

MATH 101A: ALGEBRA I
PART A: GROUP THEORY

In our unit on group theory we studied solvable and nilpotent groups paying particular attention to p -groups. There was also a heavy emphasis on categorical notions such as adjoint functors, limits and colimits.

CONTENTS

1. Preliminaries	1
1.1. basic structures	1
1.2. commutators	2
2. Solvable groups	4
2.1. two definitions	4
2.2. degree of solvability and examples	5
2.3. subgroups and quotient groups	6
3. Operations of a group on a set	8
3.1. definition and basic properties	8
3.2. examples	8
3.3. morphisms of G -sets	10
4. p -groups and the class formula	11
4.1. orbit-stabilizer formula	11
4.2. actions of p -groups	11
4.3. class formula	12
5. Nilpotent groups	13
6. Sylow Theorems	15
7. Category theory and products	18
7.1. categories	18
7.2. product of groups	18
7.3. categorical product	20
7.4. products of nilpotent groups	21
8. Products with many factors	23
8.1. product of many groups	23
8.2. weak product	24
8.3. recognizing weak products	24
8.4. products of Sylow subgroups	25
9. Universal objects and limits	27
9.1. initial and terminal objects	27
9.2. products as terminal objects	28

9.3.	general limits	28
9.4.	existence of limits	30
9.5.	examples	31
10.	Limits as functors	34
10.1.	functors	34
10.2.	the diagram category	34
10.3.	The limit is a functor	35
10.4.	adjoint functors	37
11.	Categorical limits	39
11.1.	colimits in the category of sets	39
11.2.	pull-back	40
11.3.	push-forward [push-out] of groups	40
11.4.	direct limit of groups	41
11.5.	universal property of the direct limit of groups	42
11.6.	free groups	42
11.7.	an important example	43
12.	More about free products	44
12.1.	Amalgamated products again	44
12.2.	free group as adjoint functor	46
12.3.	actions and free products	46

1. PRELIMINARIES

1.1. **basic structures.** First, I defined semigroups, monoids and groups.

Definition 1.1. A *semigroup* is a set S with an associative binary operation:

$$* : S \times S \rightarrow S$$

written $(a, b) \mapsto ab$ so that $a(bc) = (ab)c$ for all $a, b, c \in S$. A *monoid* is a semigroup M with a (two-sided) *identity* (or *neutral element*). This is defined to be an element $e \in M$ so that

$$xe = ex = x$$

for all $x \in M$. A *group* is a monoid G with *inverses*. This means $\forall x \in G, \exists y \in G$ s.t.

$$xy = yx = e.$$

$y = x^{-1}$ is called the *inverse* of x .

I gave examples similar to the following.

- (1) The set of positive integers is a semigroup under addition.
- (2) (\mathbb{Z}, \cdot) is a monoid.
- (3) $(2\mathbb{Z}, +)$ is a group.

Definition 1.2. A *homomorphism* between two groups is a mapping $\phi : G \rightarrow H$ so that

$$\phi(ab) = \phi(a)\phi(b).$$

Theorem 1.3. Any homomorphism of groups also takes identity to identity and inverse to inverse:

$$\begin{aligned}\phi(e_G) &= e_H \\ \phi(x^{-1}) &= \phi(x)^{-1}\end{aligned}$$

Definition 1.4. A subset H of a group G is called a *subgroup* if the following conditions hold:

- (1) $H \neq \emptyset$ (H is nonempty.)
- (2) $HH \subseteq H$ (H is closed under multiplication.)
- (3) $H^{-1} \subseteq H$ (H is closed under inverse.)

The notation for subgroup is $H \leq G$.

Here I am using the notation:

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

Conditions (2) and (3) can be combined into one condition:

$$HH^{-1} = H.$$

Theorem 1.5. *The image $\phi(G)$ of a homomorphism $\phi : G \rightarrow H$ is a subgroup of H .*

Proof. $K = \phi(G)$ is nonempty since it contains $\phi(e_G) = e_H$. It also satisfies the condition $KK^{-1} = K$ since

$$\phi(G) = \phi(GG^{-1}) = \phi(G)\phi(G^{-1}) = \phi(G)\phi(G)^{-1}$$

by Theorem 1.3. □

The *left cosets* of H in G are the sets

$$aH = \{ah \mid h \in H\}$$

The set of left cosets is written G/H . For example,

$$\mathbb{Z}/2\mathbb{Z} = \{2\mathbb{Z}, 2\mathbb{Z} + 1\}.$$

Definition 1.6. A subgroup $N \leq G$ is *normal* and we write $N \trianglelefteq G$ if

$$xNx^{-1} = N$$

for all $x \in G$.

Theorem 1.7. *A subgroup $N \leq G$ is normal if and only if G/N is a group under the operation*

$$(aN)(bN) = abN.$$

Definition 1.8. The *kernel* $\ker \phi$ of a homomorphism $\phi : G \rightarrow H$ is the set of all elements of G which go to the identity of H .

Theorem 1.9. $\ker \phi \trianglelefteq G$. *Furthermore any normal subgroup $N \trianglelefteq G$ is the kernel of some homomorphism $G \rightarrow H$.*

The homomorphism is just the quotient map $q : G \rightarrow G/N$ given by $q(a) = aN$.

The rest of the discussion was focused on commutators and abelian quotient groups.

1.2. commutators. A group G is called *abelian* if $ab = ba$ for all $a, b \in G$. This is the same as saying that the *commutator*

$$[a, b] := aba^{-1}b^{-1}$$

is trivial (equal to e).

To define the commutator subgroup, I needed to recall the definition of a subgroup generated by a subset.

Definition 1.10. If S is a subset of a group G then the *subgroup generated by S* , written $\langle S \rangle$, is defined to be the intersection of all subgroups of G which contain S :

$$\langle S \rangle := \bigcap_{S \subseteq H \leq G} H.$$

Definition 1.11. The *commutator subgroup* (also called the *derived subgroup*) $G' = [G, G]$ is the subgroup of G generated by all commutators $[a, b]$.

In general, if $A, B \leq G$ then $[A, B]$ is defined to be the subgroup of G generated by all commutators $[a, b]$ where $a \in A$ and $b \in B$.

Theorem 1.12. $G' = [G, G]$ is a normal subgroup of G .

This follows immediately from the following two facts.

Lemma 1.13. (1) $x \langle S \rangle x^{-1} = \langle S \rangle$
 (2) $x[a, b]x^{-1} = [xax^{-1}, xbx^{-1}]$.

This lemma has a generalization:

Lemma 1.14. Given a homomorphism $\phi : G \rightarrow H$, $a, b \in G$, $S \subseteq G$ we have:

- (1) $\phi \langle S \rangle = \langle \phi(S) \rangle$.
- (2) $\phi[a, b] = [\phi(a), \phi(b)]$.

Why is this a generalization of the previous lemma?

The main theorem about the commutator subgroup is the following.

Theorem 1.15. The image $\phi(G)$ of a homomorphism $\phi : G \rightarrow H$ is abelian if and only if the kernel of ϕ contains the commutator subgroup

For this we need the following lemma whose proof is obvious.

Lemma 1.16. S is a subset of $H \leq G$ iff $\langle S \rangle$ is a subgroup of H .

Proof of Theorem 1.15. The argument which we did in class is reversible, i.e., “iff” at every step: For any $a, b \in G$ we have

$$[\phi(a), \phi(b)] = \phi[a, b].$$

$\phi(G)$ is abelian iff the LHS is always e . But, the RHS is always equal to e iff $[a, b] \in \ker \phi$ for all $a, b \in G$ which, by Lemma 1.16, is equivalent to saying that $G' \leq \ker \phi$. \square

The theorem has the following variation as an obvious corollary:

Corollary 1.17. Suppose that $N \trianglelefteq G$. Then G/N is abelian iff $G' \leq N$.

2. SOLVABLE GROUPS

After doing a review of group theory at lightning speed we managed to get to the first topic of the course on the first day: Solvable groups.

2.1. **two definitions.** I gave two equivalent definitions of a solvable group. Here is the first definition.

Definition 2.1. A group G is *solvable* if an iterated derived subgroup $G^{(n)}$ is trivial for some positive integer n . Here $G^{(n)}$ is defined recursively as follows.

- (1) $G^{(0)} = G$
- (2) $G^{(i+1)} = (G^{(i)})' = [G^{(i)}, G^{(i)}]$.

Serge Lang uses normal towers to define solvable groups.

Definition 2.2. A *normal tower* for a group G is a sequence of subgroups:

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_n = \{e\}$$

so that each subgroup is normal in the previous one: $G_i \trianglelefteq G_{i-1}$. The quotient groups G_{i-1}/G_i are called the *subquotients* of the tower.

In general a *subquotient* of a group G is a quotient of a subgroup of G . (This is more general than a subgroup of a quotient. Why is that?) Here is Lang's definition of a solvable group.

Definition 2.3. A group is *solvable* if it has a normal tower whose subquotients are all abelian. Lang calls these *abelian towers*.

We need to show that these definitions are equivalent. One direction is obvious. The first definition implies the second. This is because the derived series:

$$G \supseteq G' \supseteq G'' \supseteq \cdots \supseteq G^{(n)} = \{e\}$$

is a normal tower with abelian subquotients. To prove the converse, we need one more corollary or Theorem 1.15.

Corollary 2.4. *Suppose that $H \leq G$, $N \trianglelefteq G$ and G/N is abelian. Then $H' \leq H \cap N$.*

Proof. Take the composition

$$\phi: H \hookrightarrow G \twoheadrightarrow G/N.$$

Since $\phi(H) \leq G/N$ is abelian,

$$H' \leq \ker \phi = \{x \in H \mid \underbrace{\phi(x)}_{x \in N} = e\} = H \cap N.$$

□

Coming back to the equivalence of definitions, suppose that G has a normal tower with abelian subquotients. Since G/G_1 is abelian, $G' \leq G_1$. Suppose by induction that $G^{(k)} \leq G_k$. We know that $G_{k+1} \trianglelefteq G_k$ with abelian quotient. The corollary tells us that

$$(G^{(k)})' = G^{(k+1)} \leq G^{(k)} \cap G_{k+1} \leq G_{k+1}.$$

Therefore $G^{(n)} \leq G_k = \{e\}$ making G solvable by the first definition.

2.2. degree of solvability and examples. We say that G is solvable of degree n if $G^{(n)} = \{e\}$ and $G^{(n-1)}$ is nontrivial.

- (1) Abelian groups are solvable of degree 1 (except for the trivial group which is solvable of degree 0).
- (2) S_3 , the symmetric group on 3 letters is solvable of degree 2.
- (3) $T(n, \mathbb{Z})$, the group of unipotent matrices with coefficients in \mathbb{Z} is solvable, but of what degree?

To prove that S_3 is solvable, take the normal tower:

$$S_3 \supseteq A_3 \supseteq \{e\}.$$

Here $A_3 = \{e, (123), (132)\}$ is the alternating group. This is a cyclic group and thus abelian and $S_3/A_3 \cong \mathbb{Z}/2$ is also abelian. So, S_3 is solvable of degree 2.

As I mentioned in class, *unipotent* means upper triangular with 1's on the diagonal. Any unipotent matrix can be written in the form $I_n + X$ where I_n is the $n \times n$ identity matrix and X is a strictly upper triangular matrix, i.e.,

$$x_{ij} = 0 \text{ unless } j \geq i + 1$$

Let U_k be the set of strictly upper triangular matrices $X = (x_{ij})$ so that

$$x_{ij} = 0 \text{ unless } j \geq i + k$$

Then

$$U_j U_k \subseteq U_{j+k}$$

Therefore, every element of U_k , $k \geq 1$ is nilpotent: $X^n = 0$. This means that

$$(I + X)^{-1} = I - X + X^2 - X^3 + \cdots + (-1)^{n-1} X^{n-1} = I - X(I + X)^{-1}.$$

Let

$$T_k = \{I + X \mid X \in U_k\}.$$

Then $T(n, \mathbb{Z}) = \{I + X \mid X \in U_1\} = T_1$.

Lemma 2.5. $T_1 \supseteq T_2 \supseteq T_3 \supseteq \cdots \supseteq T_n = \{I\}$ is a normal tower with abelian subquotients.

Proof. Take arbitrary elements $I + X, I + Y$ in T_1, T_k resp. Then

$$(I+X)(I+Y)(I+X)^{-1} = (I+X+Y+XY)(I+X)^{-1} = I+(Y+XY)(I+X)^{-1}$$

This is an element of T_k since $(Y + XY)(I + X)^{-1} \in U_k U_0 \subseteq U_k$.

Therefore, $T_k \trianglelefteq T_1$.

If $X, Y \in U_k$ then expanding $(I + X)^{-1}$ as $I - X(I + X)^{-1}$ we get:

$$(I + X)(I + Y)(I + X)^{-1} = I + Y + XY - (Y + XY)X(I + X)^{-1}.$$

$$[I + X, I + Y] = I + (XY - (Y + XY)X(I + X)^{-1})(I + Y)^{-1} \in T_{2k}$$

Therefore, T_k/T_{2k} is abelian. \square

This shows that $T(n, \mathbb{Z})$ is solvable of degree $\leq k$ if $n \leq 2^k$.

2.3. subgroups and quotient groups. We want to know that subgroups and quotient groups of solvable groups are solvable.

Theorem 2.6. *Every subgroup of a solvable group is solvable.*

Proof. If $H \leq G$ then $H' \leq G'$. This, in turn, implies that $(H')' = H^{(2)} \leq G^{(2)} = (G')'$. Eventually we get $H^{(n)} \leq G^{(n)} = \{e\}$. So, H is solvable of degree $\leq n$. \square

Theorem 2.7. *Every quotient group of a solvable group is solvable.*

Proof. I claim that

$$G/N \supseteq G'N/N \supseteq G^{(2)}N/N \supseteq \cdots \supseteq G^{(n)}N/N = \{e\}$$

is a normal tower for G/N with abelian quotients. In fact $G^{(k)}N/N$ is the image of $G^{(k)}$ under the homomorphism

$$G^{(k)} \hookrightarrow G \twoheadrightarrow G/N.$$

Therefore,

$$G^{(k)}N/N = (G/N)^{(k)}.$$

This uses the following lemma. \square

Lemma 2.8. *If $\phi : G \twoheadrightarrow H$ is an epimorphism (surjective homomorphism) then $\phi(G^{(k)}) = H^{(k)}$ for all $k \geq 0$.*

Theorem 2.9. *Suppose that $N \trianglelefteq G$ and $N, G/N$ are solvable. Then G is solvable.*

Proof. We are given that $N, G/N$ are solvable. So, we have abelian towers

$$N \supseteq N_1 \supseteq N_2 \cdots N_n = \{e\}$$

Since subgroups of G/N all have the form H/N for some $N \leq H \leq G$, the abelian tower for G/N looks like this:

$$G/N \supseteq G_1/N \supseteq G_2/N \cdots G_m/N = \{e\}$$

Where $G_m = N$. By the following lemma the abelian subquotients of this tower are

$$\frac{G^{(k)}/N}{G^{(k+1)}/N} \cong G^{(k)}/G^{(k+1)}$$

Therefore,

$$G \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_m = N \supseteq N_1 \supseteq \cdots \supseteq N_n = \{e\}$$

is a normal tower for G with abelian subquotients proving that G is solvable of degree $\leq n + m$. \square

Lemma 2.10. *If N, H are normal subgroups of G with $N \leq H$ then*

$$\frac{G/N}{H/N} \cong \frac{G}{H}.$$

Proof. Let $\phi : G/N \rightarrow G/H$ be the homomorphism given by $\phi(aN) = aNH = aH$. Then ϕ is clearly onto and $\ker \phi = \{aN \in G/N \mid aH = H\}$. But $aH = H$ iff $a \in H$. So, $\ker \phi = H/N$. The lemma follows from the equation

$$\text{image}(\phi) = \frac{\text{domain}(\phi)}{\ker \phi}$$

\square

3. OPERATIONS OF A GROUP ON A SET

3.1. definition and basic properties.

Definition 3.1. If X is a set, the *permutation group* of X , $Perm(X)$ is the group of all bijections $f : X \rightarrow X$ under composition (i.e., with composition as the group law).

For example, if $X = \{1, 2, 3\}$ then $Perm(X) = S_3$ is the permutation group on three letters.

Definition 3.2. An *action* of a group G on a set X is a homomorphism

$$\pi : G \rightarrow Perm(X)$$

sending every $g \in G$ to a permutation π_g of X .

The notation is $\pi_g(x) = gx$.

Definition 3.3. If $x \in X$, the *orbit* of x is the set

$$\pi_G(x) := \{\pi_g(x) = gx \mid g \in G\}.$$

The *stabilizer* of x in G is the subgroup

$$G_x := \{g \in G \mid gx = x\}.$$

Theorem 3.4. (1) X is a disjoint union of orbits: $X = \coprod \pi_G(x_i)$.

(2) The stabilizer of any two elements of the same orbit are conjugate.

Proof. I only proved the second part. In fact I showed that

$$G_{gx} = gG_xg^{-1}$$

by the following reversible proof:

$$a \in G_{gx} \Leftrightarrow agx = gx \Leftrightarrow g^{-1}agx = x \Leftrightarrow g^{-1}ag \in G_x \Leftrightarrow a \in gG_xg^{-1}.$$

□

In the lecture I realized that this proof assumes the following lemma.

Lemma 3.5. Two elements $x, y \in X$ lie in the same orbit of an action of G if and only if $y = gx$ for some $g \in G$.

Proof. Suppose that x, y lie in the orbit $\pi_G(z)$. Then they can be written as $x = hz, y = kz$. So, $z = h^{-1}x$ and $y = kh^{-1}x$. □

3.2. **examples.** I believe I gave three examples of group actions.

3.2.1. *example: left multiplication.* Let H be a subgroup of a group G . Then the group H acts on the set G by left multiplication:

$$\lambda_h(x) = hx.$$

What does it mean that this is an action? The definition says that this must be a homomorphism

$$\lambda : H \rightarrow \text{Perm}(G)$$

and this is just one equation:

$$\lambda_g \pi_h = \lambda_{gh}$$

Verification is trivial:

$$\lambda_g \lambda_h(x) = \lambda_g(hx) = ghx = \lambda_{gh}(x).$$

I also pointed out that right multiplication $\rho_h(x) = xh$ is not an action.

Questions:

- (1) What are the orbits of this action?
- (2) What is the stabilizer subgroup of $g \in G$?

Answers:

- (1) The orbits are the right cosets of H .
- (2) The stabilizers are all trivial.

3.2.2. *example: conjugation of elements.* Let G be any group and take the action of G on G by conjugation:

$$\gamma_g(x) = gxg^{-1}.$$

The orbits of this action are the *conjugacy classes* of elements of G :

$$\gamma_G(x) = \{gxg^{-1} \mid g \in G\}.$$

I write this as $C(x)$.

The stabilizers of the action also have a name. They are called the *centralizers* of the elements of G :

$$Z_G(x) := \{g \in G \mid gx = xg\}.$$

For example, if $G = S_3$ then

$$\begin{array}{ll} \gamma_G(e) = \{e\} & G_e = S_3 \\ \gamma_G(12) = \{(12), (23), (13)\} & G_{(12)} = \{e, (12)\} \\ \gamma_G(123) = \{(123), (132)\} & G_{(123)} = A_3 \end{array}$$

3.2.3. *example: conjugation of subgroups.* Let G be any group and let S be the set of all subgroups of G . Then G acts on S by conjugation:

$$\gamma_g(H) = gHg^{-1}.$$

What can we say about the orbits and stabilizers of this action?

The only things I could think of were the following.

- (1) A subgroup $N \leq G$ is normal if and only if its orbit under the conjugation action is a singleton: $\gamma_G(N) = \{N\}$. (A *singleton* is a set with one element.)
- (2) The stabilizer subgroup of $H \leq G$ is the *normalizer*

$$N_G(H) := \{g \in G \mid gHg^{-1} = H\}.$$

- (3) The stabilizer of a normal subgroup is the whole group: $N_G(N) = N$.

For example, if $G = S_3$ then S has six elements:

$$\{e\}, \langle(12)\rangle, \langle(13)\rangle, \langle(23)\rangle, A_3, S_3$$

These form four orbits: $\{e\}, A_3, S_3$ are normal. So, they are singleton orbits. The other three form one orbit. The stabilizers of the normal subgroups are the whole group. The stabilizer of $\langle(ab)\rangle$ is itself. (These subgroups are *self-normalizing*.)

3.3. morphisms of G -sets.

Definition 3.6. If X, Y are G -sets (sets with G -action), a *morphism* of G -sets is a mapping

$$f : X \rightarrow Y$$

which commutes with the action of G . This means that

$$f \circ \pi_g = \pi_g \circ f$$

for all $g \in G$. Equivalently, $f(gx) = gf(x)$ for all $g \in G, x \in X$. You can also write this as a commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \pi_g \downarrow & & \downarrow \pi_g \\ X & \xrightarrow{f} & Y \end{array}$$

One example of a morphism of G -sets is given by right multiplication. Take $X = Y = G$ with group H acting by left multiplication. Then for a fixed $g \in G$ right multiplication by g is a morphism of H -sets since

$$\lambda_h \circ \rho_g = \rho_g \circ \lambda_h.$$

4. p -GROUPS AND THE CLASS FORMULA

The class formula is used to prove that p -groups are nilpotent. The class formula in turn comes from the orbit-stabilizer formula.

4.1. orbit-stabilizer formula. This follows from the following theorem.

Theorem 4.1. *Suppose that G acts on a set X and $x \in X$. Then there is a bijection between the orbit $\pi_G(x)$ of x and the set of left cosets of the stabilizer G_x of x . The bijection*

$$\phi : G/G_x \xrightarrow{\cong} \pi_G(x)$$

is given by $\phi(gG_x) = gx$.

When X is finite, we get a numerical formula.

Corollary 4.2 (orbit-stabilizer formula). *If X is a G -set and $x \in X$, the size of the orbit of x is equal to the index of the stabilizer of x .*

Recall that X is a disjoint union of orbits:

$$X = \coprod_{x_i \text{ rep}} \pi_G(x_i)$$

where $x_i \in X$ are representatives of the orbits (i.e., one element from each orbit).

Corollary 4.3 (orbit-sum formula). *The number of elements in a finite G -set X is given by*

$$|X| = \sum_{x_i \text{ rep}} |\pi_G(x_i)| = \sum_{x_i \text{ rep}} |G : G_{x_i}|.$$

4.2. actions of p -groups. One example of the orbit-sum formula is given by the action of p -groups on finite sets. If P is a p -group, with $|P| = p^k$ then every subgroup has order a power of p . If P acts on a finite set X then the size of each orbit is also a power of p :

$$|\pi_P(x)| = |P : P_x| = \frac{|P|}{|P_x|} = \frac{p^k}{p^j} = p^{k-j}.$$

Notice that this is divisible by p except when it is equal to 1. The size of X is given by the orbit-sum formula:

$$|X| = \sum_{x_i} |P : P_{x_i}| = \sum_{x_i} p^{k-j_i}.$$

Now we want to separate the summands which are equal to 1 and those which are greater than 1. If $|P : P_x| = 1$ then $P_x = P$ and x is a *fixed*

point of the action. This means $g \cdot x = x$ for all $g \in P$. The other orbits have more than one element and therefore the size of these orbits is divisible by p . This gives the following theorem which we need later.

Theorem 4.4. *Suppose that P is a p -group acting on a finite set X . Then the number of elements in X is congruent modulo p to the number of fixed points of the action.*

The key use of this formula is the following.

Corollary 4.5. *Every nontrivial finite p -group has a nontrivial center.*

Proof. Suppose that P is a p -group with $p^k > 1$ elements. Then P acts on P by conjugation. The theorem says that the number of fixed points of this action is congruent to $|P| = p^k$ modulo p . In other words, p divides the number of fixed points. But $g \in P$ is a fixed point of the conjugation action if and only if $g \in Z(P)$. Therefore, p divides $|Z(P)|$ which implies that $Z(P)$ has at least p elements. \square

4.3. class formula. This is another example of the orbit-sum formula. Take a finite group G acting on the set G by conjugation. Then the orbit of $x_i \in G$ is the conjugacy class of x_i and the stabilizer is the centralizer $C_G(x_i)$ of x_i . This gives the following formula.

$$|G| = \sum_{x_i} |\gamma_G(x_i)| = \sum_{x_i} |G : C_G(x_i)|$$

This is not the class formula. We need to separate the summands which are equal to 1. By the orbit-stabilizer formula,

$$|\gamma_G(x)| = |G : G_x| = |G : C_G(x)| = 1$$

This means $C_G(x) = G$. In other words, x is central, or $x \in Z(G)$. You can also just look at the definition:

$$\gamma_G(x) = \{gxg^{-1} \mid g \in G\}.$$

This is $\{x\}$ if and only if $gxg^{-1} = x$, i.e., $x \in Z(G)$.

Every central element contributes 1 to the sum and every noncentral element contributes a number $|G : C_G(x)| \neq 1$. This gives the following.

Theorem 4.6 (class formula). *The number of elements in any finite group G is given by*

$$|G| = |Z(G)| + \sum_{x_i} |\gamma_G(x_i)| = |Z(G)| + \sum_{x_i} |G : C_G(x_i)|$$

where the x_i are representatives of the conjugacy classes in G which contain more than one element.

This formula is used to prove Corollary 4.5 but we already did that.

5. NILPOTENT GROUPS

There are two definitions of nilpotent groups. I don't remember whether I proved that they are equivalent but we can do that here.

Definition 5.1. A group G is *nilpotent* if there is a normal tower

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_n = \{e\}$$

with the following properties for all i .

- (1) $G_i \trianglelefteq G$ and
- (2) $G_{i-1}/G_i \leq Z(G/G_i)$.

A normal tower satisfying these conditions will be called a *central series*.

The smallest possible value of n is called the *nilpotency class* of G . Thus G has nilpotency class 1 if and only if it is abelian (and nontrivial).

The other definition that I gave used iterated centers $Z^{(i)}(G)$ defined recursively as follows.

- (1) $Z^{(1)}(G) := Z(G)$.
- (2) $Z^{(n+1)}(G)$ is the unique normal subgroup of G which contains $Z^{(n)}(G)$ and so that

$$\frac{Z^{(k+1)}(G)}{Z^{(k)}(G)} = Z\left(\frac{G}{Z^{(k)}(G)}\right)$$

Notice that this recursive definition implies that

$$(5.1) \quad \frac{Z^{(n)}(G)}{Z(G)} = Z^{(n-1)}\left(\frac{G}{Z(G)}\right).$$

Proposition 5.2. A group G is nilpotent if and only if $G = Z^{(n)}(G)$ for some n . Furthermore, the smallest such n is equal to the nilpotency class of G .

Proof. If $G = Z^{(n)}(G)$ then the tower

$$G = Z^{(n)}(G) \supseteq Z^{(n-1)}(G) \supseteq \cdots \supseteq Z(G) \supseteq \{e\}$$

is a central series. Therefore, G is nilpotent of class $c \leq n$.

Conversely, suppose that G is nilpotent of class c . Then we have a central series:

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_{c-1} \supseteq G_c = \{e\}$$

Claim: $G_{c-k} \leq Z^{(k)}(G)$ for all k .

If this is true then, putting $k = c$, we get that $G = Z^{(c)}(G)$, which means that $c \geq n$ where n is the smallest number satisfying

$G = Z^{(n)}(G)$. Since $c \geq n$ and $c \leq n$ we conclude that $c = n$. So, it suffices to prove the claim.

The claim holds for $k = 1$ by the assumption that G_i form a central series. Suppose by induction on k that $G_{c-k} \leq Z^{(k)}(G)$ and we have the quotient map:

$$\phi : G/G_{c-k} \rightarrow G/Z^{(k)}(G).$$

Any epimorphism has the property that it sends central elements into central elements. Therefore,

$$\phi \left(\frac{G_{c-k-1}}{G_{c-k}} \right) \leq \phi \left(Z \left(\frac{G}{G_{c-k}} \right) \right) \leq Z \left(\frac{G}{Z^{(k)}(G)} \right) = \frac{Z^{(k+1)}(G)}{Z^{(k)}}.$$

Since ϕ is the quotient map this implies that $G_{c-k-1} \leq Z^{(k+1)}(G)$. Therefore, the claim and thus the proposition holds. \square

Corollary 5.3. *A nontrivial group G is nilpotent of class c if and only if $G/Z(G)$ is nilpotent of class $c - 1$.*

Proof. This follows from the proposition and Equation (5.1). G is nilpotent of class $\leq c$ if and only if $Z^{(c)}(G) = G$. By (5.1) this is equivalent to saying that $Z^{(c-1)}(G/Z(G)) = G/Z(G)$. By the proposition the last statement is equivalent to saying that $G/Z(G)$ is nilpotent of class $\leq c - 1$. Putting these together we get the corollary. \square

Corollary 5.4. *Finite p -groups are nilpotent.*

Proof. If P is a p -group then $|P| = p^k$ for some k . If $k = 1$ then P is cyclic and therefore abelian making it nilpotent of class 1. If $k > 1$ then P has a nontrivial center $Z(P)$ and $P/Z(P)$ is nilpotent by induction. By the previous corollary this implies that P is nilpotent. \square

6. SYLOW THEOREMS

This proof is from Alperin and Bell “Groups and representations.”

Lemma 6.1. *If $p \nmid m$ then*

$$\binom{p^k m}{p^k} \equiv m \pmod{p}$$

Remark 6.2. Alperin and Bell do not bother to prove this lemma. Instead they make the following bizarre logical argument. They show that the truth value of Lemma 6.1 is equal to the truth value of the Sylow theorems for any group G with $p^k m$ elements. Since the Sylow theorems are true for the cyclic group of order $p^k m$, Lemma 6.1 must be true and therefore the Sylow theorems hold for all finite groups!!

Proof. The binomial coefficient theorem says

$$(1+x)^p = 1 + px + \binom{p}{2}x^2 + \binom{p}{3}x^3 + \cdots + px^{p-1} + x^p.$$

Modulo p this gives the formula $(1+x)^p \equiv 1^p + x^p = 1 + x^p$. By induction on k this gives

$$(1+x)^{p^k} \equiv 1 + x^{p^k}$$

Raising both sides of this equation to the m th power we get:

$$(1+x)^{p^k m} \equiv (1+x^{p^k})^m = 1 + mx^{p^k} + \binom{m}{2}x^{2p^k} + \cdots$$

The coefficient of x^{p^k} on the left is $\binom{p^k m}{p^k}$. This must be congruent to the corresponding coefficient on the right which is m . \square

Lemma 6.3. *Suppose that N is a normal subgroup of a finite group G and $p \nmid |G:N|$. Then N contains every p -subgroup of G .*

Remark 6.4. In particular this means that a p -subgroup Q of G can normalize a p -Sylow subgroup P only if $Q \leq P$. [Since $P \trianglelefteq N(P)$ and $p \nmid |N(P):P|$.]

Proof. Suppose not. Then G contains a p -subgroup P so that $P \not\leq N \cap P$. Then

$$\frac{PN}{N} \cong \frac{P}{N \cap P}$$

is a nontrivial p -subgroup of G/N . So p divides $|G/N|$. \square

Lemma 6.5. *Suppose that P is a p -group acting on a finite set X . Then the number of elements of X is congruent mod p to the number of fixed points of the action of P on X .*

Proof. The size of the orbit Px of any $x \in X$ is equal to the index of its stabilizer $P_x = \{g \in P \mid gx = x\}$ which must be a power of p :

$$|Px| = |P/P_x| = \frac{|P|}{|P_x|} = \frac{p^k}{p^\ell} = p^{k-\ell}$$

When x is not a fixed point this number is not 1 so it must be divisible by p . Thus p divides the number of non-fixed points in X . The lemma follows. \square

Theorem 6.6 (Sylow). *Suppose that G is a finite group of order $p^k m$ where $p \nmid m$. Then*

- (1) G contains a p -Sylow subgroup (i.e., a subgroup P of order p^k) and every p -subgroup of G is contained in some p -Sylow subgroup of G .
- (2) Any two p -Sylow subgroups of G are conjugate.
- (3) The set \mathcal{P} of p -Sylow subgroups of G has $|\mathcal{P}| \equiv 1 \pmod{p}$ elements.

Proof. Let \mathcal{X} be the set of all subsets $X \subseteq G$ with $|X| = p^k$ elements. Then \mathcal{X} has

$$|\mathcal{X}| = \binom{p^k m}{p^k} \equiv m \pmod{p}$$

number of elements by Lemma 6.1. Note that $\mathcal{P} \subseteq \mathcal{X}$, i.e., every Sylow- p -subgroup is an element of \mathcal{X} .

There is a left action of G on \mathcal{X} :

$$\lambda_g(X) = \{gx \mid x \in X\} \in \mathcal{X}$$

The size of the orbit of X is given by the index of its stabilizer $H = G_X$:

$$|\text{orbit}(X)| = |G/H| = |G : H|$$

where

$$H = G_X = \{g \in G \mid gX = X\}$$

But

$$HX = \bigcup_{g \in H} gX = X = \bigcup_{x \in X} Hx$$

is a union of right cosets of H so $|X| = p^k$ is a multiple of $|H|$. Thus $|G_X| = |H| = p^\ell$ where $\ell \leq k$. Thus every stabilizer is a p -subgroup of G .

The number of elements in the orbit of X is:

$$|\text{orbit}(X)| = \frac{|G|}{|G_X|} = \frac{p^k m}{p^\ell} = mp^{k-\ell} \geq m$$

We will say that the orbit of X is *small* if it has exactly m element. Otherwise it is *large*. Large orbits have size divisible by p .

Claim The orbit of X is small if and only if X is a right coset of a p -Sylow subgroup of G .

Proof: For a small orbit, the stabilizer $H = G_X$ is a p -Sylow subgroup of G and X is a right coset of H . Conversely, if H is a p -Sylow subgroup of G then $Hg \in \mathcal{X}$ and G permutes these m right cosets of H . So the orbit of H is small and H is contained in the stabilizer. But the stabilizer is a p -subgroup of G . So, it must be equal to H .

Since each $H \in \mathcal{P}$ gives a small orbit, $|\mathcal{X}| = \binom{p^k m}{p^k}$ is congruent (mod p) to $m|\mathcal{P}|$. Since m is invertible mod p this shows that $|\mathcal{P}| \equiv 1 \pmod{p}$. In particular \mathcal{P} is nonempty.

Suppose that Q is a p -subgroup of G . Then we want to find a $P \in \mathcal{P}$ so that $Q \leq P$. The group Q acts on \mathcal{P} by conjugation. By Lemma 6.5 this action must have at least one fixed point P . Then Q normalizes P so $Q \leq P$ by Remark 6.4.

It remains to show that the action of G on \mathcal{P} (by conjugation) is transitive, i.e., that \mathcal{P} is a single orbit of the G -action. If not then for every $P \in \mathcal{P}$ there is a $Q \in \mathcal{P}$ which is in a different orbit, i.e., is not conjugate to P . But the G -orbit of P is a disjoint union of Q -orbits each of which has no fixed points by Remark 6.4. Thus every G orbit has a multiple of p elements and $|\mathcal{P}|$ is a multiple of p which is a contradiction since it is $\equiv 1 \pmod{p}$. \square

Alperin and Bell point out that their proof does not use Cauchy's Theorem which they derive as a corollary.

Corollary 6.7 (Cauchy's Theorem). *If G is a finite group whose order $|G|$ is divisible by a prime p then G has an element of order p .*

Proof. We know that G has at least one p -Sylow subgroup P . Take a nontrivial element $g \in P$. Since $o(g)$ divides $|P|$, $o(g) = p^n$. So, $g^{p^{n-1}}$ is an element of G of order p . \square

Homework 3: Due next Thursday.

(3.1) Show that every subgroup of G containing $N(P)$ is self-normalizing.

(3.2) If $K \trianglelefteq G$ and P is a Sylow subgroup of K then $KN_G(P) = G$. [This follows from the fact that all conjugates of P lie in K .]

(3.3) If each Sylow subgroup of G is normal then G is the product of its Sylow subgroups.

(3.4) Show that $T(n, \mathbb{Z})$ is nilpotent.

7. CATEGORY THEORY AND PRODUCTS

I want to prove the theorem that a finite group is nilpotent if and only if it is the product of its Sylow subgroups. For this we first have to go over the product of groups. And this looks like a good time to introduce category theory.

7.1. categories. I gave the definition of a category and two examples to illustrate the definition: \mathcal{Gps} is the category of groups and \mathcal{Ens} is the category of sets.

Definition 7.1. A *category* \mathcal{C} consists of four things: $\mathcal{C} = (Ob(\mathcal{C}), Mor, \circ, id)$ where

- (1) $Ob(\mathcal{C})$ is a collection of *objects*. This collection is usually not a set. For example, $Ob(\mathcal{Gps})$ is the collection of all groups and $Ob(\mathcal{Ens})$ is the collection of all sets.
- (2) For any two objects $X, Y \in Ob(\mathcal{C})$ there is a set of *morphisms*

$$Mor_{\mathcal{C}}(X, Y)$$

which are written $f : X \rightarrow Y$. For example, in $Mor_{\mathcal{Gps}}(G, H)$ is the set of homomorphisms $\phi : G \rightarrow H$ and $Mor_{\mathcal{Ens}}(S, T)$ is the set of all mappings $f : S \rightarrow T$.

- (3) For any three objects X, Y, Z , we have a composition law:

$$Mor_{\mathcal{C}}(Y, Z) \times Mor_{\mathcal{C}}(X, Y) \rightarrow Mor_{\mathcal{C}}(X, Z)$$

sending (g, f) to $g \circ f$. Composition must be associative.

- (4) Every object $X \in Ob(\mathcal{C})$ has an *identity* $id_X \in Mor_{\mathcal{C}}(X, X)$ so that $id_Y \circ f = f = f \circ id_X$ for any $f : X \rightarrow Y$.

Note that there are only two assumptions about the structure. Namely, associativity of composition and the existence of units.

The idea of category theory is to extract elementary concepts out of difficult mathematics. We look only at composition of morphisms and forget the rest of the structure. Then we can ask: What are the properties that can be expressed only in terms of composition of morphisms? One of these is the product.

7.2. product of groups.

Definition 7.2. If G, H are groups, then the *product* $G \times H$ is defined to be the cartesian product of sets

$$G \times H = \{(g, h) \mid g \in G, h \in H\}$$

with the group law given coordinate-wise by

$$(g_1, h_1)(g_2, h_2) = (g_1g_2, h_1h_2).$$

These are several things to notice about this definition. The first is that $G \times H$ contains a copy of G, H which commute. By this I mean that there are monomorphisms (1 – 1 homomorphisms):

$$j_1 : G \rightarrow G \times H, \quad j_2 : H \rightarrow G \times H$$

given by $j_1(g) = (g, e), j_2(h) = (e, h)$. These inclusion maps have commuting images since

$$j_1(g)j_2(h) = (g, e)(e, h) = (g, h) = (e, h)(g, e) = j_2(h)j_1(g)$$

7.2.1. *internal direct product.*

Lemma 7.3. *If $\phi : G \rightarrow K, \psi : H \rightarrow K$ are homomorphisms with commuting images then there is a unique homomorphism $f : G \times H \rightarrow K$ so that $f \circ j_1 = \phi$ and $f \circ j_2 = \psi$.*

Proof. This is obvious. f must be given by $f(g, h) = j_1(g)j_2(h)$. This is a homomorphism since $[j_1(G), j_2(H)] = \{e\}$. \square

Theorem 7.4. *Suppose that G contains normal subgroups H, K so that $H \cap K = \{e\}, [H, K] = \{e\}$ and $HK = G$. Then the homomorphism $f : H \times K \rightarrow G$ given by the inclusion maps $H \hookrightarrow G, K \hookrightarrow G$ is an isomorphism.*

We say that $G = H \times K$ is the *internal direct product* in this case.

Proof. The map is given by $f(h, k) = hk$. This is surjective since $HK = G$. It is 1 – 1 since $H \cap K = \{e\}$. It is a homomorphism since $[H, K] = \{e\}$. \square

7.2.2. *universal property.* The product $G \times H$ has two other projection homomorphisms

$$p_1 : G \times H \rightarrow G, \quad p_2 : G \times H \rightarrow H$$

given by $p_1(g, h) = g, p_2(g, h) = h$. These satisfy the following “universal” property which is obvious (obviously true) and which I also explained in categorical terms.

Theorem 7.5. *Suppose that G, H, K are groups and $\phi : K \rightarrow G, \psi : K \rightarrow H$ are homomorphisms. Then there exists a unique homomorphism $f : K \rightarrow G \times H$ so that $p_1 \circ f = \phi$ and $p_2 \circ f = \psi$.*

The unique homomorphism is $f(x) = (\phi(x), \psi(x))$ and it is written $f = \phi \times \psi$.

7.3. categorical product. The last theorem is categorical since it involves only composition of homomorphism. It says that $G \times H$ is a categorical product.

Definition 7.6. Suppose that X, Y are objects of a category \mathcal{C} . Then $Z \in \mathcal{C}$ is the *product* of X and Y if there are morphisms $p_1 : Z \rightarrow X, p_2 : Z \rightarrow Y$ so that for any other object W and any morphisms $\phi : W \rightarrow X, \psi : W \rightarrow Y$ there is a unique morphism $f : W \rightarrow Z$ so that $p_1 \circ f = \phi$ and $p_2 \circ f = \psi$.

The condition can be written as a commuting diagram:

$$\begin{array}{ccc}
 & X & \\
 \phi \nearrow & & \nwarrow p_1 \\
 W & \xrightarrow{\exists! f} & Z \\
 \psi \searrow & & \swarrow p_2 \\
 & Y &
 \end{array}$$

We say that Z is the *product* of X, Y in the category \mathcal{C} and we write $Z = X \times Y$. We also call Z the *categorical product* of X and Y . Theorem 7.5 was written in such a way that it is obvious that the product of groups is the categorical product.

The next point I made was that the definition of a product defines $Z = X \times Y$ uniquely up to isomorphism.

The concept of an isomorphism is categorical:

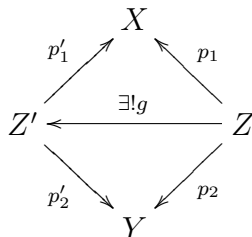
Definition 7.7. Two objects X, Y in any category \mathcal{C} are *isomorphic* and we write $X \cong Y$ if there are morphisms $f : X \rightarrow Y$ and $g : Y \rightarrow X$ so that $f \circ g = id_Y$ and $g \circ f = id_X$.

The definition of product is by a “universal condition” which forces the object Z to be unique up to isomorphism if it exists. (If the product does not exist, it suggests that the category is not large enough and perhaps we should add more objects.)

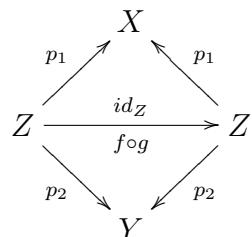
Theorem 7.8. *The product $Z = X \times Y$ is unique up to isomorphism assuming it exists.*

Proof. Suppose that Z' is another product. This means what we have morphisms $p'_1 : Z' \rightarrow X, p'_2 : Z' \rightarrow Y$ so that for any W , such as $W = Z$, any morphisms $W \rightarrow X, W \rightarrow Y$ (such as p_1, p_2) there is a unique morphism g so that $p'_i \circ g = p_i$ for $i = 1, 2$. In other words, the

following diagram commutes.



Similarly, since Z is the product, there is a unique morphism $f : Z' \rightarrow Z$ so that $p_i \circ f = p'_i$ for $i = 1, 2$. Now, take Z and Z . We have two morphisms $Z \rightarrow Z$ making the following diagram commute:



By the uniqueness clause in the definition of the product, we must have

$$f \circ g = id_Z.$$

Similarly, $g \circ f = id_{Z'}$. So, $Z \cong Z'$. □

7.4. products of nilpotent groups. Finally, I proved the following theorem which we need. I used three lemmas without proof. But, I am giving the proofs here after the proof of the theorem (and retroactively in Corollary 5.3).

Theorem 7.9. *If G, H are nilpotent groups of nilpotency class c_1, c_2 resp. then $G \times H$ is nilpotent of class $c = \max(c_1, c_2)$.*

The proof is by induction on c using the following lemma.

Lemma 7.10. *The center of $G \times H$ is $Z(G) \times Z(H)$.*

Proof. This is obvious. An element $(x, y) \in G \times H$ is central if

$$(x, y)(g, h) = (xg, yh) = (gx, hy) = (g, h)(x, y)$$

for all $g \in G, h \in H$. But this is the same as saying that $x \in Z(G)$ and $y \in Z(H)$. □

Lemma 7.11. *If $A \trianglelefteq G, B \trianglelefteq H$ then*

$$\frac{G \times H}{A \times B} \cong G/A \times H/B.$$

Proof. It suffices to find an epimorphism $\phi : G \times H \rightarrow G/A \times H/B$ with kernel $A \times B$. Such an epimorphism is given by $\phi(g, h) = (gA, hB)$. \square

Proof of Theorem 7.9. By induction on c . If $c = 1$ then G, H are both abelian. Then $G \times H$ is also abelian, which is the same as being nilpotent of class 1.

Now suppose that the theorem is true for $c-1$. Since $G/Z(G), H/Z(H)$ are nilpotent of class $c_1 - 1, c_2 - 1$, their product

$$G/Z(G) \times H/Z(H)$$

is nilpotent of class $\max(c_1 - 1, c_2 - 1) = c - 1$. By the lemmas above, this is isomorphic to $G \times H$ modulo its center. Therefore, by Corollary 5.3, $G \times H$ is nilpotent of class c . \square

8. PRODUCTS WITH MANY FACTORS

I am giving a long explanation about why finite nilpotent groups are products of Sylow subgroups:

$$G = P_1 \times P_2 \times \cdots \times P_n$$

So, I should explain first about products with many factors. There is a group theoretic definition and a categorical definition and they agree.

8.1. product of many groups. A product of finitely many groups is given by:

$$G_1 \times \cdots \times G_n = \{(g_1, \cdots, g_n) \mid g_i \in G_i \text{ for } i = 1, 2, \cdots, n\}$$

If we have an infinite family of groups G_α indexed by $\alpha \in I$ then the *product* of these groups is the cartesian product

$$\prod_{\alpha \in I} G_\alpha = \{g : I \rightarrow \cup G_\alpha \mid g(\alpha) \in G_\alpha \forall \alpha \in I\}.$$

We write $g(\alpha) = g_\alpha$ and call it the α coordinate of g . Multiplication is defined coordinatewise by

$$(fg)_\alpha = (f_\alpha)(g_\alpha)$$

This can be written as $p_\alpha(fg) = p_\alpha(f)p_\alpha(g)$. I.e., the projection maps

$$p_\alpha : \prod G_\alpha \rightarrow G_\alpha$$

are homomorphisms.

Elements of an infinite product may have infinitely many nontrivial coordinates. This sometimes causes trouble. However, it is required by category theory. The product as we just defined it satisfies the following categorical condition for obvious reasons.

Theorem 8.1. *Suppose that $G_\alpha, \alpha \in I$ is a family of groups and H is another group. Let $f_\alpha : H \rightarrow G_\alpha$ be arbitrary homomorphisms. Then there is a unique homomorphism $\phi : H \rightarrow \prod G_\alpha$ so that $p_\alpha \circ \phi = f_\alpha$ for all $\alpha \in I$.*

This means that $\prod G_\alpha$ is the product of the objects G_α of the category of groups.

Definition 8.2. If $X_\alpha, \alpha \in I$ is a family of objects in a category \mathcal{C} indexed by the set I then $Z \in \text{Ob}(\mathcal{C})$ is called the *product* of the X_α if there are morphisms $p_\alpha : Z \rightarrow X_\alpha$ for all $\alpha \in I$ so that for any other object W of \mathcal{C} and morphisms $f_\alpha : W \rightarrow X_\alpha$ there exists a unique morphism $\phi : W \rightarrow Z$ so that $p_\alpha \circ \phi = f_\alpha$ for all $\alpha \in I$.

The product of the X_α , if it exists, is unique up to isomorphism in the category \mathcal{C} because it is given by a universal condition. I will explain later a rigorous definition of “universal condition” which makes this statement obvious. What is not obvious is the definitions which become more and more complicated.

8.2. weak product. When the index set I is infinite we often consider the *weak product*

$$\prod'_{\alpha \in I} G_\alpha = \{(g_\alpha)_{\alpha \in I} \mid g_\alpha = e \text{ for all but a finite number of } \alpha\}.$$

When the groups G_α are abelian, this is called the *sum* and written $\bigoplus G_\alpha$. The weak product contains a copy of each group G_α via the monomorphism

$$j_\alpha : G_\alpha \rightarrow \prod'_{\alpha \in I} G_\alpha$$

which sends $x \in G_\alpha$ to the element $(e, \dots, e, x, e, \dots, e)$ with x in the α coordinate and e everywhere else. These elements (for different α) commute and weak product is universal with this property. I.e.:

Theorem 8.3. *Suppose that G_α are groups indexed by I and H is another group. Let $f_\alpha : G_\alpha \rightarrow H$ be homomorphisms whose images commute, i.e.,*

$$f_\alpha(x)f_\beta(y) = f_\beta(y)f_\alpha(x)$$

whenever $x \in G_\alpha, y \in G_\beta$ and $\alpha \neq \beta$. Then there exists a unique homomorphism

$$\phi : \prod'_{\alpha \in I} G_\alpha \rightarrow H$$

so that $\phi \circ j_\alpha = f_\alpha$ for all $\alpha \in I$.

Proof. An element of the infinite product (g_α) has only finitely many nontrivial coordinates $g_{\alpha_1}, g_{\alpha_2}, \dots, g_{\alpha_n}$. $\phi(g)$ must be equal to the product

$$\phi(g) = j_{\alpha_1}(g_{\alpha_1})j_{\alpha_2}(g_{\alpha_2}) \cdots j_{\alpha_n}(g_{\alpha_n})$$

where the factors can be multiplied in any order since they commute. \square

8.3. recognizing weak products. I decided to take a finite product. But the same statement is true for an infinite weak product.

Suppose that N_1, N_2, \dots, N_n are normal subgroups of a group G so that

- (1) $N_i \cap N_j = \{e\}$ for all $i \neq j$ and
- (2) $N_1 N_2 \cdots N_n = G$.

Then does this imply that G is isomorphic to the product $N_1 \times \cdots \times N_n$? The following example says not.

Take the group $G = \mathbb{Z}/2 \times \mathbb{Z}/2$. This has four elements e, a, b, ab . Each of the three nontrivial elements generates a (normal) subgroup of order 2: $N_1 = \langle a \rangle, N_2 = \langle b \rangle, N_3 = \langle ab \rangle$. These normal subgroups satisfy (1) and (2) but $N_1 \times N_2 \times N_3$ is not isomorphic to G since it has 8 elements. However, condition (1) implies that the elements of N_i and N_j commute for $i \neq j$. Therefore, Theorem 8.3 implies that there is a homomorphism

$$\phi : N_1 \times N_2 \times \cdots \times N_n \rightarrow G$$

which is the inclusion on each N_i . Condition (2) implies that ϕ is onto.

Theorem 8.4. *Suppose that N_1, N_2, \dots, N_n are normal subgroups of G satisfying conditions (1), (2) above. Then the homomorphism ϕ above is an isomorphism if and only if*

$$N_j \cap N_1 N_2 \cdots \widehat{N_j} \cdots N_n = \{e\}$$

for all j .

Proof. This new condition is certainly necessary if G is to be the product of the N_i . To show that it is sufficient, suppose it is true. Then ϕ is a monomorphism. Otherwise, there are elements $x_i \in N_i$ not all trivial so that

$$\phi(x_1, x_2, \dots, x_n) = x_1 x_2 \cdots x_n = e$$

If $x_j \neq e$ then it is the inverse of the product of the other x_i 's contradicting the new condition. \square

8.4. products of Sylow subgroups. Since p -groups are nilpotent, Theorem 7.9 implies that a product of p -groups is also nilpotent. Using Theorem 8.4 this implies the following.

Corollary 8.5. *If G is a finite group whose Sylow subgroups are normal then G is a product of its Sylow subgroups and therefore nilpotent.*

Proof. Suppose that P_1, P_2, \dots, P_n are the Sylow subgroups of G . If $P_i \trianglelefteq G$ and $|P_i| = p_i^{k_i}$ then P_i is the unique Sylow p_i -subgroup of G . Therefore the primes p_i are all distinct and $P_i \cap P_j = \{e\}$. This is condition (1). Also we must have

$$P_j \cap P_1 \cdots \widehat{P_j} \cdots P_n = \{e\}$$

since any element of P_j has order a power of p_j and elements of $P_1 \cdots \widehat{P}_j \cdots P_n$ have order prime to p_j . (Use the fact that the order of ab is least common multiple of the orders of a, b when a, b commute.) Therefore,

$$\phi : P_1 \times P_2 \times \cdots \times P_n \rightarrow G$$

is a monomorphism. But the two groups have the same number of elements. So, ϕ must be an isomorphism. \square

We need two more lemmas and HW 3.1.

Lemma 8.6. *Let $N \trianglelefteq G$ and $N \leq H \leq G$. Then*

$$N_{G/N}(H/N) = N_G(H)/N.$$

Proof. Let $N_{G/N}(H/N) = K/N$. Then $H/N \trianglelefteq K/N$ and we have a quotient map $K/N \rightarrow (K/N)/(H/N)$. The composite homomorphism:

$$K \rightarrow K/N \rightarrow \frac{K/N}{H/N}$$

has kernel H . Therefore, $H \trianglelefteq K$ and $K \leq N_G(H)$. Similarly, $H \trianglelefteq N_G(H)$ implies that $H/N \trianglelefteq N_G(H)/N$. Therefore, $N_G(H)/N$ is contained in the normalizer K/N of H/N . So $N_G(H) \leq K$. So, they must be equal. \square

Lemma 8.7. *Suppose that G is nilpotent and H is a proper subgroup of G . Then $H \neq N_G(H)$.*

Proof. We have a central series

$$G_0 \triangleright G_1 \triangleright \cdots \triangleright G_n = \{e\}.$$

I.e., $G_i \trianglelefteq G$ and $G_{i-1}/G_i \leq Z(G/G_i)$. Let i be minimal so that $G_i \leq H$. Then $G_{i-1} \not\leq H$. But

$$G_{i-1}/G_i \leq Z(G/G_i) \leq N_{G/G_i}(H/G_i) = N_G(H)/G_i$$

by the previous lemma. Therefore, $G_{i-1} \leq N_G(H)$ which shows that $H \neq N_G(H)$. \square

Theorem 8.8. *A finite group G is nilpotent if and only if it is a product of its Sylow subgroups.*

Proof. Since p -groups are nilpotent and a finite product of nilpotent groups is nilpotent, a group which is the product of its Sylow subgroups must be nilpotent. Conversely, suppose that G is nilpotent. Then we claim that every Sylow subgroup must be normal and thus G is a product of its Sylow subgroups by Corollary 8.5. To prove this claim suppose not. Then there is a Sylow subgroup P which is not normal in G . Then $N_G(P) \neq G$. But HW 3.1 says that $H = N_G(P)$ is self-normalizing. This is impossible for nilpotent groups by Lemma 8.7. \square

9. UNIVERSAL OBJECTS AND LIMITS

I explained how universal constructions were all examples of initial or terminal objects in some category. I also explained that these are equivalent if we reverse arrows.

9.1. initial and terminal objects.

Definition 9.1. Suppose that \mathcal{C} is a category. Then an object X_0 of \mathcal{C} is called *initial* if for any object Y of \mathcal{C} there is a unique morphism $f : X_0 \rightarrow Y$. Similarly, $X_\infty \in \text{Ob}(\mathcal{C})$ is called *terminal* if for any object W of \mathcal{C} there is a unique morphism $f : W \rightarrow X_\infty$. If X_0 is both initial and terminal it is called a *zero object*.

Example 9.2. In $\mathcal{E}ns$, the category of sets, the empty set \emptyset is initial and any one point set $\{*\}$ is terminal. In the category of groups $\{e\}$ is both initial and terminal. So, the trivial group is the zero object.

Theorem 9.3. *Initial and terminal objects are unique up to isomorphism if they exist.*

This is trivial but we went through it carefully because, as we will see later, it implies the uniqueness of any universal object.

Proof. Suppose that there are two initial objects X_0, X_1 . Then

$$\begin{aligned} X_0 \text{ initial} &\Rightarrow \exists! f : X_0 \rightarrow X_1 \\ X_1 \text{ initial} &\Rightarrow \exists! g : X_1 \rightarrow X_0 \\ X_0 \text{ initial} &\Rightarrow \text{Any two morphisms } X_0 \rightarrow X_0 \text{ are equal.} \end{aligned}$$

Therefore, $g \circ f = id_{X_0}$. Similarly, $f \circ g = id_{X_1}$. Therefore, $X_0 \cong X_1$. The uniqueness of terminal objects is similar (and also follows from the next theorem). \square

Definition 9.4. If \mathcal{C} is any category, its *opposite category* \mathcal{C}^{op} is “the same thing with arrows reversed.” By this I mean that

- (1) $\text{Ob}(\mathcal{C}^{op}) = \text{Ob}(\mathcal{C})$. The opposite category has the same objects. However, we put a little “op” as a superscript to indicate that we are considering the object as being in \mathcal{C}^{op} . So, if $X \in \text{Ob}(\mathcal{C})$ then X^{op} is X considered as an object of \mathcal{C}^{op} .
- (2) $\text{Mor}_{\mathcal{C}^{op}}(X^{op}, Y^{op}) = \text{Mor}_{\mathcal{C}}(Y, X)$. The morphism sets are equal. But the morphism $f : Y \rightarrow X$ in \mathcal{C} is written

$$f^{op} : X^{op} \rightarrow Y^{op}$$

in \mathcal{C}^{op} .

- (3) $id_{X^{op}} = (id_X)^{op}$. (Identities are the same.)
- (4) $f^{op} \circ g^{op} = (g \circ f)^{op}$. (Composition is reversed.)

For example, when we say that

$$f^{op} : G^{op} \rightarrow H^{op}$$

is a morphism in \mathcal{Gps}^{op} , we do not mean that we have created new objects called G^{op} and H^{op} . All this means is that we have an ordinary group homomorphism

$$f : H \rightarrow G.$$

The purpose is to change the description of the objects.

Theorem 9.5. *X is an initial (resp. terminal) object of \mathcal{C} if and only if X^{op} is a terminal (resp. initial) object of \mathcal{C}^{op} .*

9.2. products as terminal objects. If $X_\alpha, \alpha \in I$ is a family of objects in \mathcal{C} , I created a new category \mathcal{B} so that a terminal object of \mathcal{B} is the same as the product of the objects X_α in \mathcal{C} .

The *objects* of the new category \mathcal{B} consist of

- (1) an object Y of \mathcal{C} and
- (2) morphisms $f_\alpha : Y \rightarrow X_\alpha$ for all $\alpha \in I$.

I wrote the element as: $(Y, (f_\alpha)_{\alpha \in I})$. If $(Z, (g_\alpha))$ is another object in \mathcal{B} then a morphism

$$(Y, (f_\alpha)) \rightarrow (Z, (g_\alpha))$$

is defined to be a morphism $\phi : Y \rightarrow Z$ in \mathcal{C} so that $f_\alpha = g_\alpha \circ \phi$ for all $\alpha \in I$. In other words, the following diagram commutes for each α .

$$\begin{array}{ccc} & & X_\alpha \\ & \nearrow f_\alpha & \uparrow g_\alpha \\ Y & \xrightarrow{\phi} & Z \end{array}$$

Proposition 9.6. *If $(Z, (g_\alpha))$ is terminal in \mathcal{B} then $Z \cong \prod X_\alpha$.*

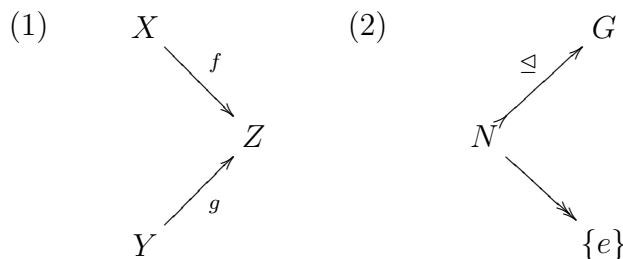
Proof. $(Z, (g_\alpha))$ is terminal implies $\phi : Y \rightarrow Z$ is unique which implies that $Z = \prod X_\alpha$. \square

9.3. general limits. A categorical product is the limit of a diagram with no arrows. We need to generalize this construction to create more general limits. We will do this now in an arbitrary category and next week we will look in the category of groups and sets and specialize to particular diagrams.

Definition 9.7. A *diagram* D in a category \mathcal{C} is a set of objects $X_\alpha, \alpha \in I$ and a set of morphisms between these objects.

I couldn't think of a good way to index the arrows in general. It depends on the diagram.

Example 9.8. Here are some important examples of diagrams.



(3) $D = \{X_1, X_2, X_3\}$ (no arrows).

(4)

$$\dots \xrightarrow{f_3} G_2 \xrightarrow{f_2} G_1 \xrightarrow{f_1} G_0$$

(5) X, Y with two morphisms $f, g : X \rightarrow Y$.

Definition 9.9. If D is a diagram in a category \mathcal{C} then the *category of objects over D* which we write as \mathcal{C}/D has objects consisting of

- (1) one object Y of \mathcal{C} and
- (2) morphisms $g_\alpha : Y \rightarrow X_\alpha$ going from Y to each object in the diagram D so that, for any morphism $f : X_\alpha \rightarrow X_\beta$ in \mathcal{C} , $f \circ g_\alpha = g_\beta$.

If $(Y, (g_\alpha))$ is a terminal object in \mathcal{C}/D , then Y is called the *limit* of the diagram D .

Being a terminal object, the limit of a diagram is unique up to isomorphism if it exists. For the special case when the diagram has no arrows, the limit is the product.

9.4. existence of limits. I showed that limits exist in the category of sets and in the category of groups. I also showed that the category of sets has colimits. Colimits in the category of groups are complicated and will be discussed next week.

Theorem 9.10. *The category of sets has arbitrary limits.*

Proof. This means that, for any diagram of sets, with objects X_α and morphisms $f : X_\alpha \rightarrow X_\beta$ (which I did not index the first day), there exists a limit of the diagram. This is given by letting $\lim X_\alpha$ be the set of all (x_α) in the Cartesian product $\prod X_\alpha$ so that $f(x_\alpha) = x_\beta$ for all morphisms $f : X_\alpha \rightarrow X_\beta$ in the diagram. \square

Theorem 9.11. *The category of sets has arbitrary colimits.*

Proof. Given a diagram D of sets with objects X_α and morphisms $f : X_\alpha \rightarrow X_\beta$, the *colimit* or *direct limit* is the set

$$\operatorname{colim} D = \coprod X_\alpha / \sim$$

which is the disjoint union of the sets X_α modulo the equivalence relation given by $x \in X_\alpha \sim f(x) \in X_\beta$.

The *disjoint union* is the union made disjoint. Formally this is the set

$$\coprod_{\alpha \in I} X_\alpha = \{(\alpha, x) \in I \times \bigcup X_\alpha \mid x \in X_\alpha\}.$$

When we say we “mod out the equivalence relation” we mean: Take the set of equivalence classes. The fact that this is the colimit is obvious. Given a mappings $g_\alpha : X_\alpha \rightarrow Y$, we get a mapping on the disjoint union. Since we require $g_\beta \circ f = g_\alpha$, the identified elements $x \in X_\alpha$ and $f(x) \in X_\beta$ map to the same element of Y , so we get an induced map of the quotient $\coprod X_\alpha / \sim$ into Y . \square

Theorem 9.12. *Arbitrary limits exist in the category of groups.*

Proof. This is easy. It is the same thing as in the category of sets. Given a diagram with groups G_α connected by morphisms $f : G_\alpha \rightarrow G_\beta$, the limit of the diagram is

$$\{(g_\alpha) \in \prod G_\alpha \mid f(g_\alpha) = g_\beta \text{ for all } f : G_\alpha \rightarrow G_\beta \text{ in the diagram}\}$$

which is the same as in the category of sets. The proof of universality is also the same. The only thing we need to check (or at least realize needs to be checked), is that this actually defines a subgroup of the product. Sometimes definitions have a hidden commutativity assumption.

So, suppose that $(g_\alpha), (h_\alpha)$ are in the inverse limit set. This means $f(g_\alpha) = g_\beta$ and $f(h_\alpha) = h_\beta$. We need to check that $(g_\alpha h_\alpha)$ is in the set:

$$f(g_\alpha h_\alpha) = f(g_\alpha)f(h_\alpha) = g_\beta h_\beta.$$

So, it works. Also, it is trivial. So, the proper thing would have been to say: “It is easy to see that this is a subgroup of the product.” \square

At this point I introduced the “forgetful functor” to explain what it means that the formula for limits is the same in the two categories. But in these notes this explanation comes later.

9.5. examples. I explained several examples to illustrate how this works.

9.5.1. double arrow. The first example (Example 9.8.5) was in answer to the question: What happens if there is more than one morphism between the same two objects.

Consider the diagram in \mathcal{Gps} with two homomorphism $f_1, f_2 : G \rightarrow H$. The limit L has the property that it has maps $p_1 : L \rightarrow G$, $p_2 : L \rightarrow H$ so that $p_2 = f_i \circ p_1$ for $i = 1$ and $i = 2$. This implies that $f_1 \circ p_1 = f_2 \circ p_1$. In other words, $p_1 : L \rightarrow G$ has image in the subgroup

$$K = \{g \in G \mid f_1(g) = f_2(g)\}$$

Since K is a subset of G , the mapping $L \rightarrow K$ is unique. This means that K is the universal object, i.e., the limit of the diagram. K is called the *equalizer* of the two arrows.

9.5.2. pull-back. The second example I gave was the more standard construction of the pull-back which is the limit of the following diagram.

$$\begin{array}{ccc} & & G \\ & & \downarrow f_1 \\ H & \xrightarrow{f_2} & K \end{array}$$

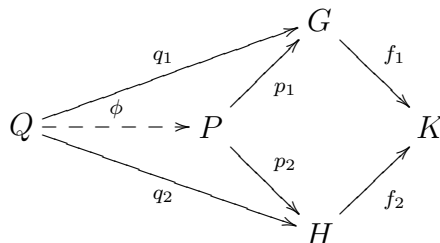
Definition 9.13. Suppose that $f_1 : G \rightarrow K, f_2 : H \rightarrow K$ are homomorphisms. Then the *pull-back* P is defined to be the subgroups of the product $G \times H$ given by

$$P = \{(g, h) \in G \times H \mid f_1(g) = f_2(h)\}$$

together with the projection homomorphisms $p_1 : P \rightarrow G, p_2 : P \rightarrow H$.

Proposition 9.14. *The pull-back is the limit of the diagram above.*

Proof. Suppose that there is another group Q with homomorphisms $q_1 : Q \rightarrow G, q_2 : Q \rightarrow H$ making the following diagram commute.



Then $f_1q_1(x) = f_2q_2(x)$ for every $x \in Q$. This is the same as saying that $\phi(x) = (q_1(x), q_2(x)) \in P$. So, ϕ in the diagram above exists and is unique. \square

According to the formula, the limit is given by

$$\lim(G \rightarrow K \leftarrow H) = \{(g, k, h) \in G \times K \times H \mid f_1(g) = k = f_2(h)\}.$$

This is isomorphic to the pull-back because the coordinate $k \in K$ is redundant. An isomorphism $P \rightarrow \lim(G \rightarrow K \leftarrow H)$ is given by

$$(g, h) \mapsto (g, f_1(g), h).$$

9.5.3. *inverse limit.* The third example I gave is the diagram that most people think of as giving the “inverse limit.”

$$\cdots \xrightarrow{f_4} G_3 \xrightarrow{f_3} G_2 \xrightarrow{f_2} G_1 \xrightarrow{f_1} G_0$$

We use the formula which says that the limit of this diagram is

$$\lim_{\leftarrow} D = \lim_{\leftarrow} G_i = \{(g_0, g_1, g_2, \cdots) \in \prod G_i \mid f_i(g_i) = g_{i-1} \forall i\}.$$

An important special case is the inverse system:

$$\xrightarrow{f_4} \mathbb{Z}/n^3 \xrightarrow{f_3} \mathbb{Z}/n^2 \xrightarrow{f_2} \mathbb{Z}/n^1$$

When $n = 10$, the elements are sequences (n_1, n_2, n_3, \cdots) so that each n_j is a j digit decimal whose last $j - 1$ digits give n_{j-1} . For example $(4, 14, 014, 9014, 19014, \cdots) \in \lim_{\leftarrow} \mathbb{Z}/10^i$. This can be viewed as an infinite decimal going to the left:

$$\cdots 19014.$$

There are two important differences between infinite decimals going to the left and those going to the right.

- (1) The expression is unique. For the usual decimals, two expressions can be equal. For example,

$$0.999999 \dots = 1.000000 \dots$$

But every element in the inverse limit is given by a unique expression.

- (2) Negative quantities are included. For decimal going to the right, we need to put a negative sign in front to get the additive inverse. For these decimals going to the left, negatives are included. For example,

$$\dots 99999. = -1$$

since, if we add 1 we get $\dots 00000$.

Definition 9.15. For any prime p the ring of p -adic numbers \mathbb{Z}_p is given by

$$\mathbb{Z}_p := \varprojlim \mathbb{Z}/p^i\mathbb{Z}.$$

But, we constructed it as a limit of additive groups. I will go over the ring structure later.

We figured out what the formula for the inverse is in \mathbb{Z}_2 . An element is an arbitrary infinite sequence of 0's and 1's:

$$\dots 01101000110.$$

To get the negative, you change all 0's to 1's and all 1's to 0's and then add 1.

$$-(\dots 01000110) = \dots 10111001 + 1 = \dots 10111010.$$

10. LIMITS AS FUNCTORS

I explained the theorem that limits are natural. This means they are functors.

10.1. functors. Given two categories \mathcal{C} and \mathcal{D} a *functor* $F : \mathcal{C} \rightarrow \mathcal{D}$ is a mapping which sends objects to objects and morphisms to morphisms and preserves the structure:

- (1) For each $X \in \text{Ob}(\mathcal{C})$ we assign an object $F(X) \in \text{Ob}(\mathcal{D})$.
- (2) For any two objects $X, Y \in \text{Ob}(\mathcal{C})$ we get a mapping of sets:

$$F : \text{Mor}_{\mathcal{C}}(X, Y) \rightarrow \text{Mor}_{\mathcal{D}}(FX, FY).$$

This sends $f : X \rightarrow Y$ to $Ff : FX \rightarrow FY$ so that the next two conditions are satisfied.

- (3) $F(id_X) = id_{FX}$. I.e., F sends identity to identity.
- (4) $F(f \circ g) = Ff \circ Fg$. I.e., F preserves composition.

The first example I gave was the *forgetful functor*

$$F : \mathcal{Gps} \rightarrow \mathcal{E}ns$$

which sends a group to its underlying set and forgets the rest of the structure. Thus

$$F(G, \cdot, e, ()^{-1}) = G.$$

The fact that there is a forgetful functor to the category of sets means that groups are sets with extra structure so that the homomorphisms are the set mappings which preserve this structure. Such categories are called *concrete*. Not all categories are concrete. For example the homotopy category \mathcal{H} whose objects are topological spaces and whose morphisms are homotopy classes of maps is not concrete or equivalent to a concrete category.

10.2. the diagram category. The limit of a diagram in the category of groups is a functor

$$\lim : \text{Fun}(\Gamma, \mathcal{Gps}) \rightarrow \mathcal{Gps}.$$

But what is the domain category?

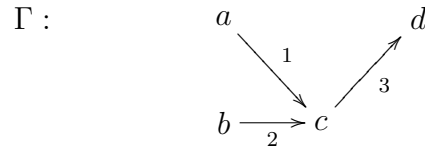
At this point I needed to make more precise how diagrams are indexed. I used the following example.

$$D : \begin{array}{ccc} X_a & & X_d \\ & \searrow f_1 & \nearrow f_3 \\ X_b & \xrightarrow{f_2} & X_c \end{array}$$

The objects of the diagram are indexed by the set $\Gamma_0 = \{a, b, c, d\}$. The arrows (morphisms) are indexed by $\Gamma_1 = \{1, 2, 3\}$. These two indexing sets are related by two mappings called *source* and *target*

$$s, t : \Gamma_1 \rightarrow \Gamma_0$$

This structure $(\Gamma_0, \Gamma_1, s, t)$ is called a *directed graph* and is drawn as follows.



For any category \mathcal{C} and any directed graph Γ , the *diagram category* $\text{Fun}(\Gamma, \mathcal{C})$ is defined to be the category whose objects are diagrams consisting of

- (1) an object X_α in \mathcal{C} for every $\alpha \in \Gamma_0$ and
- (2) a morphism $f_j : X_{s(j)} \rightarrow X_{t(j)}$ for every $j \in \Gamma_1$.

A morphism in the category

$$((X_\alpha)_\alpha, (f_j)_j) \rightarrow ((Y_\alpha)_\alpha, (g_j)_j)$$

consists of a morphism $\phi_\alpha : X_\alpha \rightarrow Y_\alpha$ for each $\alpha \in \Gamma_0$ so that the following square commutes for every $j \in \Gamma_1$ with source α and target β .

$$\begin{array}{ccc} X_\alpha & \xrightarrow{f_j} & X_\beta \\ \phi_\alpha \downarrow & & \downarrow \phi_\beta \\ Y_\alpha & \xrightarrow{g_j} & Y_\beta \end{array}$$

10.3. The limit is a functor. When we say that the limit is a functor, we mean that for any map of diagrams we should get a map of limits. For example, suppose that the graph is

$$\Gamma : \quad \cdots \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet$$

And we have two diagrams indexed by Γ :

$$D_1 : \quad \cdots \longrightarrow G_2 \longrightarrow G_1 \longrightarrow G_0$$

$$D_2 : \quad \cdots \longrightarrow H_2 \longrightarrow H_1 \longrightarrow H_0$$

Then, a morphism from D_1 to D_2 in the diagram category is a sequence of vertical arrows $\phi_i : G_i \rightarrow H_i$ making the following diagram commute.

$$\begin{array}{ccccccc} D_1 & & \cdots & \longrightarrow & G_2 & \longrightarrow & G_1 & \longrightarrow & G_0 \\ (\phi_\alpha) \downarrow & & & & \phi_2 \downarrow & & \phi_1 \downarrow & & \phi_0 \downarrow \\ D_2 & & \cdots & \longrightarrow & H_2 & \longrightarrow & H_1 & \longrightarrow & H_0 \end{array}$$

I explained how the universal property of the limit gave the induced map

$$G_\infty = \lim_{\leftarrow} G_i \rightarrow H_\infty = \lim_{\leftarrow} H_i$$

The reason is because of the following commuting diagram.

$$\begin{array}{ccccccc} G_\infty & & & & & & & & \\ & \searrow^{f_1} & & & & & & & \\ & & & & & & & & \\ & \searrow^{f_2} & & & & & & & \\ \cdots & \longrightarrow & G_2 & \longrightarrow & G_1 & \longrightarrow & G_0 & & \\ & & \phi_2 \downarrow & & \phi_1 \downarrow & & \phi_0 \downarrow & & \\ \cdots & \longrightarrow & H_2 & \xrightarrow{h_2} & H_1 & \longrightarrow & H_0 & & \end{array}$$

When we compose the vertical arrows we get the following diagram.

$$\begin{array}{ccccccc} G_\infty & & & & & & & & \\ & \searrow^{\phi_1 f_1} & & & & & & & \\ & \searrow^{\phi_2 f_2} & & & & & & & \\ \cdots & \longrightarrow & H_2 & \xrightarrow{h_2} & H_1 & \longrightarrow & H_0 & & \end{array}$$

The commutativity of the triangles (of which only one is drawn) gives, by the universal property of $\lim_{\leftarrow} H_i$, a uniquely determined homomorphism

$$\phi : G_\infty \rightarrow H_\infty$$

The explicit formula for this homomorphism is

$$\phi(g_0, g_1, g_2, \cdots) = (\phi_0(g_0), \phi_1(g_1), \phi_2(g_2), \cdots)$$

This is the unique mapping which commutes with the projection map to the n th coordinate:

$$\begin{array}{ccc} G_\infty & \xrightarrow{p_n} & G_n \\ \phi \downarrow & & \downarrow \phi_n \\ H_\infty & \xrightarrow{p_n} & H_n \end{array}$$

Theorem 10.1. *For any directed graph Γ the limit gives a functor*

$$\lim : \text{Fun}(\Gamma, \mathcal{Gps}) \rightarrow \mathcal{Gps}$$

from the category of diagrams in \mathcal{Gps} indexed by Γ to the category \mathcal{Gps} . The same holds in the category \mathcal{Ens} .

Remark 10.2. If a category \mathcal{C} has limits of diagrams indexed by Γ we need to assume that the category is *small* (i.e., the collection of objects is a set) in order for the limit to be an actual functor. The reason is that the limit is only defined up to isomorphism and you need to choose one object from each isomorphism class. Most people don't make such a fuss. (Because if you understand the problem you probably also understand the solution.)

Another theorem which I mentioned earlier is the following.

Theorem 10.3. *Limit commutes with the forgetful functor. I.e., the following diagram of categories and functors commutes.*

$$\begin{array}{ccc} \text{Fun}(\Gamma, \mathcal{Gps}) & \xrightarrow{\lim} & \mathcal{Gps} \\ F \downarrow & & \downarrow F \\ \text{Fun}(\Gamma, \mathcal{Ens}) & \xrightarrow{\lim} & \mathcal{Ens}. \end{array}$$

This theorem is a fancy way of saying that the formula for the limit is the same in the two categories.

10.4. adjoint functors. Finally, I talked about adjoint functors. This is yet another explanation of the concept of universality. In the definition of universality it says “There exists a unique ϕ .” These words should be reminiscent of the expressions:

$$(\forall x)(\exists! y)f(x) = y,$$

$$(\forall y)(\exists! x)f(x) = y.$$

These say that f is a *bijection* between two sets. If universality is a bijection between two sets then what are those two sets?

One set is $Mor_{\mathcal{C}}(Y, \lim X_{\alpha})$. This contains the morphism $\phi : Y \rightarrow \lim X_{\alpha}$. What data do you need to give you this unique morphism? You need morphisms $\phi_{\alpha} : Y \rightarrow X_{\alpha}$ for all α . These morphisms can be made into a single morphism between two diagrams:

$$\begin{array}{ccccccc} ((Y), (id)) : & \cdots & \longrightarrow & Y & \xrightarrow{id} & Y & \xrightarrow{id} & Y \\ (\phi_{\alpha}) \downarrow & & & \phi_2 \downarrow & & \phi_1 \downarrow & & \phi_0 \downarrow \\ ((X_{\alpha}), (f_i)) & \cdots & \longrightarrow & X_2 & \longrightarrow & X_1 & \longrightarrow & X_0 \end{array}$$

The diagram $((Y), (id))$ is the trivial diagram with Y at each vertex and the identity morphism on Y on each edge. A morphism in the diagram category from this diagram to the diagram $((X_\alpha), (f_i))$ is the same as a family of morphisms $\phi_\alpha : Y \rightarrow X_\alpha$ so that $f_i \circ \phi_{s(i)} = \phi_{t(i)}$:

$$\begin{array}{ccccc} & & Y & & \\ & \swarrow \phi_2 & \downarrow \phi_1 & \searrow \phi_0 & \\ \cdots & \longrightarrow & X_2 & \longrightarrow & X_1 & \longrightarrow & X_0 \end{array}$$

I pointed out that this trivial diagram $TY = ((Y), (id))$ is a functor

$$T : \mathcal{C} \rightarrow \text{Fun}(\Gamma, \mathcal{C}).$$

Definition 10.4. Two functors $F : \mathcal{C} \rightarrow \mathcal{D}$ and $G : \mathcal{D} \rightarrow \mathcal{C}$ are called *adjoint functors* if there is a natural bijection

$$\text{Mor}_{\mathcal{D}}(GX, Y) \cong \text{Mor}_{\mathcal{C}}(X, FY)$$

for all objects X in \mathcal{C} and all objects Y in \mathcal{D} . F is called the *right adjoint* of G and G is called the *left adjoint* of F .

Theorem 10.5. *The trivial diagram functor and the limit are adjoint functors for the category of groups and for any category that has limits of diagrams indexed by Γ . I.e.,*

$$\text{Mor}_{\text{Fun}(\Gamma, \mathcal{G}_{ps})}(((Y), (id)), ((X_\alpha), (f_i))) \cong \text{Mor}_{\mathcal{G}_{ps}}(Y, \lim X_\alpha).$$

11. CATEGORICAL LIMITS

(Two lectures by Ivan Horozov. Notes by Andrew Gainer and Roger Lipsett. [Comments by Kiyoshi].)

We first note that what topologists call the “limit”, algebraists call the “inverse limit” and denote by \lim_{\leftarrow} . Likewise, what topologists call the “colimit”, algebraists call the direct limit and denote \lim_{\rightarrow} .

Example 11.1. Take the inverse system

$$\cdots \rightarrow \mathbb{C}[x]/(x^n) \rightarrow \cdots \rightarrow \mathbb{C}[x]/(x^3) \rightarrow \mathbb{C}[x]/(x^2) \rightarrow \mathbb{C}[x]/(x)$$

Let $f \in \lim_{\leftarrow} \mathbb{C}[x]/(x^n) = \mathbb{C}[[x]]$. Then $f = \sum_{n \geq 0} a_n x^n$ is a formal power series in x over \mathbb{C} .

We note that $\mathbb{Z}_p = \lim_{\leftarrow} \mathbb{Z}/p^n$ [is the inverse limit of]

$$\cdots \rightarrow \mathbb{Z}/p^{n+1} \rightarrow \mathbb{Z}/p^n \rightarrow \cdots \rightarrow \mathbb{Z}/p$$

11.1. colimits in the category of sets.

Definition 11.2. Let X_α be sets indexed by $\alpha \in I$ and let $f_{\alpha,\beta} : X_\alpha \rightarrow X_\beta$ be functions with $\alpha, \beta \in I$. [Only for pairs (α, β) so that $\alpha < \beta$ in some partial ordering of I .] Then $\{f_{\alpha,\beta}, X_\alpha\}_{\alpha \in I}$ is a *directed system of sets* if, for every pair of composable morphisms $f_{\alpha,\beta} : X_\alpha \rightarrow X_\beta$, $f_{\beta,\gamma} : X_\beta \rightarrow X_\gamma$ [i.e., wherever $\alpha < \beta < \gamma$ in I], the following diagram commutes

$$\begin{array}{ccc} X_\alpha & \xrightarrow{f_{\alpha,\beta}} & X_\beta \\ & \searrow f_{\alpha,\gamma} & \downarrow f_{\beta,\gamma} \\ & & X_\gamma \end{array}$$

and, for every $\alpha, \beta \in I$ there exists a $\gamma \in I$ for which there are maps $f_{\alpha,\gamma}$ and $f_{\beta,\gamma}$ [i.e., $\alpha, \beta \leq \gamma$] as such:

$$\begin{array}{ccc} X_\alpha & \xrightarrow{f_{\alpha,\gamma}} & X_\gamma \\ & & \nearrow f_{\beta,\gamma} \\ X_\beta & & \end{array}$$

One can think of a directed system as a graph with sets as points and arrows as edges.

We can now define the direct limit on directed systems of sets by

$$\lim_{\rightarrow \alpha \in I} X_\alpha = \coprod_{\alpha} X_\alpha / \sim$$

where, for each $f_{\alpha,\beta}$ and for all $x \in X_\alpha$, we set $x \sim f_{\alpha,\beta}(x)$. Informally then, the direct limit is the set of equivalence classes induced by all functions $f_{\alpha,\beta}$. That \sim is an equivalence relation follows from the

two properties illustrated diagrammatically above. [The colimit of any diagram of sets exists. The assumption of being “directed” implies that any two elements of the colimit, represented by say $x \in X_\alpha, y \in X_\beta$, are equivalent to elements of the same set X_γ .]

11.2. **pull-back.** In the following diagram, $*$ is a “pull-back”:

$$\begin{array}{ccc} * & \longrightarrow & G \\ \downarrow & & \downarrow \alpha \\ H & \xrightarrow{\beta} & K \end{array}$$

The pull-back here is a subgroup (or subset) of $G \times H$ given by

$$G \times_K H = \{(g, h) \in G \times H \mid \alpha(g) = \beta(h)\}.$$

11.2.1. *universal property of pull-back.* If G' is a group such that

$$\begin{array}{ccccc} G' & & & & \\ & \searrow & & & \\ & & (g, h) \in G \times_K H & \longrightarrow & G \\ & \searrow & \downarrow h & & \downarrow \alpha \\ & & H & \xrightarrow{\beta} & K \end{array}$$

commutes then there exists a unique map $G' \rightarrow G \times_K H$ such that

$$\begin{array}{ccccc} G' & & & & \\ & \searrow & & & \\ & & G \times_K H & \longrightarrow & G \\ & \searrow & \downarrow & & \downarrow \\ & & H & \longrightarrow & K \end{array}$$

commutes.

11.3. **push-forward [push-out] of groups.**

$$\begin{array}{ccc} K & \longrightarrow & G \\ \downarrow & & \downarrow \\ H & \longrightarrow & G *_K H \\ & \searrow & \searrow \\ & & G \end{array}$$

In this diagram, if $K = \{e, \}$ then $G * H$ is the *amalgamated free product* of G and H given by

$$G * H = \{g_1 h_1 g_2 h_2 \cdots g_n h_n \mid g_i \in G, h_i \in H\}.$$

Note that $(g_1 h_1 g_2)^{-1} = g_2^{-1} h_1^{-1} g_1^{-1}$. [So, the set of such products is closed under the operation of taking inverse. So, $G * H$ is a group.]

More generally, $G *_K H$ is the quotient group $G * H / \sim$ where

$$(g\alpha(k)) \cdot (\beta(k)^{-1}h) \sim gh$$

and

$$hg \sim (h\beta(k)) \cdot (\alpha(k)^{-1}g).$$

Exercise. Compute $(\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$ and explain the computation.

11.4. direct limit of groups. In order to take the direct limit of groups, we require a directed system of groups:

Definition 11.3. $\{G_\alpha, f_{\alpha,\beta}\}$ is a *directed system of groups* if

i) $f_{\alpha,\beta} : G_\alpha \rightarrow G_\beta, f_{\beta,\gamma} : G_\beta \rightarrow G_\gamma$ are homomorphisms then [there is a homomorphism $f_{\alpha,\gamma} : G_\alpha \rightarrow G_\gamma$ in the system and]

$$\begin{array}{ccc} G_\alpha & \xrightarrow{f_{\alpha\beta}} & G_\beta \\ & \searrow f_{\alpha\gamma} & \downarrow f_{\beta\gamma} \\ & & G_\gamma \end{array}$$

[i.e. the diagram commutes.]

ii) For every $\alpha, \beta \in I$ there exists $\gamma \in I$ such that $f_{\alpha,\gamma}, f_{\beta,\gamma}$ are defined and

$$\begin{array}{ccc} G_\alpha & \xrightarrow{f_{\alpha\gamma}} & G_\gamma \\ G_\beta & \xrightarrow{f_{\beta\gamma}} & \nearrow \end{array}$$

We then define a new object [the weak product] $\prod'_{\alpha \in I} G_\alpha \subseteq \prod_{\alpha \in I} G_\alpha$:

Definition 11.4. For $x = \{x_\alpha\}_\alpha$ we let $x \in \prod'_{\alpha \in I} G_\alpha$ if x has only finitely many x_α coordinates which are not $e_\alpha \in G_\alpha$.

[The direct limit of a directed system of groups is then the same set as the direct limit of sets:

$$\lim_{\rightarrow \alpha \in I} G_\alpha = \prod_{\alpha \in I} G_\alpha / \sim$$

where, $x_\alpha \in G_\alpha$ is equivalent to $x_\beta \in G_\beta$ if and only if

$$f_{\alpha,\gamma}(x_\alpha) = f_{\beta,\gamma}(x_\beta)$$

for some $\gamma \in I$, with the additional structure that the product of $x_\alpha \in G_\alpha, x_\beta \in G_\beta$ is defined to be the product of their images in G_γ .]

In practice, one works with $\lim_{\rightarrow} G_\alpha$ in the following way: [Here Horozov says that each element of the direct limit is represented by a single element of a single group G_α . I wrote that as the definition.]

11.5. universal property of the direct limit of groups. If $\{G_\alpha, f_{\alpha,\beta}\}$ is a directed system of groups and $g_\alpha : G_\alpha \rightarrow H$ are homomorphisms such that

$$\begin{array}{ccc} G_\alpha & \xrightarrow{g_\alpha} & H \\ f_{\alpha\beta} \downarrow & \nearrow g_\beta & \\ G_\beta & & \end{array}$$

commutes, then there exists a unique homomorphism $g : \lim_{\rightarrow} G_\alpha \rightarrow H$ such that

$$\begin{array}{ccc} G_\alpha & \xrightarrow{h_\alpha} & \lim_{\rightarrow} G_\alpha \\ & \searrow g_\alpha & \nearrow \exists! g \\ & & H \end{array}$$

commutes for all $\alpha \in I$.

[Any element \bar{x} of the direct limit is represented by some element of some group $x_\alpha \in G_\alpha$. Then we let $g(\bar{x}) = g_\alpha(x_\alpha)$. If $x_\beta \in G_\beta$ is another representative of the same equivalence class \bar{x} then, by definition, $f_{\alpha,\gamma}(x_\alpha) = f_{\beta,\gamma}(x_\beta) = x_\gamma \in G_\gamma$ for some $\gamma \in I$. But then $g_\alpha(x_\alpha) = g_\gamma(x_\gamma) = g_\beta(x_\beta)$. So, g is well-defined.]

11.6. free groups. Let X be a set. The free group on X , $F(X)$, is defined by the following universal property: given any group G and set map $f : X \rightarrow G$, there is a unique $g : F(X) \rightarrow G$ that is a group homomorphism such that the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & G \\ & \searrow i & \nearrow \exists! g \\ & & F(X) \end{array}$$

commutes.

It remains to define $F(X)$ and say what the map i is. Suppose $X = \{x_1, x_2, \dots\}$ (note that X need not be countable; we use this subscript notation simply for ease of use). Then the *words in X* , w , are all finite sequences chosen from the set $X \cup X^{-1}$, where $X^{-1} = \{x_1^{-1}, x_2^{-1}, \dots\}$

(here the $^{-1}$ notation is purely formal). If $W = \{w\}$ is the set of such word then

$$F(X) = W / \sim$$

where \sim is the smallest possible relation so that we get a group: i.e., that for all i , both $x_i x_i^{-1}$ and $x_i^{-1} x_i$ are trivial.

Then each word in $F(X)$ has a unique reduced form, in which no further simplification induced by the above relation are possible, and $F(X)$ thus consists of all the reduced words.

11.7. an important example. $PSL_2(\mathbb{Z}) = SL_2(\mathbb{Z}) / \pm I$. This group acts on the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} \mid \Im z > 0\}$: $A \in SL_2(\mathbb{Z})$ gives a map

$$z \mapsto \frac{az + b}{cz + d} = \frac{-az - b}{-cz - d}.$$

Thus, for example, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ translates by 1, while $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is inversion.

It turns out that

$$PSL_2(\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}$$

This has something to do with fixed points in \mathbb{C} under this action. Similarly,

$$SL_2(\mathbb{Z}) = \mathbb{Z}/4\mathbb{Z} *_{\mathbb{Z}/2\mathbb{Z}} \mathbb{Z}/6\mathbb{Z}$$

12. MORE ABOUT FREE PRODUCTS

I decided to explain the example of $PSL_2(\mathbb{Z})$ more thoroughly.

12.1. Amalgamated products again. The free product formulas for $PSL_2(\mathbb{Z})$ and $SL_2(\mathbb{Z})$ are an example of the following theorem.

Theorem 12.1. *Suppose that G and H have isomorphic normal subgroups $N_1 \trianglelefteq G$, $N_2 \trianglelefteq H$. $N_1 \cong N_2 = N$. Then $N \trianglelefteq G *_N H$ and*

$$\frac{G *_N H}{N} = G/N * H/N.$$

Proof. $G *_N H$ is the push-out (colimit) of the following diagram.

$$\begin{array}{ccc} N & \longrightarrow & G \\ \downarrow & & \\ & & H \end{array}$$

G/N is also given by a universal property: It is the pushout of the diagram

$$\begin{array}{ccc} N & \longrightarrow & G \\ \downarrow & & \downarrow \\ \{e\} = 1 & \longrightarrow & G/N \end{array}$$

To show that $\frac{G *_N H}{N} \cong G/N * H/N$ we need to show that: For any group X and homomorphism $G *_N H \rightarrow X$ which is trivial on N , there exists a unique homomorphism $G/N * H/N \rightarrow X$ making the appropriate diagram commute.

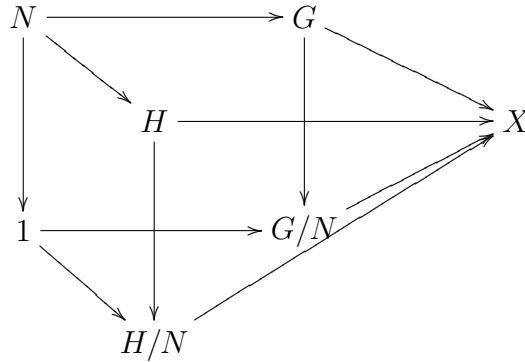
This condition is equivalent to the commuting diagram

$$\begin{array}{ccccc} N & \longrightarrow & G & & \\ & \searrow & \downarrow & \searrow & \\ & & H & \longrightarrow & X \\ & \downarrow & & \nearrow & \\ & 1 & & & \end{array}$$

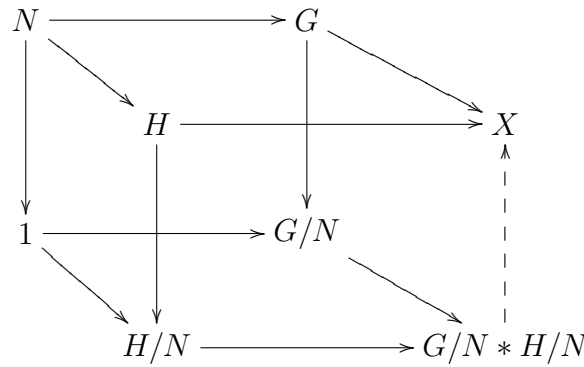
However, this commuting diagram includes the diagram

$$\begin{array}{ccc} N & \longrightarrow & G \\ \downarrow & & \downarrow \\ 1 & \longrightarrow & X \end{array}$$

So, there is an induced map $G/N \rightarrow X$ and similarly, there is an induced map $H/N \rightarrow X$ as indicated in the following diagram.



The two morphism $G/N \rightarrow X, H/N \rightarrow X$ induce a morphism from the free product $G/N * H/N \rightarrow X$:



This proves the theorem. □

I restated the theorem in terms of specific 2×2 matrices A, B

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2$$

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \quad B^2 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \quad B^3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2$$

So, $A^4 = I_2 = B^6$. The statement is

$$SL_2(\mathbb{Z}) \cong \langle A \rangle *_{\langle -I_2 \rangle} \langle B \rangle = \mathbb{Z}/4 *_{\mathbb{Z}/2} \mathbb{Z}/6$$

This means that every element of $SL_2(\mathbb{Z})$ has a unique expansion of one of the following two forms

- (1) $A^m B^{n_1} A B^{n_2} A \dots A B^{n_k}$
- (2) $A^m B^{n_1} A B^{n_2} A \dots A B^{n_k} A$

where, in both cases, each $n_i = 1$ or 2 and $m = 0, 1, 2$ or 3 .

Proof. By definition, elements of the amalgamated product $\langle A \rangle *_{\langle -I_2 \rangle} \langle B \rangle$ have the form

$$A^{m_1} B^{n_1} A^{m_2} B^{n_2} A^{m_3} \dots A^{m_k} B^{n_k}$$

subject to the condition that $A^4 = B^6 = I$ and $A^2 = B^3$. This last condition implies that, if any of the powers of A are 2 or more or any of the powers of B are 3 or more, then we can convert it into a power of the other letter and move it to the left. For example,

$$\begin{aligned} ABAB^2AB^4 &= ABAB^2AB(B^3) = ABAB^2(A^2)AB \\ &= ABA(B^3)B^2AB = \dots = (A^2)ABAB^2AB \end{aligned}$$

The only question is whether the last letter is A or B . □

12.2. free group as adjoint functor. If S is a set, let $G(S)$ be the free group generated by S . This is the set of all sequences (reduced words)

$$w = s_1^{n_1} s_2^{n_2} \dots s_k^{n_k}, \quad k \geq 0$$

where $n_i \in \mathbb{Z}$, $n_i \neq 0$ and $s_i \neq s_{i+1} \in S$. The length of w is $\ell(w) = \sum |n_i|$.

Theorem 12.2. $G : \mathcal{E}ns \rightarrow \mathcal{G}ps$ is adjoint to the forgetful functor F . I.e.,

$$\text{Hom}_{\mathcal{E}ns}(S, FH) \cong \text{Hom}_{\mathcal{G}ps}(GS, H)$$

Proof. The bijection G sends the mapping $f : S \rightarrow H$ to the group homomorphism $Gf : GS \rightarrow H$ given by

$$Gf(s_1^{n_1} s_2^{n_2} \dots s_k^{n_k}) = f(s_1)^{n_1} f(s_2)^{n_2} \dots f(s_k)^{n_k}.$$

The inverse is the restriction map; $G^{-1}(f : GS \rightarrow H) = f|_S$. □

12.3. actions and free products. As Ivan Horozov pointed out, we can tell that $PSL_2(\mathbb{Z})$ is a free product from the way that it acts on upper half-space. The following theorem, which is Exercise 54 on p.81, explains how this works.

Theorem 12.3. Suppose that $G_1, G_2, \dots, G_n \leq G$ are subgroups of G which generate G . Suppose that G acts on a set S . Suppose there are subsets $S_1, S_2, \dots, S_n \subseteq S$ and an element

$$s \in S \setminus \bigcup S_i$$

in the complement of the sets S_i with the following property. For all $g \in G_i$, $g_i \neq e$,

- (1) $g(S_j) \subseteq S_i$ for all $j \neq i$ and
- (2) $g(s) \in S_i$.

Then G is the free product of the groups G_i :

$$G = G_1 * G_2 * \cdots * G_n$$

Proof. By the universal property of the free product, there is a homomorphism

$$\phi : G_1 * G_2 * \cdots * G_n \rightarrow G$$

which is the inclusion map on each G_i . Since the groups G_i generate G , this homomorphism is onto. Thus, it suffices to show that the kernel of ϕ is trivial. So, suppose that there is an element in the kernel of ϕ . This has the form

$$g_1 g_2 \cdots g_k$$

where $g_j \neq e$ is an element of G_{i_j} and $i_j \neq i_{j+1}$.

Suppose for example that this element is $g_1 g_2 g_3$ where $g_1 \in G_5, g_2 \in G_9, g_3 \in G_4$. Then

$$g_1 g_2 g_3(s) \in G_5$$

since $g_3(s) \in S_4, g_2(g_3(s)) \in g_2(S_4) \subseteq S_9, g_1(g_2 g_3(s)) \in g_1(S_4) \subseteq S_5$. Therefore, $g_1 g_2 g_3 \neq e$. And in general, $g_1 \cdots g_k \neq e$ which implies that ϕ has a trivial kernel and is thus an isomorphism. \square

Corollary 12.4. $PSL_2(\mathbb{Z}) \cong \mathbb{Z}/2 * \mathbb{Z}/3$.

Proof. We apply the theorem to $G = PSL_2(\mathbb{Z}), G_1 = \langle A \rangle$ where $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $G_2 = \langle B \rangle$ where $B = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$. Let $S = \mathbb{H}$ be the upper half plane. Let

$$S_1 = \{z = x + iy \in \mathbb{H} \mid x < 0\}$$

Then $Az = -1/z$. So, A reverses the sign of the real part of z . Therefore, A sends $s = 1/4 + i$ and the set $\{z = x + iy \in \mathbb{H} \mid x > 0\}$ into S_1 . Let $S_2 = X \cup Y$ where

$$X = \{z \in \mathbb{H} \mid |z - 1| < 1 \text{ and } |z| \leq 1\}$$

$$Y = \{z = x + iy \in \mathbb{H} \mid x \geq 1/2 \text{ and } |z| > 1\}$$

Then $B(X) = Y, Bs \in X$ and $B(S_1) \subseteq X$. Therefore, the conditions of the theorem are satisfied and we conclude that $PSL_2(\mathbb{Z})$ is the free product of the subgroups generated by A and B . \square

