = a^{-2}b$) and $ab^{-1} = b^{-1}a^2$ (since $a^{-1}b^{-1} = b^{-1}a^{-2}$), given any word in $a, b$ we can move the positive powers of $b$ past $a^\pm 1$ to the right and the positive powers of $b^{-1}$ to the left. Thus any $g \in G$ may be written $b^{-m}a^l b^n$. Now $\theta(g) = (1/2^m, t^{n-m})$ so we see that $\theta$ is injective. So $G$ is actually isomorphic to our dyadic example, showing that it is finitely presented.

We can also do this with matrices $a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $b = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. Then $b^{-m}a^l b^n = \begin{pmatrix} 2^{n-m} & l2^{-m} \\ 0 & 1 \end{pmatrix}$

**Proposition 5.14.** There exists an infinite, finitely presented group $G$ with a finitely presented subgroup $H \leq G$ and $g \in G \setminus H$ such that $gHg^{-1} < H$ (proper inclusion).

**Proof.** Take $G$ as in Example 5.13 with $H = \langle a \rangle$. Then $bHb^{-1} = \langle a^2 \rangle$. □

In this example, we can keep conjugating by $b$ to get an infinite ascending/descending sequence of subgroups. This shows that such a group cannot have the max or min property.

### 5.3 Nilpotent groups

**Definition 5.15.** The group $G$ is nilpotent if there exists

$I = G_n \leq G_{n-1} \leq \ldots \leq G_0 = G$

such that $G_i \triangleleft G$ and $G_i / G_{i+1}$ is in the center of $G/G_{i+1}$ (a central series).

Note that abelian implies nilpotent implies solvable, but if $G$ is nilpotent and nontrivial, then $Z(G) \neq I$. So $S_3$ is polycyclic and not nilpotent. Since any cyclic group is nilpotent, this shows that nilpotence is not preserved by extensions.

**Theorem 5.16.** Nilpotence is preserved by subgroups, quotients, and direct products.

**Proof.** 1. As in 5.2, we want $(H \cap G_i) / (H \cap G_{i+1})$ to be in the center of $H/H \cap G_{i+1} \simeq HG_{i+1}/G_{i+1}$. Note that $(H \cap G_i) / (H \cap G_{i+1}) = (H \cap G_i)G_{i+1}/G_{i+1} \subset G_i / G_{i+1}$, so it commutes with all of $G/G_{i+1}$.

2. We need $G_iN / G_{i+1}N$ to be in the center of $G/G_{i+1}N$. To see this, let $xn \in G_{i+1}N$. Then $xgx^{-1}g^{-1} \in G_{i+1}$ for any $g \in G$, so

$xng(xn)^{-1}g^{-1} \in G_{i+1}N$
showing that \(x^n\) commutes with \(g\) modulo \(G_{i+1}N\).

3. Given \(G_i \leq G\) and \(H_j \leq H\), as in 5.15, we have \(G_i \times H \triangleleft G \times H\) with \(G_i \times H/G_{i+1} \times H\) in the center of \(G \times H/G_{i+1} \times H\), so we use

\[I \times I = I \times H_m \leq I \times H_{m-1} \cdots I \times H = G_n \times H \leq G_{n-1} \times H \leq \ldots \leq G \times H.\]

For \(H_1, \ldots, H_n \leq G\) define

\([H_1, H_2] = \langle [h_1, h_2] \rangle.\)

So this generalizes the usual definition of commutator. Similarly, set

\([H_1, \ldots, H_n] = [[H_1, \ldots, H_{n-1}], H_n].\)

**Definition 5.17.** For any group \(G\), the *lower central series* of \(G\) is the sequence

\[G \gneq [G, G] \gneq [[G, G], G] \gneq \ldots \gneq \gamma_{i+1}G \gneq \ldots.\]

**Lemma 5.18.** \(G\) is nilpotent if and only if the lower central series terminates at \(I\).

**Proof.** Note that each \(\gamma_iG\) is normal (in fact, characteristic) in \(G\), since commutators are characteristic. Note also that \(\gamma_iG/\gamma_{i+1}G\) is in \(Z(G/\gamma_{i+1}G)\) by definition. That shows that a terminating central series implies nilpotence.

Now suppose that \(G\) is nilpotent, and so there is some central series for \(G\):

\[I = G_n \leq G_{n-1} \leq \ldots \leq G_0 = G.\]

We claim that \(\gamma_iG \leq G_{i-1}\). By induction, we may assume this and establish the \(i+1\) case. Indeed, \(\gamma_{i+1}G = [\gamma_iG, G] \leq [G_{i-1}, G]\), which is in \(G_i\) since the quotient \(G_i/G_{i-1}\) is central in \(G/G_{i-1}\).

**Theorem 5.19 (Baer).** If \(G\) is nilpotent and finitely generated then \(\gamma_iG\) is finitely generated.

**Proof.** We proceed by induction. To see the base case and a toy computation, suppose \(\gamma_3G = I\) and \(G = \langle x_1, \ldots, x_n \rangle\) is a symmetric presentation of \(G\) (the generating set is closed under inverses). We have the following identities (true in any group):

\([y, x]x[y, z]^{-1} = [y, xz]\)
and
\[ z[x, y]z^{-1}[z, y] = [zx, y]. \]
So in \( G \), since commutators are central, we get
\[ [g, xh] = [g, x][g, h] \]
and
\[ [gx, y] = [x, y][g, y]. \]
So any element of \([G, G]\) is a finite product of things of the form \([x_j, x_k]\).
This also shows that for any \( G \), we get \( \gamma_2G/\gamma_3G \) is generated by \( \{[x_j, x_k]\} \).
Now assume that anything in \( \gamma_iG/\gamma_{i+1}G \) is a product of elements from the (finite) set
\[ \{ [[[x_{j_1}, x_{j_2}], x_{j_3}], \ldots, x_{j_r}] =: [x_{j_1}, \ldots, x_{j_r}] = i \} \]
(i.e. generated by \( i \)-depth commutators). We want to show the same for \( i + 1 \). In \( \gamma_{i+1}G/\gamma_{i+2}G \) suppose that \( g = x_ih \) Then
\[ [i[g], g] = [i[g], x_ih] = [i[g], x_i][i[g], h]x_i^{-1}. \]
By centrality, any \( i + 1 \)-commutator is in the center so in fact we have
\[ [i[g], g] = [i[g], x_i][i[g], h]. \]
This shows that anything in \( \gamma_{i+1}G/\gamma_{i+2}G \) can be written as a finite product of \([i[g], x_k]\). That says that in \( G \), any element of \( \gamma_{i+1}G \) can be written as
\[ x = \left( \prod_r [[g_1, \ldots, g_i], x_r] \right) \gamma \]
for some \( \gamma \in \gamma_{i+2}G \). Since \([g_1, \ldots, g_i] \in \gamma_iG \), the induction hypothesis implies that it can be written as a product of commutators of the \( x_i \) modulo \( \gamma_{i+1}G \), so
\[ x = \left( \prod_r \left( \prod_s [i[x_s][\beta, x_r] \right) \right)^{\gamma'} \]
where \( \beta \in \gamma_{i+1}G \) and \( \gamma' \in \gamma_{i+2}G \). This, in turn, can be rewritten as \([i[x_s][\beta, x_r] = [\beta, x_r][i[x_s], x_r] \delta \) for some \( \delta \in \gamma_{i+2}G \), hence this is equal to
\[ \prod_r \prod_s [i[x_s], x_r] \mod \gamma_{i+2}G. \]
So if \( \gamma_{n+1}G = I \) then \( \gamma_nG \) is finitely generated, and \( \gamma_{n-1}G/\gamma_nG \) is finitely generated, and finite generation is preserved by extensions so everything is. \( \Box \)
Corollary 5.20. Finitely generated nilpotent groups are polycyclic (and in particular, have max).

Proof. The terms $\gamma_i G/\gamma_{i+1} G$ are finitely generated and abelian, hence are polycyclic (and have max), and both of these properties are preserved by extensions.
Chapter 6

Finite index subgroups and virtual properties

6.1 Finite index subgroups

If a subgroup $H$ has finite index in $G$, we write $H \leq_f G$ and denote the index by $[G : H]$.

**Lemma 6.1.**

1. If $H \leq_f G$ and $H \leq J \leq G$ then $H \leq_f J \leq_f G$.

2. If $J \leq_f H \leq_f G$ then $J \leq_f G$ with $[G : J] = [G : H][H : J]$.

3. If $H \leq_f G$ and $S \leq G$, then $H \cap S \leq_f S$ with index $\leq [G : H]$, with equality holding if and only if $SH = G$. Furthermore, if $SH \leq G$ is a subgroup then $[S : S \cap H] \mid [G : H]$.

4. If $H \leq_f G$ and $J \leq_f G$ then $H \cap J \leq_f G$ with $[G : H \cap J] \leq [G : H][G : J]$.

**Proof.**

1. Obvious.

2. Suppose that $g_1, \ldots, g_k$ is a (left) transversal for $H$ in $G$ and $h_1, \ldots, h_l$ is a transversal for $J$ in $H$. Then

$$G = \bigcup_{i,j} g_i h_j J.$$ 

Check that if $g_i h_j J = g_{i'} h_{j'} J$ then then $g_i = g_{i'}, h_j = h_{j'}$, so $\{g_i h_{j'}\}$ is a transversal for $J$ in $G$.

3. Write a coset decomposition $G = g_1 H \cup \ldots \cup g_n H$ and throw away any $g_i H$ with $(g_i H) \cap S = \emptyset$ (this happens if and only if $SH \neq G$). Then $g_1(H \cap S), \ldots, g_k(H \cap S)$ are disjoint and we claim that their union is $S$. Indeed, we can replace each $g_i$ with some $s_i \in S$ in the same $H$. 

34
coset (by the nonemptiness assumption). Then if \( s \in g_i H \), we have
\( s \in g_i H \cap S = s_i H \cap S = s_i(H \cap S) \). This shows that \([S : H] \leq n\).

Now, if \( SH \) is a subgroup then \( H \leq SH \leq G \). Using (i), (ii) we find
\[ [SH : H] = [S : H \cap S]. \]

4. \([G : H \cap J] = [G : J][J : H \cap J] \leq [G : J][G : H]. \]

This immediately implies:

**Theorem 6.2** (Poincaré). A finite intersection of finite index subgroups has finite index.

**Lemma 6.3.** If \([G : H] = k\) then for any \( g \in G\), there exists \( i \) with \( 1 \leq i \leq k \) such that \( g_i \in H \). If \( H \triangleleft G \) then one can take \( i \mid k \).

**Proof.** The list \( H, gH, \ldots, g^k H \) cannot all be distinct cosets, so we have \( g^j H = g^i H \) for some \( 0 \leq i < j \leq k \). Then \( g^{j-i} \in H \), where \( 1 \leq j - i \leq k \).

If \( H \triangleleft G \), then \( gH \) has order dividing \( k = |G/H| \) by Lagrange’s theorem.

**Proposition 6.4.** Let \( \theta : G \to H \) be an onto homomorphism.

1. If \( B \leq_f H \) then \( \theta^{-1}(B) \leq_f G \), an in fact \([G : \theta^{-1}(B)] = [H : B]\).

2. If \( A \leq_f G \) then \( \theta(A) \leq_f H \) with \([H : \theta(A)][G : A]\).

**Proof.** 1. Write \( H = h_1 B \cup \ldots \cup h_k B \) a disjoint union of cosets, and take \( g_i \) with \( \theta(g_i) = h_i \) to get a transversal for the \( \theta^{-1}(B) \) in \( G \). To see this, suppose \( g \in G \) and let \( \theta(g) = h_i b \), so \( g_i^{-1} g \in \theta^{-1}(B) \), so \( g \in g_i \theta^{-1}(B) \).

No two of these cosets can intersect, since their images under \( \theta \) don’t intersect.

2. \( \theta^{-1}(\theta(A)) = KA \leq_f G \) for \( K = \ker \theta \), with \([G : KA] = [H : \theta(A)] \) by (1).

Also, \([G : KA] \mid [G : A]\).

**6.2 The Regular Representation**

Any group \( G \) acts on itself by (left) multiplication. Now let \( H \) be any subgroup and \( \mathcal{L} \) the set of left cosets of \( H \) in \( G \). The (left) regular representation \( \rho \) of \( G \) on \( \mathcal{L} \) is the action of \( G \) given by
\[ \rho(g)(xH) = gxH. \]
Note that this is transitive. The stabilizer of $H \in \mathcal{L}$ is $H \leq G$.

**Lemma 6.5.** With the notation above,
\[ \ker \rho = \bigcap_{x \in G} xHx^{-1}. \]

*Proof.* The stabilizer of $xH$ is $xHx^{-1}$ (because $xH = gxH \implies H = x^{-1}gxH$). \hfill \Box

**Definition 6.6.** Given $H \leq G$, the *core* of $H$ in $G$ is $\ker \rho$.

**Proposition 6.7.** The core of $H$ in $G$ is the largest normal subgroup of $G$ that is contained in $H$.

*Proof.* If $N \triangleleft G$, and $N \leq H$, then $N \subset xHx^{-1}$ for all $x$. \hfill \Box

The following theorem is so useful that we will refer to it as the Useful Theorem.

**Theorem 6.8.** If $H \leq_f G$ with $[G : H] = n$, then there exists $N \triangleleft_f G$ with $N \leq H$ and $[G : N] \mid n!$.

*Proof.* The left regular representation of $G$ on $G/H$ maps $G$ into $S(G/H) \cong S_n$. Therefore, $|G/\ker \rho| = |\im \rho| \mid n!$. \hfill \Box

**Theorem 6.9.** If $G$ is a finitely generated group, then for any $n \in \mathbb{N}$ there exist only finitely many finite index subgroups of index $n$ in $G$.

*Proof.* If $H$ is an index $n$ subgroup of $G$, then we get a regular representation of $G$ on $G/H$, which corresponds to a homomorphism $G \to S_n$, that is transitive. Moreover, $H$ can be recovered as the stabilizer of the coset $H$.

Since $G$ is finitely generated, there are only finitely many transitive homomorphisms to any $S_n$, hence only finitely many possibilities for $H$. \hfill \Box

**Corollary 6.10.** If $G$ is finitely generated and $H \leq_f G$, then there exists a characteristic subgroup $C \leq_f H \leq_f G$.

*Proof.* Let $C = \bigcap_{\alpha \in \Aut(G)} \alpha(H)$. Note that $H$ and $\alpha(H)$ have the same index in $G$ by 6.4, so this intersection involves only finitely many distinct subgroups by Theorem 6.9, hence has finite index by 6.2.

If $\beta \in \Aut(G)$, then $\beta(C) = \bigcap_{\alpha \in \Aut(G)} \beta\alpha(H) = C$. \hfill \Box

**Theorem 6.11.** If $H \leq_f G$, then $G$ is finitely generated (presented) $\iff$ $H$ is finitely generated (presented)
Proof. $G$ is finitely generated means that there exists a surjective homomorphism $\theta : F_k \to G$. The pre-image of $H$ is a finite index subgroup of $F_k$, hence is free by the index theorem, so $H$ is finitely generated.

Now suppose that $G = F_k/N$ is finitely presented, so $N = \langle (r_1, \ldots, r_m) \rangle$. Then $H = F_i/N$ for $N \leq F_i \leq F_k$. We know that $N$ is the normal closure of finitely elements of $G$, but we need to show that it is the normal closure of finitely many elements of $H$, since we are taking the normal closure with respect to a smaller group! Now take a right transversal $t_1, \ldots, t_n$ for $F_i$ in $F_k$ where $n = [G : H]$. Then $N$ consists of things of the form $(g_1 r_{i_1}^{a_1} g_1^{-1}) \ldots (g_j r_{i_j}^{a_j} g_j^{-1})$ for $g_1, \ldots, g_j \in F_k$.

But as any $g \in F_k$ is $ht_j$, so $N$ is the normal closure of $\{t_j r_i^j : 1 \leq i \leq m, 1 \leq j \leq n\}$.

This shows the implication $\implies$. For the converse, observe that finitely generated (finitely presented) groups are preserved by extensions (4.12), so if $H$ is finitely generated (finitely presented) then we can take $N \triangleleft G$ with $N \leq_f G$ and $N \leq_f H$ by Theorem 6.8. Now $N$ is finitely generated (presented) and $G/N$ is finite, so $G$ is finitely generated (presented).

We are really seeing the usefulness of the useful theorem. It’s much nicer to work with normal subgroups, since the quotient is finite, so we like a bridge to go from finite index subgroups to finite index normal subgroups.

Note that if $G$ has a presentation by $k$ generators and $m$ relations, then say the “deficiency” $\def$ of the presentation to be $k - m$. We proved earlier that if $\def > 0$, then the group is infinite. Theorem 6.11 has shown that if there exists a subgroup $H \leq G$ with $[G : H] = n$, then $H$ has a presentation with deficiency $n(k - 1) + 1 - nm = n(k - m) + 1 - n$. Stated more suggestively, we have produced a presentation for which

$$\def(H) - 1 = [G : H] \def(G) - 1$$

although we stress that this implicitly depends on the presentation.

A natural question to ask is if a group has proper finite index subgroups at all. Sometimes there are no such things.

Example 6.12. Let $H \leq_f \mathbb{Q}$ with index $n$. Then for all $q \in \mathbb{Q}$, we get that $n \left(\frac{2}{q}\right) \in H$ by 6.3. So $\mathbb{Q}$ has no finite index subgroups.

$\mathbb{Q}$ is infinitely generated - is there a finitely generated example?

Theorem 6.13 (Higman). The group

$$G = \langle a_1, a_2, a_3, a_4 \mid a_1 a_2 a_1^{-1} = a_2^2, a_2 a_3 a_2^{-1} = a_3^2, a_3 a_4 a_3^{-1} = a_4^2, a_4 a_1 a_4^{-1} = a_1^2 \rangle$$

has no proper finite index subgroups.
Proof. This is a clever application of Fermat’s little theorem. Note that for \( n > 1 \) and a prime \( p \mid 2^n - 1 \), the least prime factor of \( n \) is less than \( p \). Indeed, take \( r \) to be the order of 2 mod \( p \). Then \( r \mid n \), and \( r \mid p - 1 \), so \( r \mid (n, p - 1) \). Since \( r \) is obviously not 1, there is a smaller prime dividing \( n \).

Suppose that we did have a finite index subgroup \( H \leq f G \). Then 6.8 gives a nontrivial finite quotient \( G/N \). Let \( n_i \) is the order of \( a_i \) in \( G/N \). Then \( n_i \mid p_i - 1 \), so \( n_i \mid (n, p_i - 1) \). Since \( n_i \) is obviously not 1, there is a smaller prime dividing \( n_i \).

To make this example interesting, we should actually show that \( G \) is not trivial (this is why we chose four generators: with two or three, we do get the trivial group.

Proposition 6.14. The group \( G \) in 6.13 is infinite.

Proof. If \( H = \langle x, h \mid yxy^{-1} = x^2 \rangle \) then \( x, y \) have infinite order (many ways to see this, e.g. 5.13). So \( H \ast_\varphi H' \) where \( \varphi(x) = y' \) is an infinite group (this is an amalgamated product of infinite groups; in particular the factors embed) and is

\[
\langle x(\sim y'), y, z(\sim x') \mid yxy^{-1} = x^2, xzx^{-1} = z^2 \rangle.
\]

Note that \( y, z \) freely generate a copy of \( \mathbb{F}_2 \) by 4.20 because any word \( w(y, x') \) is \( a \)-reduced. Now we take another amalgamated product to get \( G \). Take 4 copies of \( H \),

\[ H_i = \langle a_i, b_i \mid b_i a_i b_i^{-1} = a_i^2 \rangle \]

Form \( H_1 \ast_\varphi H_2 \) as above, making \( a_1 = b_2 \), so a presentation is

\[ \langle a_1, b_1, a_2 \mid b_1 a_1 b_1^{-1} = a_1^2, a_1 a_2 a_1^{-1} = a_2^2 \rangle. \]

Call this \( K \), and note that \( \langle b_1, a_2 \rangle \cong F_2 \). Let \( L = H_3 \ast_\varphi H_4 \) similarly. Finally, make \( K \ast L \) for \( \theta(b_1) = a_4, \theta(a_2) = b_3 \) and you get \( G \).

[Question: for what \( n \) does this work?]

6.3 Virtual properties

Here we regard groups \( H \) and \( G \) as “basically the same” if \( H \leq f G \). We say that a group property \( P \) is “OK” if when \( G \) has \( P \) and \( H \leq f G \), then \( H \) has \( P \); in other words \( P \) passes from a group to its finite index subgroups. Note that we are not requiring that if \( H \) has \( P \) and \( H \leq f G \), then \( G \) has \( P \).

Definition 6.15. If a property \( P \) is “OK”, we say that a group \( G \) is virtually \( P \) if there exists \( H \leq f G \) where \( H \) has \( P \).
By Theorem 6.8, this is the same as saying that $G$ is $P$-by-finite, since a finite index subgroup with $P$ has a finite index subgroup which is normal in $G$.

**Example 6.16.**

1. $\mathbb{Z} \times \mathbb{Z}$ is abelian, but not cyclic.

2. (Nilpotent group, not virtually abelian) Let

$$G = \langle a, b, t | ab = ba, tat^{-1} = ab, tbt^{-1} = b \rangle$$

be the semidirect product $\mathbb{Z}^2 \langle a, b \rangle \rtimes \mathbb{Z} \langle t \rangle$. Then $b \in Z(G)$ and $G/\langle b \rangle = \mathbb{Z}^2$. Now if we had an abelian $A \leq_f G$, then we would have $t^i, a^j \in A$ for some $i, j > 0$ by 6.3 (since $A$ has finite index). But $t^i a^j t^{-i} = (ab)^j = a^j b^j \neq a^j \in G$.

3. (Polycyclic group, not virtually nilpotent) Consider the group

$$G = \langle a, b, t | ab = ba, tat^{-1} = a^2 b, tbt^{-1} = ab \rangle.$$ 

This is clearly polycyclic (quotient by $\langle a, b \rangle$ to get $\mathbb{Z}$). If $H \leq_f G$ with $H$ nilpotent, then take $h \neq e$ in $Z(H)$ and set $h = a^k b^l t^m$. Now there exists $i > 0$ with $t^i \in H$ (by finite index), thus $t^i h t^{-i} = h$. But this changes $l$ if $k \neq 0$, and it adds $a$’s if $l \neq 0$, so $h$ would be a power of $t$, which is a contradiction.

We now give a cultural aside on word growth. If $G = \langle X \rangle$ for a finite set $X$, then we define a growth function $\gamma_X(n) : \mathbb{N} \rightarrow \mathbb{N}$ by

$$\gamma_X(n) = \# \{ g \in G : g = w(X^{\pm 1,0}) \text{ for word length } \leq n \}.$$ 

If $S = X \cup X^{-1} \cup \{ e \}$, then $\gamma_X(n) = |S^n|$. We say that a finitely generated group $G$ has polynomial word growth if there exists a finite generating set $X$ (or equivalently, for all finite generating sets $X$) there exist $c, d > 0$ such that for all $n$, $\gamma_X(n) \leq cn^d$. One can then ask: what groups have polynomial word growth?

**Theorem 6.17** (Gromov, 1981). $G$ has polynomial word growth if and only if $G$ is virtually nilpotent.

### 6.4 Virtually polycyclic Groups

**Theorem 6.18.** The following are equivalent: $G$ is

1. virtually polycyclic,
2. polycyclic by finite,
3. poly($\mathbb{Z}$) by finite,
4. poly(\mathbb{Z} or finite)

Proof. 1 = 2 is immediate from the definition and 3 \implies 2 \implies 4 is obvious.

For 4 \implies 3, we have to show that we can slide finites past poly(\mathbb{Z})'s. Suppose N is poly(\mathbb{Z})s by a finite F, and G/N = \mathbb{Z}. Since \mathbb{Z} is free, this means that G = N \oplus \mathbb{Z}. If N is H-by-F where F is finite and H is polycyclic, then G is H \oplus \mathbb{Z}-by-F, and we are happy. Having established this, we can push all finite factors “to the right.”

♠♠♠ Tony: [This isn’t the argument we gave in class - what we did in class was much more complicated. Is there a mistake?]

Corollary 6.19. Virtually polycyclic groups are preserved by subgroups, quotients, extensions, and is the smallest such class containing all finite groups and \mathbb{Z}. In particular, virtually polycyclic groups all have max and finite presentation.

Proof. The property \( P = (\mathbb{Z} or finite) \) is preserved by subgroups and quotients, so poly-\( P \) is preserved by extensions as well by 5.9 and 6.18. Any poly-\( P \) group is contained in any class with the above properties. Since max and finite presentability are preserved by extensions and are possessed by finite groups and \mathbb{Z}, they are possessed by virtually polycyclic groups.

6.5 Virtually soluble groups

Corollary 6.20. The virtually soluble groups with max are exactly the virtually polycyclic groups.

Proof. Use 5.11 and 6.19.

Theorem 6.21. Virtually soluble groups are preserved by subgroups, quotients, and extensions.

Proof. If \( H \leq_f G \) with H soluble and \( S \leq G \) then \( H \cap S \leq_f S \) by 6.1, but \( H \cap S \leq H \) is soluble (subgroup of soluble group is soluble, by 5.2).

For quotients, use 6.4 and 5.2.

The third part is a little trickier. Let \( G/N \) and \( N \) be virtually soluble. Take \( M \trianglelefteq_f N \) a minimal index soluble normal subgroup (this exists because we get a finite index normal from any finite index subgroup from the Useful Theorem). If \( S \trianglelefteq_f N \), then \( S \leq M \) since \( SM \) is soluble, normal, finite index, and contains \( M \). This implies that \( M \) is characteristic in \( N \), hence normal in \( G \).

We claim that \( G/M \) is virtually soluble. If we can prove this, then the lift of a finite-index soluble group of \( M \) will be a finite-index soluble subgroup of \( G \), and we win. By assumption, \( G/N \) is soluble and \( G/N \cong \)
\( \frac{G/M}{N/M} = \frac{Q}{R} \), where \( R \) is finite with no nontrivial soluble normal subgroups (by maximality of \( M \)). Now, consider the centralizer

\[
C = C_Q(R) = \bigcap_{r \in R} C_Q(r) \trianglelefteq Q
\]

by 6.2 and orbit-stabilizer \( (R \) is finite normal, so the orbits under conjugation are finite, so the stabilizers have finite index\). Therefore, \( C \cap R \trianglelefteq R \) must be trivial, since this is an abelian normal subgroup. Thus in \( Q \), the second isomorphism theorem implies \( C \cong CR/R \leq Q/R \) is virtually soluble (subgroup of virtually soluble group), so \( G/M = Q \) is too (since it contains \( C \) with finite index).

\[\square\]

**Corollary 6.22.** Virtually soluble groups are the smallest class preserved by subgroups, quotients, extensions, that contains all abelian and finite groups.

**Proof.** Like 6.19 but using 6.21. \[\square\]

As in 5.7, if \( F_2 \leq G \) then \( G \) is not virtually soluble. Indeed, if it were virtually soluble, then we would have a finite index soluble subgroup \( S \leq_f G \) and \( F_2 \leq G \) so \( F_2 \cap S \leq_f F_2 \) with \( F_2 \cap S \leq S \) non-abelian and free, which leads to a contradiction by 5.7.

On the second example sheet, you are asked to show that not containing \( F_2 \) is a property preserved by subgroups, quotients, and extensions. There is a finitely generated example which is not virtually soluble, but has no \( F_2 \) subgroup. What about a finitely presented group which has no \( F_2 \) subgroup and is not virtually soluble? There are examples, but few: we know of only 5 constructions.

Aside: the Tits Alternative (1972) says that a finitely generated linear group (a group of \( n \times n \) matrices over some field \( F \)), or even any linear group in characteristic 0, is either virtually soluble or contains \( F_2 \).
Chapter 7

Maximal (normal) subgroups

7.1 Maximal subgroups

Definition 7.1. A proper subgroup $H < G$ is maximal if $H \leq J \leq G \implies J = H$ or $J = G$.

Don’t be deceived by the analogy to maximal ideals in commutative algebra: not every subgroup is contained in a maximal one.

Example 7.2. $\mathbb{Q}$ has no maximal subgroups: if $M < \mathbb{Q}$ for $M$ maximal, then $M \triangleleft \mathbb{Q}$ with $\mathbb{Q}/M$ has no proper nontrivial subgroups, so it’s $\mathbb{C}_p$. But $\mathbb{Q}$ has no proper finite index subgroups.

In this section we’ll be using Zorn’s lemma. We introduce some notation:

- A poset $X$ is a set with a partial relation that is reflexive, transitive, and antisymmetric.

- A subset $S$ is totally ordered if any two of its elements are related: for any $x, y \in S$, we have $x \prec y$ or $y \prec x$.

- A subset $S \subset X$ is a chain if it is totally ordered.

Theorem 7.3 (Zorn’s Lemma). If every chain $S$ of $X$ has an upper bound in $X$, then $X$ has a maximal element $m$.

One uses this to prove that every ideal is contained in a maximal ideal. What makes that work is that the element 1 in a ring is a witness to an ideal being proper, but there is no such thing for groups. (Note that Zorn’s lemma depends on the axiom of choice, and is in fact equivalent to it.)

A general principle is that if you want something to be true but it fails for $\mathbb{Q}$, you probably need to impose finite generation.
Proposition 7.4 (Neuman 1937). If $H \leq G$ and $g \in G \setminus H$ then there exists a maximal subgroup $M$ containing $H$ relative to $g$, i.e. $H \leq M < G$ and $g \notin M$ such that $M \leq L$ and $g \notin L$, then $L = M$.

Proof. Let $X$ be the poset of $J < G$ such that $H \leq J$ and $g \notin J$, ordered by $\leq$. Then for a chain $S = \{J_i\}$ we get $\bigcup_{i} J_i$ a subgroup which still doesn’t contain $g$, hence is in $X$. Now we can apply Zorn’s Lemma.

Corollary 7.5. If $G$ is finitely generated and $H \leq f G$, then $H$ is contained in a maximal subgroup of $G$.

Proof. We can write $G = \langle g_1, \ldots, g_k, h_1, \ldots, h_l \rangle$ for $g_i \notin H, h_j \in H$. Take $M_1$ to be maximal with $H \leq M_1$ relative to $g_1$. If $M_1$ is not maximal, then $M_1 < L < G$ with $g_1 \in L$ but (without loss of generality) $g_2 \notin L$. Rinse and repeat.

7.2 Infinite Simple Groups

Recall that $G$ is simple if $N \triangleleft G$ implies $N = I$ or $G$. If $G$ is simple and soluble, then $G = C_p$. Other simple groups include $A_n$ ($n \geq 5$) and $\text{PSL}(n, F)$ for $n > 2$ or $|F| > 3$ (the proof does work for infinite fields).

Example 7.6. Let $H = \bigcup_{n \geq 5} A_n \leq S(\mathbb{N})$. We claim that $H$ is simple. If $N \triangleleft H$, then $N \cap A_n \triangleleft A_n$. If $N \neq I$, then take $k$ with $N \cap A_k \neq I$. This is a normal subgroup of $A_k$, hence all of $A_k$, so $A_n \leq N$ for all $n \geq k$. That shows that $N = H$.

This $H$ is not finitely generated. If $G$ is an infinite simple group, then $G$ is not virtually soluble since $S \leq f G$ implies (by the useful theorem) that there is a finite index normal subgroup $N \leq f G$. So $G$ has no finite index subgroups (sound familiar?).

Theorem 7.7 (Higman 1951). Infinite, finitely generated simple groups exist.

Proof. Take Higman’s $G$ (the four generator, four relation example from 6.13) and suppose it has a normal subgroup $N \triangleleft G$. Since $G$ is finitely generated, we may assume that $N$ is a maximal normal subgroup. But $N \neq G$, and $N$ does not have finite index in $G$, so $G/N$ is infinite and finitely generated.

What about infinite, finitely-presented simple groups? R. Thompson came up with three examples which are (now) called “Thompson groups” $F,T,V$. These are all finitely presented. $F$ is not simple, but it is not virtually soluble and has no $F_2$-subgroups. $T,V$ are simple.
Chapter 8

Residually Finite, Hopfian, Baumslag-Solitar

8.1 Residually Finite Groups

Definition 8.1. A group is residually finite if

$$\bigcap_{H \leq G} H = I.$$ 

Proposition 8.2. The following are equivalent:

1. $G$ is residually finite.
2. $\bigcap_{N \trianglelefteq G} N = I$.
3. For all $g \in G \setminus \{e\}$, there exists a homomorphism $\theta$ onto a finite group $F$ such that $\theta(g) \neq e$.
4. For any finitely many $g_1, \ldots, g_n \neq e$, there exists a homomorphism $\theta$ onto a finite group $F$ such that $\theta(g_i) \neq e$ for all $i$.

Proof. (2) $\implies$ (1) is clear. By the Useful Theorem 6.8, $\bigcap_{N \trianglelefteq G} N \leq \bigcap_{H \leq G} H = I$, establishing (1) $\implies$ (2).

(2) $\iff$ (3): If $g \neq e$, then there exists a homomorphism $\theta: G \to F$ with $g \notin \ker \theta$ if and only if there is some normal subgroup not containing $g$.

(2) $\implies$ (4): take $\theta \to F_1 \times F_2 \times \ldots F_n$ where $\theta_i : G \to F_i$ doesn’t kill $g_i$. □

Example 8.3. Finite groups are obviously residually finite. The proposition implies $\mathbb{Z}$ is residually finite. But $\mathbb{Q}$ and the Higman group are not residually finite, since they have no proper finite index subgroups.
An infinite residually finite group has infinitely many finite-index subgroups (Poincaré implies a finite intersection of finite index subgroups has finite index), of arbitrarily high index by 6.1: if \( H = \bigcap_{i=1}^{n} H_i \), and \( J \leq_f G \), then \( H \cap J \) has even larger index than \( H \) if \( J \) doesn’t contain \( H \).

Aside on topological groups: if \( G \) is a finite topological group, we typically expect it to have the discrete topology. In general, we can do something more interesting: define the basic open sets to be the cosets of \( N \leq_f G \). This is called the “profinite topology.” This topology is Hausdorff if and only if \( G \) is residually finite. It is discrete if and only if \( G \) is finite, and indiscrete if and only if there are no proper finite index subgroups.

**Lemma 8.4.** Let \( R_G = \bigcap_{N \leq_f G} N \). Then \( G/R_G \) is residually finite.

**Proof.** Normal finite index subgroups of \( G/R_G \) are in bijection with normal finite index subgroups of \( G \) containing \( R_G \). \( G/R_G \) is universal in the following sense: if \( G \twoheadrightarrow Q \) is a surjection onto a residually finite group, then \( \theta \) factors through \( G/R_G \). Indeed, one easily checks that \( \theta(R_G) \subset R_Q \), so in particular if \( R_Q = I \), then \( R_G \) is in the kernel.

**Proposition 8.5.** 1. If \( G \) is residually finite and \( H \leq G \), then \( H \) is residually finite.

2. If \( H \) is residually finite and \( H \leq_f G \), then \( G \) is residually finite.

3. If \( G,H \) are residually finite, then \( G \times H \) is residually finite. If furthermore \( G \) is finitely generated, then \( G \times H \) is residually finitely.

**Proof.** 1. \( R_G = \bigcap_{L \leq_f G} L \) from 8.2. Then

\[
I = R_g \cap H = \bigcap_{L \leq_f G} L \cap H \geq R_H.
\]

2. For \( H \leq_f G \), we have \( L \leq_F H \) then \( L \leq_f G \) so \( R_G \leq R_L \) but \( R_G \geq R_G \) by (i).

3. For \( (g,h) \neq \text{id} \), then take \( \theta_1 : G \rightarrow F_1 \) and \( \theta_2 : H \rightarrow F_2 \) with \( \theta_1(g), \theta_2(h) \) are not both \( e \), and construct the product homomorphism \( \theta_1 \times \theta_2 : G \times H \rightarrow F_1 \times F_2 \). For \( G \times H \), we have the projection \( \theta : G \times H \rightarrow H \), so there exists \( \varphi \) with \( \varphi \theta(gh) = \varphi(h) \neq e \). Therefore, the only things we have to worry about are the \( (g,e) \). Take \( L \leq_f G \) with \( g \notin L \). If \( G \) is finitely generated, 6.10 lets us pass to a finitely index characteristic subgroup \( C \leq_f L \). Then \( C \triangleleft G \times H \), so \( CH \leq G \times H \), and \( G \cap H = I \) so \( g \notin CH \leq_f G \times H \). \( \square \)
Corollary 8.6. If $G$ is virtually polycyclic, then $G$ is residually finite.

Proof. We have $H \trianglelefteq_f G$ with $H$ being poly-$\mathbb{Z}$ by 6.18. Now if $M/N \cong \mathbb{Z}$, with $N$ being finitely generated and residually finite, then $M \cong N \times \mathbb{Z}$ by 3.5 so $M$ is residually finite by 8.5(3). Thus $H$ is residually finite, hence $G$ too by 8.4(2).

Theorem 8.7. Free groups are residually finite.

Intuitively, this should be plausible. Residually finite is equivalent to having “enough” homomorphism to finite groups to distinguish any two elements, and free groups are the “easiest” to define homomorphisms out of.

First proof. Recall 3.22, that $F = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $G = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ freely generate $F_2$. Given a reduced word for $F_2$, we have $w(F,G) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If we take a large prime $p$ such that this doesn’t reduce to the identity mod $p$, then we have produced a homomorphism to a finite group $\text{SL}_2(\mathbb{F}_p)$ in which $w$ does not map to the trivial element. So this establishes that $F_2$ is residually finite.

For any finite $n$ we have $F_n \leq_f F_2$, so these are residually finite. Now consider a general free group $F(X)$ where $X = \{x_i : i \in I\}$. Take a word $w \neq \emptyset \in F(X)$ involving only $x_i_1,\ldots,x_i_k$ (possible because any word is finite). So we have $\theta : F(X) \to F_k(\{x_i_1,\ldots,x_i_k\})$ induced by sending the rest of $X$ to $e$. Now $\theta(w) \neq e$, so we can compose with some $\varphi : F_k \to \text{finite group}$ (since $F_k$ is residually finite) to obtain a homomorphism such that $\varphi \circ \theta(w) \neq e$.

Second proof. As before, it suffices to show the result for $F_2$. Let $F_2$ be free on $a,b$, and create a reduced word $w$ using $a,b,A,B$ where $A$ and $B$ are the formal inverses of $a,b$ respectively, e.g.

$$w = AbABBBAAabA.$$ If $w$ has length $n$, we define a map $F_2 \to S(n+1)$ such that the image of $w$ is nontrivial. To do this, we just have to choose the images of $a,b$ such that the image of $w$ is nontrivial. We can guarantee this by choosing things in a way so as to make $w$ take 1 to $n+1$.

Let’s work out how to do this in the example; then the general case will be clear. Numbering

$$
\begin{array}{cccccccccccc}
11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\
\end{array}
$$

We interpret the this as describing the action of $a,b$ on the objects 1,\ldots,11. For instance, the first $A$ between 11 and 10 tells us that $A$ should take 10
to 11, i.e. \(a\) should take 11 to 10 (reading right to left, as we should for group actions). The end result will definitely take 1 to 11, as we want. This only partially defines the permutations:

\[
\begin{array}{cccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\
\end{array}
\]

We can then fill out the rest of the table in an arbitrary valid way (so that we actually end up with a permutation). The only potential problem would be if we had to put the same number twice in the same row, but this won’t happen because the word is reduced (both \(a\) and \(b\) only take \(n \mapsto n \pm 1\), so it could only happen if \(a\) was next to \(A\) or similarly for \(b\) and \(B\), which is ruled out by the reducedness of the word).

**Theorem 8.8.** If \(G_1, G_2\) are residually finite, then \(G_1 \ast G_2\) is residually finite.

**Proof.** First we do the case where \(G_1, G_2\) are actually finite. Given a reduced sequence \(g_1 \ldots g_n \in G_1 \ast G_2\) of length \(n \geq 1\), let \(X_n = \{g \in G_1 \ast G_2, 0 \leq \text{length}(g) \leq n\}\). This is manifestly a finite set. Define a permutation action on \(X_n\) in the natural way: if \(\gamma \in G_2\), then

\[
\gamma(g_1 \ldots g_k) = \begin{cases} 
g_1 \ldots g_k & \text{length(RHS)} \leq n \\
g_1, \ldots, g_k & k = n, g_1 \in G_1
\end{cases}
\]

The inverse of this permutation is \(\gamma^{-1} \in G_2\), so there is a homomorphism \(G_2 \to S(X_n)\), and similarly \(G_1 \to S(X_n)\).

Using the universal property, we can extend this to a homomorphism \(G_1 \ast G_2 \to S(X_n)\), the latter being a finite group. Now \(g_1 \ldots g_n(\emptyset) = g_1 \ldots g_n\), showing that this homomorphism is injective (into a finite group). This completes the case where \(G_1\) and \(G_2\) are finite.

Now we tackle the general case. For any finitely many elements \(g_1 \ldots g_n \in G_1 \ast G_2\), choose \(N_1 <_f G_1\) and \(N_2 <_f G_2\) such that \(g_1, \ldots, g_n \notin N_1 \cup N_2\) (this is possible by 8.2(4)). Then \(G_1 \to G_i/N_i \hookrightarrow (G_1/N_1) \ast (G_2/N_2)\), which extends (by the universal property) to

\[
G_1 \ast G_2 \to (G_1/N_1) \ast (G_2/N_2)
\]

and the image of \(g_1 \ldots g_n\) is reduced so has length \(n\). We have shown that for any \(g_1 \ldots g_n \in G_1 \ast G_2\), there is a homomorphism to a residually finite group not killing \(g_1 \ldots g_n\), so \(G_1 \ast G_2\) is residually finite. \(\square\)
8.2 Hopfian Groups

Definition 8.9. A group $G$ is Hopfian if every surjective endomorphism $\theta : G \to G$ is injective.

If $G$ is no Hopfian, then we get $G/\ker \theta \cong G$, so $G$ is isomorphic to a proper quotient of itself.

Example 8.10. 1. Obviously, all finite groups are Hopfian since injective $\iff$ surjective for finite sets.

2. $\mathbb{Z}$ is Hopfian.

3. The free group on infinitely many generators is not Hopfian, since we can kill a generator and biject the remaining generators with the original ones. (By comparison, we proved any free group is residually finite.)

Theorem 8.11. Finitely generated residually finite groups are Hopfian.

Proof. Let $\theta : G \to G$ be a surjective homomorphism. If $H \leq_f G$ has index $n$, then $\theta^{-1}(H)$ has index $n$ too. Note also that if $\theta^{-1}(H_1) = \theta^{-1}(H_2)$, then applying the surjective map $\theta$ we get $H_1 = H_2$. So the pullback map is injective on the set of index $n$ subgroups of $G$. By 6.9, this is a finite set for each $n$, so pullback induces a permutation on this finite set.

However, $\ker \theta \leq \theta^{-1}(H)$ for all $H \leq_f G$, $\ker \theta$ is in the intersection of the pullback of finite-index subgroups, which we just saw is the intersection of all the finite-index subgroups, which is trivial since $G$ is residually finite.

Corollary 8.12. If $g_1, \ldots, g_n$ generate the free group $F_n$, then they are freely generate $F_n$.

Proof. If $x_1, \ldots, x_n$ is a free basis for $F_n$, then we have a surjective map $F\{"g_1", \ldots, "g_n"\} \to F\{x_1, \ldots, x_n\}$ sending “$g_i$” to $g_i$ which is therefore an isomorphism.

Example 8.13. An infinite simple group is Hopfian but not residually finite. There are finitely generated infinite simple groups, so there is no converse.

8.3 Baumslag-Solitar Groups

Definition 8.14. The Baumslag-Solitar group $B_{m,n}$ is $\langle a, t \mid ta^m t^{-1} = a^n \rangle$ for $m, n \neq \emptyset$. 
So $B_{1,1} = \mathbb{Z} \times \mathbb{Z}$ and $B_{1,-1}$ is the Klein bottle group. You can change $m, n$ without (necessarily) changing the group: $B_{m,n} \cong B_{-m,-n} \cong B_{n,m}$ (in fact, these are the only identifications). They are HNN extensions $\mathbb{Z} = \langle a \rangle \star \varphi$ with $\varphi: \langle a^m \rangle \to \langle a^n \rangle$.

**Proposition 8.15.** $B_{m,n}$ is soluble if $|m|$ or $|n| = 1$ and contains $F_2$ otherwise.

**Proof.** If $|m|$ or $|n| = 1$, then without loss of generality we have $B_{1,n}$. (We have already seen that $B_{1,2}$ is soluble, cf. Example 5.13.) We can quotient by $\langle a \rangle$ to get $\mathbb{Z}$, showing solubility.

Otherwise, $a$ is not in the domain or image of $\varphi$. So for any reduced word $w(x,y) \in F_2$, we have $w(t,ata^{-1})$ is a reduced sequence in the HNN extension, hence nontrivial by Britton’s Lemma. Therefore, $\langle t, ata^{-1} \rangle \cong F_2 \leq B_{m,n}$.

**Theorem 8.16.** $B_{2,3}$ is not Hopfian.

**Proof.** Let $\theta(t) = t$ and $\theta(a) = a^2$. This preserves the relation, $\theta(ta^2t^{-1}) = ta^4t^{-1}$ and $\theta(a^3) = a^6$, so it is well-defined.

Is it actually surjective? We have to show that $a$ is in the image.

$$\theta(tat^{-1}a^{-1}) = ta^2t^{-1}a^{-2} = a^3a^{-2} = a.$$ 

Finally, we want to show that this has nontrivial kernel. Well,

$$\theta([tat^{-1}, a]) = [a^3, a^2] = e$$

so $[tat^{-1}, a] \in \ker \theta$. In the original group, this is $tat^{-1}ata^{-1}t^{-1}a^{-1}$ is a reduced sequence, hence nontrivial by Britton’s Lemma again.

**Theorem 8.17.** There exists a finitely generated soluble group which is not finitely presented.

**Proof.** The idea is to take a non-Hopfian group and iterate the surjective endomorphism. Consider $G = B_{2,3}$ and $\theta$ as above. Let $K_i = \ker \theta^i \triangleleft G$ with $K_i < K_{i+1}$. For $y \neq e$ with $\theta(y) = e$, we have $y = \theta^i(x)$ as $\theta^i$ is onto, so $x \in K_{i+1} \setminus K_i$.

Now form $Q := G/(\bigcup K_i)$. This is not finitely presented by 4.5: if it were, then it would be finitely presented by any set of generators, so we could take $Q = \langle a, t \mid S \rangle$ with $S = \langle s_1, \ldots, s_k \rangle$ in $G$. These lie in a finite union of the $K_i$, but we just produced something in $K_{N+1} \setminus K_N$ for any $N$.

To show solubility, observe that $G'$ is generated by $t^i at^{-i}$ for $i \in \mathbb{Z}$. But

$$\theta^i(t^i at^{-i}) = \theta^i(t^i a^2 t^{-i}) = \theta^i(a^3)$$
which commutes with $\theta^i(a)$. This shows tat $[t^i a t^{-i}, a] \mapsto e$ in $Q$. Conjugating by an appropriate power of $t$,

$$[t^j a t^{-j}, t^k a t^{-k}] \in \ker \theta^i.$$ 

Thus $\theta(G') = Q'$ is abelian, so $Q'' = I$. 

$\square$
Chapter 9

The Generalized Burnside Problem

Recall that a torsion group is one in which every element has finite order. It’s easy to produce many examples: finite groups, infinite direct sums of finite groups, $\mathbb{Q}/\mathbb{Z}$, . . .

The generalized Burnside problem asks: does there exist an infinite but finitely generated torsion group?

**Lemma 9.1.** If $G$ is finitely generated and torsion, then

1. If $G \rightarrow Q$, then $Q$ is either finite or infinite, finitely generated and torsion.

2. For $H \leq_f G$, $H$ is also finitely generated infinite torsion by 6.11.

3. $G$ is not virtually soluble.

**Proof.** Only the third part is not obvious. If $H \leq_f G$ is soluble, then $H/H'$ is finitely generated, torsion, and abelian, hence finite. Thus $H'$ is finitely generated, infinite, torsion and soluble. Continuing down the derived series, we get $H^{(n)} = I$ is a finitely generated infinite torsion.

The Burnside problem asks: if $G$ is finitely generated and there exists $k$ such that for all $g \in G$, $g^k = e$ (so $G$ has “exponent $k$”), can $G$ be infinite?

Define the “free Burnside groups” $FB(n,k) = \langle x_1, \ldots, x_n \mid w^k = e \forall w \rangle$.

Then a group $G$ is $n$-generated with exponent $k$ if and only if $FB(n,k) \rightarrow G$. So the Burnside question can be equivalently phrased: do there exist $n,k$ such that $FB(n,k)$ is infinite?

Let us mention one more problem. The restricted Burnside Problem asks: can $G$ be a positive example for the Burnside problem, which is infinite and residually finite? In terms of the free Burnside group, the universal property of the finite radical (8.4) implies that this is equivalent to $FB(n,k)/R$ being infinite for some $n,k$. 
**Theorem 9.2** (Golod ’64). There exists a finitely generated, infinite $p$-group (every element has order $p^k$ for some $k$ depending on the element).

Let $p$ be any prime.

**Definition 9.3.** In $F_n$, the $p$-value $v_p(w)$ of a non-identity word $w \in F_n$ is

$$\max\{k : w = u^{p^k} \text{ for some } u \in F_n\}.$$

**Definition 9.4.** The $p$-deficiency of a presentation is defined to be

$$\langle x_1, \ldots, x_n \mid r_1, r_2, \ldots \rangle = n - \sum_{i=1}^{\infty} \frac{1}{p^{v_p(r_i)}}$$

if it converges.

**Lemma 9.5.** Say a group $F$ acts on a set $X$ and suppose that $S \triangleleft F$ satisfies $[F : S] = p$. For $x \in X$, if there exists $g \in \text{Stab}_F(x) - S$ then $\text{Orb}_F(x) = \text{Orb}_S(x)$.

**Proof.** By the hypothesis, we have $F = S \text{Stab}_F(x)$. For $f(x) \in \text{Orb}_F(x)$, write $f = st$ where $s \in S$ and $t \in \text{Stab}_F(x)$, so $f(x) = s(x)$.

**Theorem 9.6.** For any prime $p$, and any $n \geq 2$, there exists an infinite $n$-generator $p$-group which is residually finite.

**Proof.** Suppose $\langle x_1, \ldots, x_n \mid r_1, r_2, \ldots \rangle$ is a presentation $P$ with $p$-deficiency $\geq 1$. Let $G = F_n/R$ be the group it defines. Here is the game we will play.

1. Any group with $p$-deficiency at least 1 admits a surjection to $C_p$.
2. The kernel of such a surjection is a normal subgroup $N$ with $p$-deficiency at least 1.
3. By repeating, we deduce that any group with $p$-deficiency at least 1 has an infinite descending sequence of index-$p$ subgroups, and in particular is infinite.
4. All of these subgroups contain $R_G$, so $G/R_G$ is residually finite and infinite.
5. We construct a torsion group with $p$-deficiency at least 1.

Now let’s start proving these claims. First up, we establish a surjection $G \twoheadrightarrow C_p$. Note that there is a natural map $F_n \rightarrow \mathbb{F}_p^n$. We can have $v_p(r_i) = 0$ for at most $n - 1$ relators, or else the $p$-deficiency of $P$ is $\leq n - n - \ldots$. By 4.9, if the span of $\overline{r_i}$ is not $\mathbb{F}_p^n$, then $G \hookrightarrow C_p$. But if $v_p(r_i) \geq 1$ then $\overline{r_i} = 0 \in (\mathbb{F}_p)^n$, so at most $n - 1$ relators have non-vanishing image in $\mathbb{F}_p^n$, so the span has dimension at most $n - 1$. 
Let \( N \) be the kernel of this surjection, and set \( N = S/R \) for \( R \leq S \triangleleft F_n \) with \( [F_n : S] = p \). By the proof of 6.11, \( N \) is generated by \( p(n - 1) + 1 \) elements and \( R = \langle \langle t^i r_i t^{-j} : i \in \mathbb{N}, 0 \leq j \leq p - 1 \rangle \rangle_s \) where \( \{e, t, \ldots, t^{p-1}\} \) is a transversal for \( S \) in \( F_n \) if \( t \notin S \). We could calculate the \( p \)-deficiency of this presentation crudely, but to get what we want we have to be more careful with our relations. Consider one relator \( r = r_i \in P \) and set \( k = v_p(r) \), so \( r = w^p \) for some \( w \). We have two cases:

**Case (a).** If \( w \notin S \), then by 9.5 applied with respect to the conjugation action of \( G \) and \( x = r, f = w \), we get that the conjugacy class in \( F_n \) of \( r \) is the conjugacy class in \( S \) of \( r \), so \( \langle\langle r, \ldots, t^{p-1}rt^{-1}\rangle\rangle_s = \langle\langle r \rangle\rangle_s \) since everything in there is conjugate to \( r \) in \( G \). Now, \( r = (w^p)^{k-1} \) and \( w^p \in S \) since \( S \) is normal of index \( p \).

**Case (b)** If \( w \in S \), then \( t^i r_i t^{-j} = (t^i wt^{-j})^p \).

Thus for each \( r_i \), we can either (a) throw out the \( t^i r_i t^{-j} \) and use that \( v_p(r_i) \) in \( S \) is \( v_p(r_i) \) in \( G \) minus 1, or (b) keep all the \( t^i r_i t^{-j} \) and use that \( v_p(r_i) \) in \( S \) is equal to \( v_p(r_i) \) in \( G \). This gives a presentation \( Q \) for \( N \) with \( p \)-deficiency

\[
p(n - 1) + 1 - \sum p \frac{p}{p^{v_p(r_i)}} = p(p-\text{def}_G(P) - 1) + 1.
\]

In particular, \( Q \) has \( p \)-deficiency at least 1.

Now we simply have to construct a torsion group with \( p \)-deficiency at least 1. List the nonidentity elements of \( F_n \) as \( \{w_1, w_2, \ldots\} \) and set \( P = \langle x_1, \ldots, x_n \mid w_1^p, w_2^p, w_3^p, \ldots \rangle \). This has \( p \)-deficiency at least 1 for \( n \geq 2 \), since

\[
\sum_{i=1}^{\infty} \frac{1}{p^i} < 1.
\]

What about residual finiteness? By 8.4, \( G/R_G \) is residually finite \( p \)-group which is finitely generated, and \( N_1, N_2, \ldots \geq R_G \) so \( G/R_G \) is also infinite (it also has an infinite descending sequence of subgroups with successive index \( p \)).

**Remark 9.7.** Is it just me, or is this proof insane?

The other two problems are very difficult! \( FB(n,2) \) is abelian, and \( FB(n,3 \text{ or } 4 \text{ or } 6) \) are finite. Novikov and Adyan showed that \( B(n, k) \) is infinite for all odd \( k \geq 665 \). Ol’Shanksii showed that for all large primes \( p \), there exists a finitely generated infinite group \( G \) such that if \( I < H < G \) then \( H \simeq C_p \). In particular, \( G \) has max!

Zelmanov showed that the answer to the restricted Burnside problem is no in 1994, for which he won the Fields medal (kind of a big deal).
Chapter 10

Example Sheet 1

1. Show that the property of being a torsion group is preserved under subgroups, quotients, and extensions. What about the property of being torsion-free?

The closure under subgroups and quotients is obvious. If

$$0 \to N \to G \to Q \to 0$$

is exact and $g \in G$, then the image of $g$ in $Q$ is torsion, so $g^n \in N$ for some $n$. Since $N$ is torsion, $g^n$ is torsion, hence so is $g$.

Torsion-free is not preserved under quotients, e.g. $\mathbb{Z} \to \mathbb{Z}/n$.

2. Show that every proper subgroup of $\mu_{p^\infty}$ is finitely generated.

Since $\mu_{p^\infty}$ is generated by the primitive $p^k$ roots of unity, a proper subgroup fails to contain one such for some $k$. But then it is contained in $\mu_{p^{k-1}}$, which is finite.

3. We can define the group property min in exactly the same way as max except we replace ascending by descending. What can you say about a group with min (in particular, does $\mathbb{Z}$ have min)? Have you seen any examples in this course of infinite groups with min? Of infinite finitely generated groups with min?

Despite the similarities in appearance, min is crazier than max (think Artinian vs Noetherian in commutative algebra, if that helps). $\mathbb{Z}$ doesn’t have min, since we have the infinite descending sequence $\mathbb{Z} \supset 2\mathbb{Z} \supset 4\mathbb{Z} \supset \ldots$. In particular, any group with min must be torsion (or else it contains $\mathbb{Z}$).

The group $\mu_{p^\infty}$ has min, since all its proper subgroups are finite.

♠♠♠ Tony: [what about the third part?]

4. For $F$ a field, prove that the center of $\text{GL}(n, F)$ consists of non-zero scalar matrices.
Take a generic element and compare its right and left multiplication with the elementary matrix $E_{ij}$, which is the identity except with a 1 in spot $(i, j)$. This shows that anything in the center must be diagonal, and then one compares with a diagonal matrix having distinct entries on the diagonal to deduce that it must be scalar.

Note that this needn’t be the largest abelian subgroup of $GL(n, F)$. For instance, if $L/F$ is a quadratic extension, then we have an embedding $L^* \hookrightarrow GL(2, F)$.

5. Show that any subgroup $H$ of $G = \mathbb{Z}^n$ is isomorphic to $\mathbb{Z}^m$ for some $m \leq n$, with $m = n$ if and only if $H$ has finite index in $G$.

This is a special case of the result that any submodule of a free module over a PID is free (of lower or equal rank). That argument proceeds by applying row and column operations to a presentation matrix for $H$.

There are a couple of other quick arguments that work in this case. First, one could note that $H$ is finitely generated and torsion-free, and apply the classification of finitely generated abelian groups.

Another, very elementary approach is to $G_i = \mathbb{Z}^i \subset G$, and $H_i = G_i \cap H$. Then $H_i/H_{i-1} \subset G_i/G_{i-1} \cong \mathbb{Z}$, so $H_i/H_{i-1} \cong \mathbb{Z}$ or is trivial. Since $\mathbb{Z}$ is free, there is a splitting $H_i \cong \mathbb{Z} \oplus H_i/H_{i-1}$. By induction, we get $H = \mathbb{Z}^n$ (and the $m \leq n$ comes for free). Actually, this is another approach to proving the general result over PIDs.

To see the second part, tensor the short exact sequence

$$0 \to H \to G \to G/H \to 0$$

with $\mathbb{Q}$. We will in fact have an inclusion $H \otimes \mathbb{Q} \hookrightarrow G \otimes \mathbb{Q}$ since $H$ and $\mathbb{Q}$ are both free $\mathbb{Z}$-modules, so $m \leq n$. If $m = n$, then by the right-exactness of the tensor product we have $(G/H) \otimes \mathbb{Q} = 0$, hence is torsion. It is obviously finitely generated, so it is finite.

6. (i) If $N \triangleleft G$ then show that the abelianization of $G/N$ is isomorphic to $G/G'N$ (so that the abelianization of the quotient is a quotient of the abelianization).

(ii) Find a finitely generated group with finite abelinization but with a subgroup of index 2 having infinite abelianization (so we cannot replace “quotient” with “subgroup” in (i)).

(i) Note that $[G/N, G/N]$ is generated by things of the form $[gN, hN] = [g, h]N$, so the commutator subgroup is $[G, G]N/N$. The result then follows from the third isomorphism theorem.

(ii) Let’s try to do the simplest thing possible. The simplest infinite abelian group is $\mathbb{Z}$, so let’s try $G = \mathbb{Z} \times C_2$. We don’t want this to
be a direct product, so we define the conjugation to be \( ata^{-1} = t^{-1} \).
Since \( ata^{-1}t^{-1} = t^{-1}t^{-2} \), we get \( (G/[G,G]) \leq 4 \).

Note that this group is also known as the infinite dihedral group, or \( C_2 \ast C_2 \).

7. (i) Use the Klein combination theorem to show that the subgroup of \( PSL(2, \mathbb{Z}) \) generated by \( f(z) = -1/z \) and \( g(z) = -1/(z + 1) \) is the free product \( C_2 \ast C_3 \).
(ii) Show further that \( \langle f, g \rangle = PSL(2, \mathbb{Z}) \) so \( PSL(2, \mathbb{Z}) \cong C_2 \ast C_3 \).

(i) You can consider the action on \( \mathbb{H} \), with the region \( \text{Re} \, z < 0 \) as a \( f \)-packing and a region that follows the line \( \text{Re} \, z = 0 \) up a bit and then moves slightly to the left as a \( g \)-packing.
(ii) This is a standard computation. Note that \( f \) and \( g \) correspond to the linear fractional transformations
\[
S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.
\]

One can induct on \( |c| \), and use row and column operations to reduce \( |c| \) to 0, at which point the resulting matrix is obviously a power of \( T \). More geometrically, one can see this by studying the orbits of \( \text{SL}_2(\mathbb{Z}) \) on \( \mathbb{H} \). By considering how \( \text{SL}_2(\mathbb{Z}) \) acts on the imaginary part of \( \tau \), one establishes the standard fundamental domain for the modular group, and then shows that any point has an \( \langle S, T \rangle \) orbit in this fundamental domain.

8. (i) For \( w(a,b,c) \in F_3 \) let \( \theta(w) \) be the exponent sum of \( a \) in \( w \), i.e. the number of appearances of \( a \) minus the number of appearances of \( a^{-1} \). Show that \( \theta : F_3 \to \mathbb{Z} \) is a surjective homomorphism. Let \( \theta_k : F_3 \to C_k \) be the exponent sum of \( a \) modulo \( k \). What is the rank \( r \) of \( \ker \theta_k \)? Give a set of \( r \) generators for \( \ker \theta_k \).
(ii) What are the index 2 subgroups of \( F_2 \)?

(i) The map \( \theta \) is obviously surjective. The homomorphism part is also pretty obvious; think about concatenating reduced words and observe that cancellation not affecting the image. To find the index, recall the formula that the rank of an index \( i \) subgroup is \( i(n-1)+1 \). A set of generators is \( a^i a^{-i}, a^i b a^{-i}, a^i b^i c a^{-i} \) for \( 0 \leq i \leq k - 1 \).

(ii) All index 2 subgroups are normal, so this is equivalent to describing the homomorphisms \( F_2 \to C_2 \). That, in turn, is equivalent to describing the non-zero set maps \( \{a,b\} \to \{0,1\} \) (by the universal property of free groups). We see that there are three such.

9. For \( F_2 \) free on \( a \) and \( b \), express \( a^2 b^2 a^{-2} b^{-2} \) as an element of the normal closure of \( aba^{-1} b^{-1} \).
A convenient way to do this is to write down the group $\langle a, b \mid ab = ba \rangle$. (This makes it obvious that $a^2b^2a^{-2}b^{-2}$ should be in the normal closure.) Then, starting with $a^2b^2a^{-2}b^{-2}$, apply the relation $ab = ba$ repeatedly to get $e$. If you carefully track your steps, and replace the relation by inserting the relator $aba^{-1}b^{-1}$, then you get an expression of the desired form, although it is painful.

10. Given an automorphisms $\alpha$ of a group $N$ and the integers $\mathbb{Z} = \langle t \rangle$, we can form the semidirect product $N \rtimes_{\phi} \mathbb{Z}$ for the homomorphism $\phi : \mathbb{Z} \to \text{Aut}(N)$ sending $t^k$ to $\alpha^k$. We write $N \rtimes_{\alpha} \mathbb{Z}$ for this.

(i) Show that for each $i \in \mathbb{N}$, the group $N \rtimes_{\alpha} \mathbb{Z}$ has an index $i$ subgroup of the form $N \rtimes_{\alpha^i} \mathbb{Z}$.

(ii) Suppose $\alpha$ has finite order $d$. Then what can we say about the index $d+1$ subgroup (or even the index $kd+1$ subgroup for any $k \in \mathbb{N}$) obtained in (i)?

(i) The group $N \rtimes_{\alpha^i} \mathbb{Z}$ can be realized as $N \times_{\alpha} i\mathbb{Z}$, which obviously has index $i$ in $N \rtimes_{\alpha} \mathbb{Z}$.

(ii) This subgroup is actually isomorphic to $N \rtimes_{\alpha} \mathbb{Z}$, since $\alpha^{kd+1} = \alpha$. So the group is isomorphic to a proper subgroup of itself.

11. Let $F_2$ be the free group of rank 2 on $a, b$. What group is obtained when we form the HNN extension by taking the isomorphism $\phi$ from $\langle a \rangle$ to $\langle b \rangle$ given by $\phi(a^k) = b^k$?

The HNN extension can be presented as $\langle a, b, t \mid tat^{-1} = b \rangle = \langle a, t \rangle \cong F_2$. 
Chapter 11

Example Sheet 2

1. Show that the group defined by the presentation
\[ \langle a, b \mid aba^{-1} = b^2, bab^{-1} = a^2 \rangle. \]
is trivial.

(ii) Do the same for
\[ \langle a, b \mid ab^2a^{-1} = b^3, ba^2b^{-1} = a^3 \rangle. \]

(iii) If \( G \) is the group
\[ \langle a, b, c \mid aba^{-1} = b^2, bcb^{-1} = c^2, cac^{-1} = a^2 \rangle \]
than show that \( G \) can be generated by just \( a \) and \( b \). Conclude that the second derived group \( G'' \) is trivial. As it has no proper finite index subgroups (by 6.13), what do we conclude about \( G \)?

(i) We have \( ab = b^2a \) and \( ba = a^2b \) by the relations. Then \( ab^2a = a^2b = ba \), so \( ab = e \). Putting that above, we get \( a = b = e \).

(ii) From the relation \( ab^2a^{-1} = b^3 \), we have
\[ a^3b^8a^{-3} = b^{27}. \]

From the relation \( ba^2b^{-1} = a^3 \) we have
\[ a^3b = ba^2 \]
\[ a^{-3} = ba^{-2}b^{-1}. \]

Substituting the first one above gives
\[ ba^2b^7a^{-3} = b^{27} \implies a^2b^7a^{-3} = b^{26}. \]
Substituting the second one gives

\[ a^2b^8a^{-2} = b^{27}. \]

On the other hand, the first relation gives

\[ a^2b^8a^{-2} = b^{24}. \]

Therefore, \( b^3 = 1 \). Plugging that into the first relation gives \( ab^2 = a \), which forces \( b^2 = 1 \) as well, so \( b = 1 \) and \( a^2 = a^3 \implies a = 1 \).

(iii) Tony: For the second part, \( H = \langle a, b \mid [a, b] = b \rangle \) surjects onto \( G \), and \( H'' = 1 \) implies \( G'' = 1 \). So \([a, b] = b, [c, a] = a, [b, c] = c\). That would imply \( G = G' = G'' = 1 \).

2. Show that the property of not containing a non-abelian free group is preserved under subgroups, quotients, and extensions.

The assertion for subgroups is obvious. Quotients: if \( G \twoheadrightarrow Q \) and \( F_2 \subset Q \), say that this \( F_2 \) is generated by \( a, b \). We can pick any two pre-images \( \tilde{a}, \tilde{b} \) of \( a, b \) and map \( F_2 \rightarrow G \) by sending \( a, b \) to \( \tilde{a}, \tilde{b} \). This must be an isomorphism of \( F_2 \) onto its image subgroup, since after composing with the surjection we get the identity.

Extensions: suppose that

\[ 0 \rightarrow N \rightarrow G \rightarrow Q \rightarrow 0. \]

if \( F_2 \simeq H \leq G \), then \( H \cap N \) is free and hence isomorphic to \( I \) or \( \mathbb{Z} \). If it’s \( I \), then \( H \) maps isomorphically onto its image in \( Q \), but \( Q \) doesn’t contain \( F_2 \) by hypothesis. Therefore, \( H \cap N \cong \mathbb{Z} \) generated by some \( w(a, b) \). Then \( N \cap H \lhd H \), but picking a cyclically reduced representative \( w = wu^{-1} \) shows that this is impossible.

3. (i) Let \( G \) be the direct product of \( N \) copies of the vector space \( \mathbb{Z}/2\mathbb{Z} \).

Show that every proper subgroup of \( G \) is contained in a maximal subgroup.

(ii) Find a proper subgroup of \( \mathbb{Q} \times C_2 \) that is contained in a maximal subgroup and one that is not.

(i) Let \( G = \prod_{N} \mathbb{Z}/2\mathbb{Z} \). This is a vector space over \( \mathbb{F}_2 \), and any subgroup \( H \) is a subspace. We can then apply Zorn’s lemma to obtain a basis extending one for \( H \), and pick a maximal subspace which is spanned by all but one basis element.

(ii) \( \mathbb{Q} \times \{0\} \subset \mathbb{Q} \times C_2 \) is maximal, but e.g. \( \{0\} \times C_2 \) is not maximal (any maximal subgroup containing it would project to a maximal subgroup of \( \mathbb{Q} \)).
4. We define the Frattini subgroup \( \text{Frat}(G) \) to be the intersection of all the maximal subgroups of \( G \) and we say that \( g \in G \) is a non-generator if whenever \( g \in X \subseteq G \) and \( \langle X \rangle = G \), then \( \langle X - \{g\} \rangle \) is also equal to \( G \). Prove that \( g \) is a non-generator if and only if \( g \in \text{Frat}(G) \).

Let \( g \in G \) be a non-generator. Pick \( M \) a maximal subgroup. Then \( \langle M, g \rangle = M \) or \( G \). If it were \( G \), then \( g \) is a generator, so we must have \( \langle M, g \rangle = M \), i.e. \( g \in M \).

In the other direction, suppose \( g \in \text{Frat}(g) \). If there exist a set \( X \) such that \( \langle X \rangle = G \) and \( \langle X - g \rangle \neq G \), then Zorn’s lemma implies that there is a maximal subgroup \( M \) containing \( \langle X - g \rangle \). Therefore, \( g \notin M \), a contradiction.

5. Show that if \( G = K \ast L \) for \( K, L \neq I \) then \( \text{Frat}(G) = I \).

Without loss of generality, write \( g = k_1l_1 \ldots k_nl_n \) with \( k_1 \) nontrivial. If \( g \neq k_1 \), then the group \( \langle gKg^{-1}, L \rangle \) doesn’t contain \( K \), but if we throw in \( g \) then it does.

If \( g = k_1 \), then pick \( l \in L - I \) and consider \( \langle glKl^{-1}g^{-1}, L \rangle \).

6. If \( m \in \mathbb{N} \) then prove that the Baumslag-Solitar group

\[
B_{m,m} = \langle a, t \mid ta^mt^{-1} = a^m \rangle.
\]

is residually finite by expressing it as a semidirect product.

We want to realize this presentation as a semidirect product. A semidirect product comes from the action of one group on another, and the idea is that the conjugation action should move the generators around in such a way that the \( m \)th iteration is the identity. So we try

\[
F(x_1, \ldots, x_m) \rtimes \mathbb{Z}
\]

with action \( ax_1a^{-1} = x_2, ax_2a^{-1} = x_3, \ldots, ax_ma^{-1} = x_1 \). This certainly maps to \( B \), sending \( x_1 \) (say) to \( a \). To check that the inverse is well-defined, we just have to check that the relation is satisfied, which is true by construction.

7. (i) Extract from the proof of Theorem 6.21 the result that if \( G \) is finite by soluble then it is soluble by finite. Adapt this to get the same statement with soluble replaced by nilpotent.

(ii) Show that if \( G \) is finitely generated and finite by \( \mathcal{P} \), where \( \mathcal{P} \) is respectively abelian, nilpotent, polycyclic, then \( G \) is \( \mathcal{P} \) by finite.
(i) There exists $F \triangleleft G$ finite such that $G/F$ is soluble. By orbit-stabilizer, $C_G(f) \triangleleft f G$ for all $f \in F$. Therefore, $C := C_G(F) = \bigcap_{f \in F} C_G(f)$ the centralizer has finite index in $G$.

$$C/C \cap F \cong CF/F \leq G/F$$

and $G/F$ is soluble, hence $CF/F$ is soluble (subgroup of soluble is soluble), hence $C/C \cap F$ is soluble. Also, $C \cap F = Z(F)$, which is finite abelian hence soluble. That shows that $C$ is soluble, so we are happy.

Virtually polycyclic implies residually finite (8.6). So in all cases, $P$ implies residually finite. Suppose $F \triangleleft G$ is a finite normal subgroup and $G/F$ has $P$. For all $f \in F$ there exists $N_f \triangleleft f G$ such that $f \notin N_f$ since $G$ is residually finite. Then

$$N = \bigcap_{f \in F} N_f \triangleleft f G$$

and $N \cap F = 1$, so $N = N/N \cap F \cong NF/F \triangleleft G/F$ so $N$ has $P$.

8. Suppose $G$ is finitely generated.

(i) If $g^2 = e$ for all $g \in G$ then $G$ is finite. Why?

(ii) If now $g^3 = e$ for all $g \in G$ then show that for all $g, h \in G$ we have $[g, hgh^{-1}] = e$. Then by taking $G = \langle g_0, g_1, \ldots, g_n \rangle$ and considering $G/\langle \langle g_0 \rangle \rangle$ show that $G$ is finite by induction.

(i) $G$ is abelian, and finitely generated, and torsion, hence finite.

(ii) By induction, $G/\langle \langle g_0 \rangle \rangle$ is finite. We claim that $\langle \langle g_0 \rangle \rangle$ is abelian, i.e. $[g, hgh^{-1}] = e$. Well, (using the fact that all cubes are trivial)

$$g(hgh^{-1})g^{-1}(hg^{-1}h^{-1}) = ghgh^{-1}(h^{-1}gh^{-1}h^{-1})$$

$$= ghgh^{-2}gh^{-2}$$

$$= ghghg = e.$$

Now, we want to show that $\langle \langle g_0 \rangle \rangle$ is finitely generated. This follows from the fact that it has finite index in a finitely generated group. (By the way, why is that true? A group is finitely generated if and only if it is the fundamental group of a finite graph, and a finite index subgroup is the fundamental group of a finite covering space, which is again a finite 1-dimensional finite graph.)

9. There exists an infinite group $G$ and a large prime $p$ such that every proper non-trivial subgroup of $G$ has order $p$. 
(i) Show that $G$ is finitely generated.

(ii) Show that $G$ is simple.

If we replace the above condition on the subgroups of $G$ with: every element has order $p$, then do (i) or (ii) still hold?

(i) Any non-identity element of $G$ generates a subgroup. If it all of $G$, then we are done (although this can’t be the case, because $G \neq \mathbb{Z}$). Otherwise, take an element not in this group: it generates a subgroup of order more than $p$, hence all of $G$.

(ii) Now, suppose we have a nontrivial proper normal subgroup. Then the quotient is again an infinite group, and pulling back a proper subgroup gives a proper subgroup of $G$ having order bigger than $p$.

An infinite direct product of $\mathbb{Z}/p$ gives a counterexample for the weaker condition where every element has order $1$ or $p$. 