Presentation of Groups

Derek Holt
Notes by Florian Bouyer

23rd April 2013

Contents

1 Introduction

1.1 Motivation

This course is about group presentations, $\langle X|R\rangle$ where X is the set of generators and R the set of relations which are equations between words in A^*, with $A = X \cup X^{-1}$.

Example. $X = \{x, y\}$ and $R = \{x^5 = 1, y^2 = 1, y^{-1}xy = x^{-1}\}$. Usually write as $\langle x, y|x^5 = y^2 = 1, y^{-1}xy = x^{-1}\rangle$.

A presentation $\langle X|R\rangle$ defines a group, which is roughly the "largest group" which is generated by X such that all equations in R holds in G. In the above example we can show any group $G = \langle x, y\rangle$ with $x^5 = y^2 = 1, y^{-1}xy = x^{-1}$ has at most 10 elements, and dihedral group D_{10} is unique group of order 10. So we can say $G \cong D_{10}$. The advantage of this way of defining groups:

1. For many groups, it is the most compact definition, particularly useful for systematically enumerating small groups.
2. Many groups from algebraic topology arises naturally in this form, so we need to study them.

Disadvantage: In general it is impossible to analyse a group by a presentation. For example we cannot decide if $\langle X|R\rangle$ is finite, trivial or even abelian. It has been proven that there are no algorithm to decide this. There exist specific presentations $\langle X|R\rangle$ with X, R finite for which we cannot decide whether an element $g \in G$, given as $w \in (X \cup X^{-1})^*$ is the identity 1_G (this is know as the word problem).

1.2 Preliminaries

Notation.

- If G is a group, $H \subseteq G$ then H is a subset of G. $H \leq G$ then H is a subgroup of G. $H \lhd G$ then H is a normal subgroup of G.
- C_k is the cyclic group of order k.
- Infinite cyclic group is $\mathbb{Z}$ (under $+$)
- S_k is the symmetric group of degree k on $\{1, 2, \ldots, k\}$
- A_k is the alternating group degree k
- $H \leq G$, $G = \sqcup Hg_i$, so g_i are right coset representative. Call $\{g_i|i \in I\}$ a *right transversal* of H in G.
- If $H \lhd G$, $\{Hg_i|i \in I\}$ forms the quotient group G/H.
- Group homomorphism: $\theta : G \to H$ such that $\theta(g_1g_2) = \theta(g_1)\theta(g_2)$
- Group monomorphism: If θ is injection if and only if $\ker(\theta) = \{1\}$
- Group epimorphism: if θ is surjection, $\operatorname{im}\theta = H$
- Group isomorphism: θ bijective
- Group endomorphism if $G = H$
- Group automorphism: isomorphism with $G = H$
- $\operatorname{Aut}(G) = \{\theta|\theta : G \to G$ automorphism $\}$ is a group under composition

The isomorphism Theorem.

1. *Let $\theta : G \to H$ is a homomorphism and $K = \ker(\theta)$. Then $\overline{\theta} : G/K \to \operatorname{im}\theta$ defined by $\overline{\theta}(gK) = \theta(g)$ is an isomorphism. So $G/K \cong \operatorname{im}\theta$*
2. *If $M \leq G, N \lhd G$ then $MN/N \cong M/(N \cap M)$* (Recall: if $A \leq G$ and $B \leq G$ then AB not always a subgroup, but it is if $A \lhd G$ or $B \lhd G$ and it is normal if $A \lhd G$ and $B \lhd G$.
3. *If $M \leq N \lhd G$ with $M \lhd G$ then $\frac{G/M}{N/M} \cong G/N$*

Proof. 1. See Algebra II course

2. Define $\theta : M \to MN/N$ by $\theta(m) = mN$. Then $\operatorname{im}\theta = MN/N$ and $\ker(\theta) = M \cap N$. Then the result follows from the 1st isomorphism theorem (part 1 of this theorem)

3. Define $\theta : G \to \frac{G/N}{N/M}$ by $\theta(g) = (gM)(N/M)$ (i.e. $G \to G/M \to \frac{G/M}{N/M}$). Then we have $\operatorname{im}\theta = \frac{G/M}{N/M}$ and $\ker\theta = N$. Then the result follows from the 1st isomorphism theorem.

□

Notation. (non-standard) If G is a group permutation of Ω (i.e., $G \leq \operatorname{Sym}\Omega$). Given $\alpha \in \Omega$, $g \in G$ then we write the image of α under g as α^g rather than $g(\alpha)$. This means that gh means g followed by h (not h then g)

Example. $\Omega = \{1,2,3,4\}$, $g = (1,2,3)$ and $h = (3,4)$. Then $2^g = 3, 3^h = 4$ so $2^{gh} = (2^g)^h = 4$. Note $gh = (1,2,4,3)$ while $hg = (1,2,3,4)$.

1.3 Generators of Groups

Let A be a set. A *word* over A is a finite string $w = a_1a_2\dots a_l$ with each $a_i \in A$. The *length* of w is $l = l(w) = |w|$. We allow $l = 0$, this is the *empty word*, which we denote by ϵ.

Definition 1.1. Let $X \subseteq G$ with G a group. We define *subgroup* $\langle X\rangle$ of G generated by X in two ways:

1. $\langle X\rangle$ =intersection of all subgroups of G that contains X, i.e., $\langle X\rangle = \cap_{H\leq G, X\subseteq H} H$

2. Let $X^{-1} = \{x^{-1} | x \in X\}$. $A = X \cup X^{-1}$. We define A^* to be the set of all words over A. Elements of A^* represent elements of G, it is closed under concatenation and inversion. So it is a subgroup of G. We define $\langle X\rangle = A^*$. The empty word represents 1_G

Lemma 1.2. *The two definition of $\langle X\rangle$ are equivalent*

Proof. Let $X \subseteq G$. Let H_1 and H_2 be $\langle X\rangle$ according to definition 1. and 2.

If $g \in H_2$, then $g = a_1a_2\dots a_l$, $l \geq 0$ $a_i \in X \cup X^{-1}$. Then since subgroups are closed under multiplication and inversion, g is contained in any $H \leq G$ with $X \subseteq H$. Hence $g \in H_1$, i.e., $H_2 \subseteq H_1$.

But H_2 is a subgroup of G containing X, so $H_1 \subseteq H_2$. Hence $H_1 = H_2$. □

Example. 1. $G = (\mathbb{Z}, +)$ and $X = \{12, 18\}$. Then by definition 2, $\langle X\rangle = \{12a + 18b | a, b \in \mathbb{Z}\} = \langle 6\rangle = 6\mathbb{Z}$

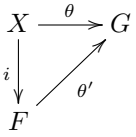
2. $G = S_4$, $X = \{a, b\}$ with $a = (1,2,3)$ and $b = (2,3,4)$. Now $a, b \in A_4$ so $\langle X\rangle \leq A_4$ (by the 1st definition), and $ab = (1,3)(2,4)$ while $ba = (1,2)(3,4)$ and $abba = (1,4)(2,3)$. So $\{1, ab, ba, abba\} \leq \langle X\rangle$ so $4|\,|\langle X\rangle|$ and $3 = |a|\,|\,|\langle X\rangle|$ so $12|\,|\langle X\rangle|$,so $\langle X\rangle = A_4$.

2 Free Groups

These are groups $\langle X|\,\rangle$ with $R = \emptyset$ with no relations.

We define them using their principal property. If V, W are vector spaces, V with basis B. Any map $\theta : B \to W$ uniquely determines a linear map $\theta' : V \to W$ with $\theta'(b) = \theta(b)\,\forall b \in B$. Free groups have a similar property (but all vector spaces have bases, but not all groups are free)

Definition 2.1. Let F be a group and $X \subseteq F$. Then F is *free* on X if for any group G and any map $\theta : X \to G$ $\exists!$ homomorphism $\theta' : F \to G$ with $\theta'(x) = \theta(x)\,\forall x \in X$, i.e the diagram

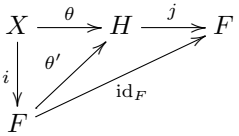


commutes. I.e., θ has a unique *extension* $\theta' : F \to G$. I.e. $\theta = \theta'$ where $i : X \to F$ is the *insertion/inclusion* map defined by $i(x) = x\,\forall x \in X$

We will prove existence later, first we prove a few properties.

Proposition 2.2. *Let F be free on X. Then $F = \langle X \rangle$ (i.e., X generates F)*

Proof. Assume F is free on X and let $H = \langle X \rangle \leq F$.

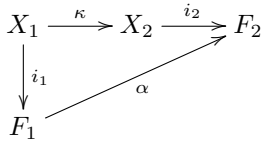


Let $\theta : X \to H, i : X \to F$ be insertion maps (i.e., $i(x) = \theta(x) = x$). Then there exists $\theta' : F \to H$ with $\theta' i = \theta$. Let $j : H \to F$ be insertion map, then $\theta' j$ and id_F are both extension of $j\theta : X \to F$. So by the uniqueness part of the definition we have $\theta' j = \mathrm{id}_F$. Since id_F is surjective, we have j is surjective hence $H = F$ □

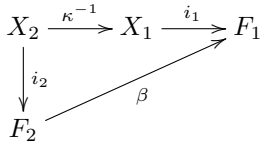
Proposition 2.3. *Let F_1 be free on X_1 and F_2 be free on X_2. Then $F_1 \cong F_2$ if and only if $|X_1| = |X_2|$. (In particular any two free groups on X are isomorphic)*

Proof. "⇒": Let G be any non-trivial finite group (such as C_2). The number of maps $X_1 \to G$ is $|G|^{|X_1|}$. Since each map uniquely determines a homomorphism, we get $|G|^{|X_1|} = |\hom(F_1, G)|$. Similarly we find $|G|^{|X_2|} = |\hom(F_2, G)|$. Now $F_1 \cong F_2 \Rightarrow |\hom(F_1, G)| = |\hom(F_2, G)|$ (exercise show this). This means $|G|^{|X_1|} = |G|^{|X_2|} \Rightarrow |X_1| = |X_2|$, which is clear if $|X_i|$ are finite. (In fact for infinite cardinal numbers $2^\alpha = 2^\beta \Rightarrow \alpha = \beta$ is independent of the axioms of set theory). In fact it can be proved that for X infinite, F free on X implies $|F| = |X|$ so result is true anyway. (proof omitted).

"⇐": Assume $|X_1| = |X_2|$ and let $\kappa : X_1 \to X_2$ be a bijection.



where i_1, i_2 are insertions. So $i_2\kappa$ extends to $\alpha : F_1 \to F_2$



so similarly, $i_1\kappa^{-1}$ extends uniquely to $\beta : F_2 \to F_1$. For $x \in X_1$, $\beta\alpha(x) = \beta i_2\kappa(x) = \beta\kappa(x) = i_1\kappa\kappa^{-1}(x) = i_1(x) = x$. So $\beta\alpha : F_1 \to F_2$ extends identity map $X_1 \to F_1$. So by uniqueness $\beta\alpha = \mathrm{id}_{F_1}$. Similarly we find $\alpha\beta = \mathrm{id}_{F_2}$. So α, β are isomorphism.

□

We now prove existence of free groups:

$F = \langle X| \ \rangle$ should be generated by X, so elements represented by strings in $x, x^{-1}, x \in X$, i.e. elements of A^* where $A = X \cup X^{-1}$. All words should be distinct? Except we want $xx^{-1} = x^{-1}x = 1$. Let X be any set, let X^{-1} be a set with $|X| = |X^{-1}|$ and $X^{-1} \cap X = \emptyset$, denote elements of X^{-1} by $\{x^{-1}|x \in X\}$. Define $(x^{-1})^{-1} = x \, \forall x \in X$. Let $A = X \cup X^{-1}$. A word in A^* is called reduced if it contains no subwords xx^{-1} or $x^{-1}x$ for $x \in X$. For example $X = \{x, y\}$, $xy^{-1}xxy^{-1}y^{-1}$ is reduced while $xyx^{-1}xyyx^{-1}$ is not

Proposition 2.4. *For any set X, there exists a free group F_X on X.*

Proof. Define X^{-1} and A as above, and let F_X be the set of reduced words in A^*. Make F_X into as follows. Let $\alpha, \beta \in F_X$. We define $\alpha\beta$ by concatenating α and β and deleting any strings xx^{-1} or $x^{-1}x$ in the middle.

Example. $X = \{x, y\}$, $\alpha = xyx^{-1}$ and $\beta = x^{-1}$, then $\alpha\beta = xyx^{-2}$. If $\alpha = xyx^{-1}y^{-1}$ and $\beta = yxyx$ then $\alpha\beta = xy^2x$. If $\alpha = xyx$, $\beta = x^{-1}y^{-1}x^{-1}$ then $\alpha\beta = \epsilon$.

We have an identity element, ϵ. Given α, we get α^{-1} by reversing α and replacing every letter by its inverse. Finally we need to show associativity. Let $\alpha, \beta, \gamma \in F_X$, let $l = |\alpha|, m = |\beta|$ and $n = |\gamma|$. When we multiply $\alpha\beta$, let r be the length of suffix of α that is cancelled (could have $r = 0$). Loose prefix of length r from β, so $|\alpha\beta| = l + m - 2r$. Similarly let s be the length of suffix of β that is cancelled in $\beta\gamma$, then $|\beta\gamma| = m + n - 2s$.

Case 1. $r + s \leq m$. The cancelled prefix and suffix of β do not intersect. Let $\alpha = ab^{-1}, \beta = bcd$ and $\gamma = d^{-1}e$, where b, d are cancelled parts and $|c| \geq 0$. So $\alpha\beta = acd$, $(\alpha\beta)\gamma = ace$. On the other hand $\beta\gamma = bce$, $\alpha(\beta\gamma) = ace$ as required.

Case 2. $r+s > m$. Then the cancelled prefix and suffix of β overlaps. So let $\alpha = ac^{-1}b^{-1}, \beta = bcd$ and $\gamma = d^{-1}c^{-1}e$. Then we have $\alpha\beta = ad$ and $(\alpha\beta)\gamma = ac^{-1}e$. On the other hand $\beta\gamma = be$, $\alpha(\beta\gamma) = ac^{-1}e$ as required.

So F_X is a group. We now show that it is free on X. Let G be any group and $\theta : X \to F$ a map. If $\theta' : F_X \to G$ is a homomorphism extending θ we must have $\theta'(x^{-1}) = \theta'(x)^{-1} \, \forall x \in X$. We also must have $\theta'(a_1 a_» \ldots a_k) = \theta'(a_1) \ldots \theta'(a_k) \, \forall a_1 a_2 \ldots a_k \in F_X$. So θ' is unique. But defining θ' like this, does define a homomorphism $\theta' : F_X \to G$ which extends θ. Hence F_X is free on X. □

Notation. For a set X, F_X denotes F_X as defined in the above proof. (So any free group on X is isomorphic to F_X by Proposition 2.3)

Definition 2.5. Let F_X be the free group on X. Then $|X|$ is called the *rank* of F_X

Example.

1. If $X = \emptyset$, then $F_X = \{\epsilon\}$ the trivial group.
2. $X = \{x\}$, then $F_X = \{x^n | n \in \mathbb{Z}\} \cong \mathbb{Z}$, the infinite cyclic group.
3. $X = \{x, y\}$, then F_X is "big group" which is non-abelian.

Lemma 2.6. *If F is free on X and $\theta : F \to F'$ is an isomorphism, then F' is free on $\{\theta(x)|x \in X\}$.*

Proof. Easy exercise □

Proposition 2.7. *If $X \subseteq G$, then G is free on X if and only if $G = \langle X \rangle$ and all reduced words in $(X \cup X^{-1})^*$ represent distinct elements of G.*

Proof. "⇐": Define $\theta : X \to G$ by $\theta(x) = x$. This extends uniquely to $\theta' : F_X \to G$. So $G = \langle X \rangle$ implies θ' is surjective and assumption on reduced words being distinct in G, implies θ' is injective. Hence θ' is an isomorphism, so G is free on X.

"⇒" $G = \langle X \rangle$ by Proposition 2.2. By Proposition 2.3, all free groups on X are isomorphic, and there exist an isomorphism $\theta : F_X \to G$ with $\theta(x) = x \, \forall x \in X$. So elements of F_X, the reduced words in A^*, have distinct images in G.

□

Proposition 2.8. *Any group G is isomorphic to a quotient group of a free group.*

Proof. Choose any set $X \subseteq G$ with $G = \langle X \rangle$ (could even choose $X = G$). The map $\theta : X \to G$ extends to $\theta' : F_X \to G$. Now $G = \langle X \rangle$ implies θ' is surjective. So by the first isomorphism theorem $G = \text{im}(\theta') \cong F_X / \ker(\theta')$ □

Definition 2.9. $w = a_1 a_2 \dots a_n \in A^*$ is called *cyclically reduced,* if w is reduced and either $n = 0$ or $a_i \neq a_n^{-1}$.

Proposition 2.10. *If F is a free group and $1 \neq w \in F$, then $|w| = \infty$, i.e., F has no elements of finite order. (F is* torsion free*)*

Proof. Let $F = F_X$. Let $w \in F$, $w \neq 1$. Note if w is cyclically reduced then so is w^n, (since $a_n \neq a_1^{-1}$). Hence $w^n \neq \epsilon$, so $|w| = \infty$.

In general, we can write $w = a_1 a_2 \dots a_r (a_{r+1} \dots a_{l-r}) a_r^{-1} \dots a_2^{-1} a_1^{-1}$ where $l = |w|$ and $a_{r+1} \neq a_{l-r}^{-1}$. (Note the whole of w can not cancel like this or w would not be reduced). Hence $w = \beta\alpha\beta^{-1}$ with α cyclically reduced and $\alpha \neq 1$. Then $w^n = \beta\alpha^n\beta^{-1}$ with α^n cyclically reduced. Hence $w^n \neq 1$, so $|w| = \infty$. □

Proposition 2.11. *Let F be a free group. Let $a, b \in F$, then $ab = ba$ if and only if $\exists u \in F$ with $a = u^h$, $b = u^k$ for some $h, k \in \mathbb{Z}$. (i.e., a, b commutes if and only if they are powers of a common element u)*

Proof. "⇐": Clear since $ab = ba = u^{h+k}$

"⇒": Let $F = F_X$, $a = a_1 a_2 \dots a_l$ and $b = b_1 b_2 \dots b_m$. Proof by induction on $l + m$ and without loss of generality assume $l \leq m$

If $l = 0$, then $a = 1$ so $a = b^0$ and $b = b^1$, so the result is true with $u = b$.

Hence we can assume that $l > 0$. Let r be the length of the suffix of a that cancels when calculating ab. So $0 \leq r \leq l$ and $l(ab) = l + m - 2r$. Since $ba = ab$, we get $l(ba) = l + m - 2r$, so also get cancellation length r in ba.

Case 1. $r = 0$, i.e., no cancellation. Here $ab = a_1 \dots a_l b_1 \dots b_m$ and $ba = b_1 \dots b_m a_1 \dots a_l$ are equal as words. Since $l \leq m$, $a_i = b_i$ for $1 \leq i \leq l$. So a is a prefix of b, hence let $b = ac$ with $l(c) = m - l$. Since $ab = ba$, we have $b \in C_F(a) = \{x \in F | xa = ax\}$ (the *centraliser*). Also $a \in C_F(a)$, since $C_F(a)$ is a subgroup of F, so $c = a^{-1}b \in C_F(a)$. Hence $ac = ca$. Since $l(c) < l(b)$, we can use induction on ac to get $\exists u$ such that $a = u^h, c = u^k$. So $b = ac = u^{h+k}$ and we are done.

Case 2. $r = l$, i.e., the whole of a cancels is ab. So $b = a^{-1}c$ for some c with $l(c) = m - l$. Then we are back in case 1

Case 3. $0 < r < l$. In this case we have $ab = a_1 \dots a_{l-r} b_{r+1} \dots b_m = ba = b_1 \dots b_{m-r} a_{r+1} \dots a_l$. Since $r < l$, we have $a_1 = b_1$ and $b_m = a_l$. Since $0 < r$, there is some cancellation, so $a_l = b_1^{-1}$ and similarly $b_m = a_1^{-1}$. Putting all this together we get $a_l = b_1^{-1} = a_1^{-1} = b_m$, i.e., $a = a_1 \alpha a_1^{-1}$ and $b = a_1 \beta a_1^{-1}$, where $l(\alpha) = l - 2$ and $l(\beta) = m - 2$. Now $ab = ba \Rightarrow a_1\alpha a_1^{-1} a_1 \beta a_1^{-1} = a_1 \beta a_1^{-1} a_1 \alpha a_1^{-1} \Rightarrow \alpha\beta = \beta\alpha$. So by induction, we have that $\exists u$ such that $\alpha = u^h$ and $\beta = u^k$. Then $a = a_1 u^h a_1^{-1} = (a_1 u a_1^{-1})^h$ and $b = (a_1 u a_1^{-1})^k$, and we are done using $a_1 u a_1^{-1}$

□

3 Subgroups of Free Groups

The main result of this section is that subgroups of free groups are free. [Can easily deduce Proposition 2.10,2.11 directly from this]. There exists two algebraic proofs: Schreier and Nielsen. Really need them both for advance work, but in this course we will just do Schreier's proof. Both are in Johnson's Book. There also exists proofs from Algebraic Topology.

Definition 3.1. A relation $\leq$ on a set A is a *well ordering* if it is a total ordering ($\forall a, b \in A$, we have $a < b$,$b < a$ or $a = b$) such that each subset of A has a least element

Example. $\leq$ on $\mathbb{N}$ is well ordering, but $\leq$ on $\mathbb{Z}$,$\mathbb{Q}$ are not.

Axiom. *1. Any set can be well ordered*

2. Axiom of choice

3. Zorn's Lemma

These three axioms are all equivalent (using basis set theory) and independent of basic axioms of set theory. Moth mathematicians assume them.

So we will assume our set $A = X \cup X^{-1}$ has a well ordering $\leq$. (In most of our examples X will usually be finite, so this is clearly true)

Definition 3.2. We define the *lenlex ordering* of A^* as follows. Let $a = a_1 \dots a_l$ and $b = b_1 \dots b_m$ in A^*. We say $a < b$ if either:

1. $l < m$
2. $l = m$ and for some $i < l$ we have $a_j = b_j$ for $j < i$ but $a_{i+1} < b_{i+1}$.

Exercise. Check that the lenlex ordering is a well-ordering.

Note. The *lex* ordering (dictionary) is not a well ordering. $A = \{x, y\}$ and $x < y$ then $\{x^k y | k \geq 0\}$ has no least element.

Lenlex has the following properties (exercise):

$$\forall v, w \in A^*, u < w \Rightarrow \begin{cases} vx < wx \\ xv < xw \end{cases} \quad \forall x \in A \quad (\dagger)$$

Any well-ordering of A^* that satisfies (†) can be used in the following theory.

Definition 3.3. Let F be a group (not necessarily free) generated by X. Let $E \leq F$ and $A = X \cup X^{-1}$. Let $U \subset A^*$ then U is called a *Schreier transversal* of E in F if it is a right transversal and it is prefix-closed, i.e., if $a_1 a_2 \dots a_l \in U \Rightarrow a_1 a_2 \dots a_{l-1} \in U$.

So in particular $\epsilon \in U$ as representative of coset E. Also all words in$U$ must be reduced, since otherwise we would get two different words representing the same group element ($yxx^{-1} \in U \Rightarrow y \in U$)

Proposition 3.4. *Every $E \leq F$ has a Schreier transversal. We can define one by choosing well-ordering of A^* satisfying (†) and taking least element of each coset Eg as its representative in U.*

Proof. Define U as in the statement. Let $a_1 a_2 \dots a_l \in U$. If $a_1 \dots a_{l-1} \notin U$ then there exists $b_1 b_2 \dots b_m \in Ea_1 a_2 \dots a_{l-1}$ with $b_1 \dots b_m < a_1 \dots a_{l-1}$. By (†) $b_1 \dots b_m a_l < a_1 \dots a_{l-1} a_l$ contradicting $a_1 a_2 \dots a_l \in U$ □

Example. Let $F = F_X$ and $X = \{x, y\}$. Let $G = \langle g \rangle$, be cyclic of order 6. By definition of free groups $\exists \theta : F \to G$ with $\theta(x) = g^2, \theta(y) = g^3$. Now $g = \theta(yx^{-1}) \in \text{im}(\theta)$, so $\text{im}(\theta) = G$. Take $E = \ker(\theta)$. Take $F/E \cong G \Rightarrow |F \cdot E| = |G| = 6$. Note $Ea = Eb$ if and only if $ab^{-1} \in E \Rightarrow \theta(a) = \theta(b)$. So a transversal consist of 6 elements with distinct images under θ. Could choose $U = \{\epsilon, x, y, x^2, xy, x^2 y\}$, then under θ the elements are $1, g^2, g^3, g^4, g^5, g^7 = g$. This is a Schreier Transversal (but by guessing it worked). More systematically we can compute U using Proposition 3.4 by considering $\theta(a)$ for increasing a under the ordering. $A = \{x, y, x^{-1}, y^{-1}\}$ need ordering on A, say $x < x^{-1} < y < y^{-1}$. Using lenlex:

$w \in A^*$	ϵ	x	x^{-1}	y	y^{-1}	x^2	xy	xy^{-1}	x^{-2}	$x^{-1}y$
$\theta(w)$	1	g^2	g^4	g^3	g^3	g^4	g^5	g^5	g^2	g

delete repeated $\theta(w)$ and get $U = \{\epsilon, x, x^{-1}, y, xy, x^{-1}y\}$

Let U be a Schreier transversal of E in F. For $g \in E$, let $\overline{g}$ be the unique element of $Eg \cap U$. Now $\overline{g} \in Eg \Rightarrow g\overline{g}^{-1} \in E\,(*)$. For $u \in U$, $x \in F$ we have $\overline{\overline{ux}x^{-1}}$ is in the same coset as $\overline{ux}x^{-1}$ in the same coset as u. That is $\overline{\overline{ux}x^{-1}}, u \in Eu \cap U$, so $\overline{\overline{ux}x^{-1}} = u\ (**)$. Define $Z = \{ux\overline{ux}^{-1} | u \in U, x \in X\}$. By $(*)$, $Z \subseteq E$.

Lemma 3.5. *Let Z be as define above, then $Z^{-1} = S$ where $S = \{ux^{-1}\overline{ux^{-1}}^{-1} | u \in U, x \in X\}$*

Proof. Let $g \in Z^{-1}$, then $g = (ux\overline{ux}^{-1})^{-1} = \overline{ux}x^{-1}u^{-1}$. Let $v \in \overline{ux} \in U$. By $(**)$ $\overline{vx^{-1}} = u \Rightarrow g = vx^{-1}\overline{vx^{-1}}^{-1} \in S$.

Conversely, let $g = ux^{-1}\overline{ux^{-1}}^{-1} \in S$. Then $g^{-1} = \overline{ux^{-1}}xu^{-1}$, let $v = \overline{ux^{-1}}$,$\overline{vx} = u$ by $(**)$ so $g = vx\overline{vx}^{-1} \in Z^{-1}$. □

Notation. For $u, v \in A^*$, $u =_F v$ means u, v define same elements of G

Proposition 3.6. *With the above notation $E = \langle Z \rangle$*

Proof. Let $h \in E$ so $h = a_1a_2 \dots a_l$ with $a_i \in A$. Define $u_0 = \epsilon$ and $u_i = \overline{a_1a_2 \dots a_l}$ for all $1 \leq i \leq l$. So $h \in E \Rightarrow u_l = \overline{h} = \epsilon$. Now $h =_F (u_0a_1u_1^{-1})(u_1a_2u_2^{-1}) \dots (u_{l-1}a_lu_l^{-1})$ since $u_0 = u_l = \epsilon$. We have $u_{i+1} = \overline{u_ia_i}$ by definition of u_i. So $h =_F (u_0a_1\overline{u_0a_1}^{-1})(u_1a_2\overline{u_1a_2}^{-1}) \dots (u_{l-1}a_l\overline{u_{l-1}a_l}^{-1})$ which (by Lemma 3.5) is a product of elements of Z and Z^{-1}, hence in $(Z \cup Z^{-1})^*$ □

Example 3.7. With the above notation Z is called the *set of Schreier generators* of E (this depends on U)

Corollary 3.8. *A subgroup of finite index in a finitely generated group is itself finitely generated.*

Proof. X, U are finite, then $|Z| = |X| \cdot |U|$ hence Z is finite. □

Example. Going back to the previous example, we had $U = \{\epsilon, x, x^{-1}, y, xy, x^{-1}y\}$. We now calculate Z.

$\theta(u)$	u	ux	uy	$\overline{ux}$	$\overline{uy}$	$ux\overline{ux}^{-1}$	$uy\overline{uy}^{-1}$
1	ϵ	x	y	x	y	1	1
g^2	x	x^2	xy	x^{-1}	xy	x^3	1
g^4	x^{-1}	1	$x^{-1}y$	ϵ	$x^{-1}y$	1	1
g^3	y	yx	y^2	xy	ϵ	$yxy^{-1}x^{-1}$	y^2
g^5	xy	xyx	xy^2	xy^{-1}	x	$xyxy^{-1}x$	xy^2x^{-1}
g	$x^{-1}y$	$x^{-1}yx$	$x^{-1}y^2$	y	x^{-1}	$x^{-1}yxy^{-1}$	$x^{-1}y^2x$

To find $\overline{ux}$, calculate $\theta(ux)$, find elements of U with same image under θ. For example $\theta(x^2) = g^4 = \theta(x^{-1})$ and $x^{-1} \in U$. Hence we have $Z = \{1, x^3, yxy^{-1}x^{-1}, y^2, xyxy^{-1}x, xy^2x^{-1}, x^{-1}yxy^{-1}, x^{-1}y^2x\}$.

To express $xy^{-1}x^{-1}y \in E$ as word in $Z \cup Z^{-1}$. We follow the proof of Proposition 3.6.

$$\begin{pmatrix} u_0 & a_1 & \overline{u_0a_1}^{-1} = u_1^{-1} \\ \epsilon & x & x^{-1} \end{pmatrix} \begin{pmatrix} u_1 & a_2 & u_2^{-1} \\ x & y^{-1} & y^{-1}x^{-1} \end{pmatrix} \begin{pmatrix} u_2 & a_3 & u_3^{-1} \\ xy & x^{-1} & y^{-1} \end{pmatrix} \begin{pmatrix} u_3 & a_4 & u_4^{-1} \\ y & y & \epsilon^{-1} \end{pmatrix} = (xy^2x^{-1})^{-1}(yxy^{-1}x^{-1})^{-1}(y^2)$$

We don't want 1 in our generating set. So define $Y = \{ux\overline{ux}^{-1} | u \in U, x \in X, ux \neq \overline{ux}\}$. We still have $Y^{-1} = \{ux^{-1}\overline{ux^{-1}}^{-1} | u \in Y, x \in X, ux^{-1} \neq \overline{ux^{-1}}\}$ We still have $E = \langle Y \rangle$ since $Y = Z \setminus \{1\}$

For the remainder of the section let F be free on X and we assume $F = F_X$. We will prove E is free on Y.

Lemma 3.9. *Let $ua\overline{ua}^{-1} \in Y \cup Y^{-1}$ (with $u \in U$ and $a \in A = X \cup X^{-1}$). Then in the word $ua\overline{ua}^{-1}$ the letter a does not cancel.*

Proof. Let $u, \overline{ua} \in U$ so they are reduced words. Hence if a cancels, it cancels with final letter of u or the first letter of $\overline{ua}^{-1}$.

Case 1. Let $u = a_1a_2 \dots a_l$ with $a_l = a^{-1}$. Then $ua = a_1a_2 \dots a_{l-1} \in U$ by the prefix closure condition. This means $ua = \overline{ua}$ so $ua\overline{ua}^{-1} \notin Y \cup Y^{-1}$. This is a contradiction.

Case 2. Let $\overline{ua} = a_1 \dots a_l$ with $a_l = a$. Then $\overline{ua}a^{-1} = a_1 \dots a_{l-1} \in U$. But $\overline{\overline{ua}a^{-1}} = u$ by $(**)$ so $\overline{ua}a^{-1} = u \Rightarrow ua = \overline{ua}$. This is again a contradiction.

□

Remark. By this lemma the words $ua\overline{ua}^{-1}$ in Y are reduced as words, hence they are distinct, that is different u, a give different words.

Lemma 3.10. *Let $ua\overline{ua}^{-1}, vb\overline{vb}^{-1} \in Y \cup Y^{-1}$ with $ua\overline{ua}^{-1} \neq_F (vb\overline{vb}^{-1})^{-1}$. Then in product $u\underline{a}\overline{ua}^{-1}v\underline{b}\overline{vb}^{-1}$ neither of the underlined a or b cancels.*

Proof. By the last lemma, a cannot cancel with u or $\overline{ua}^{-1}$, b not with b or $\overline{vb}^{-1}$. So for one of them to cancel:

Case 1. all of $\overline{ua}^{-1}$ would cancel with v or

Case 2. all of v would cancel with $\overline{ua}^{-1}$

We deal with Case 1. as Case 2. is similar. Let $\overline{ua} = a_1 \dots a_l, v = b_1 \dots b_m$. In Case 1. we have $l \leq m$.

First suppose $l < m$, then $a_i = b_i$ for $1 \leq i \leq l$, so a cancels with b_{l+1}, i.e., $b_{l+1} = a^{-1}$. Hence $\overline{ua}a^{-1} = a_1 \dots a_l b_{l+1} = b_1 \dots b_l b_{l+1} \in U$. So $\overline{ua}a^{-1} = u \Rightarrow \overline{ua} = ua$. Contradiction as before.

Next suppose $l = m$. So $a_i = b_i$ for $1 \leq i \leq l$ and $\overline{ua} = v$, $a = b^{-1}$. So $\overline{vb} = \overline{\overline{ua}b} = \overline{\overline{ua}a^{-1}} = u$. So $(ua\overline{ua}^{-1})^{-1} = \overline{ua}a^{-1}u^{-1} = vb\overline{vb}^{-1}$, contradicting our assumption. □

Corollary 3.11. *Let $b_i = u_i a_i \overline{u_i a_i}^{-1} \in Y \cup Y^{-1}$ for $1 \leq i \leq l$, where $b_i \neq b_{i+1}^{-1}$. Then in the product $b_1 b_2 \dots b_l$ none of the letters a_i cancel, so $b_1 b_2 \dots b_l \neq_F 1$.*

Proof. This follows immediately. □

Theorem 3.12 (Main Theorem)**.** *If F is free on X, $E \leq F$ and U a Schreier transversal of E in F then E is free on the set Y as defined above. Furthermore if U, X are finite, $|U| = n$ and $|X| = r$, then $|Y| = (r-1)n+1 = nr-(n-1)$.*

Proof. We have $E = \langle Y \rangle$ by Proposition 3.6. So by Proposition 2.7 it is enough to prove distinct words in $Y \cup Y^{-1}$ define distinct elements of E.

So let $b_1 \dots b_l, c_1 \dots c_m$ be distinct reduced words in $Y \cup Y^{-1}$ with $b_1 \dots b_l =_F c_1 \dots c_m$. So $b_1 \dots b_l c_m^{-1} \dots c_1^{-1} =_F 1$ by Corollary 3.11. Now $b_1 \dots b_l c_m^{-1} \dots c_1^{-1}$ is now reduced in $Y \cup Y^{-1}$. So we must have $b_l = c_m, b_{l-1} = c_{m-1}$ etc. and we get $b_1 \dots b_l = c_1 \dots c_m$ (as words), which is a contradiction.

Now assume $|U| = n$ and $|X| = r$ are finite. By the remark above, all elements of Y are distinct, so $|Y| = nr - t$ where t is the number of pairs (u, x) with $ux =_F \overline{ux}$. Let $v = a_1 a_2 \dots a_l \in U \setminus \{\epsilon\}$. If $a_l = x \in X$ then $u = a_1 \dots a_{l-1} \in U$ and $ux = v \in U$ so $ux = \overline{ux}$. Otherwise $a_l = x^{-1} \in X^{-1}$. Then $ux = \overline{ux}$ with $u = v$. So each $v \in U \setminus \{\epsilon\}$ gives rise to a (u, x) with $ux = \overline{ux}$. Conversely if $ux = \overline{ux}$, $ux\overline{ux}^{-1} =_F 1$ then x must cancel against end of u, or beginning of $\overline{ux}$. So this (u, x) arises in one of the two ways above.

So the pairs (u, x) with $ux = \overline{ux}$ are in 1-1 correspondence with $U \setminus \{\epsilon\}$, so $t = |U \setminus \{\epsilon\}| = n - 1$. Hence $|Y| = nr - (n-1) = n(r-1) + 1$. □

Example. Carrying on the previous example, we have $|\langle g \rangle| = 6$, so $n = 6, r = 2$ and 5 elements of Z where trivial. We had $|Y| = 7$

If $|F : E|$ infinite, is Y infinite? No, you can take $E = \{1\}$, then $|Y| = 0$. Or a slightly less trivial example, let $X = \{x, y\}$ and $E = \langle x \rangle$, then $|Y| = 1$.

Proposition 3.13. *Let F be free, $E \leq F, |F : E| = \infty$ and suppose there exists $\{1\} \neq N \lhd F$ with $N \leq E$. Then Y is infinite.*

Proof. Let U be a Schreier Transversal of E in F. Let $1 \neq w = a_1 \dots a_l \in N \leq E$. For $u \in U$, then $Euw = Euwu^{-1}u = Eu$ since $uwu^{-1} \in N \leq E$. So $\overline{uw} = u \neq uw$ (as words), so $uw \notin U$. Choose the least k such that $ua_1 \dots a_k \notin U$. So $ua_1 \dots a_{k-1} \in U$ but $ua_1 \dots a_k \notin U$, since U is infinite there exists $1 \leq k \leq l$ and an infinite subset $V \subseteq U$ with $ua_1 \dots a_{k-1} \in U$ and $ua_1 \dots a_k \notin U$ for all $u \in V$. So let $u_k = ua_1 \dots a_{k-1}$for $u \in U$ then $\{u_k a_k \overline{u_k a_k}^{-1} | u \in V\}$ is an infinite subset of Y. Hence Y is infinite. □

Example. Let F be free on $X = \{x, y\}$. Let $G = \langle g \rangle$ infinite cyclic. Define $\theta : F \to G$ with $\theta(x) = \theta(y) = g$.So $\text{im}(\theta) = G$ and let $E = \ker(\theta)$. So $|F : E| = |G| = \infty$. Since $\theta(x^i) = g^i$, $i \in \mathbb{Z}$, we can take $U = \{x^i | i \in \mathbb{Z}\}$. So $ux = \overline{ux} \, \forall u \in U$. Then $Y = \{uy\overline{uy}^{-1} | u \in U\} = \{x^i y x^{-i-1} | i \in \mathbb{Z}\}$.

4 Presentation of Groups

Definition 4.1. Let $G = \langle X \rangle$ and $A = X \cup X^{-1}$. A *relator* of G is a word $w \in A^*$ with $w =_G 1$.

So in free groups, relators are like $xyxx^{-1}y^{-1}x^{-1}$

Definition 4.2. A *relation* of G is an equation $w_1 =_G w_2$ with $w_1, w_2 \in A^*$.

Both definition are quite related as $w_1 = w_2$ is a relation of G if and only if $w_1 w_2^{-1}$ is a relator of G

Definition 4.3. Let $R \subseteq G$. The *normal closure* $\langle R^G \rangle$ is the intersection of all normal subgroups of G that contain R, i.e., the smallest normal subgroup of G that contains R.

Lemma 4.4. *We have* $\langle R^G \rangle = \langle g^{-1}rg | g \in G, r \in R \rangle$

Proof. Denote $H_1 := \langle R^G \rangle$ and $H_2 := \langle g^{-1}rg | g \in G, r \in R \rangle$.

We must have $g^{-1}rg \in N$ for any $N \lhd R$ with $R \subseteq N$. So $H_2 \leq H_1$.

For the other way, we have $H_1 \leq G$, $R \subset H_2$, from the definition $g^{-1}H_2 g = H_2 \, \forall g \in G$, so $H_2 \lhd G$. Hence $H_1 \leq H_2$. □

Definition 4.5. Let $F = F_X$ be the free group on a set X. Let $R \subseteq F$. Then the group defined by the *presentation* $\langle X | R \rangle$ is F/N with $N = \langle R^F \rangle$.

So elements of $G = \langle X | R \rangle$ are cosets Nw with $w \in A^*$ where $A = X \cup X$. Normally we just write $w \in G$ not wN . But note that w now has three meanings:

1. A word in A^*
2. An element of F
3. An element of G

So $w_1 =_G w_2$ means w_1, w_2 are the same elements of G, i.e., $w_1 N = w_2 N \iff w_1 w_2^{-1} \in N$. For elements $r \in R$ we have $r \in N$, so $r =_G 1$, hence elements of R are relators of G. They are called the *defining relators* of G. So we can think of $\langle X | R \rangle$ as "the largest group" generated by X in which elements of R are relators.

Example.

1. Let $X = \{x, y\}$ and $R = \{x^4, y^3, (xy)^2\}$. We write $\langle X | R \rangle$ as $\langle x, y | x^4, y^3, (xy)^2 \rangle$. We can also use equivalent relations, as $\langle x, y | x^4, y^2 = y^{-1}, xy = y^{-1}x^{-1} \rangle$.

   $w_1 = w$" is a presentation is defined to be the same as writing $w_1 w_2^{-1}$

2. $\langle X | \, \rangle = F_X$ (i.e., $R = \emptyset$)
3. $\langle x | x^k, k \neq 0 \rangle = \langle x \rangle / \langle x^k \rangle \cong C_{|k|}$. The finite cyclic group of order $|k|$.

Presentation is finite if X, R are finite

Proposition 4.6. *All groups have presentations and finite groups have finite presentations.*

Proof. Choose $X \subset G$ with $G = \langle X \rangle$. Let $F = F_X$ be the free group of X. By the definition of free group $\theta : X \to G$ with $\theta(x) = x \, \forall x \in X$ extends to $\theta' : F \to G$. So $G \cong F/N$ where $N = \ker(\theta)$. Choose $R \subset N$ such that $N = \langle R \rangle$. Then $G = \langle X | R \rangle$

If G is finite, choose X finite. Then $|F \cdot N| = |G|$ =finite so can choose R is finite by Corollary 3.8. □

Note. We have chosen R with $N = \langle R \rangle$, but we only need $N = \langle R^G \rangle$. So usually a smaller set R will work.

Fundamental Theorem of Presentation of Groups. *Let $G = \langle X | R \rangle$,$H$ any groups and $\theta : X \to H$ a map. For $x \in X$ define $\theta(x^{-1})$ by $\theta(x)^{-1}$. Suppose that for all $r = a_1 \dots a_l \in A^*$ with $r \in R$, we have $\theta(a_1)\theta(a_2)\dots\theta(a_l) =_G 1$. Then θ extends uniquely to a homomorphism $\theta' : G \to H$.*

Proof. If θ extends at all then we must have $\theta(x^{-1}) = \theta(x)^{-1}$ and $\theta(a_1 \dots a_l) = \theta(a_1) \dots \theta(a_l)$, so θ' is unique is it exists.

Let F be free on X. Then θ extends to $\psi : F \to H$. The hypothesis on θ says $\psi(r) =_G 1 \, \forall r \in R$, so $R \subset N = \ker(\psi)$. Now $N \lhd G \Rightarrow \langle R^F \rangle \leq N$, so ψ induces a well defined map $\theta' : R/\langle R^F \rangle \to H$ by $\theta'\left(\langle R^F \rangle w\right) = \psi(w)$ with $w \in F$. So θ' extends θ as required. □

Note. The hypothesis on θ is also necessary for θ to extend to $\theta' : G \to H$ (exercise)

In general we cannot say much about a group defined by $\langle X|R\rangle$.

A general approach is to manipulate the relators to get an upper bound on $|G|$ and they use the Fundamental Theorem to find an epimorphism $\theta : G \to H$ and hence prove $|G| \geq |H|$. If $|H|$ is the same as the upper bound, then θ is an isomorphism.

Example. Let $G = \langle x, y, |x^n, y^2, (xy)^2\rangle$ with $n > 1$. Consider $w \in A^*$ with $w \in G$, so w is a string in x, x^{-1}, y, y^{-1}. Now $x^n \in R$ implies $x^{-1} =_G x^{n-1}$, so we can replace any x^{-1} in w by x^{n-1}. Similarly, we have $y =_G y^{-1}$ so replace y^{-1} in y. Finally $(xy)^2 =_G 1$ means $yx =_G x^{-1}y^{-1} =_G x^{n-1}y$. So we can replace w by a word of the form $x^k y^l$ with $k, l \geq 0$. Since $x^n =_G 1$ and $y^2 =_G 1$ we can assume $0 \leq k < n$ and $0 \leq l < 2$. (This is a *normal form* for group elements). So we have that $|G| \leq 2n$.

For H we choose D_{2n}, with is the rotations and reflections of a regular n-gon. Let $g = (1, 2, \ldots, n)$ be the rotation. Let $h = (2, n)(3, n-1), \ldots$ be the reflection fixing 1. We can see that $g^n = 1$ and $h^2 = 1$. We calculate $gh = (1, n)(2, n-1)(3, n-2)\ldots$, so $(gh)^2 = 1$. Hence we can apply the Fundamental Theorem to $\theta : X \to H$ with $\theta(x) = g$, $\theta(y) = h$, hence θ extends to $\theta' : G \to H$. Since $g, h \in \mathrm{im}(\theta')$ we have $G = \langle g, h\rangle \subseteq \mathrm{im}(\theta')$, so $|G| \geq |H| = 2n$. Hence $|G| = 2n$, so θ' is an isomorphism.

Proposition 4.7. *Let $G = \langle X\rangle, H \leq G$ and $S = \cup_{i=1}^{r} Hg_i$ for some $g_i \in G$ and $g_1 = 1$. If $g_i a \in S \forall a \in A$ $(= X \cup X^{-1})$ then $G = S$. If we know $|x|$ is finite for all $x \in X$, then enough to assume $g_i a \in S \forall a \in X$.*

Proof. Let $g \in G$, we want to prove $g \in S$. Let $g = a_1 \ldots a_l, a_i \in A$. We use induction on l.

If $l = 0$ then $g = 1 = g_1 \in S$.

If $l > 0$, by induction $a_1 \ldots a_{l-1} = hg_i \in S$ for some $g \in H, g_i$. By assumption $g_i a_l = h' g_j \in S$, so $g = hg_i a_l = hh' g_j \in S$

If all $|x| \in X$ have finite order ($x^n = 1 \Rightarrow x^{-1} = x^{n-1}$) , we can write $g = a_1 a_2 \ldots a_l$ with $a_i \in X$ □

Example. $G = \langle x, y|x^3, y^3, (xy)^2\rangle$. Choose $H = \langle y\rangle$. We have $A_5 = \langle g, h\rangle$ where $g = (1, 2, 3), h = (2, 3, 4)$. Then $g^3 = 1, h^3 = 1, gh = (1, 3)(2, 4)$, hence $(gh)^2 = 1$. Define $\theta : X \to A_4$ by $\theta(x) = g$ and $\theta(y) = h$. Then by the Fundamental Theorem θ extends to $\theta : G \to A_4$. Now $\mathrm{im}(\theta) = A_4$, so $|G| \geq 12$.

We want to prove $|G| = 12$, i.e., $|G \cdot H| = 4$. So we want to find g_1, g_2, g_3, g_4 so we can apply the last proposition. Since $\theta(H) = \langle(2, 3, 4)\rangle = \mathrm{Stab}_{A_4}(1)$, $\theta(g_i)$ should be coset representations of $\theta(H)$ in A_4. We want $1^{\theta(g_i)} = i$, so we can choose $g_1 = 1, g_2 = x, g_3 = xy, g_4 = xy^{-1}$. So by Proposition 4.7, we want to prove that $g_i x, g_i y \in S, 1 \leq i \leq 4$ where $S = \cup_{i=1}^{4} Hg_i$. (Since $x^3 = y^3 = 1 \Rightarrow |x|, |y|$ are finite)

g_i	x	y
1	$x \in Hx$	$y \in H$
x	$x^2 = x^{-1} = yxy \in Hxy$	$x \in Hxy$
xy	$xyx = y^{-1} \in H$	$xy^2 = xy^{-1} \in Hxy^{-1}$
xy^{-1}	$xy^{-1}x = y^{-1}xy^{-1} \in Hxy^{-1}$	$x \in Hx$

So $G = S, |G| \leq 12$, hence θ is an isomorphism and $G \cong A_4$

Definition 4.8. The *commutator* is $[x : y] := x^{-1}y^{-1}xy$.

Note that $[x, y] = 1$ if and only if $xy = yx$.

Example. Let $G = \langle x, y|[x, y]\rangle = \langle x, y|xy = yx\rangle$. Using $yx = xy$, we get $G = \{x^a y^b | a, b \in \mathbb{Z}\}$. We want to show that $x^a y^b$ is normal form for group elements.

Let $H \cong \mathbb{Z}^2$, free abelian of rank 2. Then $H = \langle g\rangle \times \langle h\rangle = \{g^a h^b | a, b \in \mathbb{Z}\}$ (Writing H multiplicatively). Define $\theta : X \to H$ by $\theta(x) = g, \theta(y) = h$. Now $\theta([x, y]) = g^{-1}h^{-1}gh = 1_H$ so θ extends to $\theta : G \to H$. And $\theta(x^a y^b) = g^a h^b$. Clearly bijection so isomorphism.

Similarly $\langle x_1, \ldots, x_n|[x_i, x_j] : 1 \leq i < j \leq n\rangle \cong \mathbb{Z}^n$ (free abelian of rank n). This is also an example of a direct product, for which we have:

Proposition 4.9. *Let $G = \langle X|R\rangle, H = \langle Y|S\rangle$ with $X \cap Y = \emptyset$. Let $[X, Y] = \{[x, y]|x \in X, y \in Y\}$ and $T = R \cup S \cup [X, Y]$. Define $F = \langle X \cup Y|T\rangle$, then $G \times H \cong K$.*

Proof. $G \times H = \{(g, h)|g \in G, h \in H\}$. Define $\theta : X \cup Y \to G \times H, \theta(x) = (x, 1), \theta(y) = (1, y)$. Let $r = a_1 \ldots a_l \in T, a_i \in X \cup Y \cup X^{-1} \cup Y^{-1}$. If $r \in R, r =_G 1$ so $\theta(r) = (1, 1) = 1_{G \times H}$. Similarly $r \in S \Rightarrow \theta(r) = 1_{G \times H}$. If $r \in [x, y]$ then $r = x^{-1}y^{-1}xy, x \in X, y \in Y$ so we get $\theta(r) = (x^{-1}, 1)(1, y^{-1})(x, 1)(1, y) = (1, 1) = 1_{G \times H}$. So θ extends to $\theta : K \to G \times H$ by the Fundamental Theorem. It is clearly surjective since $G = \langle X\rangle$ and $H = \langle Y\rangle$.

It remains to show that $\ker(\theta) = 1$. Since $xy = yx, x \in X, y \in Y$ any element of K can be written as $g = a_1 a_2 \dots a_l b_1 \dots b_m$ with $a_i \in X \cup X^{-1}$ and $b_i \in Y \cup Y^{-1}$. Then $\theta(g) = (a_1 \dots a_l, b_1 \dots b_m)$, so if $g \in \ker(\theta)$ then $a_1 \dots a_l =_G 1$ and $b_1 \dots b_m =_H 1$. So $a_1 \dots a_l \in \langle R^{F_X} \rangle$ and $b_1 \dots b_m \in \langle S^{F_Y} \rangle$, so both in $\langle T^{F_{X \cup Y}} \rangle$. So $a_1 \dots a_l =_K 1 =_K b_1 \dots b_m$. □

4.1 Tietze Transformations

Some group presentation "clearly" isomorphic. For example $\langle x, y | x^3, y^2 \rangle \cong \langle x, y | x^2, y^3 \rangle$ or $\langle x, y | y = x^2 \rangle \cong \langle x | \, \rangle$ (eliminate y). Tietze Transformation enable us to justify such manipulations of presentations.

Proposition 4.10. *Let $G = \langle X | R \rangle = F/N$, $N - \langle R^F \rangle$.*

1. *If r is a relator of G, then $G = \langle X | R \cup \{r\} \rangle$*
2. *If $y \notin X$ and $w \in A^*$. Then $G \cong \langle X \cup \{y\} | R \cup \{yw^{-1}\} \rangle =: G'$, where the isomorphism induces the identity on X.*

Proof. 1. If $r \in R$ then $r \in N$ by definition. So $N = \langle (R \cup \{r\})^F \rangle$.

2. Since relators in R are also relators of G'. So the map $\theta : X \to G'$ with $\theta(x) = x \, \forall x \in X$ satisfies $\theta(r) =_{G'} 1$, so by the Fundamental Theorem it extends to $\theta : G \to G'$.

   Define $\theta' : X \cup \{y\} \to G$ by $\theta'(x) = x \, \forall x \in X$ and $\theta'(y) = w$. Again $\theta'(r) =_G 1 \, \forall r \in R$, and $\theta'(yw^{-1}) = \theta'(y)\theta'(w^{-1}) = ww^{-1} = 1$ (since w is a word in $(X \cup X^{-1})^*$). So θ' maps relators of G' to 1_G. So by the Fundamental Theorem, it extends to $\theta' : G' \to G$.

   Look at $\theta \circ \theta' : G' \to G'$ and $\theta' \circ \theta : G \to G$, they both extends the identity map on generators $X \cup \{y\}$ of G' and X of G. So by the uniqueness of the Fundamental Theorem, we have $\theta\theta' = I_{G'}$ and $\theta'\theta = 1_G$, so both isomorphism.

□

Definition 4.11. We define four types of *Tietze Transformation* on $G = \langle X | R \rangle$.

R^+: Add a relator: If r is a relator of G, then replace $\langle X | R \rangle$ by $\langle X | R \cup \{x\} \rangle$

R^-: Remove a relator: If there exists $r \in R$ with $\langle (R \setminus \{r\})^G \rangle = \langle R^G \rangle$, then replace $\langle X | R \rangle$ by $\langle X | R \setminus \{r\} \rangle$

X^+: Add a new generator: For any $w \in A^*$, replace $\langle X | R \rangle$ by isomorphic group $\langle X \cup \{y\} | R \cup \{yw^{-1}\} \rangle$

X^-: Remove a generator: If there exists $r \in R$ with $r = yw^{-1}$ for some $y \in X$, such that w and all other $s \in R$ do not contain y or y^{-1}, then replace G by isomorphic group $\langle X \setminus \{y\} | R \setminus \{r\} \rangle$.

This is mostly used in combinators.

Example. 1. Let $G = \langle X | R \cup \{r\} \rangle$, $r =_G 1 \Rightarrow r^{-1} =_G 1$, so $G = \langle X | R \cup \{r, r^{-1}\} \rangle$. Now $r^{-1} =_G 1 \Rightarrow r =_G 1$, so r is redundant, so $G = \langle X | R \cup \{r^{-1}\} \rangle$. Call these two moves $R^\pm$, replace relator by inverse

2. Similarly $G = \langle X | R \cup \{r\} \rangle = \langle X | R \cup \{g^{-1}rg\} \rangle$ (since $r = 1 \iff g^{-1}rg = 1$ for any $g \in G$). So can use $R^\pm$ to replace relator by conjugate. Often used for cyclic conjugates. Replace xyz by yzx $(= x^{-1}(xyz)x)$.

Whenever some generator y or y^{-1} appears just once in some relators we can use $R^\pm$ followed by X^- to eliminate y.

3. $\langle x, y, z | (xz)^2, (yz)^3, xyz \rangle$. We can eliminate x using $xyz =_G 1 \iff z = y^{-1}x^{-1}$. Since $x = y^{-1}x^{-1}$ we have

$$\begin{aligned} (xz)^2 = 1 &\iff (xy^{-1}x^{-1})^2 = 1 \\ &\iff xy^{-2}x^{-1} = 1 \\ &\iff y^{-2} = 1 \\ &\iff y^2 = 1 \end{aligned}$$

   so can use $R^\pm$ to replace $(xy)^2 = 1$ by $y^2 = 1$. Similarly $(yz)^3 = 1 \iff x^3 = 1$. So by $R^\pm$, $G = \langle x, y, z | y^2, x^3, zxy \rangle$. Now we can apply X^- to get $G \cong \langle x, y | y^2, x^3 \rangle$

Examples 1-3 were to illustrate how we use them, while Examples 4-5 are "real" examples.

4. $\langle x,y|xyxyx = yxyxy\rangle \cong_{X^+} \langle x,y,a|xyxyx = yxyxy, a = xy\rangle$, now use RX^- to eliminate $y = x^{-1}a$, so $G \cong \langle x,a|aax = x^{-1}aaa\rangle \cong_{X^+} \langle x,a,b|a^2x = x^{-1}a^3, b = xa^2\rangle$. Eliminate $x = a^2b$ by RX^-, $G \cong \langle a,b|b = b^{-1}a^5\rangle = \langle a,b|b^2 = a^5\rangle$.

5. For $l,m,n \geq 1$ define $D = (l,m,n) = \langle x,y|x^l, y^m, (xy)^n\rangle$. (Note that $D(2,2,n) \cong D_{2n}$). Introduce $a = xy$ then eliminate $y = x^{-1}a$. So $\langle x,a|x^l, x^{-1}a)^m, a^n\rangle$. Replace x by x^{-1}, $\langle x,a|x^{-l}, (xa)^m, a^n\rangle = \langle x,a|x^l, (xa)^m, a^n\rangle = D(l,n,m)$. We also have clearly $D(l,m,n) \cong D(m,l,n)$. Since (m,n) and (l,m) generate S_3 on $\{l,m,n\}$ we have $D(l,m,n) \cong D(l',m',n')$. For any permutation l',m',n' of l,m,n (so we can assume $l \leq m \leq n$ if we want to)

The following is a basic result, but less useful than it appears because is practice we might not know in whether $\langle X|R\rangle \cong \langle Y|S\rangle$. Therefore it does not enable you to decide this. (It has been proved to be undecidable in general)

Proposition 4.12. *Let $G \cong \langle X|R\rangle \cong \langle Y|S\rangle$. Then by using a sequence of Tietze Transformation, we can transform $\langle X|R\rangle$ to $\langle Y|S\rangle$.*

Proof. $G \cong \langle X|R\rangle \cong \langle Y|S\rangle$. Think of X,Y as subsets of G. Think of X,Y as subsets of G. $G = \langle X\rangle = \langle Y\rangle$, so elements of X can be written as words in $(Y \cup Y^{-1})^*$ and vice versa. Write this $X = X(Y)$, $Y = Y(X)$, so

$$\begin{aligned}
G &\cong \langle X|R\rangle \\
&\cong \langle X|R(X)\rangle \\
&\cong \langle X \cup Y|R(X) \cup \{Y = Y(X)\}\rangle && X^+ \\
&= \langle X \cup Y|R(X) \cup \{Y = Y(X)\} \cup \{X = X(Y)\}\rangle && R^+ \\
&= \langle Y|R(X(Y)) \cup \{Y = Y(X(Y))\rangle && R^- \\
&= \langle Y|R(X(Y)) \cup \{Y = Y(X(Y)) \cup S(Y)\rangle && R^+ \\
&= \langle Y|S(Y)\rangle && R^-
\end{aligned}$$

□

4.1.1 A Presentation of symmetric group S_n

S_n acts on $\{1,2,\ldots,n\}$. Let $\tau_i = (i,i+1)$, for $1 \leq i \leq n-1$.

Lemma 4.13. $S_n = \langle \tau_i|1 \leq i \leq n-1\rangle$

Proof. Well known. □

What are the relations? We have $\tau_i^2 = 1$, $(\tau_i\tau_{i+1})^3 = 1$ since $\tau_i\tau_{i+1} = (i,i+2,i+1)$. Finally $(\tau_i\tau_j)^2 = 1$ (equivalently $\tau_i\tau_j = \tau_j\tau_i$) for $|i-j| > 1$.

Proposition 4.14. *Let $G_n = \langle X|R_1 \cup R_2 \cup R_3\rangle$ with $X = \{x_1,\ldots,x_{n-1}\}$, $R_1 = \{x_i^2|1 \leq i \leq n-1\}$, $R_2 = \{(x_ix_{i+1})^3|1 \leq i \leq n-2\}$ and $R_3 = \{(x_ix_j)^2|i < j, |j-i| > i\}$. Then $G_n \cong S_n$ with isomorphism $x_i \mapsto \tau_i$.*

Note. This presentation defines a Coxeter group. The general definition of them is $\langle x_1,\ldots,x_n|x_i^2, (x_ix_j)^{m_{ij}}\ 1 \leq i,j \leq n\rangle$ with $2 \leq m_{ij} \leq \infty$ and $m_{ij} = m_{ji}$. The case $n = 2$ gives the Dihedral groups. They include Weyl groups studied in Lie Algebra. Out presentation for S_n is the Weyl group of type A_{n-1}.

Proof. By the Fundamental Theorem, the map $\theta : x_i \mapsto \tau_i$ extends to a homomorphism $\theta : G_n \to S_n$, which is surjective by Lemma 4.13. So $|G_n| \geq |S_n| = n!$.

Claim: $|G_n| \leq n!$ (which will prove the result)

Use induction on n. The case $n = 1,2,3$ there is nothing to show ($S_3 = \langle x_1,x_2|x_1^2, x_2^2, (x_1x_2)^3\rangle$)

So assume $|G_n| \leq n!$ and we prove $|G_{n+1}| \leq (n+1)!$. Define $H = \langle x_2,\ldots,x_n\rangle \leq G_{n+1}$. The relations of G_n are satisfied by generators of H (renumbered as $x_1,\ldots,x_{n-1}$). So by induction $|H| \leq n!$. So we want to prove $|G : H| \leq n+1$. Note $\theta(H) \cong S_n$ is the stabilizer of 1 in S_{n+1}. So if $g_0,\ldots,g_n$ coset representation we want $\theta(g_0),\ldots,\theta(g_n)$ to map 1 to $1,2,\ldots,n+1$. So we can choose $\theta(g_i) = \tau_1\ldots\tau_i$. Then $1^{\theta(g_i)} = 1^{\tau_1\cdots\tau_i} = i+1$. Define $g_0 = 1, g_i = x_2x_2\ldots x_i$, i.e., $g_i = g_{i-1}x_i$. Define $S = \cup_{i=0}^n Hg_i$, it is enough to prove $S = G_n$. So by Proposition 4.7 it is enough to prove $g_ix_j \in S\ \forall i,j$.

Case 1. $j > i+1$: We have $g_ix_j = x_jg_i \in S$ (since $x_j \in H$)

Case 2. $j = i+1$: We have $g_i x_j = g_{i+1} \in S$

Case 3. $j = i$: We have $g_i x_j = g_{i-1} \in S$

Case 4. $j < i$: We prove by induction of $i-j$ that $g_i x_j = x_{j+1} g_i \in S$ since $x_{j+1} \in H$.

Base case: $i-j = 1$ ($\Rightarrow i \geq 2$ and $j = i-1$). $g_i x_j = g_{i-2} x_{i-1} x_i x_{i-1} = g_{i-2} x_i x_{i-1} x_i = x_i g_{i-2} x_{i-1} x_i = x_i g_i = x_{j+1} g_i$

Induction step: If $i = j > 1$ then $g_i x_j = g_{i-1} x_i x_j = g_{i-1} x_j x_i = x_{j+1} g_{i-1} x_i = x_{j+1} g_i$.

So we can apply Proposition 4.7 to get $G_n = S$, hence $|G_n : H| \leq n+1$, hence $|G_n| \leq (n+1)!$. Hence θ is an isomorphism. □

4.1.2 Presentation of $(\mathbb{Q}, +)$

The group $(\mathbb{Q}, +)$ is not finitely generated. To see this let $H = \langle g_1, \ldots, g_k \rangle \leq (\mathbb{Q}, +)$. Then $g_i = m_i/n_i$ for some $m_i \in \mathbb{Z}$ and $n_i \in \mathbb{Z}_{>0}$. Elements of H all have denominators at most $\text{lcm}(n_1, \ldots, n_k)$, so we cannot have $H = \mathbb{Q}$.

$(\mathbb{Q}, +)$ is generated by $\{1/n | n \in \mathbb{Z}_{>0}\}$. Also by $\{1/n! | n \in \mathbb{Z}_{>0}\}$ since any $m/n = m(n+1)! \cdot 1/n!$

Proposition 4.15. *Let* $G = \langle x_i (i \geq \mathbb{Z}_{>0}) | x_n^n = x_{n-1} (n > 1) \rangle$. *Then* $G \cong (\mathbb{Q}, +)$.

Note. G is multiplicative, while $(\mathbb{Q}, +)$ is additive

Proof. Define $\theta : \{x_i\} \to (\mathbb{Q}, +)$ by $\theta(x_n) = 1/n!$. Then $\theta(x_n^n) = n/n! = 1/(n-1)! = \theta(x_{n-1})$. So by the Fundamental Theorem θ extends to homomorphism $\theta : G \to (\mathbb{Q}, +)$ surjective, since $(\mathbb{Q}, +)$ is generated by $\{1/n!\}$.

We now need to prove $\ker(\theta) = 1$. Given generators x_n, x_m of G, $n > m$, x_m is some power of x_n so $x_n x_m = x_m x_n$. So G is abelian. So for each $g \in G$, $g = x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$ for some n, $k_i \in \mathbb{Z}$. Since $x_j^j = x_{j-1}$ for $j > 1$, we can replace x_j^j by x_{j-1} and assume $0 \leq k_j < j$ for $j > 1$ (and $k_n \neq 0$). So $\theta(g) = \frac{k_1}{1!} + \frac{k_2}{2!} + \cdots + \frac{k_n}{n!}$. Suppose $1 \neq g$, $\theta(g) = 0$ (i.e., $g \in \ker(\theta)$). So $(n-1)! \left(\frac{k_1}{1!} + \cdots + \frac{k_{n-1}}{(n-1)!} \right) + \frac{k_n}{n} = m + \frac{k_n}{n} = 0$ with $m \in \mathbb{Z}$. But if $n > 1$ then $\frac{k_n}{n} \notin \mathbb{Z}$ since $0 < k_n < n$. Hence $n = 1$, $g = x_1^{k_1}$ and $\theta(g) = k_1 = 0$, hence $k_1 = 0$ and $g = 1$. So $\ker(\theta) = 1$. □

4.2 Groups Acting on Sets (Recap)

We use right actions (as opposed to left actions used in Algebra II) For a group G acting on a set Ω, we write α^g for the action of $g \in G$ on $\alpha \in \Omega$. We can define an *action* as a homomorphism $\phi : G \to \text{Sym}(\Omega)$ where we write α^g instead of $\alpha^{\phi(g)}$ [$\phi(g)$ is a permutation of Ω]. The *kernel* of the action $\ker(\phi) = \{g \in G | \alpha^g = \alpha \, \forall \alpha \in \Omega\}$. An action is *faithful* if $\ker(\phi)$ is trivial, if and only if, $G \cong \text{im}(\phi)$. An action is *transitive* if there is a single orbit Ω, i.e., $\forall \alpha, \beta \in \Omega \, \exists g \in G, \alpha^g = \beta$.

Two actions $\phi_1 : G \to \text{Sym}(\Omega_1)$, $\phi_2 : G \to \text{Sym}(\Omega_2)$ are *equivalent* if there exist a bijection $\tau : \Omega_1 \to \Omega_2$ with $\tau(\alpha)^{\phi_2(g)} = \tau(\alpha^{\phi_1(g)}) \, \forall g \in G$.

4.2.1 Coset Actions

Let G be a group, $H \leq G$. Let $\Omega = \{Hg : g \in G\}$ be the set of distinct right cosets. Define action of G on Ω by $(Hk)^g = Hkg$ (by right multiplication). So $\phi(g) : Hk \to Hkg$, check that $\phi(g), \phi(g^{-1})$ are inverses of each other, so $\phi(g) \in \text{Sym}(\Omega)$. It is clearly a homomorphism. Given $Hk_1, Hk_2 \in \Omega$, $(Hk_1)^{k_1^{-1} k_2} = Hk_2$, so it is transitive.

$$\begin{aligned}
\ker(\phi) &= \{g \in G | Hkg = Hk \, \forall k \in G\} \\
&= \{g \in G | kgk^{-1} \in H \, \forall k \in G\} \\
&= \{g \in G | g \in k^{-1} H k \, \forall k \in G\} \\
&= \bigcap_{k \in G} k^{-1} H k \\
&=: \text{Core}_G(H)
\end{aligned}$$

This is the largest normal subgroups of G contained in H. So $H \lhd G$ if and only if $\text{Core}_G(H) = H$. So we have $1 \leq \text{Core}_G(H) \leq H \leq \langle H^G \rangle \leq G$.

5 Coset Enumeration

(Todd - Coxeter 1930)

Given $G = \langle X|R\rangle$ with X, R finite and a finite set $Y \leq A^*$ generating $H = \langle Y\rangle \leq G$. Algorithm tries to prove that $|G : H|$ is finite and to compute the coset action of G on the cosets of H. If $|G : H|$ is infinite it will not halt. If $|G : H|$ is finite it will succeed but there is no upper bound on the time taken.

At any time we have a set Ω of positive, $\Omega = \{i_1, \ldots, i_n\}$ positive integers, which represent cosets of H in G. For each $i \in \Omega$ there is an associated word. For each $i \in \Omega$ there is an associated word $a_i \in A^*, i \hookrightarrow a_i$ where $i =$coset Ha_i. We always have $1 \in \Omega, a_1 = \epsilon$. So $1 = H$. For $i \in \Omega, x \in A$, ix may or may not be defined. If defined $ix \in \Omega$. (This means $Ha_ix = Ha_j$ where $ix = j$). We always have

(R1) $ix = j$ if and only if $jx^{-1} = i$.

For $a_1a_2 \ldots a_k \in A^*$ with $a_i \in A$. Then $ia_1 \ldots a_k$ is defined recursively to $(ia_1 \ldots a_{k-1})a_k$ provided everything is defined.

(R2) If $i \hookrightarrow a_i \in A^*$ with $i \in \Omega$, then $1a_i$ is defined and equal to i.

If some ix is not define then we can choose some $j > 0$, $j \notin \Omega$, replace Ω by $\Omega \cup \{j\}$ and define $ix = j, jx^{-1} = i$ and put $a_j = a_ix$. So $1a_j = 1a_ix = ix = j$ (Hence (R1) and (R2) remains true)

Example. Let $G = \langle x, y|x^2, y^3, xyxy\rangle$ and $H = \langle xy\rangle$.

1. We have a table with a single row for each generator of H. For each relator, we have a table with one row for each elements of Ω. Start with $\Omega = \{1\}$, $1 \hookrightarrow \epsilon$.

	x	y	
1			1

	x	x			y	y	y			x	y	x	y	
1			1	1				1	1					1

2. Now let $\Omega = \{1, 2\}$, $2 \hookrightarrow x$. Definitions $1x = 2$, so $2x^{-1} = 1, 1y^{-1} = 2$: Deduction $2y = 1$ (from (a)), $2x = 1$ (from (b)). (Note _ stands for, by definition)

	x		y	
1	_	2	$\overset{(a)}{=}$	1

	x		x			y		y		y			x		y		x		y	
1		2	$\overset{(b)}{=}$	1	1				2		1	1		2		1		2		1
2		1		2	2		1				2	2		1						2

3. Now let $\Omega = \{1, 2, 3\}$, $3 \hookrightarrow y$. Definitions: $1y = 3$. Deductions: $3y = 2$ (from (c)), $3x = 3$ (from (d))

	x		y	
1	_	2	$\overset{(a)}{=}$	1

	x		x			y		y		y			x		y		x		y	
1		2	$\overset{(b)}{=}$	1	1	_	3	$\overset{(c)}{=}$	2		1	1		2		1		2		1
2		1		2	2		1		3		2	2		1		3	$\overset{(d)}{=}$	3		2
3		3		3	3		2		1		3	,3		3		2		1		3

Now ix is defined for all $i \in \Omega, x \in A$. All tables are complete, so the process stops. We have shown that $|G : H| = 3$ with $i \to ia$ defining an action ϕ of G on $\Omega = \{1, 2, 3\}$ defined by $x \mapsto (1, 2), y \mapsto (1, 2, 3)$. So $\phi : G \to \mathrm{Sym}(\Omega)$, hence $\mathrm{im}(\phi) = \mathrm{Sym}(3)$ which has order 6. Hence $|G| \geq 6$. Since $H = \langle xy\rangle$ with $(xy)^2 = 1$, $|H| \leq 2$, hence $|G| \leq 6$. So $|G| = 6$.

We also have:

(R3) If $i \hookrightarrow a_i, j \hookrightarrow a_j$ and $ix = j$, then $Ha_ix = Ha_j$.

Proposition 5.1. (R3) *remains true during the process.*

Proof. $ix = j$ is either a definition or a deduction. If it is a definition then we define a_j so to make (R3) true.
So assume that it is a deduction, from a row of a table.

Case 1. From row k of relator $x_1x_2 \dots x_s$, of a relator table.

	x_1	x_2	$\dots$		$x_t = x$		x_{t+1}	$\dots$	x_s	
k	all defined			i	deduced $ix = j$	j	all defined			k

Can assume by induction that (R3) is true for all equations $i'x' = j'$ that are know prior to deduction of $ix = j$. So $Ha_kx_1 \dots x_{t-1} = Ha_i$, $Ha_jx_{t+1} \dots x_s = Ha_k$. So $Ha_jx_{t+1} \dots x_sx_1 \dots x_{t-1} = Ha_i$. Hence $x_1x_2 \dots x_s =_G 1 \Rightarrow x_{t+1} \dots x_sx_1 \dots x_t =_G 1$ (cyclic conjugate). So $Ha_ix_t = Ha_jx_{t+1} \dots x_sx_1 \dots x_{t+1}x_t = Ha_j$

Case 2. Deduction as above, but from a subgroup generator, so $k = 1$ and $Ha_k = H$. So $x_1x_2 \dots x_s \in H$. As $Ha_ix = Ha_ix_t = Hx_1 \dots x_t = Hx_s^{-1} \dots x_{t-1}^{-1} = Ha_j$, since $x_1x_2 \dots x_s \in H$.

□

Theorem 5.2. *Suppose the process terminates with $H = \langle y \rangle \leq G = \langle X|R \rangle$. Then there exists an action θ of G on Ω with $i^{\theta(x)} = i^x = j \iff ix = j$. The action is equivalent to the coset action of G on right cosets of Ω, with equivalence τ, with $\tau(i) = Ha_i$ for $i \in \Omega$.*

Proof. When the process stops all ix are defined. So define $\theta(x) : \Omega \to \Omega$ by $i^{\theta(x)} = i^x = ix$, for all $x \in A$. Now (R1) says $\theta(x^{-1}) = \theta(x)^{-1}$, so $\theta \in \mathrm{Sym}(\Omega)$. The fact that relators tables are full, says exactly that all relators r, $k^{\theta(r)} = k$ for all $k \in \Omega$. So $\theta(r)$ is the identity and by the Fundamental Theorem θ extends to $\theta : G \to \mathrm{Sym}(\Omega)$.

Define $\tau : \Omega \to \{Hg : g \in \Omega\}$ by $\tau(i) = Ha_i$ with $i \hookrightarrow a_i$. Show that τ is a surjection. Let $g \in G$, $g = x_1x_2 \dots x_s$ (for $x_i \in A$) . $1g = 1x_1x_2 \dots x_s = i$ for some $i \in \Omega$. By (R3), $Hg = Ha_i$ with $i \hookrightarrow a_i$. So $\tau(i) = Ha_i = Hg$. So τ is indeed a surjection.

Now to show τ is injective. Suppose $\tau(i) = \tau(j)$ $(i \hookrightarrow a_i, j \hookrightarrow a_j)$, then $Ha_i = Ha_j$, so $a_ia_j^{-1} \in H = \langle y \rangle$. So $a_ia_j^{-1} = y_1 \dots y_t$ with $y_i \in Y \cup Y^{-1}$. From subgroups tables, $1y_i = 1 \forall y_i \in Y$. So $1a_ia_j^{-1} = 1^{y_1 \dots y_t} = 1$. Hence $1a_i = 1a_j \Rightarrow i = j$ by (R2).

By (R3) we have $ix = j \Rightarrow Ha_ix = Ha_j$. Hence τ is an equivalence of actions. □

Example. Let $G = \langle x, y | x^3, y^4, (xy)^2 \rangle = D(3, 4, 2)$ and $H = \langle y \rangle$, so we have one subgroup table

	y	
1	=	1

- Starting table of the relators:

	x		x		x			y		y		y		y			x		y		x		y	
1						1	1		1		1		1		1	1							1		1

- Next define $1x = 2, 2x = 3$

	x		x		x			y		y		y		y			x		y		x		y	
1		2		3	=	1	1		1		1		1		1	1		2	=	3		1		1
2		3		1		2	2		3						2	2		3						2
3		1		2		3	3						2		3	3		1		1		2		3

- Next defined $3y = 4$ and $4y = 5$

	x		x		x			y		y		y		y			x		y		x		y	
1		2		3	=	1	1		1		1		1		1	1		2	=	3		1		1
2		3		1		2	2		3		4		5	=	2	2		3		4	=	5		2
3		1		2		3	3		4		5		2		3	3		1		1		2		3
4		5				4	4		5		2		3		4	4		5		2		3		4
5						5	5		2		3		4		5	5						4		5

- Finally define $5x = 6$

	x		x		x			y		y		y		y			x		y		x		y	
1		2		3	=	1	1		1		1		1		1	1		2	=	3		1		1
2		3		1		2	2		3		4		5	=	2	2		3		4	=	5		2
3		1		2		3	3		4		5		2		3	3		1		1		2		3
4		5		6	=	4	4		5		2		3		4	4		5		2		3		4
5		6		4		5	5		2		3		4		5	5		6	=	6		4		5
6		4		5		6	6		6		6		6		6	6		4		5		6		6

Then we are done as all tables are complete and all definition have been made. So $|G:H| = 6$ and since $|H| \leq 4$, we have $|G| \leq 24$. We can also show $|G| \geq 24$ by the map $\phi : G \to S_4$ with $\phi(x) = (1,2,3)$, $\phi(y) = (1,3,2,4)$ and $\phi(xy) = (1,4)$. This easily shows that $\mathrm{im}(\phi) = S_4$ so $|G| \geq 24$, hence $G \cong S_4$. So ϕ is an isomorphism.

Coincidences: Sometimes we get a deduction $ix = j$ where we already know $ix = k$ with $k \neq j$, (or $kx = j$ for $k \neq i$). This means coset Ha_j and Ha_k (respectively Ha_i, Ha_k) are equal. This is a deduction $j = k$.

If we find $Ha_j = Ha_k$ with $k > j$, then we replace all k by j in tables which often leads to more coincidences.

Example. Let $G = \langle a, b | a^{-1}ba = b^3, b^{-1}ab = a^3 \rangle$ and $H = \langle a \rangle$. Then the subgroup table is just

	a	
1	=	1

- Setting up the table:

	a^{-1}		b		a		b^{-1}		b^{-1}		b^{-1}			b^{-1}		a		b^{-1}		a^{-1}		a^{-1}		a^{-1}	
1		1										1	1						1		1		1		1

- Define $1b = 2$

	a^{-1}		b		a		b^{-1}		b^{-1}		b^{-1}			b^{-1}		a		b^{-1}		a^{-1}		a^{-1}		a^{-1}	
1		1		2		1				2		1	1						1		1		1		1
2												2	2		1		1		2						2

- Defined $1b^{-1} = 3$

	a^{-1}		b		a		b^{-1}		b^{-1}		b^{-1}			b^{-1}		a		b^{-1}		a^{-1}		a^{-1}		a^{-1}	
1		1		2		1		3		2		1	1		3	=	3		1		1		1		1
2												2	2		1		1		2						2
3		3		1		1		2		1		3	3		1		1		3		3		3		3

- But we are deducing $2b = 1$, but $3b = 1$, so $2 = 3$. Hence replace 3 by 2

	a^{-1}		b		a		b^{-1}		b^{-1}		b^{-1}			b^{-1}		a		b^{-1}		a^{-1}		a^{-1}		a^{-1}	
1		1		2		1		2		2		1	1		2	=	2		1		1		1		1
2		2		1		1		2		1		2	2		1		1		2		2		2		2
2		2		1		1		2		1		2	2		1		1		2		2		2		3

So we didn't need the last row, and the table is complete. We get $|G:H| = 2$ and $b \to (1,2)$ and $a \to \mathrm{id}$. Since $|H:G| = 2$ we have $H \lhd G$, so $b^2 \in H = \langle a \rangle$ .So $b^2a = ab^2$ (since H is abelian). Hence $a^{-1}ba = b^3$ implies $a^{-1}b^2a = (a^{-1}ba)^2 = b^6$, but $a^{-1}b^2a = b^2$, we get $b^6 = b^2$ and hence $b^4 = 1$. By symmetry $a^4 = 1$, so $|H| \leq 4$ and $|G| \leq 8$. To prove $|G| \geq 8$, let H be the multiplicative group of complex 2×2 matrices generated by $g = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $h = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. Then $g^2 = h^2 = -I_2$ and $g^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $h^{-1} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$. We have $g^{-1}hg = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} g = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = h^3$. Similarly $h^{-1}gh = g^3$. So we get $\phi : G \to H$, since $|g| = 4$ and $h \notin \langle g \rangle$, $|H| > 4$, since $|H|\,|\,|G|\,|8$ we must have $|G| = |H| = 8$, so θ is an isomorphism. This group is called the quaternion Q_8.

It can be shown that (R1),(R2),(R3) remain true after processing coincides. So Theorem 5.2 is still true.

We have not discussed how we can decide which new definition to make. By hand the natural choice is to fill small gaps in tables to get deductions. But this can result in long relations being ignored, which could be bad.

There are two main systematic strategies used in programming.

1. Choose the first $i \in \Omega$ for which ia is undefined for some $a \in A = X \cup X^{-1}$ and define it. This means we define cosets Ha_i is order of increasing a_i in the lenlex order.

2. Go through relator tables in order filling in gaps as you go (even for long gaps)

Generally 1. is better for work by hand. Also 2. leads to many more coincidences, but it is easier to program and runs fast on computers for routine examples. Another advantage of 2. is that each row of each relator table only needs to be scanned once while with one 1. need to keep revisiting row. It is important that strategy satisfies

(R4) $\forall i \in \Omega, a \in A$, ia will eventually be defined.

(R4) is satisfied with 1 but not necessarily with 2 so we must occasionally use 1 $(xx^{-1}, x^{-1}x)$

Theorem 5.3. *If $|G : H|$ is finite and definition strategy satisfies* (R4) *then it will eventually finish.*

Proof. With a coincidence $i = j$, $i > j$ meaning $Ha_i = Ha_j$, we eliminate larger number i, so numbers in Ω representing Ha_i can only decrease so must eventually stabilise as Ha_k. From then on, $k \in \Omega$ and remains there.

Let $\overline{\Omega}$ be the set of stabilised numbers in Ω. (Note: We have no way of knowing during the run whether a number is in $\overline{\Omega}$. It is in $\overline{\Omega}$ if it would not change any more however long the process ran.)

Case 1. $\overline{\Omega}$ is finite. Then once all elements that will ever be in $\overline{\Omega}$ are in $\overline{\Omega}$, then $\overline{\Omega}$ stabilise. By (R4) eventually all ia are defined and in $\overline{\Omega}$, with $i \in \overline{\Omega}$, $a \in X$. At that point, $\Omega = \overline{\Omega}$ (since all elements of Ω can be reached by definitions from cosets) and the process stops

Case 2. $\overline{\Omega}$ is infinite. It is still true that ia are eventually defined and in $\overline{\Omega}$ for all $i \in \Omega$, $a \in A$ by (R4). (R1), (R2), (R3) still all hold. So Theorem 5.2 applies to action of G on $\overline{\Omega}$, so this action is equivalent to coset action. Hence $|\overline{\Omega}| = |G : H|$ contradicting $|G : H|$ finite.

□

6 Presentation of Subgroups

6.1 Digression

Definition 6.1. For $g, h \in G$, we define the *commutator* of g, h to be $[g,h] = g^{-1}h^{-1}gh$.

Note that $[g,h] = 1$ if and only if $gh = hg$. Also $[g,h]^{-1} = [h,g]$.

Definition 6.2. The *Commutator Subgroup* (or *Derived Group*) of G denoted $[G,G]$ (or G') is the group $\langle [g,h] | g,h \in G \rangle$. i.e., the subgroup of G generated by the commutators.

There exists G where not all elements of $[G,G]$ are commutators, for example in the free group on a,b,c,d we have $[a,b][c,d]$ is not a commutator.

Definition 6.3. A subgroup H of G is *characteristic* if $\alpha(H) = H \, \forall \alpha \in \mathrm{Aut}(G)$ (write $H \mathrm{char} G$)

Since for $g \in G$ the map $c_g : G \to G$ defined by $h \mapsto g^{-1}hg$ is in $\mathrm{Aut}(G)$, we have $H \mathrm{char} G$ implies $H \lhd G$.

Note that for $\alpha \in \mathrm{Aut}(G)$ we have $\alpha([g,h]) = [\alpha(g), \alpha(h)]$. So α permutes the set of commutators so $[G,G] \mathrm{char} G$ (and hence $[G,G] \lhd G$).

We have $G/[G,G]$ is abelian since $g, h \in G, [g,h] \in [G,G]$ so $[g,h] = 1$ in $G/[G,G]$.

Theorem 6.4. *For any $N \lhd G$, we have G/N is abelian if and only if $[G,G] \leq N$*

Proof. We have

$$\begin{aligned} G/N \text{ abelian} &\iff [gN, hN] = 1 \, \forall g, h \in G \\ &\iff [g,h] \in N \, \forall g, h \in G \\ &\iff \langle [g,h] | g,h \in G \rangle \leq N \\ &= [G,G] \end{aligned}$$

□

Proposition 6.5. *Let $G = \langle X | R \rangle$, then $G/[G,G] \cong \langle C | R \cup C \rangle$ where $C = \{[x,y] | x,y \in X\}$*

Proof. Let F be free on X, $N = \left\langle R^F \right\rangle$, $M = \left\langle (R \cup C)^F \right\rangle$. So $G = F/N$ and let $\overline{G} = \langle X | R \cup C \rangle = F/M$. Since $[x,y] \in M \, \forall x, y \in X$ the generators of $\overline{G}$ all commutes. Hence $\overline{G}$ is abelian. By the third isomorphism theorem we have

$$\overline{G} = F/M \cong \frac{F/N}{M/N} = \frac{G}{M/N}$$

which is abelian. So $[G,G] \leq M/N$.

$M = \left\langle g^{-1}kg | g \in H, k \in R \cup C \right\rangle$ for $k \in R$ we have $g^{-1}kg \in N$. So M/N is generated by the elements $g^{-1}kgN$, $g \in F, k \in C$. So $g^{-1}kg$ is a commutator, hence so is $g^{-1}kgN$. So M/N is generated by commutators, hence $M/N \leq [G,G]$.

Hence $M/N = [G,G]$, and so $G/[G,G] \cong G/(M/N) \cong \overline{G}$. □

Example. If $G = \left\langle x, y | x^2, y^3, (xy)^6 \right\rangle$, then $G/[G,G] = \left\langle x, y | x^2, y^3, (xy)^6, [x,y] \right\rangle$.

6.2 Presentation of subgroups

Let $G = \langle X | R \rangle, H = \langle Y \rangle$ with $X, R, H, |G : H|$ all finite. We describe the Reidemeister - Schreier algorithm to derive a presentation of H. Recall, let U be a Schreier Transversal of H in G, $u \in G$, $\overline{u}$ =coset representation of u. Then $Z = \{ux\overline{ux}^{-1} | u \in U, x \in X, ux \neq \overline{ux}\}$ is a Schreier Generator of H.

This method gives presentation of H with generators of Z. A variant gives a presentation on user supplied generating set Y, this is more complicated.

The method requires coset representation of G on right cosets of H. Theorem 5.2 says that coset enumeration gives this. But in our first example, we just work it out without coset enumeration. We describe the process while working through the example.

Example. Let $G = \langle x, y | x^2, y^3, (xy)^6 \rangle = D(2,3,6)$. We define $H = [G : G]$, as we saw before $G/H \cong \langle x, y | x^2, y^3, (xy)^6, xy = yx \rangle$. In G/H the relator $(xy)^6 = x^6 y^6 = 1$, so it is redundant. So $G/H \cong \langle x, y | x^2, y^3, xy = yx \rangle \cong C_2 \times C_3 \cong C_6$ by Proposition 4.10. So $|G : H| = 6$. We will write down the coset action (which is the same as the regular representation of G/H).

Define $\theta : G \to \mathrm{Sym}(6)$ by $\theta(x) = (1,2)(3,4)(5,6)$ and $\theta(y) = (1,3,5)(2,4,6)$. [We use the facts that $|\theta(x)| = 2, |\theta(y)| = 3, \theta(x)$ and $\theta(y)$ fixes no points, and the fact $\theta(x)\theta(y) = \theta(y)\theta(x)$] So we have $\theta(xy) = (1,4,5,2,3,6)$, $\theta(xy)^4 = \theta(y), \theta(xy)^5 = \theta(x)$. So $\mathrm{im}(\theta) = \langle \theta(xy) \rangle$, hence $|\mathrm{im}(\theta)| = 6$. Since it is abelian we have $G/\ker(\theta) \cong \mathrm{im}(\theta)$, so $[G : G] \leq \ker(\theta)$. Hence $|G/H| = |G/\ker(\theta)| = 6$, so $\ker(\theta) = H$.

Exercise. Show θ is equivalent to coset representation on cosets of H

Write action of θ down as a table:

	x	y	x^{-1}	y^{-1}
1	$\underline{2}$	$\underline{3}$	2	$\underline{5}$
2	1	$\underline{4}$	1	$\underline{6}$
3	4	5	4	1
4	3	6	3	2
5	6	1	6	3
6	5	2	5	4

Scan the rows, and underline new numbers, regard the underline of numbers as definitions of these numbers. (Its important to do this even if you have definitions already from coset enumerations, since otherwise might not get Schreier transversal). So we have the definitions: $1x = 2, 1y = 3, 1y^{-1} = 5, 2y = 4$ and $2y^{-1} = 6$. We are also going to underline (twice) the inverse of the definitions:

	x	y	x^{-1}	y^{-1}
1	$\underline{2}$	$\underline{3}$	2	$\underline{5}$
2	1	$\underline{4}$	$\underline{\underline{1}}$	$\underline{6}$
3	4	5	4	$\underline{\underline{1}}$
4	3	6	3	$\underline{\underline{2}}$
5	6	$\underline{\underline{1}}$	6	3
6	5	$\underline{\underline{2}}$	5	4

For the other entries, such as $2x = 1$, we get associated Schreier generators. Things of the entries are of the from $ux = \overline{ux}$, Schreier generators $ux\overline{ux}^{-1}$. Call these $a, b, c, \ldots$ so that $ux\overline{ux}^{-1} = a$ implies $ux = a\overline{ux}$. So $2x = 1$ becomes $2x = a \cdot 1$, with a a Schreier generator.

	x	y	x^{-1}	y^{-1}
1	$\underline{2}$	$\underline{3}$	2	$\underline{5}$
2	$a1$	$\underline{4}$	$\underline{\underline{1}}$	$\underline{6}$
3	$b4$	$c5$	4	$\underline{\underline{1}}$
4	$d3$	$e6$	3	$\underline{\underline{2}}$
5	$f6$	$\underline{\underline{1}}$	6	3
6	$g5$	$\underline{\underline{2}}$	5	4

So we have $\{a, b, c, d, e, f, g\} = Z$. Recall that $Z^{-1} = \{ux^{-1}\overline{ux^{-1}}^{-1} | u \in U, x \in X\}$, so we can fill these in. For example $2x = a1$ implies $1x^{-1} = a^{-1}2$.

	x	y	x^{-1}	y^{-1}
1	$\underline{2}$	$\underline{3}$	$a^{-1}2$	$\underline{5}$
2	$a1$	$\underline{4}$	$\underline{\underline{1}}$	$\underline{6}$
3	$b4$	$c5$	$d^{-1}4$	$\underline{\underline{1}}$
4	$d3$	$e6$	$b^{-1}3$	$\underline{\underline{2}}$
5	$f6$	$\underline{\underline{1}}$	$g^{-1}6$	$c^{-1}3$
6	$g5$	$\underline{\underline{2}}$	$f^{-1}5$	$e^{-1}4$

Now all entries are either underlined or have a letter.

Theorem 6.6. *Let $G = \langle X | R \rangle = F/N$ with F free on X and $N = \langle R^F \rangle$. Let $H = E/N \leq G$. Let U be a Schreier Transversal of E in F (or equivalently of H in G, since we are using the same letters $x \in X$ to denote elements of G and F). Then $N = \langle S^E \rangle$ where $S = \{uwu^{-1} | u \in U, w \in R\}$ (so $|S| = |G : H| \times |R|$).*

Proof. $N = \langle R^F \rangle$, with $R^F = \{gwg^{-1} | g \in F, w \in R\}$. Each $g =_F hu$ with $h \in E, u \in U$. So $R^F = \{h(uwu^{-1})h^{-1} | h \in E, w \in R, u \in U\} = S^E$. So $N = \langle R^F \rangle = \langle S^E \rangle$ □

Corollary 6.7. *Let Y be the set of Schreier generators of H in G (or equivalently of E in F). Then $H \cong \langle Y|S(Y)\rangle$ where $S(Y)$ is the set $S = \{uwu^{-1}|u \in U, w \in R\}$, but rewritten as words in $(Y \cup Y^{-1})^*$.*

Idea of Proof. E is free on Y by Theorem 3.12, and $N = \langle S^E \rangle$ by the previous theorem. So $\langle Y|S(Y)\rangle$ is a presentation of $H = E/N$. (Warning this is not a formal proof, because we've not justified the change of generators S to $S(Y)$. The correctness depends on U being a Schreier transversal, otherwise $\langle Y|S(Y) \cup Y = X(Y(X))\rangle$) $\square$

So looking at row u of relator table tracing it through using table gives us r

	x		x	
1		2		$a1$
2		$a1$		$a2$
3		$b4$		$bd3$
4		$d3$		$db4$
5		$f6$		$fg5$
6		$g6$		$gf6$

relators from $w = x^2$ are a, a, bd, db, fg, gf. If rows of table are cyclic conjugate then so are the resulting relators, so only need one of them.

	y		y		y	
1		3		$c5$		$c1$
2		4		$e6$		$e2$

Note that the rest of the rows (3 to 6) are cyclic conjugates. The tracing for $1(xy)^6$ is as follows:

	x		y		x		y		x		y		x	
1		2		4		$d3$		$dc5$		$dcf6$		$dcf2$		$dcfa1$

	y		x		y		x		y
$defa1$		$dcfa3$		$dcfab4$		$dcfabe6$		$dcfabeg5$	

Note that all numbers arise in the above row, so all other rows gives cyclic conjugates of same relators. Usually the resulting presentation of H can be simplified

$H = \langle a, b, c, e, d, f, g|a, bd, fg, c, e, dcfabeg\rangle$

$\cong \langle b, d, f, g|bd, fg, dfbg\rangle$ eliminating relators on their own

$\cong \langle b, f|b^{-1}fbf^{-1}\rangle$ eliminating d, g using $d = b^{-1}, g = f^{-1}$

$\cong \langle b, f|bf = fb\rangle$

$\cong \mathbb{Z}^2$

So $G = D(2, 3, 6)$ is infinite, since $[G : G] \cong \mathbb{Z}^2$ (and we say $|G/[G : G]| = 6$

Example. (Question 3 from 2011 Exam). Let $G = \langle a, b|a^3, b^5, (ab)^2\rangle = D(3, 5, 2)$ and let $H = \langle b, ab^{-1}ab^2a^{-1}\rangle$.

- First we start with coset enumeration.

  Subgroup Table:

	b			a		b^{-1}		a		b		b		a^{-1}	
1	=	1	1		2		5		6	=	5		2		1

  Relator Tables:

	a		a		a			b		b		b		b		b			a		b		a		b	
1		2		3	=	1	1		1		1		1		1		1		1		2	=	3		1		1
2		3		1		2	2		3		4		6		5		2		2		3		4		5	=	2
3		1		2		3	3		4		6		5		2		3		3		1		1		2		3
4		5		6	=	4	4		6		5		2		3		4		4		5		2		3		4
5		6		4		5	5		2		3		4	=	6		5		5		6		5		6		5
6		4		5		6	6		5		2		3		4		6		6		4		6		4		6

  And the Table of Definitions and Deductions we made on the way:

Definitions		$1a = 2, 2a = 3$	$3b = 4, 4a = 5$	$5a = 6$
Deductions	$1b = 1$	$3a = 1, 2b = 3$	$5b = 2$	$6a = 4, 6b = 5, 4b = 6$

- Now we make a table to define U and Y

	a	b	a^{-1}	b^{-1}
1	$\underline{2}$	$t1$	$v^{-1}3$	$t^{-1}1$
2	$\underline{3}$	$u3$	$\underline{\underline{1}}$	$x^{-1}5$
3	$v1$	$\underline{4}$	$\underline{\underline{2}}$	$u^{-1}2$
4	$\underline{5}$	$w6$	$y^{-1}6$	$\underline{\underline{3}}$
5	$\underline{6}$	$x2$	$\underline{\underline{4}}$	$z^{-1}6$
6	$y4$	$z5$	$\underline{\underline{5}}$	$w^{-1}4$

Call the Schreier generators, t, u, v, w, x, y, z.

- Now we trace the relators:

	a		a		a	
1		2		3		$v1$
4		5		6		$y4$

(No need to do $2, 3$ as they appear in first row, and no need to do $5, 6$ as they appear in the second row)

	b		b		b		b		b	
1		$t1$		$t^2 1$		$t^3 1$		$t^4 1$		$t^5 1$
2		$u3$		$u4$		$uw6$		$uwz5$		$uwzx2$

(No need to do $3, 4, 5, 6$ as they appear in the second row)

	a		b		a		b	
1		2		$u3$		$uv1$		$uvt1$
2		3		4		5		$x2$
5		6		$z5$		$z6$		$z^2 5$
6		$y4$		$yw6$		$ywy4$		$(yw)^2 6$

(No need to do 3, as $3a$ appears in the first row. $4a$ appears in the second row, so ignore 4 as well)

Hence we have that

$$\begin{aligned} H &\cong \langle t, u, v, w, x, y | v, y, t^5, uvwzx, uvt, x, z^2, (yw)^2 \rangle \\ &\cong \langle t, u, w, z | t^5, uwz, ut, z^2, w^2 \rangle \\ &\cong \langle t, w, z | t^5, t^{-1}wz, z^2, w^2 \rangle \\ &\cong \langle t, w | t^5, w^2, (w^{-1}t)^2 \rangle \\ &\cong \langle t, w | t^5, w^2, (wt)^2 \rangle \\ &\cong D(5, 2, 2) \\ &\cong D_{10} \end{aligned}$$

So $|H| = 10$, and since $|G : H| = 6$ we have $|G| = 60$, hence $D(2, 3, 5)$ is finite. In fact $D(2, 3, 5) \cong A_5 \cong \mathrm{PSL}_2(5)$.

Example. An example of computing presentation on given generators of subgroups (not examinable)

Let $G = \langle x, y |, x^3, y^5, (xy)^2 \rangle$, and $H = \langle xy, x^{-1}y^{-1}xyx \rangle$. Label generators of H, $a = xy$, $b = x^{-1}y^{-1}xyx$. For this, keep a, b in tables while doing coset enumeration.

Generators:

	x		y			x^{-1}		y^{-1}		x		y		x	
1		2	=	$a1$	1		3		4	=	$b4$		$b3$		$b1$

Relators:

	x		x		x			y		y		y		y		y	
1		2		3	=	1	1		$b^{-1}4$		$b^{-1}3$		$b^{-1}5$	=	$a^{-1}2$		1
2		3		1		2	2		$a1$		$ab^{-1}4$		$ab^{-1}3$		$ab^{-1}5$		2
3		1		2		3	3		5		$ba^{-1}2$		$b1$		4		3
4		$b4$		$b^2 4$		$b^3 4$	4		3		5		$ba^{-1}2$		$b1$		4
5		$ab^{-1}5$		$(ab^{-1})^2 5$		$(ab^{-1})^3 5$	5		$ba^{-1}2$		$b1$		4		3		5

	x		y		x		y	
1		2		$a1$		$a2$		$a^2=1$
2		3		5	=	$ab^{-1}5$		2
3		1	=	$b^{-1}4$		4		3
4		$b4$		$b3$		$b1$		4
5		$ab^{-1}5$		2		3		5

Definitions and Deductions:

Definitions	$1x=2$, $2x=3$	$3y^{-1}=4$	$3y=5$
Deductions	$3x=1$,$2y=a1$	$4x=b4$, $1y=b^{-1}4$	$5y=ba^{-1}2$,$5x=ab^{-1}5$
relators	$a^2=1$		$b^3=1$, $(ab^{-1})^3$

So we have $H=\langle a,b|a^2,b^3,(ab^{-1})^3\rangle \cong \langle a,b|a^2,b^3,(ab)^3\rangle = D(2,3,3)$ which has order 12. So $|G|=60$.
Dealing with coincidence with this method is much harder and has additional technicalities.

Proposition 6.8. *Let $G=\langle X|R\rangle$ with X,R finite and $|R|<|X|$, then $G/[G,G]$ is infinite and hence so is G.*

Proof. Let $C=\langle t\rangle$,$|t|=\infty$ be the infinite cyclic group. Consider a map $\theta: X\to C$, let $X=\{x_1,\dots,x_m\}$, $\theta(x_i)=t^{\lambda_i}$ for some $\lambda_i\in\mathbb{Z}$. By the Fundamental Theorem, θ extends to $\theta: G\to C$ if $\theta(r)=1\,\forall r\in R$. Note that $\theta(r)=1\,\forall r\in R$ if and only if the associated system of homogeneous linear equation in $\lambda_1,\dots,\lambda_n$ as solution (for example if $r=x_2x_1^{-1}x_3^2x_2^3x_1^{-1}$ then $\theta(r)=t^{-2\lambda_1+4\lambda_2+2\lambda_3}$, so $\theta(r)=1$ if and only if $-2\lambda_1+4\lambda_2+2\lambda_3=0$). If $|X|<|R|$, then we have fewer equations than variables, so by Linear Algebra, there exists a non-zero solutions $\lambda_1,\dots,\lambda_m\in\mathbb{Q}$. Multiply by a constant to get a solution $\lambda_i\in\mathbb{Z}$. So there exists a non-trivial homomorphism $\theta: G\to C$ with $\mathrm{im}(\theta)\le C$, so $\mathrm{im}(\theta)$ is infinite. So $G/\ker(\theta)$ is infinite and abelian, hence $[G,G]\le\ker(\theta)$, so $G/[G,G]$ is infinite. □

Lemma 6.9. *Let $H\le G$, $\Omega=\{Hg|g\in G\}$, coset action of G on Ω. Then for any $g\in G$, $Hk\in\Omega$ we have $(Hk)^g=Hk$ if and only if $g\in H$. That is $Stab_G(Hk)=H$.*

Proof. We have $(Hk)^g=Hk$, if and only if, $kgk^{-1}\in H$, if and only if, $g\in k^{-1}Hk=H=Hkg$ since $H\lhd G$. □

Theorem 6.10. *Let $G=\langle x_1,\dots,x_r|w_1^{m_1},w_2^{m_2},\dots,w_s^{m_s}\rangle$ where $\sum_{i=1}^s\frac{1}{m_i}\le r-1$. Assume also that there is a group P and a homomorphism $G\to P$ in which $|\theta(w_i)|=m_i$ for $1\le i\le s$. Then G is infinite.*

Proof. Assume $\mathrm{im}(\theta)=P$. If P is infinite then G is clearly infinite, so assume P is finite

Let $H=\ker(\theta)$ so $|G:H|=n$ is finite. Apply Reidemeister - Schreier to get $H=\langle E|S(Y)\rangle$, where $|E|=(r-1)n+1$ by Theorem 3.12. $|S(Y)|=|G:H||R|=ns$, so $|S(Y)|>|E|$. By we can eliminate cyclic conjugates form $S(Y)$. Consider row c for the relator table for $w_i^{m_i}$

	w_i	...	w_i	...	w_i	
$c=c_1$		v_2c_2	...	$v_{m_i}c_{m_i}$		vc_1

so v is a relator. Since $|\theta(w_i)|=m_i$ by assumption, $\theta(w_i^l)\ne1$ for $l<m_i$, so $w_i^l\notin H$ for $1\le l<m_i$. So by Lemma 6.9, for any coset c_j we have $c_j^{w_i^l}\ne c_j$ for $1\le l<m_i$. So in the row for c, the coset numbers $c_1,c_2,\dots,c_{m_i}$ are all different, so these m_i rows just give cyclic conjugates of relator v for c_1. So we only need to keep n/m_i of relators from $w_i^{m_i}$. So can reduce to presentation of H with $(r-1)n+1$ generators $n\sum_{i=1}^s\frac{1}{m_i}$ relators which by assumption is less than $n(r-1)$. So by Proposition 6.8, we have $H/[H,H]$ is infinite, so G is infinite. □

Example. Let $G=D(2,4,5)=\langle x,y|x^2,y^4,(xy)^5\rangle$. We have $\frac12+\frac14+\frac15<1$, so to prove infinite we need to find $\theta: G\to P$. Can look for permutations, $a=(1,2)$, $b=(2,3,4,5)$ then $ab=(1,3,4,5,2)$. So can take $P=\langle(1,2),(2,3,4,5)\rangle=S_5$

Let $G=\langle x,y,z,|x^3,y^3,z^3,(xy)^3,(yz)^3,(xz)^3\rangle$,then $\sum\frac{1}{m_i}=2\le3-1$. So for P, take $P=\langle a,b,c\rangle$ where $a=(1,2,3)$, $b=(4,5,6)$ and $c=(7,8,9)$. So G is infinite.

6.3 The groups $D(l,m,n)$

The groups $D(l,m,n)=\langle x,y|x^l,y^m,(xy)^n\rangle$. Assume $2\le l\le m\le n$. Let $G=\langle a,b,c|a^2,b^2,c^2,(ab)^l,(bc)^m,(ac)^n\rangle$. The subgroup $\langle ab,bc\rangle$ has index 2 in G and is isomorphic to $D(l,m,n)$ (easy proof with Reidemeister - Schreier). G is a 3 generators *Coxetor group* and can be studied as reflection groups. Take the triangular tessellation of "plane" (Euclidean, elliptic or hyperbolic) using triangles with angles $\pi/l,\pi/m,\pi/n$. Then a,b,c are reflections in sides of some fixed triangle. It can be proved that group generated by a,b,c is isomorphic to group G defined by presentation. The three cases are:

Elliptic Case When $1/l + 1/m + 1/n > 1$. The only cases are $(2,2,n) = D_{2n}$; $(2,3,3) = A_4$; $(2,3,4) = S_4$ and $(2,3,5) = A_5$. So they are all finite. The plane is the surface of a sphere, so the tessellation is finite. (For picture look at Triangle Groups)

Euclidean Case When $1/l + 1/m + 1/n = 1$. This is the "normal" plane and the only cases are $(3,3,3)$, $(2,4,4)$ and $(2,3,6)$. The groups act regularly (transitive, trivial stabiliser) on triangles

Hyperbolic Case When $1/l + 1/m + 1/n < 1$. The plane is the hyperbolic plane.

So it is not hard to prove (using some hyperbolic geometry) that $D(l,m,n)$ is infinite if and only if $1/l+1/m+1/n \leq 1$. But we will prove this algebraically. We have already considered all $1/l + 1/m + 1/n > 1$ cases as examples.

Theorem 6.11. *$D(l,m,n)$ is infinite when $1/l + 1/m + 1/n \leq 1$.*

Proof. By Theorem 6.10, it is enough to find a group P containing elements x, y with $|x| = l$, $|y| = m$ and $|xy| = n$. We can always find finite permutation with these orders. For example, with $l = 5, m = 7$and $n = 9$ we can take $x = (1,2,3,4,5)$, $y = (3,4,5,6,7,8,9)$ then $xy = (1,2,4,6,7,8,9,3,5)$. It is hard to describe in general this process.

In fact we construct P as a quotient group of matrices. Let K be a field of characteristic not 2. Let $\mathrm{SL}_2(K)$ be as usual the multiplicative group of 2×2 invertible matrices with entries in K and determinant 1. Let $Z = \{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\}$, note $1 \neq -1$ since $\mathrm{char}(K) \neq 2$. Then $Z \lhd \mathrm{SL}_2(K)$ (in fact Z is the centre of $\mathrm{SL}_2(K)$. Define $\mathrm{PSL}_2(K) = \mathrm{SL}_2(K)/Z$.

Lemma 6.12. $\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$ *is the only element of order* 2 *in* $\mathrm{SL}_2(K)$

Proof. Let $\begin{pmatrix}a & b\\ c & d\end{pmatrix} = A$ with $\mathrm{ord}(A) = 2$. Then $A^2 = I$ implies $a^2 + bc = bc + d^2 = 1$, so $b(a+d) = c(a+d) = 0$ hence either $b = c = 0$ or $a + d = 0$. If $a + d = 0$, then $d = -a$, but $\det(A) = ab - bc = 1$ implies $-a^2 - bc = 1$ contradicting $a^2 + bc = 1$ (using $1 \neq -1$). So $b = c = 0$ and hence $a^2 = 1 = d^2$, hence either $a = d = 1$ or $a = d = -1$ □

Lemma 6.13. *If $A \in \mathrm{SL}_2(K)$ with $\mathrm{ord}(A) = 2r$ for some $r \geq 1$ then order of AZ in $\mathrm{PSL}_2(K)$ is r*

Proof. Since $\mathrm{ord}(A) = 2r$ then $\mathrm{ord}(A^r) = 2$. So $A^r = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix} \in Z$ and $A^s \notin Z$ for $1 \leq s < r$, so the result follows. □

Let $A = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in S$, then $\mathrm{Tr}(A) = a + d$ and the characteristic polynomial of $A = (a-x)(d-x) - bc = x^2 - (a+d)x + (ad - bc) = x^2 - \mathrm{Tr}(A) + 1$. So when $\mathrm{Tr}(A) \neq \pm 2$ we have distinct eigenvalues.

Lemma 6.14. *Let $A, B \in S$ and $\mathrm{Tr}(A) = \mathrm{Tr}(B) \neq \pm 2$. Then $\mathrm{ord}(A) = \mathrm{ord}(B)$*

Proof. Now $\mathrm{Tr}(A) = \mathrm{Tr}(B)$ implies A and B have the same characteristic polynomial and eigenvalues (possibly in an extension field L of K) are distinct, so by Linear Algebra, A, B are similar, i.e., there exists $P \in \mathrm{SL}_2(L)$ such that $P^{-1}AP = B$. So A, B are conjugate in $\mathrm{SL}_2(L)$. So $\mathrm{ord}(A) = \mathrm{ord}(B)$. □

Lemma 6.15. *For all $l, m, n \geq 2$ and assume that K contains primitive $(2l)^{\mathrm{th}}$, $(2m)^{\mathrm{th}}$ and $(2n)^{\mathrm{th}}$ root of unity. Then there exists $A, B \in S$ with $\mathrm{ord}(A) = 2l$, $\mathrm{ord}(B) = 2m$ and $\mathrm{ord}(AB) = 2n$.*

Proof. Let λ, μ be primitive $(2l)^{\mathrm{th}}$ and $(2m)^{\mathrm{th}}$ roots of unity respective. Since $2l, 2m \geq 4$ we have $\lambda + \lambda' \neq \pm 2$ and $\mu + \mu^{-1} \neq \pm 2$. (Note λ root of unity and $\lambda + \lambda^{-1} = \pm 2$ then $\lambda = \pm 1$). So let $A = \begin{pmatrix}\lambda & 0\\ 1 & \lambda^{-1}\end{pmatrix}$ and $B = \begin{pmatrix}\mu & \alpha\\ 0 & \mu^{-1}\end{pmatrix}$ for some α. Then $\mathrm{Tr}(A) = \lambda + \lambda^{-1}$, so by Lemma 6.14 $\mathrm{ord}(A) = 2l$. Similarly $\mathrm{ord}(B) = 2m$ for any $\alpha \in K$. Now $AB = \begin{pmatrix}\lambda\mu & \lambda\alpha\\ \mu & \alpha + \mu^{-1}\lambda^{-1}\end{pmatrix}$. Let ν be a primitive $(2n)^{\mathrm{th}}$ root of unity. By Lemma 6.14 if $\mathrm{Tr}(AB) = \nu + \nu^{-1}$, then $\mathrm{ord}(AB) = 2n$. So we just choose $\alpha = \nu + \nu^{-1} - \lambda\mu - \lambda^{-1}\mu^{-1}$. □

Now by Lemma 6.13, in $\mathrm{PSL}_2(K)$ element $AZ, BZ, ABZ = (AZ)(BZ)$ have orders l, m, n. So we have the required group P and Theorem 6.11 follows from Theorem 6.10. □

If we choose $K = \mathbb{C}$, then $\langle AZ, BZ \rangle$ will usually be infinite. Can we find a finite group P with elements α, β.

Theorem 6.16. *Let $l, m, n \geq 2$, then there exists a finite group P, with $\alpha, \beta \in P$ and $|\alpha| = l, |\beta| = m, |\alpha\beta| = n$.*

Proof. To prove this we just need a finite field K, such that $\mathrm{char} K \neq 2$ and containing primitive $(2l)^{\text{th}}$, $(2m)^{\text{th}}$ and $(2n)^{\text{th}}$ root of unity. Recall that for each prime power $q = p^n$, there exists a unique finite field $\mathbb{F}_q$ of order q with characteristic p, and the multiplicative group $\mathbb{F}_q \setminus \{0\}$ is cyclic of order $q-1$. So $\mathbb{F}_q$ will have the required roots of unity if and only if $2l, 2m, 2n$ divides $q-1$. By Theorem of Dirichlet (from Number Theory), there exists infinitely many primes $p \equiv 1 \mod 2lmn$. But we don't actually need this theorem, if we choose any odd prime p, coprime to $2lmn$ and let e be the order of p in $U(2lmn)$ (unit group), then $p^e \equiv 1 \mod 2lmn$. So choose $q = p^e$. □

7 Baumslaq - Solitar Groups

Let $G(m,n) = \langle x, y | y^{-1}x^m y = x^n \rangle$, $m, n \neq 0$. $G(2,3)$ was the first example (1962) of a finitely generated non-Hopfian group, i.e., there exists $1 \neq N \lhd G$ with $G/N \cong G$. Easy to find non-finitely generated example. The groups with $m = 1$ have an easier structure. So we will consider those first. Note that the case $m = n = 1$ is the free abelian group, so we assume $n > 1$.

Consider $G = (1,n) = \langle x, y | y^{-1}xy = x^n \rangle$. Define $x_r = y^{-r}xy^r$ for any $r \in \mathbb{Z}$ (so $x_0 = x$ and $x_1 = x^n$).

Lemma 7.1. *We have $x_r^n = x_{r+1}$ for all $r \in \mathbb{Z}$. More generally $x_r^{n^k} = x_{r+k}$, $k \geq 0$.*

Proof. $x_r^n = (y^{-r}xy^r)^n = y^{-r}x^n y^r = y^{-r}y^{-1}xyy^r = x_{r+1}$.

Second statement is by induction on k. So $x_r = x^{n^r}$. But not $x_{-1} = yxy^{-1}$ is not a power of x. □

Let $N = \langle x_r | r \in \mathbb{Z} \rangle$.

Corollary 7.2. *N is abelian.*

Proof. If $s > r$, then x_s is a power of x_r, so $[x_r, x_s] = 1$. □

Corollary 7.3. *$N = \langle x^G \rangle \lhd G$*

Proof. Recall $\langle x^G \rangle = \langle g^{-1}xg | g \in G \rangle$. For any $r, x_r = y^{-r}xy^r \in \langle x^G \rangle$, so $N = \langle x_r \rangle \leq \langle x^G \rangle$.

For the other direction, we need to prove that for all $g \in G$, we have $g^{-1}xg$ is a product of the x_r . Write g as a word in $x^{\pm 1}$, $y^{\pm 1}$. We do a prove by induction on length of the word.

If Length is zero, then g is the identity, so $g^{-1}xg = x = x_0$

Let $g = wa$ with $a \in \{x^{\pm 1}, y^{\pm 1}\}$. Then $g^{-1}xg = a^{-1}(w^{-1}xw)a$, which by induction is $a^{-1}(x_{r_1} \dots x_{r_k})a$. If $a = x^{\pm 1}$ then $a^{-1}x_r a = x_r$ for all r, hence $g^{-1}xg = w^{-1}xw$ and we are done. If $a = y^{\pm 1}$, then $y^{-1}x_r y = x_{r+1}$ and $yx_e y^{-1} = x_{r-1}$, so we still get $g^{-1}xg =$ product of x_r. □

Now $G/N = G/\langle x^G \rangle$, which is what we get if we add relator x to presentation of G. That is $G/\langle x^G \rangle \cong \langle x, y | y^{-1}xy = x^n, x = 1 \rangle \cong \langle y | \, \rangle$ infinite cyclic. So we have N is abelian (but not finitely generated) and G/N infinite cyclic, so G is *metabelian group*. We need a norm form for group elements. Using $x_r y = yx_{r+1}, x_r y^{-1} = y^{-1}x_{r-1}$ we can move y, y^{-1} to the left of a word to get in the form $g = y^k x_{r_1}^{\pm 1} x_{r_2}^{\pm 1} \dots$. Since x_s is a power of x_r for each $s > r$, each x_{r_i} is a power of x_r with $r = \min r_i$. So we get $g = y^k x_r^s$ for some $s \in \mathbb{Z}$. Since $x_r^n = x_{r+1}$ we can assume that $n \nmid s$.

Example. Let $n = 3$, $g = yxyxy^{-1} = y\underbrace{x_0 y}\underbrace{x_0 y} = yy\underbrace{x_1 y^{-1}}x_{-1} = yyy^{-1}x_0 x_{-1} = yx_{-1}^4$.

Proposition 7.4. *Each $g \in G$ has a unique expression as $y^k x_r^s$ with $k, r, s \in \mathbb{Z}$ and $n \nmid s$.*

Proof. Let $H = \mathrm{GL}_2(\mathbb{Q})$ (= multiplicative group of 2×2 invertible matrices with entries in $\mathbb{Q}$)

Define $\theta : \{x, y\} \to H$ by $\theta(x) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $\theta(y) = \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix}$. Then

$$\begin{aligned} \theta(y^{-1}xy) &= \begin{pmatrix} 1/n & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1/n & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix} \\ &= \theta(x)^n \end{aligned}$$

So by the Fundamental Theorem, θ extends to a homomorphism $\theta : G \to H$.

Now

$$\begin{aligned} \theta(x_r) &= \theta(y^{-r}xy^r) \\ &= \begin{pmatrix} n^{-r} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} n^r & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ n^r & 1 \end{pmatrix} \end{aligned}$$

So

$$\begin{aligned}\theta(y^k x_r^s) &= \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix}^k \begin{pmatrix} 1 & 0 \\ n^r & 1 \end{pmatrix}^s \\ &= \begin{pmatrix} n^k & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ sn^r & 1 \end{pmatrix} \\ &= \begin{pmatrix} n^k & 0 \\ sn^r & 1 \end{pmatrix}\end{aligned}$$

So for distinct r, k, s with $s \nmid n$, the distinct elements $y^k x_r^s$ have distinct images under θ. □

Note $\theta(N)$ is generated by $\theta(x_r) = \left\{ \begin{pmatrix} 1 & 0 \\ n^r & 1 \end{pmatrix} | r \in \mathbb{Z} \right\} = \left\{ \begin{pmatrix} 1 & 0 \\ sn^r & 1 \end{pmatrix} | r, s \in \mathbb{Z} \right\}$ which is isomorphic to the subgroup $\{sn^r | s \in \mathbb{Z}, r \in \mathbb{Z}\}$ of $(\mathbb{Q}, +)$, which is the set of rationals $\frac{a}{b}$ such that primes dividing b are those dividing n.

Now we move onto the general case $G(m, n)$ with $m, n \geq 2$. We still have $\theta : G \to \mathrm{GL}_2(\mathbb{Q})$ defined by $\theta(x) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \theta(y) = \begin{pmatrix} n/m & 0 \\ 0 & 1 \end{pmatrix}$, but θ is not a monomorphism.

Digression HNN (Higman, Neumann, Neumann)extensions

Let $H = \langle X | R \rangle$ be a group and $\langle z_1, \ldots, z_k \rangle, \langle z'_1, \ldots, z'_k \rangle$ be isomorphic subgroups of H with isomorphic $z_i \mapsto x'_i$. The associated HNN extension is $G = \langle X, y | R, y^{-1} z_i y = z'_i \rangle$. It can be proved for example that: The natural map $H \to G$ defined by $X \to X$ is a monomorphism.

$G(m, n)$ is HNN extension, $z_1 = x^m, z'_1 = x^n$ so $\langle z_1 \rangle, \langle z'_1 \rangle$ are both infinite cyclic, and $H = \langle x | \rangle$. HNN are used in proof of unsolvability of word problem for group presentations. Also used for constructing examples with strange properties. Such as infinite groups in which all non-identity elements are conjugate.

We look for a normal form for group elements in $G(m, n)$. (This generalises easily to arbitrary HNN extensions). We still use $x^m y = y x^n$ and $x^n y^{-1} = y^{-1} x^m$ to move y, y^{-1} to the left whenever possible. In a subword $x^k y$, we write $k = t + mu$ with $0 \leq t < m$. Then $x^k y = x^{t+mu} y = x^t y x^{nu}$. Similarly for $x^k y^{-1}$ we can write $k = t + nu$ with $0 \leq t < n$ and get $x^k y^{-1} = x^t y^{-1} x^{mu}$. So doing this as much as possible brings the word into the form $g = x^{t_1} y^{u_1} \ldots x^{t_r} y^{u_r} x^v$ where:

1. $t_i, u_i, v, r \in \mathbb{Z}$ and $r > 0$
2. $t_i \neq 0$ for all $i > 1$ and $t_1 = 0$ if $r = 0$
3. $u_i \neq 0$ for all i
4. $u_i > 0$ then $0 \leq t_i < m$
5. $u_i < 0$ then $0 \leq t_i < n$

Note that 1. ,2. and 3. we could do in any groups, while 4. and 5. we can impose using substitutions above arising from group relators. We conjecture that this is a normal form for group elements.

Theorem 7.5. *Each $g \in G$ has unique expression satisfying* 1. -5. *above.*

Remark. This method of proof can be used in general to prove that a conjectured normal form for a group really is a normal form. Useful if there is no representation of the group that can be used. We construct a representation as group of permutations of set of normal forms, which will be equivalent to regular group action.

Proof. Let Ω be the set of words satisfying 1. -5. We construct an action of G on Ω . First want an action of generators which should correspond to multiplying on the right. So for $\alpha \in \Omega, a \in \{x^{\pm 1}, y^{\pm 1}\}$ we want α^a to be the normal form of αa.

For $\alpha = x^{t_1}y^{u_1}\dots x^v$ as above, let $\beta = x^{t_1}y^{u_1}\dots x^{t_{r-1}}y^{u_{r-1}}$ or $\beta = \epsilon$ is $r \leq 1$. Let $t = t_r, u = u_r$ (with $t = u = 0$ if $r = 0$). So $\alpha = \beta x^t y^u x^v$, write $v = am + b$, $0 \leq b < m$, and $v = cn + d$,$0 \leq d < n$, then we define

$$\begin{aligned}
\alpha^x &= \beta x^t y^u x^{v+1} \\
\alpha^{x^{-1}} &= \beta x^t y^u x^{v-1} \\
\alpha^y &= \begin{cases} \beta x^t y^u x^b y x^{an} & b \neq 0 \\ \beta x^t y^{u+1} x^{an} & b = 0, u \neq -1 \\ \beta x^{t+an} & b = 0, u = -1 \end{cases} \\
\alpha^{-y} &= \begin{cases} \beta x^t y^u x^d y^{-1} x^{cm} & d \neq 0 \\ \beta x^t y^{u-1} x^{cm} & d = 0, u \neq 1 \\ \beta x^{t+cm} & d = 0, u = 1 \end{cases}
\end{aligned}$$

We defined maps $\theta(a) : \Omega \to \Omega$ for $a \in \{x^{\pm 1}, y^{\pm 1}\}$. To prove $\theta(a) \in \mathrm{Sym}(\Omega)$ (that is $\theta(a)$ is a bijection), we prove that $\theta(x)$, $\theta(x^{-1})$ and $\theta(y)$, $\theta(y^{-1})$ are inverse maps. This is clear in the case $\theta(x), \theta(x^{-1})$. So we prove $\theta(y)\theta(y^{-1}) = I_\Omega$, i.e., $(\alpha^y)^{y^{-1}} = \alpha \, \forall \alpha \in \Omega$. (The proof for $(\alpha^{y^{-1}})^y = \alpha$ is similar). Consider 3 cases for α^y

Case 1. This is Case 3 for $(\alpha^y)^{y^{-1}}$. We get $(\alpha^y)^{y^{-1}} = \beta x^t y^u x^{b+am} = \alpha$.

Case 2. This is Case 2 for $(\alpha^y)^{y^{-1}}$ (or Case 3 if $u = 0$ with $r = t = 0$). In either case we have $(\alpha^y)^{y^{-1}} = \beta x^t y^u x^{am} = \alpha$.

Case 3. If $t \neq 0$ this is Case 1 for $(\alpha^y)^{y^{-1}}$. Then we have $(\alpha^y)^{y^{-1}} = \beta x^t y^{-1} x^{am} = \alpha$. If $t = 0$, then $r = 1$ so $\alpha = y^{-1}x^{am}$. Hence $\alpha^y = x^{an}$ and so $(\alpha^y)^{y^{-1}} = y^{-1}x^{am} = \alpha$.

So we have that $(\alpha^y)^{y^{-1}} = \alpha$ in all cases (similarly for $(\alpha^{y^{-1}})^y = \alpha$). So $\theta(a) \in \mathrm{Sym}(\Omega)$ for $a \in \{x^{\pm 1}, y^{\pm 1}\}$. To use the Fundamental Theorem to prove θ extends to an action on Ω, i.e., $\theta : G \to \mathrm{Sym}(\Omega)$, we have to check $\theta(y)^{-1}\theta(x)^m\theta(y) = \theta(x)^n$, equivalently $\theta(x)^m\theta(y) = \theta(y)\theta(x)^n$. So we must check $\alpha^{x^m y} = \alpha^{yx^n} \, \forall \alpha \in \Omega$. Again split into the three cases for α^y

Case 1. $\alpha = \beta x^t y^u x^{am+b}$. So $\alpha^{x^m} = \beta x^t y^u x^{(a+1)m+b}$, so we are still in Case 1 for $\alpha^{x^m y}$. Hence $\alpha^{x^m y} = \beta x^t y^u x^b y x^{(a+1)n} = \alpha^{yx^n}$.

Case 2. $\alpha = \beta x^t y^u x^{am}$. So $\alpha^{x^m} = \beta x^t y^u x^{(a+1)m}$ and we are still in Case 2. Hence $\alpha^{x^m y} = \beta x^t y^{u+1} x^{(a+1)n} = \alpha^{yx^n}$

Case 3. $\alpha = \beta x^t y^{-1} x^{am}$. So $a^{x^m} = \beta x^t y^{-1} x^{(a+1)m}$ and we are still in Case 1. Hence $\alpha^{x^m y} = \beta x^{t+(a+1)n} = \alpha^{yx^n}$.

So by the Fundamental Theorem, we do have an action of G on Ω. Let α be a normal form word. Then it is easy to see that applying the group element α represented by α to $\epsilon \in \Omega$, we get $\epsilon^\alpha = \alpha$. So distinct element $\alpha, \beta \in \Omega$ have distinct images under θ (since $\epsilon^\alpha \neq \beta = \epsilon^\beta$) we must have $\alpha \neq_G \beta$ and we are done. □

Same kind of proof can be used more generally to get normal form for HNN extensions.

Now let $G = G(2,3) = \langle x, y | y^{-1}x^2y = x^3 \rangle$ and $r = x^{-1}y^{-1}xyx^{-1}y^{-1}xyx^{-1} = [x,y]^2x^{-1}$. We want to put r into normal form. $x^{-1}y^{-1} = x^2x^{-3}y^{-1} = x^2y^{-1}x^{-2}$. So

$$\begin{aligned}
r &= \underline{x^{-1}y^{-1}}xyx^{-1}y^{-1}xyx^{-1} \\
&= x^2y^{-1}\underline{x^{-1}y}x^{-1}y^{-1}xyx^{-1} \\
&= x^2y^{-1}xy\underline{x^{-4}y^{-1}}xyx^{-1} \\
&= x^2y^{-1}xyx^2y^{-1}\underline{x^{-3}y}x^{-1} \\
&= x^2y^{-1}xyx^2y^{-1}xyx^{-7}
\end{aligned}$$

Definition 7.6. A group G is *non-Hopfian* if there exists $1 \neq N \lhd G$ such that $G/N \cong G$

Theorem 7.7. *$G(2,3)$ is non-Hopfian. (In general $G(m,n)$ is non-Hopfian if and only if there exists primes p, q with $p|m$,$q|n$ and $p \nmid n$, $q \nmid m$)*

Proof. Let $G = G(2,3)$, $r = [x,y]^2x^{-1} = x^2y^{-1}xyx^2y^{-1}xyx^{-7} \neq \epsilon$. So $r \neq_G 1$. Let $N = \langle r^G \rangle$,so $N \neq 1$. To get presentation of G/N, we just add r as extra relation to G. So

$$\begin{aligned}
G/N &\cong \langle x,y | y^{-1}x^2y = x^3, x = [x,y]^2 \rangle \\
&\cong \langle x,y,w | y^{-1}x^2y = x^3, x = w^2, w = [x,y] \rangle \\
&\cong \left\langle y,w | y^{-1}w^4y = w^6, \underbrace{w = [w^2,y]}_{w = w^{-2}y^{-1}w^2y} \right\rangle \\
&\cong \langle y,w | y^{-1}w^4y = w^6, y^{-1}w^2y = w^3 \rangle \\
&\cong \langle y,w, | y^{-1}w^2y = w^3 \rangle \\
&\cong G
\end{aligned}$$

since $y^{-1}w^2y = w^3$ implies the relation $y^{-1}w^4y = w^6$.

□

8 The Burnside Problem

Burnside 1902: Given a finitely generated group G with all elements having finite order (i.e. G a torsion group), is G finite? (Certainly true for Abelian groups by the Fundamental Theorem of Finitely Generated Abelian Groups)

The general answer is no: In 1964 an example was given by Gringorchuk

Harder question is: Given a finitely generated group G with all element having finite order, and there is an upper bound on the order, equivalently assume there exists n such that $g^n = 1$ for all $g \in G$. Then is G finite?

The answer is no for large enough n ($n \geq 8000$?). The answer is yes for $n = 2, 3, 4, 6$. So there are still a lot of n which is unknown for.

For large enough primes p, there exists *Tarski Monsters*: $G = \langle x, y \rangle$ with G infinite and $g^p = 1$ for all $g \in G$ and only subgroups of G are $\langle g \rangle$ for $g \in G$.

Definition 8.1. We define the *Burnside Group* $B(r, n)$ as follows: $B(r, n) = \langle x_1, \ldots, x_r | w^n = 1 \, \forall w \in A^* \rangle$

The above question is now for which r, n is $B(r, n)$ finite. Not known for $B(2, 5)$.

Restricted Burnside Problem: Is there a largest finite group G with r generators and $g^n = 1 \, \forall g \in G$. Equivalently does $B(r, n)$ have largest finite quotient.

The answer is Yes, done by Zelmanov (1990?), there exists largest finite quotient $RB(r, n)$. In particular $|RB(2, 5)| = 5^{34}, |RB(3, 5)| = 5^{2282}$ and $|RB(2, 7)| = 7^{20416}$.

Theorem 8.2. *$B(r, 2)$ is abelian of order 2^r*

Proof. Let $G = B(r, 2)$, $g, h \in G$. Consider $(gh)^2 = ghgh = 1$ but $g^2h^2 = gghh = 1$. So we get $gh = hg$ and G is abelian.

So $G = \langle x_1, \ldots, x_r | x_i^2, [x_i, x_j] \rangle \cong \langle x_1 \rangle \times \cdots \times \langle x_r \rangle \cong C_2^r$ (by Proposition 4.9) □

Cauchy's Theorem. *If G is a finite group $p||G|$ with p prime, then G has an element of order p*

Proof. Let $S = \{(x_1, \ldots, x_p) | x_i \in G, x_1 x_2 \ldots x_p = 1\}$. Since $x_p = (x_1 \ldots x_{p-1})^{-1}$, we can choose $x_1, \ldots, x_{p-1}$ arbitrarily so $|S| = |G|^{p-1}$. Hence $p||S|$. Note $(x_1, \ldots, x_p) \in S$ implies $(x_2, x_3, \ldots, x_p, x_1) \in S$. So we have an action of cyclic group $\langle g \rangle = C_p$ on S with $g \cdot (x_1, \ldots, x_p) = (x_2, \ldots, x_p, x_1)$. So orbit of action has size 1 or p (since p is prime). Now an orbit of size 1 will be $(x, x, \ldots, x), x^p = 1$. There exists at least one of these $x = 1$. But since $p||S|$ we have p|number of orbit of size 1. So there exists $x \neq 1$ with $x^p = 1$. □

Theorem 8.3. *$B(r, 3)$ is finite of order at most $3^{3^{r-1}}$*

Proof. Induction on r.

For $r = 1$, we clearly have $|B(1, 3)| = 3 = 3^{3^{1-1}}$

Let $r > 1$. Note for $g, h \in G = B(r, 3)$ we have $(gh)^3 = 1$ implies $ghg = h^{-1}g^{-1}h^{-1}$. Let $H = \langle x_1, \ldots, x_{r-1} \rangle \leq G$. By induction, we know H is finite with $|H| \leq 3^{3^{r-2}}$. For any $g = G$ let $g = u_1 z^{\pm 1} u_2 \ldots z^{\pm 1} u_m$ with $u_i \in H$ and $z = x_r$ (so $z^3 = 1$). Do this with the smallest m. If we had zuz we can turn that into $u^{-1}zu^{-1}$ which reduces m. So we have that z and z^{-1} must alternate in the word. Note also that $zuz^{-1}vz = zuzzvz = u^{-1}z^{-1}u^{-1}v^{-1}z^{-1}v^{-1}$ reduces m. So we have $m \leq 3$ and hence G is finite. Also $uz^{-1}vzw \to uz^{-1}vz^{-1}z^{-1}w \to uv^{-1}zv^{-1}z^{-1}w$, so elements of G are

- $g = u_1$ (we have $|H|$ of them)
- $g = u_1 x^{\pm 1} u_2$ (we have $2|H|^2$ of them)
- $g = u_1 z u_2 z^{-1} u_3$ (we have $|H|^3$ of them)

So $|G| \leq |H| + 2|H|^2 + |H|^3$. By Cauchy's Theorem, we have $|H| = 3^m$ for some m. So $|G| \leq 3^m + 2 \cdot 3^{2m} + 3^{3m} < 3^{3m+1}$. But $|G|$ is also a power of 3, so $|G| \leq 3^{3m}$. By induction $m \leq 3^{r-2}$ so $|G| \leq 3^{3^{r-1}}$. □

In fact $|B(r, 3)| = 3^{m(r)}$ where $m(r) = r + \binom{r}{2} + \binom{r}{3} = O(r^3)$. The bounds agree for $|B(1, 3)| = 3$ and $|B(2, 3)| = 3^3$. But $|B(3, 3)| \leq 3^9$ when in fact $|B(3, 3)| = 3^7$.

Let G be a group, define lower central series $\gamma^i(G)$ as follows: $\gamma^1(G) = G$ and $\gamma^{r+1}(G) = [G, \gamma^r(G)]$. G is of nilpotent class C if $\gamma^{C+1}(G) = 1$ ($\gamma^C(G) \neq 1$), so note that Class 1 means abelian. [Note G finite and G nilpotent

is the same as all Sylow subgroups normal). $B(r,3)$ is nilpotent of class 3. $|\gamma^2(G)| = 3^{\binom{r}{2}+\binom{r}{3}}, |\gamma^3(G)| = 3^{\binom{r}{3}}$ and $\gamma^4(G) = 1$.

We can also define the derived series of G as follows, $G^{(0)} = G$ and $G^{(r+1)} = [G^{(r)}, G^{(r)}]$. We have $G^{(r)} \leq \gamma^{(r+1)}(G)$ for all $r \geq 0$. G is soluble/solvable of derived length r if $G^{(r+1)} = 1$ and $G^{(r)} \neq 1$. Nilpotent groups are soluble.

The groups $B(r,4)$ are finite, order $2^{k(r)}$ for some $k(r)$. No know formula known fro $k(r)$. WE have $k(1) = 4, k(2) = 12, k(3) = 69, k(4) = 422$ and $k(5) = 2728$. $B(r,4)$ is nilpotent and soluble but its derived length tends to ∞ as r tends to ∞.

Theorem 8.4. *$B(r,4)$ is finite for all $r \geq 1$*

This will follow from:

Theorem 8.5. *Let $G = \langle K, z\rangle$ with $K \leq G$, K finite, $z^2 \in K$ and $g^4 = 1$ for all $g \in G$. Then G is finite.*

Proof of Theorem8.4 (using Theorem 8.5). We do this by induction on r.

For $r = 1$, we have $|B(1,4)| = 4$

Let $r > 1$, $G = \langle x_1, \ldots, x_r\rangle$, let $H = \langle x_1, \ldots, x_{r-1}\rangle$then H is finite by induction. So let $L = \left\langle H, x_r^2\right\rangle$. Then applying Theorem 8.5 with $K = H$ and $z = x_r$, we have L is finite. Then apply Theorem 8.5 again to $K = L$ and $z = x_r$ to get that G is finite. □

Proof of Theorem 8.5. Since $z^2 \in K$, we have $z^{-2} \in K$, since $z^{-1} = zz^{-2}$ with $z^{-2} \in K$, we see that in a word we can replace powers of z by z^0 or z^1 times an element of K. So for any $g \in G$ we have $g = u_0 z u_1 z \ldots u_{m-1} z u_m$ with $u_i \in K$ and $u_1, \ldots, u_{m-1} \neq 1$. Do this with m as small as possible.

Since $(u_i x)^4 = 1$ for $1 \leq i \leq m-1$ we have $zu_i z = u_i^{-1} z^{-1} u_i^{-1} z^{-1} u_i^{-1} = u_i^{-1} z u' z u_i^{-1}$ for some $u' \in K$. This does not reduces m, but it replaces u_{i-1} by $u_{i-1}u_i^{-1}$. Also $g = \ldots u_{i-1} z u_i z \cdots = \ldots u_{i-1} u_i^{-1} z u' z \ldots$. So if $u_{i-1} = u_i$ then we reduce m.

If $m \geq 3$, we can replace u_{m-2} by element in the set $U_{m-2} = \{u_{m-2}, u_{m-2}u_{m-1}^{-1}\}$. Since $u_{m-1} \neq 1$, elements of U_{m-2} are distinct and non-trivial, since otherwise we could reduce m.

Similarly if $m \geq 4$, we can replace u_{m-3} by elements in the set $U_{m-3} = \{u_{m-3}, u_{m-3}t^{-1}, t \in U_{m-2}\}$. All such elements are distinct, since t are distinct and non-trivial so $|U_{m-3}| = 1 + |U_{m-2}| = 2$. None are trivial or we could reduce m.

So by induction if $m \geq i+1$, we can replace u_{m-i} by any elements of $U_{m-i} = \{u_{m-i}, u_{m-i}t^{-1}, t \in U_{m-i+1}\}$. By induction the elements are distinct and non-trivial. So $|U_{m-i}| = i$.

So $i = |U_{m-i}| \leq |K| - 1$, putting $i = m-1$, gives $m \leq |K|$. So we have a bound on m. Hence G is finite as we get $|G| \leq \sum_{i=0}^{|K|+i} |K|^i$, since there are at most $|K|^{i+1}$ words with $m = i$. Using Geometric Series argument, we have $|G| \leq |K|^{|K|+2}$. □

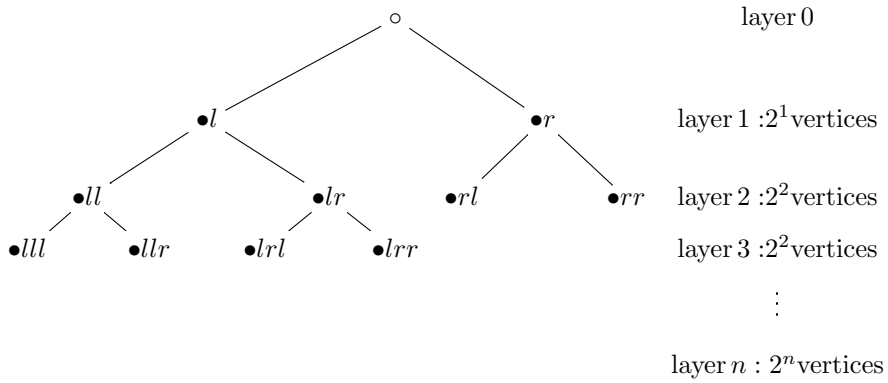
By working harder and using proven properties of U_{m-i}, we can get $|G| \leq |K|^{|K|}$ when $|K| \geq 4$. This gives a very bad inaccurate bound on $|B(r,4)|$, e.g., $|B(2,4)| \leq (4^4)^{(4^4)} = 2^{2048}$ but $|B(2,4)| = 2^{12}$.

$B(2,6)$ is finite, but the proof is longer and works by induction by splitting z^6 into square and a cube. Note for $n = 5$, we cannot split up, $zuz = u^{-1}z^{-1}u^{-1}z^{-1}u^{-1}z^{-1}u^{-1}$ which increases number of z, so no techniques for reducing word length.

8.1 The Grigorchuk Group

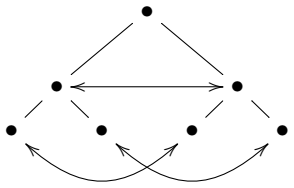
First example of an infinite, finitely generated group in which all elements have a finite order. All elements have order 2^k but exists such elements for all k. (So it doesn't prove any $B(r,n)$ infinite)

Let T be infinite rooted binary tree



At every vertex label v, there is a copy T_v of T rotted at v.

An automorphism of T is a permutation of vertices that maps edges to edges. It must fix root vertex since this has just two neighbours, all other have three. So it must fix the set $\{l, r\}$ of neighbours of roots, i.e. fixes the set of vertices at level 1 - so either fixes l and r or interchange them. By easy induction it fixes the set of vertices at layer n for all $n \geq 0$. Let id_T be the identity map of T and π_T interchanges the corresponding vertices in T_l and T_r



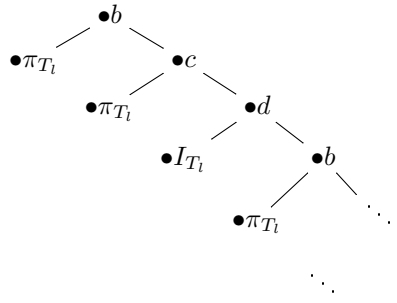
i.e. $\pi_T = (l, r)(ll, rl)(lr, rr)(lll, rll)\dots$

The group G to be constructed is defined as $G = \langle a, b, c, d\rangle$ with $a = \pi_T$, b, c, d all fix l and r so can be defined by specifying their actions on T_l and T_r which is as follows:

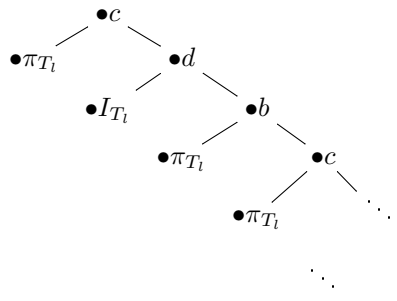
	T_l	T_r
b	a_{T_l}	c_{T_r}
c	a_{T_l}	d_{T_r}
d	I_{T_l}	b_{T_r}

This is a recursive definition. More explicitly

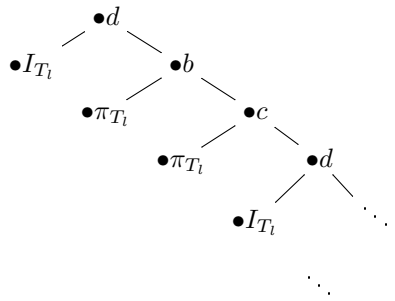
- b



- c



- d



So the recursive definition allows us to calculate actions of b, c, d on any vertex. Clearly $a^2 = 1$. Since for b, c, d all vertices lie in a subtree T_v, which action is a_{T_v} or id_{T_v} we must have $b^2 = c^2 = d^2 = 1$. Furthermore these actions of (b, cd) on T_v are $(\pi, \pi, I), (\pi, I, \pi)$ or (I, π, π), we must have $bc = d$ on all such subtrees. So $bc = d$, in fact $\{1, b, c, d\}$ is a Klein 4 groups and $G = \langle a, b, c\rangle$.

Consider aba: $T_r \to T_l \underset{a}{\to} T_l \to T_r$, so aba acts as a_{T_r} on T_r.

	action on T_l	action on T_r	image under ϕ
b	a	c	(a, c)
c	a	d	(a, d)
d	id	b	(id, b)
aba	c	a	(c, a)
aca	d	a	(d, a)
ada	b	id	(b, id)

Lemma 8.6. *Let $H = \langle b, c, d, aba, aca, ada\rangle$,then $|G : H| = 2$. In fact H is the stabiliser G_l of l and r*

Proof. Clearly $H \leq G_l$, $a \notin G_l$ so $a \notin H$, hence $|G : H| > 1$. So it is enough to prove $G = H \cup Ha$. We use Proposition 4.7. So $G = \langle a, b, c\rangle$ and $S = H \cup Ha$, $g_1 = 1, g_2 = a$

	a	b	c
g_1	$a \in Ha$	$b \in H$	$c \in H$
g_2	$1 \in H$	$ab = (aba)a \in Ha$	$ac = (aca)a \in Ha$

So $S = G$. □

Define a homomorphism $\phi : H \to G \times G$ by $\phi(h) = (\text{action of } h \text{ on } T_l, \text{action of } h \text{ on } T_r)$.

Proposition 8.7. *G is infinite*

Proof. Let ρ_1, ρ_2 be projections of $G \times G$ onto its direct factors, then $a, b, c, d \in \mathrm{im}(\rho_1 \circ \phi)$ (and also in $\mathrm{im}(\rho_2 \circ \phi)$). If G was finite, then $\mathrm{im}(\rho_1 \circ \phi) \geq |G|$, but domain $\rho_1 \circ \phi = H$ and $|H| < |G|$ which is a contradiction. □

Proposition 8.8. *Every $g \in G$ has finite order (a power of 2)*

Proof. Let $g \in G$. Write g as a word of length k, so $g = g_1 \dots g_k$ with $g_i \in \{a, b, c, d\}$. We use induction on k

$k = 0$: $g = 1$ which is fine

$k = 1$, then $g = a, b, c$ or d and $g^2 = 1$

$k > 1$: If bc is a subword, we can replace by d, similarly bd is replaced by c and cd by b. So a, and a letter in $\{b, c, d\}$ alternate in word. Since conjugate elements have same order, if $g_1 = g_k$ then $|g| = |g_2 \dots g_{k-1}|$ which is a power of 2 by induction, so we can assume $g_1 \neq g_k$. If for example $g_1 = b$ and $g_k = c$, then $bgb = g_2 \dots g_{k-1}d$ which is shorter, hence we can assume a is at one end and b, c or d at the other, and by conjugating by a, we can assume $g_1 = a$. So $g = ax_1ax_2 \dots ax_n$ where $k = 2n$ and $x_i \in \{b, c, d\}$.

Case 1. n is even: $g = (ax_1a)x_2(ax_3a)x_4 \dots (ax_{n-1}a)x_n$. So $g \in h$. A product of generators of H, of length n, so $\phi(g) = (w_1, w_2)$ where w_i are words of length $n < k = 2n$. So by induction $|w_1|, |w_2|$ are powers of 2, and hence so $|\phi(g)|$. But ϕ is clearly injective, so $|g|$ is a power of 2.

Case 2. n is odd: $(g \notin H)$, then we have $g^2 = (ax_1a)a_2(ax_3a)x_4 \dots (ax_na)x_1(ax_2) \dots (ax_{n-1}a)x_n$. So each x_i and each ax_ia occurs once in product. Hence $|g^2| = 4n$ with $2n$ bracketed terms in H, $\phi(g^2) = (g_1, g_2), g_1, g_2 \in G$, with $|g_1| = |g_2| = 2n$. So cannot immediately apply induction to g_1, g_2. So we get three more cases

1. If some $x_i = d$, then $\phi(x_i) = (1, b)$ and $\phi(ax_ia) = (b, 1)$ so there is a 1 in words for g_1, g_2. So $|g_1|, |g_2| < 2n$ by induction have order power of 2, hence so does g^2, hence so does g.
2. If some $x_i = c$, then $\phi(x_i) = (a, d)$ and $\phi(ax_ia) = (d, a)$. So both g_1, g_2 involve a d and by case 2.1 we have $|g_1|, |g_2|$ are power of 2, so is g
3. All $x_i = b$, in which case $\phi(x_i) = (a, c)$, $\phi(ax_ia) = (c, a)$ and by case 2.2 we have $|g_1||g_2|$ are a power of 2, and hence so is $|g|$

□

Example. Let $g = ab$, $g^2 = (aba)b$, so $\phi(g^2) = (c, a)(a, c) = (ca, ac)$

$$g^2 = \begin{array}{ccc} & \bullet & \\ ca & & ac \end{array}$$

We have $(ca)^2 = c(aca)$ and $\phi((ca)^2) = (a, d)(d, a)(ad, da)$, $\phi(ac)^2 = (da, ad)$ so

$$g^4 = \begin{array}{ccc} & \bullet & \\ (ca)^2 & & (ca)^2 \end{array} = \begin{array}{ccccccc} & & & \bullet & & & \\ & & \bullet & & \bullet & & \\ ad & & da & & da & & ad \end{array}$$

Finally $\phi(ad^2) = \phi((ada)d) = (b, 1)(1, b) = (b, b) = \phi((da)^2$

$$g^8 = \begin{array}{ccccccc} & & & \bullet & & & \\ & & \bullet & & \bullet & & \\ ad & & da & & da & & ad \end{array} = \begin{array}{ccccccccccccccc} & & & & & & & \bullet & & & & & & & \\ & & & & \bullet & & & & & & \bullet & & & & \\ & & \bullet & & & & \bullet & & & & & & \bullet & & \\ b & & b & & b & & b & & b & & b & & b & & b \end{array}$$

Since $b^2 = 1$, we have $g^{16} = 1$, so $|g| = 16$. Note $b \in H$, $\phi(b) = (a, c)$.

Theorem 8.9. *G has elements of order 2^n for all $n \geq 0$.*

Proof. Let $g = ab$, let $x = g^2, \phi(x) = (ca, ac)$. Let $K = \langle x^G \rangle = \langle g^{-1}xg | g \in G \rangle$. So $x^c = cxc \in K, \phi(x^c) = \phi(cxc) = (a, d)(ca, ac)(a, d) = (ac, dab)$.

$xx^c \in K$, $\phi(xx^c) = (ca, ac)(ac, dab) = (1, x)$. Conjugating by a we get $(x, 1) \in K$. Hence for all $g \in G$ we have

$$h = \begin{array}{ccc} & \bullet & \\ g & & g' \end{array} \in H$$

for some g'.

$$h^{-1}xh = \begin{array}{ccc} & \bullet & \\ g^{-1}xg & & 1 \end{array} \in K$$

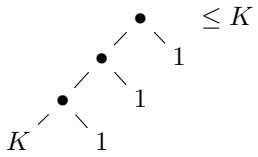
, this is true for all $g \in G$. Since $K = \langle g^{-1}xg | g \in G \rangle$, we have

$$\begin{array}{ccc} & \bullet & \\ K & & 1 \end{array} \leq K$$

and

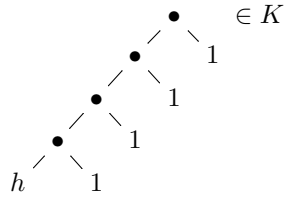
$$\begin{array}{ccc} & \bullet & \\ & \bullet & 1 \\ K & & 1 \end{array} \leq K$$

and

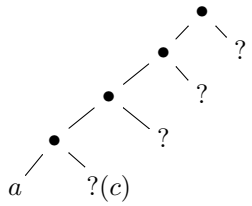


etc.

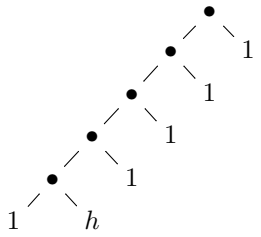
We now prove K has elements of order 2^n for all $n \geq 1$. This is true for $n = 0, 1, 2, 3$ ($|x| = 8$). For induction, let $h \in K$ with $|h| = 2^n$. By above

$\overline{h} =$  $\in K$

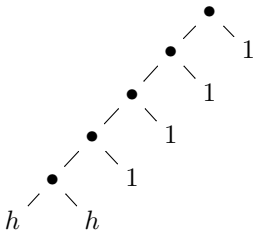
, we saw that

$g^8 =$ 

so

$g^8\overline{h}g^8 =$ 

so

$(\overline{h}g^8)^2 = \overline{h}g^8\overline{h}g^8 =$ 

Since $|h| = 2^n$ by assumption, so $|(\overline{h}g^8)| = 2^n$, so $|\overline{h}g^8| = 2^{n+1}$ completing the induction. □

Let $G = \langle X \rangle$, X finite. Let $g = a_1 a_2 \dots a_n$ with $a_i \in A^* = (X \cup X^{-1})^*$. For $g \in G$, define length of $|g|$ to be the minimum n with $g = a_1 \dots a_n$, i.e., the length of the shortest word for g (it depends on X). Define the growth function $\lambda_{G,X} : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ by $\lambda_{G,X}(n) = |\{g \in G | |g| \leq n\}|$ (i.e., the ball of radius n around 1).

For free group of rank k, $|x| = k$, $\lambda_{G_k}(n) = 1 + 2k + \sum_{i=2}^{n}(2k)(2k-1)^{i-1} = O((2k-1)^n)$, its an exponential function of n. So G has exponential growth (this property is independent of X)

Let $G = \langle g_1, \dots, g_k | [g_i, g_j] = 1 \rangle$, free abelian group of rank k. Then $\lambda_{G,X}(n) = O(n^k)$, polynomial in n. We say G has polynomial growth (again this property is independent of X)

Theorem 8.10 (Hard (Gromov))**.** *Let G be finitely generated, G has polynomial growth if and only if there exists $H \leq G$ with $|G : H|$ finite and H nilpotent.*

Grigorchuk's group was the first example with intermediate growth, that is $\lambda_{G,X}$ less than all polynomial in n, but less than all exponential functions in n.