Math 57 Notes

Gaurav Goel

Summer 2020, Session 1

Contents

1 Introduction and Acknowledgements

This is an introductory course on group theory, with a category-theoretic flavor. The primary text for the course is Abstract Algebra by Dummit and Foote. It is based partly on the course Math 55 taught at Harvard each year. I make no claims to the originality of the content presented, and I want to thank my teachers, Professor Harris in particular, from whom I learned the present material. Any mistakes in this material, and there will be many, are due to me, and to me alone. Please keep me notified about any and all mistakes you spot.

2 A Crash Course in Logic and Set Theory

2.1 Preliminaries

We take as primitive the three concepts:
(a) The notion of ***containment***, i.e. what we mean by $a \in A$.
(b) The existence of the ***empty set***, denoted by $\varnothing$.
(c) For any statement p, either it is ***true***, denoted by $p \equiv \mathsf{T}$, or ***false***, denoted by $p \equiv \mathsf{F}$, but not both.

The following are logical connectives:
(a) Conjunction: the logical ***and***, denoted by $\wedge$.
(b) Disjunction: the logical ***or***, denoted by $\vee$. In logic, this is always the *inclusive* or.
(c) Exclusive disjunction: the logical ***xor***, denoted by $\oplus$.
(d) Negation: the logical ***not***, denoted by $\neg$. The negation is involutary: $\neg(\neg p) \equiv p$.
(e) Implication: the logical ***if-then***, denoted by $\Rightarrow$. By definition, $p \Rightarrow q \equiv (\neg p) \vee q$.
(f) Equivalence: the logical ***if and only if*** or ***iff***, denoted by $\Leftrightarrow$. Note $p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$.

Claim 1 (De Morgan's Laws) — These are the two claims:
(a) The negation of a disjunction is the conjunction of the negations, i.e. $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$.
(b) The negation of a conjunction is the disjunction of the negations, i.e. $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$.

Example 1
For $x \in \mathbb{R}$, the statement $(x^2 < 0 \Rightarrow x = 23) \equiv \mathsf{T}$. Such a statement is called ***vacuously true***.

For an implication $s \equiv p \Rightarrow q$, we call $\neg p \Rightarrow \neg q$ the ***reverse*** of s, $q \Rightarrow p$ the ***converse*** of s, and $\neg q \Rightarrow \neg p$ the ***contrapositive*** of s. It is a standard result that $p \Rightarrow q \equiv \neg q \Rightarrow \neg p$, i.e. an implication is true iff its contrapositive is.

The following are logical quantifiers:
(a) The ***universal quantifier***, denoted by $\forall$. By definition, $\forall a \in A : p(a) \equiv (a \in A) \Rightarrow p(a)$.
(b) The ***existential quantifier***, denoted by $\exists$. By definition, $\exists a \in A : p(a) \equiv \neg(\forall a \in A : \neg p(a))$.
(c) The ***uniqueness quantifier***, denoted by $\exists!$. By definition,

$$\exists! a \in A : p(a) \equiv (\exists a \in A : p(a)) \wedge ((a, b \in A : p(a) \wedge p(b)) \Rightarrow a = b).$$

The following are standard:
(a) Non-containment: $a \notin A$ iff $\neg(a \in A)$.
(b) Set-difference: $a \in A \setminus B$ iff $a \in A \wedge a \notin B$.
(c) Subset: Set A is ***contained in*** or a ***subset of*** set B, written $A \subseteq B$, if $\forall a \in A : a \in B$.
(d) Proper subset: $A \subset B$ (or $A \subsetneq B$) if $(\forall a \in A : a \in B) \wedge (\exists b \in B \setminus A)$.
(e) Union: $A \cup B := \{a : a \in A \vee a \in B\}$.
(f) Intersection: $A \cap B := \{a : a \in A \wedge a \in B\}$.
(g) Symmetric Difference: $A \Delta B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$.

Definition 1. For a collection $\mathscr{A}$ of sets, the ***union*** $\bigcup \mathscr{A} = \bigcup_{A \in \mathscr{A}} A := \{a : (\exists A \in \mathscr{A} : a \in A)\}$, and for any *nonempty* collection $\mathscr{A}$, the ***intersection*** $\bigcap \mathscr{A} = \bigcap_{A \in \mathscr{A}} A := \{a : (\forall A \in \mathscr{A} : a \in A)\}$.

Think about why we had to include *nonempty*. Have you heard of the concept of a ***universal set*** ξ?

2.2 Cartesian Products, Relations and Functions

Define the ***ordered pair*** $(a, b) := \{\{a\}, \{a, b\}\}$. Therefore, $(a, b) = (c, d) \Leftrightarrow a = c \wedge b = d$. Then, define the ***cartesian product*** $A \times B := \{(a, b) : a \in A \wedge b \in B\}$.

Definition 2. A ***relation*** R between elements of set A and set B is any subset $R \subseteq A \times B$. If $(a, b) \in R$, we say a is ***related to*** b, and write aRb.

We often have to deal with three special kinds of relations: functions, equivalence relations, and order relations.

Definition 3. A relation $f \subseteq A \times B$ is said to be a ***function*** or a ***mapping***, written $f : A \to B$, if
(a) (Definition) $\forall a \in A : \exists b \in B : afb$, and
(b) (Uniqueness of Definition) $\forall a \in A, \forall b, b' \in B : afb \wedge afb' \Rightarrow b = b'$.

Usually, for functions, we denote afb by $f(a) = b$, and b is called the ***value*** of f at a.[1] For a function $f : A \to B$, A is called the ***domain of definition*** or simply ***domain*** of f, and B is called the ***codomain*** of f. For a subset $S \subseteq A$, the set $f(S) := \{f(s) : s \in S\} \subseteq B$ is called the ***image*** of S under f. The set $f(A) \subseteq B$ is called the ***image*** of f.

Example 2

The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is distinct from the function $g : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ given by $g(x) = x^2$. For instance, the latter function has an ***inverse***, but the former doesn't.

For any subset $T \subseteq B$, the set $f^{-1}(T) = \{a \in A : f(a) \in T\} \subseteq A$ is called the ***pre-image*** of T. For any element $b \in B$, the pre-image $f^{-1}(\{b\})$ is, by abuse of notation, written as $f^{-1}(b)$, and is called the ***fiber*** over b. This will be important when we later talk about fibered products and fibered sums in a category.

Definition 4. Two important classes of functions:
(a) A function f is said to be ***injective***, written $f : A \hookrightarrow B$, if $\forall a, a' \in A : f(a) = f(a') \Rightarrow a = a'$.
(b) A function f is said to be ***surjective***, written $f : A \twoheadrightarrow B$, if $\forall b \in B : \exists a \in A : f(a) = b$.
(c) A function f is said to be ***bijective***, written $f : A \xrightarrow{\sim} B$, if it is both injective and surjective.

The following is a handy characterization of bijective functions:

Claim 2 — A function $f : A \to B$ is bijective iff there is a function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.

Proof. If f is bijective, then take $g = f^{-1}$. For the converse, suppose that f and g are as described. If $f(a) = f(a')$, then applying g to both sides, $g \circ f(a) = g \circ f(a') \Rightarrow a = a'$ so that f is injective. If $b \in B$, then $g(a) \in A : f \circ g(a) = b$, so f is surjective. ∎

The set of all functions $f : A \to B$ is denoted by B^A. (Think about why.) For the set $2 := \{0, 1\}$, the set 2^A is called the ***power set*** of A, written often as $\wp(A)$.

Example 3

The functions $\pi_1 : A \times B \to A$ given by $\pi_1(a, b) = a$ and $\pi_2 : A \times B \to B$ given by $\pi_2(a, b) = b$ are called the ***projection maps*** onto the first and second factors respectively. If none of A and B are empty, then these are surjective.

Think about what happens in the case either of A and B is empty.

2.3 Binary Relations

[1] Observe that we are allowed to write $f(a) = b$ without violating the transitivity of equality only because of condition (b). Something like $\ln : \mathbb{C}^\times \to \mathbb{C}$ is not actually a function, and that's why you get problems like $0 = \ln 1 = 2\pi\imath$ if you try to write $\ln z = w$. In this case, ln is what is called a ***multi-valued*** or ***generalized*** function. Since such generalized functions arise primarily in complex analysis, we will not study them here.

Definition 5. A ***binary relation*** $\star$ on a set A is a relation $\star \subseteq A^2 := A \times A$.

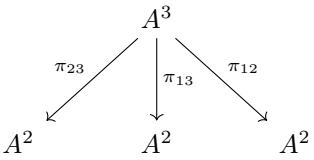
Example 4

The ***diagonal*** $\Delta := \{(a, a) : a \in A\} \subseteq A^2$ is the binary relation corresponding to equality.

2.3.1 Equivalence Relations

Definition 6. A binary relation $\sim$ on A is an ***equivalence relation*** if it satisfies the following axioms:
(a) (Reflexivity) $\forall a \in A : a \sim a$.
(b) (Symmetry) $\forall a, b \in A : a \sim b \Rightarrow b \sim a$.
(c) (Transitivity) $\forall a, b, c \in A : (a \sim b \wedge b \sim c) \Rightarrow a \sim c$.
If $a \sim b$, then a is said to be ***equivalent*** to b.

Equivalently, an equivalence relation is a subset $\Phi \subseteq A^2$ s.t.
(a) The diagonal $\Delta \subseteq \Phi$.
(b) It is preserved under the involution $\sigma : A^2 \to A^2$ that swaps factors $(a, b) \mapsto (b, a)$, i.e. $\sigma(\Phi) = \Phi$.
(c) If we have the projection maps as shown:



then $\pi_{13}(\pi_{12}^{-1}\Phi \cap \pi_{23}^{-1}\Phi) \subseteq \Phi$.

If $\sim$ is an equivalence relation on A, then for any a, the set $\{x \in A : x \sim a\}$ is called the ***equivalence class*** of a under $\sim$. If $\mathcal{C}$ is an equivalence class, any element $a \in \mathcal{C}$ is called a ***representative*** of the class $\mathcal{C}$. Two equivalence classes are either disjoint or equal, and so they form a ***partition*** of A.

Definition 7. A ***partition*** of A is any collection $\mathscr{A} = \{A_i : i \in I\}$ s.t.
(a) (Nonempty Subsets) $\forall i \in I : \varnothing \subsetneq A_i \subseteq A$.
(b) (Cover) $\bigcup_{i \in I} A_i = A$.
(c) (Pairwise Disjoint) $\forall i, j \in I : i \neq j \Rightarrow A_i \cap A_j = \varnothing$.
If $\mathscr{A} = \{A_i : i \in I\}$ is a partition of A, then we write $A = \coprod_{i \in I} A_i$.

Example 5

If $\omega_1, \omega_2 \in \mathbb{C}$, then the relation $z \sim w$ if $\exists m, n \in \mathbb{Z} : z - w = m\omega_1 + n\omega_2$ is an equivalence relation on $\mathbb{C}$, written $z \equiv w \pmod{\Lambda}$, where $\Lambda = \mathbb{Z}\langle \omega_1, \omega_2 \rangle$. Two complex numbers $\omega_1, \omega_2 \in \mathbb{C}$ are said to be ***linearly independent*** over $\mathbb{R}$ if for $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1\omega_1 + \lambda_2\omega_2 = 0 \Rightarrow \lambda_1 = \lambda_2 = 0$. If ω_1 and ω_2 are linearly independent over $\mathbb{R}$, then they induce a partition of $\mathbb{C}$ into parallelograms.

2.3.2 Order Relations and Posets

Definition 8. A binary relation $\preceq$ on a set A is called a ***(non-strict) partial order relation*** if it satisfies the following axioms:
(a) (Reflexivity) $\forall a \in A : a \preceq a$.
(b) (Antisymmetry) $\forall a, b \in A : (a \preceq b) \wedge (b \preceq a) \Rightarrow a = b$.
(c) (Transitivity) $\forall a, b, c \in A : (a \preceq b \wedge b \preceq c) \Rightarrow a \preceq c$.
An ordered pair $(A, \preceq)$ of a set A along with a specified partial order $\preceq$ on A is called a ***partially ordered set*** or a ***poset***.

Definition 9. A partial order $\preceq$ on a set A is called a ***total order*** if in addition to the above it satisfies the axiom:
(d) (Comparability) $\forall a, b \in A : a \preceq b \vee b \preceq a$.
An ordered pair $(A, \preceq)$ of a set and a total order on A is called a ***totally ordered set***.

Example 6
$(\mathbb{R}, \leqslant)$ is a totally ordered set. For any set A, the power set $\wp(A)$ ordered by inclusion $\subseteq$ is a poset. As soon as $|A| > 2$, $(\wp(A), \subseteq)$ is *not* a totally ordered set.

2.4 Infinite Cartesian Products

Let $\mathscr{A}$ be a nonempty collection of sets. An ***indexing function*** is a surjection $f : I \twoheadrightarrow \mathscr{A}$ for some set I, called the ***index set***. Given $i \in I$, denote the set $f(i)$ by A_i. Then $\mathscr{A} = \{A_i : i \in I\}$ is called an ***indexed family of sets***.

The most common index sets are the sets of the form $[n] := \{1, 2, \cdots, n\}$ and the set $\mathbb{N}$ of all positive integers. Note that in this course, we let $\mathbb{N} := \mathbb{Z}_{>0}$. Note that some people define it with $\geqslant 0$; it is *mostly* a matter of preference. We then define a ***tuple***.

Definition 10. Let $n \in \mathbb{N}$. Given a set X, define an ***n-tuple*** of elements of X to be a function $\mathbf{x} : [n] \to X$. If $\mathbf{x}$ is an n-tuple, we denote $\mathbf{x}(i)$ by x_i, and call it the i^{th} ***coordinate*** of $\mathbf{x}$, and write $\mathbf{x} = (x_1, \ldots, x_n)$.

Definition 11. Let $\mathscr{A} = \{A_1, \ldots, A_n\}$ be a family of sets indexed by $[n]$. Let $X = \bigcup_{i=1}^{n} A_i = \bigcup \mathscr{A}$. Then the ***cartesian product*** of this family, denoted by $\prod_{i=1}^{n} A_i$ or $A_1 \times \cdots \times A_n$ to be the set of n-tuples $(x_1, \cdots, x_n)$ s.t. $\forall i \in [n], x_i \in A_i$.

This motivates the following general definition:

Definition 12. Let $\mathscr{A}$ be a family of sets indexed by I. Let $X = \bigcup_{i \in I} A_i$. Then the ***cartesian product*** of this family, denoted by $\prod_{i \in I} A_i$ is the set of functions $\mathbf{x} : I \to X$ s.t. $\forall i \in I, x_i := \mathbf{x}(i) \in A_i$. We define the projection maps $\pi_\alpha : \prod_{i \in I} A_i \to A_\alpha$ by $\mathbf{x} \mapsto x_\alpha$.

As we shall later see, this is the categorical product in the category (Set).

2.5 The Axiom of Choice

Everyone's heard of the Axiom of Choice. Everyone's scared of it. No one actually knows what it's actually about. Let's figure out. It turns out, it is closely related to the infinite cartesian products we were talking about.

Let us set up again. Let $\mathscr{A}$ be a nonempty collection of nonempty sets. The cartesian product $\prod \mathscr{A}$ can equivalently be thought of as the set of functions $c : \mathscr{A} \to \bigcup \mathscr{A}$ such that $\forall A \in \mathscr{A}, c(A) \in A$. But who's to say that any such function exist at all? Well, duh: obviously, there are such functions–right? For finite collections of finite sets: the answer is unamibiguously yes. This is called the ***axiom of finite choice***, and it accepted unequivocally by all mathematicians. Yet for infinite collections and sets, the answer is not so clear. This innocent-seeming question, as it turns out, has far-reaching philosophical implications. At the turn on the 20th century, when there was a turmoil for establishing mathematics on a firmer foundation, Zermelo and Fraenkel came up with a system of axioms for set theory, which we denote by ZF. However, as fate would have it, it was later shown that this system of axioms cannot answer the above question in the positive or negative–that this system admits models in which the above question can have either answer, and as such can neither prove nor disprove the existence of such functions. Therefore, we have to consider an additional *axiom*:

Axiom (The Axiom of Choice)

Let $\mathscr{A}$ be a collection of nonempty sets. Then there is a ***choice function*** $c : \mathscr{A} \to \bigcup \mathscr{A}$ s.t. $\forall A \in \mathscr{A} : c(A) \in A$.

The Zermelo-Fraenkel system of axioms, along with the new axiom of choice, is denoted by ZFC. It turns out to be a much more powerful system than ZF alone. The reason this axiom is so controversial is that if you accept it, you also have to accept a series of unintuitive results and unanticipated consequences that follow from it. The most (in)famous of these is the Banach-Tarski "Paradox". There are too many equivalent formulations of this axiom to discuss here, but we will talk about the two that are the most common and easiest to apply. Suffice it to say that any time you're making an "arbitrary choice of infinitely many elements," you're probably invoking the Axiom of Choice.

A related concept is well-ordering:

Definition 13. Let $(A, \leqslant)$ be a totally ordered set and $\varnothing \subsetneq B \subseteq A$ any subset. An element $b_0 \in B$ is called the ***smallest element*** of B if $\forall b \in B : b_0 \leqslant b$. A totally ordered set $(A, \leqslant)$ is said to be ***well-ordered*** by $\leqslant$ if every nonempty subset *contains* a smallest element.

Observe that, by antisymmetry, if a smallest element exists, then it is unique, so we may speak of *the* smallest element.

Example 7

It can be proven by induction that $(\mathbb{N}, \leqslant)$ is well-ordered. (Try it!) On the other hand, $(\mathbb{R}, \leqslant)$ or even $(\mathbb{Q}, \leqslant)$ is not well-ordered because the interval $(0, 1)$ in either does not *contain* a smallest element.

A completely unexpected equivalent formulation of the Axiom of Choice is:

Proposition 1 (Well-Ordering Principle)

If A is any set, then there is a total order relation on A that is a well-ordering.

The following is taken verbatim from *Munkres*: "This theorem was proved by Zermelo in 1904, and it startled the mathematical world. There was considerable debate as to the correctness of the proof; the lack of any constructive procedure for well-ordering an arbitrary uncountable set led many to be skeptical. When the proof was analyzed closely, the only point at which it was found that there might be some question was a construction involving an infinite number of arbitrary choices, that is, a construction involving–the choice axiom." The proof that the AoC implies the WOP is rather long, so we omit it. It can be found in the exercises in Munkres.

Let's look at another equivalent proposition formulated by Hausdorff in 1914.

Proposition 2 (Hausdorff Maximum Principle)

Let $(A, \preceq)$ be a poset. Then there is a maximal totally ordered subset of A. In other words, there is a subset $B \subseteq A$ that is totally ordered by $\preceq$ and such that it is not contained in a larger totally ordered subset of A.

One can give an intuitive proof: pick an arbitrary element of A and throw it in B. Next pick any element of A and check if it is comparable with the element in your box: if so, put it in; else, throw it away. Continue till you are done checking all elements of A. Then every element not in the box will be noncomparable with at least one element in the box, because that is why it was thrown. Of course, the problem with this "proof" is: how do you know if you are done checking? That is where the well-ordering principle comes in!

Proof. If A is empty, we are done. Hence, assume A is nonempty. By the well-ordering principle, we may bijectively index A by a nonempty well-ordered set J, writing $A = \{a_\alpha : \alpha \in J\}$. Define a function $\chi : J \to \{0, 1\}$ by saying $\chi(\alpha) = 0$ if we "toss a_α away" and 1 if we "put a_α in the box." More

formally, by well-ordering, J has a smallest element, say α_0. Define $\chi(\alpha_0) = 1$ and for every $\alpha > \alpha_0$, set $\chi(\alpha) = 1 \Leftrightarrow a_\alpha$ is comparable under $\preceq$ to every element of $\{a_\beta | \beta < \alpha \wedge \chi(\beta) = 1\}$. Then we claim that $B = \{a_\alpha : \chi(\alpha) = 1\}$ is a maximal simply ordered subset of A. To show this, assume that $\alpha \neq \beta : \chi(\alpha) = \chi(\beta) = 1$; WLOG assume that $\beta < \alpha$. Then $\beta < \alpha \wedge \chi(\beta) = 1$ along with $\chi(\alpha) = 1$ implies that a_α is comparable to a_β–this proves that any two elements of B are comparable, so that B is totally ordered. On the other hand, it is maximal because if $\exists \eta \in J : a_\eta$ is not in B but comparable to every element of B, then $\chi(\eta) = 0$ AND a_η is comparable to every element of $\{a_\beta | \beta < \eta \wedge \chi(\beta) = 1\} \subseteq B$ so that $\chi(\eta) = 1$, a contradiction. ■

Another important formulation is Zorn's Lemma. We start with one confusing pair of terms:

Definition 14. Let $(A, \preceq)$ be a poset.
(a) An element $m \in A$ is called a ***maximal element*** or a ***non-dominated element*** if $\forall a \in A : m \preceq a \Rightarrow m = a$.
(b) An element $u \in A$ is called an ***upper bound*** of A if $\forall a \in A : a \preceq u$.

Observe that, by antisymmetry, every upper bound is maximal, but the converse it not necessarily true. The key difference between these terms lies in the fact that a maximal element need not be related to all elements, and consequently there can be more than one maximal element. However, an upper bound is necessarily related with all elements by definition. This is highlighted by the fact that is A is totally ordered by $\preceq$, then these two notions coincide.

With this, the formulation of the Axiom of Choice that we use most often is:

Lemma 1 (Zorn's Lemma)
Let $(A, \preceq)$ be a poset. If every subset of A that is totally ordered by $\preceq$ has an upper bound in A, then A has a maximal element.

Proof. By the Maximum Principle, A has a maximal totally ordered subset B. By hypothesis, $\exists u \in A : \forall b \in B : b \preceq u$. We claim that u is a *maximal element* of A: if $a \in A : u \preceq a$ and $u \neq a$, then $a \notin B$, but because a is comparable to u, it is comparable to every element of B. Then the set $B \cup \{a\}$ would be a larger totally ordered subset of A, contradicting the maximality of B. ■

Another really useful equivalent formulation (whose equivalence is immediate by taking $\preceq$ to be inclusion $\subseteq$) is given as follows:

Lemma 2 (Kuratowski's Lemma)
Let $\mathscr{A}$ be a collection of sets such that for all subcollections $\mathscr{B} \subseteq \mathscr{A}$ that are totally ordered by $\subseteq$, we have $\bigcup \mathscr{B} \in \mathscr{A}$. Then $\mathscr{A}$ has an element that is properly contained in no other.

Finally, we mention only that all of these statements are, in fact, equivalent to each other and to the following:

Lemma 3
Every vector space V has a basis.

Proof Sketch. Consider the collection $\mathscr{A}$ of sets of linearly independent elements, and apply Kuratowski's Lemma. This maximal element of $\mathscr{A}$ is the required basis. ■

Example 8
A basis for $\mathbb{R}$ as a vector space over $\mathbb{Q}$ is called a ***Hamel basis***.

Joke. The Axiom of Choice is obviously true, the Well-Ordering Principle obviously false, and who can say about Zorn's Lemma?

Why is this a joke? A detailed discussion of this and more topics can be found in *Halmos* or *Munkres*. We shall limit our discussion of this topic here, and return to this later if needed.

2.6 Cardinalities of Sets

We are quite familiar with the theory of the cardinality of finite sets. This section is devoted to the theory of cardinality of larger sets. The basic tool at our disposal is:

Claim 3 — If $f : A \hookrightarrow B$ is an injection, then the map $\tilde{f} : A \to f(A)$ by $a \mapsto f(a)$, i.e. the map formed by restricting the image of f, is a bijection. Therefore, we may consider A to be a subset of B by the identification $a \sim f(a)$.

We are now ready to speak of cardinalities.

Definition 15. If there is an injection $f : A \hookrightarrow B$, then we say that the cardinality of A is at most that of B, and write $|A| \leqslant |B|$. If there is a surjection $f : A \twoheadrightarrow B$, then we say that the cardinality of A is at least that of B, and write $|A| \geqslant |B|$. If there is a bijection $f : A \xrightarrow{\sim} B$, then we say that A and B have the same cardinality, and write $|A| = |B|$.

Think about what we could mean by $|A| < |B|$.

Example 9

If $B \subseteq A$ is any subset, then $|B| \leqslant |A|$ and $|A| \geqslant |B|$.

It is clear, that if $f : A \hookrightarrow B$ is an injection, then the function $g : B \to A$ defined by sending $f(a) \mapsto a$ and other elements arbitrarily is a surjection. For finite sets, if $f : A \twoheadrightarrow B$ is a surjection, then by arbitrarily choosing an element of each fiber, we get an injection $B \hookrightarrow A$. However, it is not obvious for infinite sets that if $f : A \twoheadrightarrow B$ is a surjection, then there is an injection in the opposite direction–and indeed, this needs the Axiom of Choice (in fact, this is one of the equivalent formulations). Further, it is not obvious at all if $|A| \leqslant |B|$ and $|B| \leqslant |A|$ that $|A| = |B|$: and that is the content of the Schröder-Bernstein Theorem.

Theorem 1 (Schröder-Bernstein)
If $B \subseteq A$ is any subset, and $f : A \hookrightarrow B$ is an injection, then $|A| = |B|$.

Proof. Define the map $h : A \to B$ by

$$h(a) = \begin{cases} f(a), & \text{if } \exists n \geqslant 0 : a \in f^n(A \setminus B), \\ a, & \text{otherwise.} \end{cases}$$

We claim that $h : A \xrightarrow{\sim} B$. To show injectivity, the only case that needs thought is when $f(a) = b$ for some $a \in f^n(A \setminus B)$ and $b : \forall n \geqslant 0 : b \notin f^n(A \setminus B)$, which is not possible. To show surjectivity, if $b \in B$ then either $\forall n \geqslant 0, b \notin f^n(A \setminus B)$ so $h(b) = b$ works OR $\exists n \geqslant 0 : b \in f^n(A \setminus B)$. Since $b \in B$, $n \geqslant 1$. ■

Corollary 1.1 — If $f : A \hookrightarrow B$ and $g : B \hookrightarrow A$ are injections, then $|A| = |B|$.

Proof. By Claim 3, g allows us to identify B with the subset $g(B) \subseteq A$. We're done since $|B| = |g(B)|$. ■

This means that our definitions about cardinalities are consistent with the usual usage of the symbols $\leqslant, \geqslant$ and $=$.

Example 10

The map $f : \mathbb{Z} \to \mathbb{N}$ by $f(n) = \begin{cases} 2n, & n \geqslant 1, \\ 1 - 2n, & n \leqslant 0 \end{cases}$ is a surjection. Therefore, $|\mathbb{Z}| = |\mathbb{N}|$.

We now define the three most basic types of cardinalities:

Definition 16. Recall that for $n \in \mathbb{N}$, $[n] := \{1, 2, \cdots, n\}$.
(a) A set A is said to be ***finite*** if $\exists n \in \mathbb{N} : |A| = |[n]|$. Otherwise, it is said to be ***infinite***.
(b) A set A is said to be ***countably infinite***, or simply ***countable***, if $|A| = |\mathbb{N}|$.
(c) A set A is said to be ***at most countable*** if it is either finite or countably infinite. Otherwise, it is said to be ***uncountable***.

Example 11

The cardinality of the set of natural numbers is denoted by $\aleph_0$. In other words, when we say $|A| = \aleph_0$, we mean that A is countably infinite. We denote the cardinality of the set $\mathbb{R}$ by $\mathfrak{c}$, for ***continuum***.

The following are standard results that we'll not prove here, and leave as (fun) exercises to the reader.
(a) A subset of an at most countable set is at most countable.
(b) A countable union of at most countable sets is at most countable.
(c) A finite product of at most countable sets is at most countable. (Warning: a countable product of countable sets need not be countable!)

We prove some basic results about cardinalities.

Theorem 2 (Countability of Rationals)
$|\mathbb{Q}| = \aleph_0$.

Proof. The map $f : \mathbb{Q} \to \mathbb{N}$ given by sending $0 \mapsto 1$ and sending $q \neq 0$, written in lowest terms as $q = \pm a/b$ with $a, b \in \mathbb{N}$, to $f(q) = 2^{\operatorname{sign} q + 1} 3^a 5^b$ is an injection. Finish by Theorem 1. ■

The following is a more sophisticated version of the standard Cantor's "diagonal argument":

Theorem 3 (Cantor's Theorem)
If A is any set, the $|A| < |\wp(A)|$.

Proof. The map $f : A \to \wp(A)$ by $a \mapsto \{a\}$ is an injection. Suppose $g : A \twoheadrightarrow \wp(A)$. Then let $S = \{a \in A : a \notin g(a)\} \in \wp(A)$. By surjectivity, $\exists s \in A : g(s) = S$. But then $s \in S \Leftrightarrow s \notin S$, a contradiction. ■

This seemingly innocent but powerful statement allows us to establish the existence of uncountable sets, and in fact of a tower of infinite sets, each with strictly large cardinality than the previous one.

Corollary 3.1 — The reals are uncountable. In other words, $\aleph_0 < \mathfrak{c}$. In fact, $\mathfrak{c} = 2^{\aleph_0}$.

Proof. Observe that by the bijection $t \mapsto \frac{t-1/2}{t(1-t)}$, $|\mathbb{R}| = |(0,1)|$. But by the maps $t \mapsto \frac{t-1/2}{2}$ in both directions and by Schröder-Berstein, it is clear that $|(0,1)| = |[0,1]|$. The map $\wp(\mathbb{N}) \to [0,1]$ given by $S \mapsto \sum_{s \in S} 2^{-s}$ is a bijection. Therefore, $|\mathbb{N}| < |\wp(\mathbb{N})| = |\mathbb{R}|$ by Cantor. ■

We end with a nice characterization of infinite sets. Again, this is not something we prove here, but something that can be found in all standard set-theory books, like *Halmos*.

Claim 4 — Let A be a set. The following are equivalent:
(a) There is an injection $f : \mathbb{N} \hookrightarrow A$.
(b) There is a bijection of A with a proper subset of itself.
(c) A is infinite.

3 Groups

3.1 Introduction and Examples

Definition 17. Let G be a set.
(a) A ***binary operation*** $\star$ on a set G is a function $\star : G \times G \to G$. For $a, b \in G$, denote $\star(a, b)$ by $a \star b$.
(b) A binary operation $\star$ on G is ***associative*** if $\forall a, b, c \in G : (a \star b) \star c = a \star (b \star c)$.
(c) A binary operation $\star$ on G is ***commutative*** if $\forall a, b \in G : a \star b = b \star a$.

Suppose $\star$ is a binary operation on G and $H \subseteq G$. If $\forall a, b \in H : a \star b \in H$, then H is said to be ***closed*** under the operation $\star$.

Definition 18. An ordered pair $(G, \star)$ of a set G along with a specified binary operator $\star$, called a ***law of composition***, on G is called a ***group*** if it satisfies the following axioms:
(a) (Associativity) The operation $\star$ is associative.
(b) (Existence of Identity) $\exists e \in G : \forall a \in G : a \star e = e \star a = a$.
(c) (Existence of Inverse) $\forall a \in G : \exists a^{-1} \in G : a \star a^{-1} = a^{-1} \star a = e$.
If the pair $(G, \star)$ satisfies the following additional axiom:
(d) (Commutativity) $\forall a, b \in G : a \star b = b \star a$,
then the group is called ***abelian*** or ***commutative***. A group which does not satisfy (d) is called ***nonabelian*** or ***noncommutative***.

Often, when the group operation is understood, we abuse terminology, and call the set G a group. Further, we call the cardinality $|G|$ of the underlying set the ***order*** of the group.

Example 12
The ***trivial*** group $G = \{e\}$ is the unique group of order 1.

Example 13
The set $\{\mathsf{T}, \mathsf{F}\}$ with law of composition $\wedge$ is not a group, but with law of composition $\oplus$ is a group.

Example 14
$(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +)$ are all groups, but $(\mathbb{N}, +)$ is not. $(\mathbb{Q} \setminus \{0\}, \cdot), (\mathbb{R} \setminus \{0\}, \cdot), (\mathbb{C} \setminus \{0\}, \cdot)$ are groups, but $(\mathbb{Z} \setminus \{0\}, \cdot)$ is not. For the sake of brevity, for $F = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, we let $F^\times := F \setminus \{0\}$.

We introduce the following bit of terminology that we won't use much:

Definition 19. Let G be a set and $\star$ a binary operation on G.
(a) A structure $(G, \star)$ is called a ***semigroup*** if it satifies (a), but not necessarily (b) or (c).
(b) A structure $(G, \star)$ is called a ***monoid*** if it satisfies (a) and (b), but not necessarily (c).

$(\mathbb{N}, +)$ is a semigroup , whereas $(\mathbb{Z} \setminus \{0\}, \cdot)$ and $(\mathbb{N}_0, +)$ are monoids.

Example 15
The set of all symmetries (isometries) of a plane figure X forms a group under composition, denoted by $\operatorname{Sym} X$. For any set X, a bijection $\sigma : X \to X$ is called a ***permutation***, and the set of permutations $\sigma : X \to X$ forms a group under composition, denoted by S_X or $\mathfrak{S}_X$ or $\operatorname{Perm}(X)$. For $X = [n]$, the group $S_n := S_{[n]}$, also written $\mathfrak{S}_n$, is called the ***symmetric group*** on n letters. It has order $n!$. As soon as $n \geqslant 3$, S_n is nonabelian.

Example 16

For $n \in \mathbb{N}$ the group $\mathbb{Z}/n\mathbb{Z}$ is a group under the usual addition modulo n; it is called the ***cyclic group*** of order n.
For any prime p, the set $\mathbb{Z}_{(p)} := \{\frac{a}{b} \in \mathbb{Q} : (a, b) = 1, p \nmid b\}$ is a group under addition.
In fact–both of these are more than just groups under addition: they are also closed under multiplication. Such a structure is called a ***ring***. We'll have more to say about rings later.

Example 17

For any $n \in \mathbb{N}$, the set of elements of $\mathbb{Z}/n\mathbb{Z}$ that are coprime to n form a group under multiplication. (This is because if $(a, n) = 1$, then the map $\varphi_a : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ given by left-multiplication by a is a bijection.) This group is denoted by $(\mathbb{Z}/n\mathbb{Z})^\times$ and has order $\varphi(n)$.

Example 18

The set of $n \times n$ real matrices and non-zero determinant form a group under multiplication, called the ***general linear group*** of ***degree*** n ***over*** $\mathbb{R}$, and denoted by $\mathrm{GL}_n(\mathbb{R})$. As soon as $n \geqslant 2$, $\mathrm{GL}_n(\mathbb{R})$ is *not* abelian. The group $\mathrm{GL}_n(\mathbb{C})$ is defined similarly.

Example 19

The ***circle group*** $\mathbb{S}^1$ is the immensely important group $\{z \in \mathbb{C} : |z| = 1\}$ under usual multiplication.

From the definitions, the following are straightforward:

Claim 5 — If G is a group under $\star$, then:
(a) (Uniqueness of Identity) The identity element e is unique.
(b) (Uniqueness of Inverses) For each $a \in G$, $\exists! a^{-1} \in G$. This allows us to define a map $\mathrm{inv} : G \to G$ by $a \mapsto a^{-1}$.
(c) (inv is an Involution) $\forall a \in A, (a^{-1})^{-1} = a$.
(d) (Contravariant Nature of inv) $\forall a, b \in G, (a \star b)^{-1} = (b^{-1}) \star (a^{-1})$.
(e) (Generalized Associativity) $\forall a_1, \cdots, a_n \in G$, the value of $a_1 \star a_2 \star \cdots \star a_n$ is independent of the order of parenthetization.
(f) (Cancellation Laws) For any $a, b, u, v \in G$, $a \star u = a \star v \Rightarrow u = v$ and $u \star b = v \star b \Rightarrow u = v$.

In general, we say a group is written ***multiplicatively*** if we "pretend" that $\star$ is multiplication $(\cdot)$. Then, we omit the operation symbol. Some authors also write 1 for the identity element e, but we will avoid that convention here; further, if we need to emphasize what group the identity element belongs to, we will write it as e_G. Finally, we write x^n for $\underbrace{x \cdot x \cdots x}_{n \text{ times}}$, and similarly for negative powers.

If $(G, \star)$ is abelian, then it is customary to write G ***additively***, i.e. we "pretend" that $\star$ is addition $(+)$. Then we write 0 for e and $-a$ for a^{-1}. Finally, we write na for $\underbrace{a + a + \cdots + a}_{n \text{ times}}$, and similarly for negative n. Observe that under this convention, $0a = 0$: the zero on the right is the identity element of the group, but zero on the left is *not* the identity element of the group–it simply reflects the fact that we're adding nothing. This notation is usually not used for nonabelian groups.

Henceforth, unless specified otherwise, we write all groups multiplicatively.

Proof. This can be found in Dummit and Foote, Chapter 1.
(a) If e, e' are identity elements then $e = ee' = e'$.
(b) Assume b, c are inverses of a; then $b = be = b(ac) = (ba)c = ec = c$.
Further, (c) and (d) are trivial, (e) can be shown by induction, and (f) is easy by multiplication by inverses. ∎

Definition 20. If $(G, \star)$ is a group,and if $H \subseteq G$ is closed under $\star$ and preserved under inv, then H is said to be a ***subgroup*** of G, written $H \leqslant G$. In other words, $H \leqslant G$ iff $\forall a, b \in H, a \star b \in H$ and $\forall a \in H, a^{-1} \in H$.

Example 20

For any group G, the trivial subgroup $\{e_G\}$ and the whole group G are subgroups of G.

A subgroup $H \leqslant G$ is ***proper***, written $H < G$, if $H \subsetneq G$. It is ***nontrivial*** if $\{e_G\} < H$.

Example 21

For any $n \in \mathbb{N}, (n\mathbb{Z}, +) \leqslant (\mathbb{Z}, +)$.
If T is any triangle in the plane, then rotations form a subgroup of $\operatorname{Sym} T$.

Example 22

The subset of $n \times n$ real matrices with determinant 1 form a subgroup of $\mathrm{GL}_n(\mathbb{R})$, called the ***special linear group*** of ***degree*** n ***over*** $\mathbb{R}$, and denoted by $\mathrm{SL}_n(\mathbb{R})$. Note that while the set of invertible integer matrices "$\mathrm{GL}_n(\mathbb{Z})$" is not a group under multiplication, $\mathrm{SL}_n(\mathbb{Z})$ very well is. (We'll actually see later that for any ring R, the group $\mathrm{GL}_n(R)$ is actually defined slightly differently.)

Example 23

The group $\mathcal{Z} = \{z \in \mathbb{C} : \exists n \in \mathbb{N} : z^n = 1\}$ with law of composition the usual multiplication is called the group of ***roots of unity*** in $\mathbb{C}$. Then $\mathcal{Z} \leqslant \mathbb{S}^1$.

With these definitions, it is natural to ask: what are all the subgroups of a given group? In general, this is a difficult question to answer. However, we can do it rather easily for some special cases.

Claim 6 — The only subgroups of $(\mathbb{Z}, +)$ are the $(n\mathbb{Z}, +)$.

Proof. If $H \leqslant \mathbb{Z}$ is trivial, then $H = 0\mathbb{Z}$. If H is nontrivial, it has some positive element. By the Well-Ordering Principle, we may choose a *smallest* positive element, say n. Then if H contained an element m s.t. $n \nmid m$, then by the Euclidean algorithm, we could derive a contradiction to the minimality of n. Therefore, $H = n\mathbb{Z}$. Finally notice that $(-n)\mathbb{Z} = n\mathbb{Z}$. ∎

The following is a handy characterization of subgroups.

Claim 7 (The Subgroup Criterion) — If $H \subseteq G$, we have $H \leqslant G$ iff $H \neq \varnothing$ and $\forall g, h \in H : gh^{-1} \in H$. If H is finite, it suffices to check that H is closed under the group operation.

Proof. If $H \leqslant G$, then the claim holds. Conversely, $H \neq \varnothing \Rightarrow \exists g \in H \Rightarrow e = gg^{-1} \in H$. Then $g^{-1} = eg^{-1} \in H$, so that H is preserved under inv. Finally, $gh = g(h^{-1})^{-1} \in H$, so it is closed under $\star$. If H is finite, then for any $g \in H$, the set $\{e, g, g^2, \cdots\}$ is finite, so $\exists i > j : g^i = g^j$. By cancellation, $g^{i-j-1} = g^{-1} \in H$, so H is automatically preserved under inv . ∎

3.2 Generators and Cyclic Groups

Observe that if G is a group and $H, H' \leqslant G$ are subgroups, then so is $H \cap H'$. This leads us to the following important concept:

Definition 21. Given a group G and any subset $S \subseteq G$, the smallest subgroup of G containing S is called the ***subgroup generated by*** S. It is denoted by $\langle S \rangle$.

Abstractly, $\langle S \rangle = \bigcap_{\substack{H \leqslant G \\ H \supseteq S}} H$. However, a more useful construction is $\langle S \rangle = \{a_1 a_2 \cdots a_k : a_i \in S \cup S^{-1}\}$. This is based on the concept of a ***word*** in G, which is a mapping from a sequence $(a_1, \cdots, a_k)$ of elements of G to the composition $a_1 a_2 \cdots a_k \in G$. We will have more to say about words later.

If $S = \{g_1, \cdots, g_n\}$, then we write $\langle g_1, \cdots, g_n \rangle := \langle S \rangle$. If G is an abelian group written additively, then $\langle S \rangle = \{\sum_{i=1}^{n} \lambda_i g_i : n \in \mathbb{N}, \lambda_i \in \mathbb{Z}, g_i \in S\}$ is simply the subgroup of all integer linear combinations of the elements of S. Let's start with a cute and simple application to number theory:

Lemma 4 (Bézout's Lemma)
If $a, b \in \mathbb{Z}$ are two integers, not both zero, and $d = (a, b)$, then $\exists x, y \in \mathbb{Z}$ s.t. $d = ax + by$.

The standard proof uses reverse-tracing the Euclidean Algorithm. Group Theory gives us a slicker

Proof. Consider the subgroup $\langle a, b \rangle \leqslant \mathbb{Z}$ consisting of all integer linear combinations $ax + by$ for $x, y \in \mathbb{Z}$. By Claim 6, $\exists d \in \mathbb{Z}$ s.t. $\langle a, b \rangle = \langle d \rangle$. Since $\langle a, b \rangle$ is not trivial, $d \neq 0$. By symmetry, we may choose $d > 0$. We claim that then $d = (a, b)$. Since $a, b \in \langle a, b \rangle$, and since $\forall c \in \langle a, b \rangle = \langle d \rangle : d \mid c$, in particular $d \mid a \wedge d \mid b \Rightarrow d \mid (a, b)$. Conversely, $d \in \langle d \rangle = \langle a, b \rangle \Rightarrow \exists x, y \in \mathbb{Z} : ax + by = d$. But then $(a, b) \mid ax + by = d$. Since both d and (a, b) are positive, and each divides the other, $d = (a, b)$. ∎

A lot of the action of the group generated by a single element can be captured rather succinctly. We are led to the notion of a ***cyclic*** group.

Definition 22. A group G is said to be ***cyclic*** if $\exists g \in G$ s.t. $G = \langle g \rangle$. In this case, G is said to be ***generated*** by g, and g is said to be a ***generator*** of G.

A generator of G need not be unique: for instance, both $+1$ and -1 generate $\mathbb{Z}$. It is immediate from the definition that all cyclic groups are abelian.

Example 24
$(\mathbb{Z}, +)$ is cyclic because it is generated by 1. For any $n \in \mathbb{N}$, the group $(\mathbb{Z}/n\mathbb{Z}, +)$ is cyclic because it is generated by $\bar{1}$. $(\mathbb{R}, +)$ is *not* cyclic.

Example 25
If for some $n \in \mathbb{N}$, the group $(\mathbb{Z}/n\mathbb{Z})^\times$ is cyclic with generator a, then a is called a ***primitive root*** modulo n.

Let's add one more term to our vocabulary as we move along. This'll come in handy later.

Definition 23. For a group G and $g \in G$, the ***order*** of g, written $|g|$, is the smallest positive integer n (if it exists) s.t. $g^n = e$. If no such positive integer exists, then g is said to have ***infinite order***.

Some authors prefer to use the terminology "zero order" in stead of "infinite order," and for good reason, as we shall later see. Observe that an element of a group has order 1 iff it is the identity. Observe that for any $g \in G$, the order of g is just the order of the subgroup $\langle g \rangle$ of G generated by g.

Example 26
In $(\mathbb{Z}/4\mathbb{Z}, +)$, the element 1 has order 4, whereas the element 2 has order 2. In the group $\mathbb{Z}$, every nonzero element has infinite order.

3.3 Homomorphisms

Let us now consider relationships between groups. The fundamental way to understand relationships between groups are maps between them that "respect their group structure." This is a theme you'll see showing up rather frequently in mathematics.

Definition 24. Let $(G, \star)$ and $(H, \diamond)$ be groups. A map $\varphi : G \to H$ is called a ***homomorphism of groups*** if it respects group structures, i.e. $\forall g, h \in G : \varphi(g \star h) = \varphi(g) \diamond \varphi(h)$. The set of all homomorphisms $\varphi : G \to H$ is denoted by $\mathrm{Hom}(G, H)$.

The condition be expressed compactly by saying that a $\varphi \in \mathrm{Hom}(G, H)$ iff the following diagram commutes:

$$\begin{array}{ccc} G \times G & \xrightarrow{\varphi \times \varphi} & H \times H \\ \downarrow{\scriptstyle \star} & & \downarrow{\scriptstyle \diamond} \\ G & \xrightarrow{\varphi} & H \end{array}$$

It should be noted that:

(a) When the group operations for G and H are not explicitly written, the homomorphism condition simply becomes $\varphi(gh) = \varphi(g)\varphi(h)$.

(b) When G and H are more than just groups, e.g. rings, fields, modules, etc., the notation $\mathrm{Hom}(G, H)$ takes on additional, more specicif meaning. In the coming sections, unless otherwise specified, $\mathrm{Hom}(G, H)$ refers only to the set of group homomorphisms.

(c) For any group G, $|\mathrm{Hom}(G, \{e\})| = |\mathrm{Hom}(\{e\}, G)| = 1$.

Example 27

For any $n, m \in \mathbb{Z}$ s.t. $n \mid m$, we get the natural "reduction mod n" homomorphissm $\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. The special case $m = 0$ is simply the reduction $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$.

Example 28

The map $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ is a homomorphism.

Example 29

Given any group G and a fixed element $g \in G$, the natural map $\mathbb{Z} \to G$ taking $n \mapsto g^n$ is a homomorphism.

Here are some immediate consequences of the definition:

Claim 8 — Let $\varphi \in \mathrm{Hom}(G, H)$. Then:
(a) $\varphi(e_G) = e_H$.
(b) $\forall g \in G : \varphi(g^{-1}) = \varphi(g)^{-1}$.

Proof. For any $g \in G$, $\varphi(g) = \varphi(ge_G) = \varphi(g)\varphi(e_G)$, so by cancellation we get (a). Similarly, we have $e_H = \varphi(e_G) = \varphi(gg^{-1}) = \varphi(g)\varphi(g^{-1})$, so by uniqueness of inverses, we get (b). ∎

We have some more vocabulary to familiarize ourselves with.

Definition 25. Let $\varphi \in \mathrm{Hom}(G, H)$.
(a) The ***kernel*** of φ is the fiber over e_H, i.e. $\ker \varphi := \{g \in G : \varphi(g) = e_H\}$.
(b) The ***image*** of φ is the set-theoretic image, i.e. $\mathrm{im}\, \varphi = \{h \in H : \exists g \in G : \varphi(g) = h\}$.

Observe that if $\varphi : G \to H$ is *any* homomorphism, we right away get two subgroups: $\ker\varphi \leqslant G$ and $\operatorname{im}\varphi \leqslant H$. There is a very handy characterization of the nature of φ just from these:

Claim 9 — A homomorphism $\varphi : G \to H$ is injective iff $\ker\varphi = \{e_G\}$; it is surjective iff $\operatorname{im}\varphi = H$.

Proof. The second claim is a tautology. If φ is injective, then by Claim 8(a), $\ker\varphi = \{e_G\}$. Conversely, let $\ker\varphi = \{e_G\}$. Then if for $g, h \in G : \varphi(g) = \varphi(h)$, then by Claim 8(b), $\varphi(gh^{-1}) = \varphi(g)\varphi(h)^{-1} = e_H$ so that $gh^{-1} \in \ker\varphi \Rightarrow gh^{-1} = e_G \Rightarrow g = h$; this shows that φ is injective. ■

Let's look at some examples.

Example 30

In Example 27 above, the kernel is $n(\mathbb{Z}/m\mathbb{Z})$.

Example 31

In Example 28 above, the kernel $\ker\det = \mathrm{SL}_n(\mathbb{R}) \leqslant \mathrm{GL}_n(\mathbb{R})$.

Example 32

In Example 29 above, let $\operatorname{ord}(g) \leqslant \mathbb{Z}$ denote the kernel. Then $\operatorname{ord}(g) \leqslant \mathbb{Z}$, and by Claim 6, $\operatorname{ord}(g) = \langle n\rangle$ for some n. If we choose our $n \geqslant 0$, then this n is precisely the order of g in G.

In fact, if you observe carefully, $\ker\varphi$ is more than just a subgroup of G: it has the special property that $\forall g \in G, \forall k \in \ker\varphi : \varphi(gkg^{-1}) = \varphi(g)\varphi(k)\varphi(g)^{-1} = \varphi(g)e_H\varphi(g)^{-1} = e_H$, i.e. $gkg^{-1} \in \ker\varphi$. This can be written compactly as $\forall g \in G : g(\ker\varphi)g^{-1} = \ker\varphi$. A subgroup $N \leqslant G$ is said to be ***normal*** if $\forall g \in G : gNg^{-1} = N$; this is denoted by $N \trianglelefteq G$. It is clear that every subgroup of an abelian group is normal, but that is not necessarily always the case with nonabelian groups. We will have a lot more to say about normal subgroups.

3.4 Isomorphisms and Automorphisms

We now have a way to think of what it means for two groups to have the "same structure."

Definition 26. A bijective homomorphism $\varphi : G \to H$ is called an ***isomorphism***.

If $\varphi : G \to H$ is an isomorphism then the inverse map $\varphi^{-1} : H \to G$ is also an isomorphism. This allows us to establish an equivalence relation on the set of all groups: if there is an isomorphism $\varphi : G \to H$, then G and H are said to be ***isomorphic***, and this is denoted by $G \cong H$. Intuitively, this means that the groups G and H are the same, just labelled differently.

Example 33

Consider the groups $\mathbb{Z}/3\mathbb{Z}$ and the group G of rotations an equilateral triangle in the plane. Then G has order 3, and is generated by the single element $\mathrm{rot}\,(2\pi/3)$. The map $\varphi : \mathbb{Z}/3 \to G$ given by $1 \mapsto \mathrm{rot}\,(2\pi/3)$ is an isomorphism.

Think about why specifiying simply $\varphi(1)$ is sufficient to define the map φ. In general, if G is a cyclic group and H is any group, then a $\varphi \in \mathrm{Hom}(G, H)$ is completely determined by the image of a generator of G. This is because a cyclic group is a quotient of a free group on one generator–more on that below.

The following is another way to think about isomorphisms that generalizes more easily.

Claim 10 — A $\varphi \in \mathrm{Hom}(G, H)$ is an isomorphism iff $\exists \psi \in \mathrm{Hom}(H, G) : \psi \circ \varphi = \mathrm{id}_G$ and $\phi \circ \psi = \mathrm{id}_H$.

Example 34

The homomorphisms $\exp : (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$ and $\log : (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +)$ are inverses to each other, so that $(\mathbb{R}, +) \cong (\mathbb{R}_{>0}, \cdot)$.

The following are straightforward consequences of having the same structure:

Claim 11 — If $G \cong H$, then:
(a) $|G| = |H|$.
(b) G is abelian iff H is.
(c) $\forall g \in G : |g| = |\varphi(g)|$.

Using Claim 3, we come up with and tuck away the following lemma, which will be useful later.

Lemma 5

If $\varphi : G \hookrightarrow H$ is an injective homomorphism, then $\tilde{\varphi} : G \to \mathrm{im}\, \varphi$ is an isomorphism. In particular, G is isomorphic to a subgroup of H.

In a sense, a broad goal of Group Theory is to classify all groups. We don't want to count multiple times the groups that have essentially the same structure, so we say rather that the broad goal of Group Theory is to classify all groups ***up to isomorphism***. In general, this is a difficult task. However, it can be done rather easily in special cases:

Theorem 4

Let G be a cyclic group. If G is infinite, $G \cong \mathbb{Z}$. If $|G| < \infty$, then $G \cong \mathbb{Z}/|G|\mathbb{Z}$.

Proof. Let $g \in G$ be any generator. Consider the homomorphism of Example 29; it is surjective by definition. If G is infinite, then the powers g^n are all distinct, so that the map is injective; this means that it is an isomorphism. If G is finite, then from Example 32, we know that $|g| = |G|$; and it is clear that in that case we get a natural bijective homomorphism $\mathbb{Z}/|g|\mathbb{Z} \to G$. ■

If the last part of the argument wasn't very convincing, don't worry. We'll treat it more formally later. The cyclic group of order n is sometimes denoted by C_n or Z_n, and the infinite cyclic group is sometimes denoted by C_∞ or Z_∞. Usually when denoted like this, we write Z_n or Z_∞ multiplicatively.

An interesting question to think about is: if $G \cong H$, how many distinct isomorphisms $\varphi : G \to H$ do we have? In general, there may be more than one. For instance, in Example 33, both the maps $1 \mapsto \mathrm{rot}(2\pi/3)$ and $1 \mapsto \mathrm{rot}(4\pi/3)$ are isomorphisms. Suppose we know $\varphi : G \to H$ is an isomorphism, and let $\psi : G \to H$ be another one. Then $\xi := \psi^{-1} \circ \varphi : G \to G$ is an isomorphism of a group onto itself; conversely, if $\xi : G \to G$ is an isomorphism, then $\psi := \varphi \circ \xi^{-1} : G \to H$ is another isomorphism. Therefore, considering the number of isomorphisms $G \to H$ is the same as considering the number of isomorphisms $G \to G$. This leads us to the following definition:

Definition 27. Let G be a group.
(a) A homomorphism $\varphi : G \to G$ is said to be an ***endomorphism***. The set of endomorphisms is denoted by $\mathrm{End}(G) := \mathrm{Hom}(G, G)$.
(b) An isomorphism $\varphi \in \mathrm{End}(G)$ is said to be an ***automorphism***. The set of all automorphisms is denoted by $\mathrm{Aut}(G)$.

Observe that if G and H are arbitrary groups, then there is no notion of composition on $\mathrm{Hom}(G, H)$. However, we *can* impose a binary operation on $\mathrm{End}(G)$ by composition: taking $(\psi, \varphi) \in \mathrm{End}(G)^2$ to $\psi \circ \varphi \, \mathrm{End}(G)$. This binary operation is naturally associative, and also has a natural identity element: id_G. However, in this structure, inverses need not exist: for instance, if $e \in \mathrm{End}(G)$ is the null homomorphism

that takes everything to e, then it has no inverse, i.e. $\nexists \psi \in \mathrm{End}(G) : \psi \circ e = \mathrm{id}_G$. Therefore, in general, $(\mathrm{End}(G), \circ)$ is only a monoid. However, as you've probably guessed by now, $(\mathrm{Aut}(G), \circ)$ *is* a group, and it is called the ***automorphism group*** of G. For any group G, the automorphism group $\mathrm{Aut}(G)$ packages a lot of information about the structure of G. We'll talk more about automorphisms and automorphism groups later.

3.5 Groups of Small Order

Let's try to classify some groups of small order upto isomorphism. Here's some handy vocabulary:

Definition 28. Let $G = \{g_1, \cdots, g_n\}$ be a finite group with $g_1 = e$. The ***multiplication table*** or ***group table*** or ***Cayley table*** or the ***Cayley matrix*** of the group G is the $n \times n$ matrix whose i, j entry is $g_i g_j$.

For a finite group, the Cayley table contains all the information about the structure of the group. Since the size of the table increases as the square of the size of the group, it is conceptually (and computationally) feasible to use Cayley tables to understand group structure only for groups of relatively small orders. The following are elementary observations:

Claim 12 — Let G be a finite group.
(a) The first row and column of the Cayley matrix are copies of $\{g_1, \cdots, g_n\}$.
(b) By the cancellation property, every row and column of the Cayley matrix contains each element of the group exactly once. This means that the Cayley matrix is a ***Latin square***.
(c) G is abelian iff the Cayley matrix of G is symmetric.

Example 35

The trivial group $G = \{e\}$ has the Cayley matrix:

$\star$	e
e	e

Example 36

Let's look at groups of order 2. Then $G = \{e, a\}$ for some $a \neq e$. What can a^2 possibly be? Well, $a^2 \in G$ so $a^2 = e$ or $a^2 = a$; by the cancellation property, the latter would mean $a = e$, which is not the case. Therefore, $a^2 = e$. This determines the Cayley table completely:

$\star$	e	a
e	e	a
a	a	e

Therefore, there is a unique group of order 2. In other words, $\mathbb{Z}/2\mathbb{Z} \cong \mathfrak{S}_2 \cong (\{T, F\}, \oplus)$ and so on.

Example 37

Let's look at groups of order 3. Then $G = \{e, a, b\}$. Observe that by the cancellation property, ab can't be a or b, so it must be e. Therefore, $b = a^{-1}$. We can now fill in the remaining spots by using Claim 12(b). This leads us to the complete table, showing that $a^3 = b^3 = e$.

$\star$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Therefore, again, there is a unique group of order 3–$\mathbb{Z}/3\mathbb{Z}$.

Example 38

You will show on the homework that there are exactly two nonisomoprhic groups of order 4–$\mathbb{Z}/4\mathbb{Z}$ and another group K_4 called the ***Klein 4-group***.

$\mathbb{Z}/4\mathbb{Z}$	e	a	a^2	a^3
e	e	a	a^2	a^3
a	a	a^2	a^3	e
a^2	a^2	a^3	e	a
a^3	a^3	a	a^2	a^2

K_4	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

The two groups are nonisomorphic because $\mathbb{Z}/4\mathbb{Z}$ contains an element of order 4, whereas every element of K_4 has order 2. Observe, in particular, that both the groups of order 4 are abelian. We will see later when we study direct products that $K_4 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

4 Building Vocabulary

4.1 Cosets and Quotient Groups

The primary question that we want to investigate in this section is: when is a given subgroup $H \leqslant G$ normal? As we've seen, if G is abelian, then every subgroup $H \leqslant G$. When G is not abelian, the answer is more subtle, and naturally leads us to the theory of quotient groups.

Let's start with a concrete example. Fix $n \in \mathbb{N}_0$. We've seen that there is an equivalence relation on $\mathbb{Z}$ given by saying $a \sim b$ iff $-a + b \in n\mathbb{Z}$; in that case we write $a \equiv b \pmod{n}$. The set of equivalence classes is denoted by $\mathbb{Z}/n\mathbb{Z}$ and (luckily) has a nice group structure of its own. This group is a way of abstracting away the nonessential information, and carrying with us only the remainder of an integer when divided by n. This generalizes rather nicely to abitrary groups, except we have to be careful about the order (since groups, in general, are not abelian). Start off by defining:

Definition 29. Let G be a group and $H \leqslant G$ any subgroup.
(a) For any $a \in G$, the set $aH := \{ah : h \in H\}$ is called a ***left coset*** of H in G. The set of left cosets of H in G is denoted by G/H.
(b) For any $a \in G$, the set $Ha := \{ha : h \in H\}$ is called a ***right coset*** of H in G. The set of right cosets of H in G is denoted by $H\backslash G$.

For any $a \in G$, it is clear that $aH = H \Leftrightarrow Ha = H \Leftrightarrow a \in H$. Notice that $aH = bH$ does not mean $a = b$; it simply means that $a^{-1}b \in H$. Similarly, $Ha = Hb \Leftrightarrow ab^{-1} \in H$. The cosets eH and He are denoted simply by H. We are now ready for the generalization.

Definition 30. Given a subgroup $H \leqslant G$, define an equivalence relation on G called ***left equivalence modulo*** H given by $a \equiv_l b \pmod{H}$ iff $a^{-1}b \in H$. Since $a^{-1}b \in H \Leftrightarrow aH = bH$, the equivalence classes are precisely the left cosets of H in G.

We know from the general theory of equivalence relations that G/H is a partition of G. Note that in general, the left cosets of a subgroup are not subgroups because they do not contain the identity element e; the only left coset that is indeed a subgroup is H itself. The following claim is the heart of the theory:

Theorem 5
If $H \leqslant G$, then all the left cosets of H in G have the same cardinality. In particular, the function $\varphi_a : H \to aH$ taking $g \mapsto ag$ is a bijection.

The proof is trivial: the map is injective because of the cancellation property and surjective by definition. This in itself already has numerous consequences:

Corollary 5.1 (Lagrange's Theorem) — If G is a finite group and $H \leqslant G$, then $|H|$ divides $|G|$.

Proof. G is partitioned into left cosets and by Theorem 5 they have the same size, so $|G| = |G/H| \cdot |H|$. ∎

Corollary 5.2 — If G is a finite group and $g \in G$, then $|g|$ divides $|G|$. Further, $\forall g \in G : g^{|G|} = e$.

Proof. Apply the above to the subgroup $\langle g \rangle \leqslant G$. ∎

Corollary 5.3 (Euler's Theorem) — If $n \in \mathbb{N}$ and $a \in \mathbb{Z} : (a, n) = 1$, then $a^{\varphi(n)} \equiv 1 \pmod{n}$.

Proof. Apply the above to the group $(\mathbb{Z}/n\mathbb{Z})^\times$. ∎

Corollary 5.4 (Fermat's Little Theorem) — If p is a prime and $a \in \mathbb{Z} : p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$.

Proof. Apply the above to the case $n = p$. ∎

You know your theory is going in the right direction if what people call theorems are now simple corollaries to your theorems! We also have a heads up on our classification of groups:

Corollary 5.5 — If G is a group of prime order p, then $G \cong Z_p$.

Proof. If $g \in G$ is any nonidentity element, then $1 < |g|$ divides $|G| = p$. Therefore, $|g| = p$ and $G = \langle g \rangle$. In fact, we've shown something stronger: if G is a group of prime order, then any nonidentity element generates G. ■

Of course, we can choose to do the exactly identical thing and define right equivalence modulo H by saying $a \equiv_r b \pmod{H}$ iff $ab^{-1} \in H$. The equivalence classes under $\equiv_r$ are the right cosets of H in G. Clearly, when G is abelian, then these two notions coincide. Even when G is nonabelian, we have:

Lemma 6

If G is any group (possibly infinite) and $H \leqslant G$ any subgroup, then $|G/H| = |H\backslash G|$.

Proof. Consider the map $G/H \to H\backslash G$ given by taking $aH \mapsto Ha^{-1}$. Note that this map is well-defined, i.e. does not depend on the choice of the representative a of the coset aH because $aH = bH \Leftrightarrow a^{-1}b \in H \Leftrightarrow a^{-1}(b^{-1})^{-1} \in H \Leftrightarrow Ha^{-1} = Hb^{-1}$. By the exact same argument, the map $H\backslash G$ given by $Hb \mapsto b^{-1}H$ is well-defined and an inverse to the given map. Therefore, the map is a bijection. ■

This number is called the ***index*** of H in G and is denoted by $|G : H|$. If G is finite, then it is clear from Lemma 5.1 that $|G/H| = |G|/|H|$; this doesn't exactly make sense if G is infinite. Infinite groups can have subgroups of finite or infinite index, for e.g. $\langle 0 \rangle$ has infinite index in $\mathbb{Z}$, whereas $\langle n \rangle$ has index n.

Note that in the case $G = \mathbb{Z}$ and $H = n\mathbb{Z}$, the left cosets G/H themselves form a group, with the natural law of composition that makes the projection map $\pi : G \to G/H$ a homomorphism. When exactly does this happen?

Theorem 6

For any $H \leqslant G$, the set of left cosets G/H themselves form a group with law of composition $aH \cdot bH = abH$ iff $H \trianglelefteq G$. In this case, G/H is called the ***quotient group*** of G by H, and the natural ***quotient map*** $\pi : G \to G/H$ given by $a \mapsto aH$ is a homomorphism.

Recall that H is a normal subgroup, i.e. $H \trianglelefteq G$, iff $\forall g \in G : gHg^{-1} = H \Leftrightarrow \forall g \in G : gH = Hg$.

Proof. Suppose H is normal; we have to check that the law of composition is well-defined, i.e. we have to check that $\forall a, a', b, b' \in G : aH = a'H \wedge bH = b'H \Rightarrow abH = a'b'H$. Now $aH = a'H \Rightarrow a^{-1}a' \in H$ and $bH = b'H \Rightarrow b^{-1}b' \in H$. Since $H \trianglelefteq G$, $b^{-1}(a^{-1}a')b \in H$ so that $b^{-1}a^{-1}a'bb^{-1}b' = (ab)^{-1}a'b' \in H$. Conversely, if G/H is a group with group law as given, then π is a homomorphism, so $H = \ker \pi \trianglelefteq G$. ■

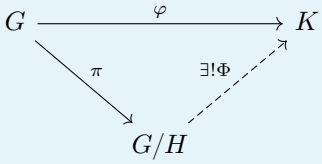
This allows us to give a complete answer to our initial question:

Corollary 6.1 — A subgroup $H \leqslant G$ is normal iff it is the kernel of a (surjective) homomorphism.

Proof. First, observe that if $\varphi : G \to K$ is any homomorphism, then $\tilde{\varphi} : G \to \operatorname{im} \varphi$ is a surjective homomorphism with the same kernel. We've seen that $H = \ker \varphi \trianglelefteq G$. Conversely, if $H \trianglelefteq G$, then H is the kernel of the quotient homomorphism $\pi : G \twoheadrightarrow G/H$. ■

Henceforth, for the sake of brevity, whenever we say "cosets," we mean *left* cosets, unless specified otherwise. There is another way to look at the key concept of quotient groups that has a more category-theoretic flavor.

Theorem 7 (Characteristic Property of Quotient Groups)
Suppose $H \trianglelefteq G$, and $\varphi \in \mathrm{Hom}(G, K)$ such that $H \leqslant \ker\varphi$, then $\exists!\,\Phi \in \mathrm{Hom}(G/H, K)$ s.t. the following diagram commutes:



Further,
(a) If φ is surjective, then so is Φ.
(b) If $H = \ker\varphi$, then Φ is injective.

Proof. Since we want the diagram to commute, we must have $\Phi : aH \mapsto \varphi(a)$. This is well-defined because $aH = a'H \Rightarrow aa'^{-1} \in H \Rightarrow \varphi(aa'^{-1}) = e_K \Rightarrow \varphi(a) = \varphi(a')$. This is a homomorphism because $\Phi(aH \cdot bH) = \Phi(abH) = \varphi(ab) = \varphi(a)\varphi(b) = \Phi(aH)\Phi(bH)$. Finally, this determines Φ uniquely. If φ is surjective, then the commutativity of the diagram implies that Φ is surjective. If $H = \ker\varphi$, then $\Phi(aH) = \Phi(bH) \Rightarrow \varphi(a) = \varphi(b) \Rightarrow \varphi(a^{-1}b) = e_K \Rightarrow a^{-1}b \in \ker\varphi = H \Rightarrow aH = bH$. ■

This characterstic property leads us naturally to an isomorphism theorem.

4.2 First Isomorphism Theorem, Exact Sequences, Simple Groups

Theorem 8 (First Isomorphism Theorem)
If $\varphi : G \to K$ is any homomorphism, then $G/\ker\varphi \cong \mathrm{im}\,\varphi$.

Note that this is the analogue of the rank-nullity formula for vector spaces.

Proof. If $\varphi : G \to K$ is any homomorphism with kernel $\ker\varphi$, then $\tilde{\varphi} : G \to \mathrm{im}\,\varphi$ is a *surjective* homomorphism with the same kernel $\ker\varphi$. Taking $H = \ker\varphi$ in Theorem 7, $\Phi : G/H \xrightarrow{\sim} \mathrm{im}\,\varphi$. ■

Example 39
Let's return to an argument we left before. Let G be a cylic group with generator g; then we get a natural map $\varphi : \mathbb{Z} \to G$ by $n \mapsto g^n$, and this is surjective by definition. If G is infinite, then the map is injective and so is an isomorphism. If G is finite, then $\ker\varphi = \langle|g|\rangle$, so by the First Isomorphism Theorem, $\mathbb{Z}/\langle|g|\rangle \cong G$.

Often when we have a series of such homomorphisms, a compact way to express such relationships is by using *exact sequences*.

Definition 31. Suppose we have a sequence of groups with homomorphisms:

$$\cdots \xrightarrow{\varphi_{i-2}} G_{i-1} \xrightarrow{\varphi_{i-1}} G_i \xrightarrow{\varphi_i} G_{i+1} \xrightarrow{\varphi_{i+1}} \cdots$$

This sequence is said to be ***exact*** at G_i if $\mathrm{im}\,\varphi_{i-1} = \ker\varphi_i$. (Note that this implies but is not implied by $\varphi_i \circ \varphi_{i-1}$ being the null homomorphism.) A sequence is said to be exact if it is exact at each term in the sequence.

Example 40
Let A, B, C be any groups.
(a) The sequence $\{e\} \to A \xrightarrow{\varphi} B$ is exact at A iff φ is injective.
(b) The sequence $B \xrightarrow{\psi} C \to \{e\}$ is exact at C iff ψ is surjective.

These sequences, which seem to be rather special, show up frequently enough in mathematics to deserve a name, and are the fundamental language of a branch called homological algebra. The smallest and simplest example of an exact sequence is the "short exact sequence," which, depending on how you count it, has either 3 or 5 terms. (Prof. Harris's words.)

Definition 32. For groups A, B, C and homomorphisms $\varphi \in \mathrm{Hom}(A, B)$ and $\psi \in \mathrm{Hom}(B, C)$, the sequence

$$\{e\} \to A \xrightarrow{\varphi} B \xrightarrow{\psi} C \to \{e\}$$

is called a ***short exact sequence*** if φ is injective, $\operatorname{im}\varphi = \ker\psi$, and ψ is surjective. This can be written equivalently as

$$A \overset{\varphi}{\hookrightarrow} B \overset{\psi}{\twoheadrightarrow} C.$$

Example 41

The following are examples of short exact sequences.
(a) $\mathbb{Z}/2\mathbb{Z} \overset{\iota_1}{\hookrightarrow} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \overset{\pi_2}{\twoheadrightarrow} \mathbb{Z}/2\mathbb{Z}$.
(b) $\mathbb{Z}/2\mathbb{Z} \overset{\phi}{\hookrightarrow} \mathbb{Z}/6\mathbb{Z} \overset{\psi}{\twoheadrightarrow} \mathbb{Z}/3\mathbb{Z}$, where $\phi : \bar{1} \mapsto \bar{3}$ and $\psi : \bar{1} \mapsto \bar{1}$.

Example 42

This example is for those familiar with the theory of quadratic residues. Let p be a prime, and let $\mathbb{F}_p^\times := (\mathbb{Z}/p\mathbb{Z})^\times$ denote the set of nonzero remainders modulo p, i.e. $\{1, 2, \cdots, p-1\}$ (the terminology will be explained later). Consider the group $(\mathbb{F}_p^\times, \cdot)$ and its subgroup $(\mathbb{F}_p^\times)^2$ of quadratic residues. If $\psi : \mathbb{F}_p^\times \to \{\pm 1\}$ is the Legendre symbol $\psi : \bar{a} \mapsto \bar{a}^{(p-1)/2} = \left(\frac{a}{p}\right)$, then the sequence

$$\{1\} \to (\mathbb{F}_p^\times)^2 \hookrightarrow \mathbb{F}_p^\times \overset{\psi}{\twoheadrightarrow} \{\pm 1\} \to \{1\}$$

is a short exact sequence.

In a sense, short exact sequences are all we need to deal with, because of the following result whose proof is evident:

Claim 13 — A sequence

$$\cdots \xrightarrow{\varphi_{i-2}} G_{i-1} \xrightarrow{\varphi_{i-1}} G_i \xrightarrow{\varphi_i} G_{i+1} \xrightarrow{\varphi_{i+1}} \cdots$$

is exact at G_i iff $\forall i$:

$$\{e\} \to \operatorname{im}\varphi_{i-1} \hookrightarrow G_i \overset{\varphi_i}{\twoheadrightarrow} \ker\varphi_{i+1} \to \{e\}$$

is a short exact sequence.

This is all the vocabulary we need for what we do in this course; more about exact sequences can be found in a course on module theory or homological algebra. The following lemma explains our obsession with exact sequences:

Lemma 7

If $\{e\} \to A \xrightarrow{\varphi} B \xrightarrow{\psi} C \to \{e\}$ is a short exact sequence, then $A \cong \operatorname{im}\varphi$ and $B/\operatorname{im}\varphi \cong C$. Conversely, if $\varphi : G \to K$ is any homomorphism, then $\{e\} \to \ker\varphi \hookrightarrow G \twoheadrightarrow G/\ker\varphi \to \{e\}$ is a short exact sequence, with the third term $G/\ker\varphi \cong \operatorname{im}\varphi$.

In less mathematical terms, this means exactly that if $\{e\} \to A \xrightarrow{\varphi} B \xrightarrow{\psi} C \to \{e\}$ is a short exact sequence, then by naturally identifying A with $\operatorname{im}\varphi \leqslant B$ we get $B/A \cong C$. In other words, B can be "broken up" into the component pieces coming from A and C, and how these two pieces interact completely determines the structure of B. This is very much analogous to the direct sum of vector spaces: if

U, V are vector spaces, then $\{e\} \to U \xrightarrow{\iota_1} U \oplus V \xrightarrow{\pi_2} V \to \{e\}$ is a short exact sequence, and, essentially, this is the only way to combine these spaces together. However, when A and C are groups, there are many different ways to "combine" these groups together, and the way these groups A and C interact give us insights into the structure of the combination B.

In other words, given a group G, we can start to analyse the structure of G by:
(a) Finding all normal subgroups $H \trianglelefteq G$, and
(b) Considering how the groups H and G/H "stitch together" to make G.

All groups have nontrivial proper subgroups–if G is cyclic, then we understand the subgroups of G completely; as soon as G is not cyclic, it has a nontrivial proper cyclic subgroup generated by any nonidentity element. However, the same is not true of normal subgroups–as we've seen above, the more normal subgroups a group has, the more complex its structure can be. In a sense, then, the most "fundamental" or "atomic" groups are the ones which have no nontrivial proper normal subgroups at all!

Definition 33. A group G is called ***simple*** if the only normal subgroups of G are $\{e_G\}$ and G.

The following is the key property of simple groups:

Claim 14 — If G is simple, then any homomorphism $\varphi : G \to K$ is either the null map or injective.

In other words, if G is simple, then any nontrivial homomorphism from G to any other group K must embed it into K. Therefore, we can analyse the structure of K by considering homomorphisms from known simple groups to it. In general, it is hard problem to determine whether a group is simple.

Example 43
For any prime p, the group Z_p is simple because it has nontrivial proper subgroups at all.

Example 44
We will show later that for $n \geqslant 3$, $\mathfrak{S}_n$ is not simple. In particular, there is a subgroup $\mathfrak{A}_n$ of $\mathfrak{S}_n$ that is of index 2 (and hence normal). It is nontrivial result that for $n \geqslant 5$, $\mathfrak{A}_n$ is simple, and the case $n = 5$ is the reason for the unsolvability of the general quintic.

There are three more named isomorphism theorems that we will return to later. First, let's talk about notation for group structure.

4.3 Direct Products, Free Products and Group Presentations

4.3.1 Direct Products

We know that if G is a group and $H \trianglelefteq G$, then we can create a "smaller" group G/H. Let's look at some ways to build bigger groups from smaller ones.

Definition 34. Suppose $G_1, \cdots, G_n$ are groups. Then the cartesian product $\prod_{i=1}^{n} G_i$ whose elements are n-tuples $(g_1, \cdots, g_n)$ with $g_i \in G_i$, endowed with the law of composition given pointwise is called the ***direct product*** of the groups G_i.

Example 45

When all the G_i's are the same, $G_i = G$, then $\prod_{i=1}^{n} G_i$ is written G^n. For example, taking $G = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ with law of composition given by addition, we get the familiar spaces $\mathbb{Z}^n, \mathbb{Q}^n, \mathbb{R}^n, \mathbb{C}^n$.

Example 46

If $n, m \in \mathbb{N}$ such that $(n, m) = 1$, then $Z_n \times Z_m \cong Z_{mn}$. (Prove this!)

If G_i are groups, then $|\prod_{i=1}^{n} G_i| = \prod_{i=1}^{n} |G_i|$, and if any G_i is infinite, then so is their direct product. The following are elementary observations:

Claim 15 — Let $G_1, \cdots, G_n$ be groups.
(a) The natural projection maps $\pi_\alpha : \prod_{i=1}^{n} G_i \to G_\alpha$ are group homomorphisms.
(b) For $i \in [n]$, G_i can be identified as $G_i \cong \{e_{G_1}\} \times \{e_{G_2}\} \times \cdots \times G_i \times \cdots \times \{e_{G_n}\} \trianglelefteq \prod_{i=1}^{n} G_i$.
(c) Under the natural identification above, if $i \neq j$, $g_i \in G_i$ and $g_j \in G_j$, then $g_i g_j = g_j g_i$. In other words, elements of different constituent groups commute with each other.

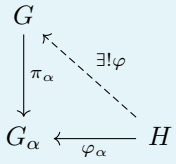
It is clear what the definition for arbitrary cartesian products should be.

Definition 35. Let $\mathscr{G} = \{G_i\}_{i \in I}$ be a family of groups indexed by some set I. Then the direct product or direct sum of $\mathscr{G}$, denoted by $\prod_{i \in I} G_i$ is the cartesian product of $\mathscr{G}$ endowed with the law of composition given pointwise.

The direct product of groups satisfies, and is completely characterized by, its universal property:

Theorem 9 (Characteristic Property of Direct Products)

Suppose $\mathscr{G} = \{G_i\}_{i \in I}$ is a family of groups indexed by set I. Then a direct product $G := \prod_{i \in I} G_i$ is a group with projection homomorphisms $\pi_\alpha : G \to G_\alpha$ such that given any group H and homomorphisms $\varphi_\alpha \in \mathrm{Hom}(H, G_\alpha), \exists! \varphi \in \mathrm{Hom}(H, G)$ s.t. $\forall \alpha \in I : \pi_\alpha \circ \varphi = \varphi_\alpha$, i.e. the following diagram commutes:



Further, this universal property characterizes G upto unique isomorphism that preserves projections, so that we may use the definite article "the" when talking about the direct product of a family of groups.

Proof. Recall that the cartesian product is defined as the set of functions $\mathbf{x} : I \to \bigcup_{i \in I} G_i$ from the indexing set to the union of the underlying sets G_i of the groups $(G_i, \star_i)$ s.t. $\forall \alpha \in I : x_\alpha = \mathbf{x}(\alpha) \in G_\alpha$.

The projection maps are simply the maps $\pi_\alpha : \mathbf{x} \mapsto x_\alpha$. In the present scenario, we may endow these with a group structure by saying that the composition $\mathbf{x} \star \mathbf{y}$ is the function $I \to \bigcup_{i\in I} G_i$ that takes any $\alpha \in I$ to $x_\alpha \star_\alpha y_\alpha$. Uniqueness upto unique isomorphism follows from the standard universal property argument in the category (Grp). ■

Let us examine closely the observation in Claim 15(c). It says that under the natural identifications of $G_i, G_j \trianglelefteq G$ the elements of the groups G_i and G_j for $i \neq j$ commute with each other. There is another, more elaborate, product of groups in which we don't have this requirement. It is called the ***free product*** of groups. Let's build up to that.

Correction. Given a family of groups $\{G_i\}_{i\in I}$, the constructions of the direct product $\prod_{i\in I} G_i$ and the direct sum $\bigoplus_{i\in I} G_i$ coincide only for finite families. In general, the direct sum of groups is defined to be the subgroup of the direct product in which all but finitely many elements are the identity. For instnce, $\prod_{i=0}^{\infty} \mathbb{Z}$ is the group of infinite sequences $(a_1, a_2, \cdots, a_n, \cdots)$ of integers, whereas the direct sum $\bigoplus_{i=0}^{\infty} \mathbb{Z} \cong (\mathbb{Z}[x], +)$ contains only sequences with all but finitely many terms zero. These concepts clearly coincide for finite I; as we shall see later, the direct sum is NOT the categorical sum in the category (Grp).

4.3.2 Free Products

Let's start with a slightly different concept:

Definition 36. Let $\{X_i\}_{i\in I}$ be an indexed family of sets. The ***coproduct*** or the ***disjoint union*** of this family, denoted by $\coprod_{i\in I} X_i := \bigcup_{i\in I} \{i\} \times X_i$. In case of a finite family, we simply write $X_1 \coprod X_2 \coprod \cdots \coprod X_n$. The elements of this disjoint union are elements of the form (i, x_i) for $x_i \in X_i$.

Example 47

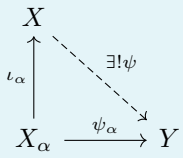
If $X_1 = X_2 = \{a, b, c\}$, then $X_1 \coprod X_2 = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}$. Intuitively, the disjoint union treats the components as if they were disjoint objects, even if they're originally not. In other words, we can think of $X_1 \coprod X_2 := \{a, b, c, a', b', c'\}$, where the prime denotes the elements from the second set.

Example 48

$\mathbb{R} \cup \mathbb{R} = \mathbb{R}$ but $\mathbb{R} \coprod \mathbb{R}$ contains the r in the first $\mathbb{R}$ and r' in the second $\mathbb{R}$ s.t. $1 \neq 1'$ etc.

Theorem 10 (Characteristic Property of Coproduct)

Let $\{X_i\}_{i\in I}$ be an indexed family of sets. The coproduct $X = \coprod_{i\in I} X_i$ is a set with inclusion maps : $\iota_\alpha : X_\alpha \to X$ such that given any set Y and any functions $\psi_\alpha : X_\alpha \to Y$ there is a unique function $\psi : X \to Y$ such that $\forall \alpha \in I : \psi \circ \iota_\alpha = \psi_\alpha$, i.e. the following diagram commutes:



Further, this universal property characterizes the coproduct upto unique isomorphism that preserves inclusions.

Proof. The $\iota_\alpha : X_\alpha \to X$ is $x_\alpha \mapsto (\alpha, x_\alpha)$. Then ψ is given by $\psi(\alpha, x_\alpha) = \psi_\alpha(x_\alpha)$ for $x_\alpha \in X_\alpha$. ■

When X_1 and X_2 are inherently disjoint, then we have a very natural identification $X_1 \cup X_2 \cong X_1 \coprod X_2$. We want to do something similar for groups.

Definition 37. Let $\mathscr{G} = \{G_i\}_{i\in I}$ be an indexed family of groups. Define a ***word*** in $\mathscr{G}$ to be a finite sequence of elements of the coproduct of the underlying sets $G = \coprod_{i\in I} G_i$, i.e. a word is an m-tuple of the form $(g_1, \cdots, g_m)$ where each g_i is an element of some G_α. In other words, the set of words in $\mathscr{G}$ is $\mathcal{W}(\mathscr{G}) = \bigcup_{m=0}^{\infty} G^m$.

The sequence of length zero is called the ***empty word***, and is denoted by $()$. Define a binary operation on $\mathcal{W}(\mathscr{G})$ by concatenation: $(g_1, \cdots, g_m)(h_1, \cdots, h_k) = (g_1, \cdots, g_m, h_1, \cdots, h_k)$. This is associative and has two-sided identity $()$, but there are no inverses. We've not used the group structures of G_i at all.

Definition 38. An ***elementary reduction*** is an operation of one of the following forms:

$$(g_1, \cdots, g_i, g_{i+1}, \cdots, g_m) \mapsto (g_1, \cdots, g_i g_{i+1}, \cdots, g_m) \text{ if } \exists \alpha : g_i, g_{i+1} \in G_\alpha;$$

$$(g_1, \cdots, g_{i-1}, e_\alpha, g_{i+1}, \cdots, g_m) \mapsto (g_1, \cdots, g_{i-1}, g_{i+1}, \cdots, g_m).$$

Define an equivalence relation on words $W \sim W'$ if there is a sequence $W_0 = W, W_1, \cdots, W_n = W'$ where for each i, either W_i is obtained from W_{i-1} by an elementary reduction or W_{i-1} is obtained by an elementary reduction of W_i. The set of equivalence classes is called the coproduct or ***free product*** of the groups $\{G_i\}$ and is denoted by $\coprod_{i\in I}(G_i, \star_i)$ or $\ast_{i\in I} G_i$.

Claim 16 — Given an indexed family of groups $\{G_i\}$, their free product is a group under composition by concatenation.

Proof. First we need to check that concatenation respects the equivalence relation, but that is clear: if $W \sim W'$ then for any word V, $VW \sim VW'$ and similarly for right concatenation. The class of $()$ is the identity, and for a word $W = [(g_1, \cdots, g_m)]$, the class of $(g_m^{-1}, \cdots, g_1^{-1})$ is the inverse W^{-1}. ■

What is not easy to show however is that the reduced word representing any given equivalence class is unique. We want to show that:

Proposition 3

Every element of $\ast_{i\in I} G_i$ is represented by a unique reduced word.

Proof Sketch. This amounts to constructing a canonical reduction algorithm from the set $\mathcal{W}$ of words to the subset $\mathcal{R} \subseteq \mathcal{W}$ of reduced words; i.e. a map $r : \mathcal{W} \to \mathcal{R}$ with the property that:
(a) $r|_{\mathcal{R}} = \mathrm{id}_{\mathcal{R}}$, and
(b) If $W \sim W'$ then $r(W) = r(W')$.
The exact construction is tedious, but can be found in the special notes uploaded if you're interested. ■

For each group G_α there is a canonical injective homomorphism $\iota_\alpha : G_\alpha \to \ast_{i\in I} G_i$ defined by $g_\alpha \mapsto (g_\alpha)$. We usually identify G_α with its image and avoid writing parenthesis. With this, we've succeeded in making some sense out of product of elements of different groups such that elements of different constituent groups need not commute with each other.

All of this is a lot of theoretical work, so let's look at some concrete examples to understand what's going on.

Example 49

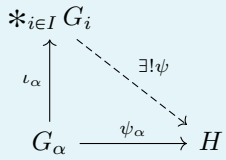
The free product $Z_\infty * Z_\infty$ can be described as follows: let a be a generator for the first group and b for the second one. Then the elements of $Z_\infty * Z_\infty$ are elements of the form $a^{n_1}b^{m_1}a^{n_2}b^{m_2}\cdots a^{n_k}b^{m_k}$ where $n_i, m_j \in \mathbb{Z}$. For instance, some elements would be $g = a^2b^3a^{-3}b^4$ and $h = b^{-3}a^{-2}b^2a^{-2}$; then $gh = a^2b^3a^{-3}b^4 \cdot b^{-3}a^{-2}b^2 = a^2b^3a^{-3}ba^{-2}b^2a^{-2}$ and $hg = b^{-3}a^{-2}b^2a^{-2}a^2b^3a^{-3}b^4 = b^{-3}a^{-2}b^4a^{-3}b^4$. Clearly, this group is nonabelian.

Example 50

Similarly, the free product $Z_2 * Z_2$ contains elements of the form $\alpha\beta\alpha\beta$ etc. and composition by concatenation. For instance, $(\alpha\beta\alpha\beta)(\beta\alpha\beta) = \alpha$ whereas $(\beta\alpha\beta)(\alpha\beta\alpha\beta) = \beta\alpha\beta\alpha\beta\alpha\beta$.

Theorem 11 (Characteristic Property of Free Products)

Suppose $\mathscr{G} = \{G_i\}_{i\in I}$ is a family of groups indexed by a set I. Then a free product $\ast_{i\in I} G_i$ is a group with inclusion homomorphisms $\iota_\alpha : G_\alpha \to \ast_{i\in I} G_i$ such that given any group H and homomorphisms $\psi_\alpha \in \mathrm{Hom}(G_\alpha, H), \exists!\psi \in \mathrm{Hom}(\ast_{i\in I} G_i, H)$ s.t. $\forall \alpha \in I : \psi \circ \iota_\alpha = \psi_\alpha$, i.e. the following diagram commutes:



Further, this universal property characterizes the free product upto unique isomorphism that preserves inclusions.

Now we use this theory of free products to develop a class of groups called the free groups.

4.3.3 Free Groups and Presentations

Definition 39. Let S be a set. Define the ***free group generated by*** S, denoted by $F(S)$, as follows:

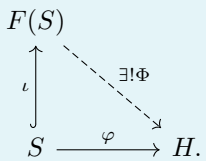
(a) If $S = \varnothing$, then let $F(\varnothing) = \{e\}$.
(b) If $|S| = 1$, i.e. $S = \{\sigma\}$ for a single element, then define $F(\{\sigma\}) = \{\sigma^n : n \in \mathbb{Z}\} \cong Z_\infty$ to be the group of formal powers of σ under the usual composition.
(c) If $|S| > 1$, then define $F(S) := \ast_{\sigma\in S} F(\{\sigma\})$.

Example 51

For any $n \in \mathbb{N}$, the ***free group on*** n ***generators***, denoted by F_n or $\mathfrak{F}_n$ is simply the group $F([n])$. In other words, $F_n \cong \ast_{i=1}^n Z_\infty$.

Theorem 12 (Characteristic Property of Free Group)

Let S be any set. Then the free group generated by S is a group $F(S)$ with an inclusion map $\iota : S \hookrightarrow F(S)$ such that if H any group and $\varphi : S \to H$ any function, then $\exists!\Phi \in \mathrm{Hom}(F(S), H)$ extending φ, i.e. s.t. $\varphi = \Phi \circ \iota$, i.e. s.t. the following diagram commutes:



Proof. Set maps $\varphi : S \to H$ correspond bijectively to collections of homomorphisms $\varphi_\sigma : F(\{\sigma\}) \to H$ via $\phi_\sigma(\sigma^n) = \phi(\sigma)^n$. Apply Theorem 11. ∎

We are now ready to talk about group presentations. Recall that just as the intersection of subgroups is a subgroup, it is easy to see that the intersection of normal subgroup is normal.

Definition 40. Let G be any group and $R \subseteq G$ any subset. Then the ***normal closure of*** R ***in*** G, denoted R^{nc} or sometimes (ambiguously) $\overline{R}$, is defined to be the smallest normal subgroup of G containing R. In other words, $\overline{R} = \bigcap_{\substack{H \trianglelefteq G \\ H \supseteq S}} H$.

Observe that $R = \overline{R}$ iff $R \trianglelefteq G$.

Definition 41. A ***group presentation*** is an ordered pair (S, R) usually denoted by $\langle S|R\rangle$, where S is any set and R any subset of $F(S)$. The elements of S are called ***generators*** and the elements of R are called ***relators*** or ***relations***. This defines a group by $\langle S|R\rangle = F(S)/\overline{R}$.

Intuitively, in passing to the quotient by $\overline{R}$, we are "modding out" by elements of R, i.e. declaring that they are the identity.

Definition 42. Let G be an arbitrary group. A ***presentation*** of G is a group presentation $\langle S|R\rangle$ with an isomorphism $\langle S|R\rangle \cong G$.

If such an isomorphism exists, it is uniquely determined by specifying which element of G corresponds to each generator in S. Often, the isomorphism is understood or irrelevant, and we say $\langle S|R\rangle$ is a presentation of G.

Definition 43. If G admits a presentation $\langle S|R\rangle$ with S finite, then G is said to be ***finitely generated***. If it admits a presentation with both S, R finite, then it is said to be ***finitely presented***.

If G is finitely presented and $S = \{s_1, \ldots, s_n\}$ and $R = \{r_1, \ldots, r_n\}$, then we usually write $\langle s_1, \ldots, s_n | r_1, r_2, \ldots, r_n\rangle$ for $\langle S|R\rangle$. $\langle s_1, \ldots, s_n | r_1 = q_1, \ldots, r_n = q_n\rangle$ is alternative notation for $\langle s_1, \ldots, s_n | r_1 q_1^{-1}, \ldots, r_n q_n^{-1}\rangle$.

Example 52

The following are presentations of some familiar groups:

(a) The free group on $\{a_1, \cdots, a_n\}$ has the presentation $F(\{a_1, \cdots, a_n\}) = \langle a_1, \cdots, a_n|\varnothing\rangle$. In particular, $\mathbb{Z} = F_1$ has the presentation $\langle \alpha|\varnothing\rangle$.

(b) The group $\mathbb{Z} \times \mathbb{Z}$ has presentation $\langle \beta, \gamma | \beta\gamma\beta^{-1}\gamma^{-1}\rangle$. This can equivalently be written as $\langle \beta, \gamma | \beta\gamma = \gamma\beta\rangle$.

(c) The cyclic group $\mathbb{Z}/n\mathbb{Z}$ has presentation $\mathbb{Z}/n\mathbb{Z} \cong \langle \alpha | \alpha^n\rangle$.

(d) The group $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ has presentation $\langle \beta, \gamma | \beta^m, \gamma^n, \beta\gamma\beta^{-1}\gamma^{-1}\rangle$.

Example 53

The following are presentations you may not have seen before:

(a) The group K_4 has presentation $\langle a, b | a^2 = b^2 = (ab)^2\rangle$, and so does $Z_2 \times Z_2$. Therefore, $K_4 \cong Z_2 \times Z_2$.

(b) The group $\mathfrak{S}_3$ has presentation $\mathfrak{S}_3 \cong \langle \sigma, \tau | \sigma^3, \tau^2, \tau\sigma\tau^{-1}\sigma\rangle$.

(c) The quaternion group Q_8 is defined by $Q_8 = \{1, -1, i, -i, j, -j, k, -k\}$ with the usual quaternion product rules. This has presentation $Q_8 \cong \langle i, j | i^4, i^2 j^{-2}, ij^{-1}ij\rangle$.

Given an arbitrary presentation, it may be difficult to tell when two elements of a group are equal. Even in simple presentations, there may be hidden or implicit relations that may cause unseen collapsing.

Example 54

The group $\langle \alpha | \alpha^5, \alpha^{17} \rangle$ is trivial. Similarly, but perhaps not as obviously,

$$G = \langle u, v | u^4 = v^3 = e, uv = v^2 u^2 \rangle$$

is trivial.

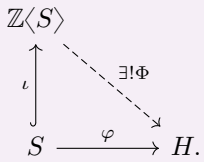
Topologists Tietze and Dehn around 1910 posed the following two problems: the ***isomorphism problem*** is to decide, given two finite presentations, whether the resulting groups are isomorphic; and the ***word problem*** is to decide, given a finite presentation, whether a word in S is equal to the identity. It was shown independently by Adian (in 1955) and Rabin (in 1958) that there is no algorithm for solving these problems that is guaranteed to give an answer in finite time, i.e. these problems are algorithmically undecidable. This laid rest to a famous problem in topology called the homeomorphism problem: given two spaces, there is no algorithm that is guaranteed to tell us if these spaces are homeomorphic in finite time. However, we've seen of course that this *is* possible in certain special cases, like we did for $\mathfrak{S}_3$.

4.3.4 Free Abelian Groups

There is another construction, very similar to the free group, called the ***free abelian group*** generated by a set. Let's quickly have a look at that. In this section, all our groups are abelian, so we write the group operation additively.

For this section, we take a different approach: we *define* the free abelian group as the group satisfying a certain universal property (this would prove that such a construction, if it existed, would be unique), and then construct it. So let's start with the characteristic property.

Definition 44 (Characteristic Property of Free Abelian Group)**.** Let S be any set. Then a ***free abelian group*** generated by S is an abelian group $\mathbb{Z}\langle S \rangle$ with an inclusion map $\iota : S \hookrightarrow \mathbb{Z}\langle S \rangle$ such that if H is any abelian group and $\varphi : S \to H$ any function, then $\exists! \Phi \in \mathrm{Hom}(\mathbb{Z}\langle S \rangle, H)$ extending φ, i.e. s.t. $\varphi = \Phi \circ \iota$, i.e. s.t. the following diagram commutes:



By the standard universal property argument in the category (Ab), such a construction, if it exists, is unique upto a unique isomorphism preserving ι, so that we may refer to *the* free abelian group generated by a set S.

Theorem 13

For any set S, the free abelian group generated by S exists, and can be thought of as the group of all formal finite integer linear combinations of elements of S.

Proof. A formal finite integer linear combination (ffilc) of elements of S is a finite sum of the form $\sum_{i=1}^{n} \lambda_i s_i$ where $n \in \mathbb{N}_0, \forall i : \lambda_i \in \mathbb{Z}, s_i \in S$. (More formally, a ffilc of elements of S is a function $f : S \to \mathbb{Z}$ with finite support, i.e. such that $f(s) = 0$ for all but finitely many $s \in S$.) The set of all such ffilc's of elements of S forms a natural abelian group under pointwise addition, and satisfies the characteristic property of the free abelian group. ∎

Example 55

For any set S, $F(S) \cong \mathbb{Z}\langle S \rangle$ iff $|S| \leqslant 1$.

Proposition 4

For a finite set $S = \{\sigma_1, \cdots, \sigma_n\}$, $\mathbb{Z}\langle S\rangle \cong \mathbb{Z}^n$ via the map $\sum_{i=1}^n \lambda_i\sigma_i \mapsto (\lambda_1, \cdots, \lambda_n)$. In general, for any set S, $\mathbb{Z}\langle S\rangle \cong \bigoplus_{\sigma\in S} \mathbb{Z}\langle\{\sigma\}\rangle$.

Example 56

Let $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ be the complex plane along with the point at infinity. Then we define the ***divisor group*** $\mathrm{Div}(\hat{\mathbb{C}})$ to be the free abelian group generated by the elements of $\hat{\mathbb{C}}$, i.e. $\mathrm{Div}(\hat{\mathbb{C}}) := \mathbb{Z}\langle\hat{\mathbb{C}}\rangle$. It consists of formal linear combinations like $2(1)+3(\imath)-7(2+4\imath)+5(\infty)$. We have a natural surjective homomorphism called the degree map $\deg : \mathrm{Div}(\hat{\mathbb{C}}) \to \mathbb{Z}$ given by $\sum_{i=1}^n n_i(P_i) \mapsto \sum_{i=1}^n n_i$. The kernel of this map is called the subgroup of ***principal divisors*** on $\hat{\mathbb{C}}$ and is denoted by $\mathrm{Div}^0(\hat{\mathbb{C}})$. The ***divisor class group*** or ***Picard group*** of $\hat{\mathbb{C}}$ is defined to be the quotient $\mathrm{Pic}(\hat{\mathbb{C}}) := \mathrm{Div}(\hat{\mathbb{C}})/\mathrm{Div}^0(\hat{\mathbb{C}})$. By the First Isomorphism Theorem, $\mathrm{Pic}(\hat{\mathbb{C}}) \cong \mathbb{Z}$. These are objects of study in the algebraic geometry of Riemann surfaces.

This is as far as we'll go with our theory of free abelian groups, but watch out for 1 (one) HW problem that needs them.

4.4 A Few More Groups of Small Order

We're done classifying groups of order $\leqslant 5$, and we know all of them are abelian. Let's now look at groups of order 6.

Theorem 14

If G is a group of order 6, then either $G \cong Z_6$ or $G \cong \mathfrak{S}_3$.

Proof. We do this is steps.

(a) Let $g \neq e \in G$. By Lagrange, $|g| \in \{2, 3, 6\}$. If $|g| = 6$, then $G \cong Z_6$. Hence, assume that all elements of G have order either 2 or 3.

(b) We show now that G must have an element of order 3. Assume that all elements of G have order 2. By the HW problem, G is abelian. Let $g \neq e \in G$. Then $|\langle g\rangle| = 2$ so that $\exists h \in G \smallsetminus \langle g\rangle$. Then $H := \langle g, h\rangle \leqslant G$ is a subgroup with presentation $\langle g, h | g^2 = h^2 = e, gh = hg\rangle \cong K_4$; this is a subgroup of G of order 4, which is not possible by Lagrange's Theorem. This contradiction shows that G must contain an element of order 3.

(c) Let $\sigma \in G$ with $|\sigma| = 3$. Then $\langle\sigma\rangle \leqslant G$ is a subgroup of index 2, and so by another HW problem, $\langle\sigma\rangle \trianglelefteq G$. Let $\tau \in G \smallsetminus \langle\sigma\rangle$. Then by normality, $\tau\sigma\tau^{-1} \in \{e, \sigma, \sigma^2\}$. We split into subcases:

(i) Assume $\tau\sigma\tau^{-1} = e$; then $\sigma = e$, a contradiction.

(ii) Assume $\tau\sigma\tau^{-1} = \sigma$; this means that σ and τ commute. We show that τ must have order 2. Assume that τ had order 3; then $\langle\sigma, \tau | \sigma^3, \tau^3, \sigma\tau\sigma^{-1}\tau^{-1}\rangle \cong Z_3 \times Z_3$ would be a subgroup of G of order 9, and this is impossible. Therefore, τ must have order two. But then $G = \langle\sigma, \tau | \sigma^3 = \tau^2 = e, \sigma\tau = \tau\sigma\rangle \cong Z_3 \times Z_2 \cong Z_6$ has the element $\sigma\tau$ of order 6, a contradiction to hypothesis.

(iii) The only option left is $\tau\sigma\tau^{-1} = \sigma^2$, and this gives us the group presentation $G = \langle\sigma, \tau | \sigma^3 = \tau^2 = e, \tau\sigma\tau^{-1} = \sigma^2\rangle \cong \mathfrak{S}_3$.

■

As you can imagine, the larger the order of the group is, the more difficult it becomes to classify subgroups. We know that there is only group of order 7, namely Z_7, but already we've run into an order we don't know how to work on: we don't yet have the tools to classify all groups of order 8 or 9. That'll have to wait another week or two. Till then, let's look at some other common families of groups.

4.5 Dihedral, Symmetric, Alternating Groups

In this section, we familiarize ourselves with some standard groups.

4.5.1 Dihedral Groups

Definition 45. The ***dihedral group*** D_{2n} of order $2n$ is the group of symmetries of a regular n-gon in the Cartesian plane. It has the group presentation: $\langle r, s | r^n, s^2, srs^{-1}r \rangle$.

The group presentation is sometimes written instead as $\langle \sigma, \tau | \sigma^n, \tau^2, \tau\sigma\tau^{-1}\sigma \rangle$. It is not hard to see that this group presentation defines a group of order $2n$ (see DF) and this is precisely the group of symmetries of a regular n-gon. Further, each element of the group can be written uniquely as $s^i r^j$ for $0 \leqslant i \leqslant 1, 0 \leqslant j \leqslant n$, and these interact by the same flipping idea $sr^j = r^{-j}s$. This is called the dihedral group because D_{2n} is the group of rotations of a dihedron with base a regular n-gon as its base.

4.5.2 Symmetric Groups

We've seen that given any set Ω, the set of "permutations on Ω," i.e., bijections $\sigma : \Omega \to \Omega$, forms a group under composition, called the ***symmetric group on*** Ω**,** and is denoted by $\mathfrak{S}_\Omega$ or S_Ω. We first observe that this group essentially depends only on the size of Ω.

Claim 17 — If $|\Delta| = |\Omega|$, then $\mathfrak{S}_\Delta \cong \mathfrak{S}_\Omega$. In particular, if $\theta : \Delta \to \Omega$ is any bijection, then the map $\mathrm{conj}_\theta : \mathfrak{S}_\Delta \to \mathfrak{S}_\Omega$ given by $\mathrm{conj}_\theta(\sigma) = \theta \circ \sigma \circ \theta^{-1}$ is an isomorphism.

Note that in this case each step of the conjugation takes us to an element of a different set. Apart from this difference, the proof is identical to the one you did on your HW. In essence, conjugation is simply a "change of labels" operation. Note, in particular, the similarity to change of basis matrices from linear algebra.

Proof. For any $\sigma \in \mathfrak{S}_\Delta, \mathrm{conj}_\theta(\sigma) \circ \mathrm{conj}_\theta(\sigma^{-1}) = \mathrm{id}_\Omega$, so that indeed $\mathrm{conj}_\theta(\sigma) \in \mathfrak{S}_\Omega$. The map $\mathrm{conj}_{\theta^{-1}} : \mathfrak{S}_\Omega \to \mathfrak{S}_\Delta$ is a two-sided inverse to conj_θ, so that conj_θ is a bijection. Finally, for $\sigma, \tau \in \mathfrak{S}_\Delta, \mathrm{conj}_\theta(\sigma\circ\tau) = \mathrm{conj}_\theta(\sigma) \circ \mathrm{conj}_\theta(\tau)$, so that conj_θ is a homomorphism. ■

The converse is also true.

Claim 18 — If Δ and Ω are any sets, then $\mathfrak{S}_\Delta \cong \mathfrak{S}_\Omega \Rightarrow |\Delta| = |\Omega|$.

This is significantly harder to prove, and the general case can be found in DF. It's much easier in the finite case.

Claim 19 — If Δ and Ω are nonempty *finite* sets, then $\mathfrak{S}_\Delta \cong \mathfrak{S}_\Omega \Rightarrow |\Delta| = |\Omega|$.

Proof. For a finite set X of cardinality n, by Claim 5.6, $|\mathfrak{S}_X| = |\mathfrak{S}_n| = n!$, i.e. $|\mathfrak{S}_X| = |X|!$. Therefore, $\mathfrak{S}_\Delta \cong \mathfrak{S}_\Omega \Rightarrow |\mathfrak{S}_\Delta| = |\mathfrak{S}_\Omega| \Rightarrow |\Delta|! = |\Omega|! \Rightarrow |\Delta| = |\Omega|$. ■

Therefore, in essence, we need only look at the groups $\mathfrak{S}_n$ for $n \in \mathbb{N}$. For $n = 1$, $\mathfrak{S}_1 \cong \{e\}$, and for $n = 2$, $\mathfrak{S}_2 \cong \mathbb{Z}/2\mathbb{Z}$, and the only nonidentity permutation is the one that switches the two elements. We've already seen one notation for elements of $\mathfrak{S}_n$: for instance,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathfrak{S}_3$$

is the element that sends $1 \mapsto 2, 2 \mapsto 3$ and $3 \mapsto 1$. As you can imagine, this notation gets really cumbersome as n increases. Moreover, and more importantly, this notation gives us no idea how to compute higher powers or compositions efficiently. For instance, if

$$\xi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix} \in \mathfrak{S}_5,$$

can you tell me immediately what ξ^{100} is? If you enjoy coding, try to implement an algorithm that composes permutations.

To talk about higher symmetric groups, let's first familiarize ourselves with more efficient notation.

Definition 46. For a given $n \in \mathbb{N}$, a ***cycle of length*** k (or a $k-$***cycle***) in $\mathfrak{S}_n$ is a permutation of the form

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_{k-1} & a_k \\ a_2 & a_3 & \cdots & a_k & a_1 \end{pmatrix},$$

i.e. a permutation that cyclically permutes the integers $a_1, \cdots, a_k \in [n]$, and leaves the other $n-k$ integers fixed. The cycle above is notated simply by $(a_1 a_2 \cdots a_k)$ or sometimes by $(a_1, a_2, \cdots, a_k)$ (especially if the integers a_j have more than 1 digit). Note that the cycle $(a_1 a_2 \cdots a_k)$ is the same as the cycle $(a_j a_{j+1} \cdots a_{k-1} a_k a_1 \cdots a_{j-1})$ for any $1 \leqslant j \leqslant k$.

Example 57

In cycle notation, the identity permuation is simply $()$. The σ above is $(123) = (231) = (312) \neq (213)$. The ξ above is NOT a cycle.

Two cycles σ and τ are said to be ***disjoint*** if the sets of integers they cyclically permute are.

Theorem 15 (Cycle Decomposition)

The following three claims underly the key significance of this notation:
(a) Any permutation can be written essentially uniquely as a product of disjoint cycles.
(b) It is easy to compute powers of cycles.
(c) Disjoint cycles commute.

By "essentially uniquely" we mean to say that the decomposition is unique upto the order of the disjoint cycles (which doesn't matter by part (c)) and upto the notational ambiguity mentioned in the definition above.

The claims in (b) and (c) are obvious. For instance, if $\sigma = (a_1 a_2 \cdots a_k)$, then observe that $\sigma^k = ()$. Therefore, given any m, to find σ^m write $m = qk + r$ for $0 \leqslant r < k$; if $r = 0$, then $\sigma^m = ()$, and if $0 < r < k$, then $\sigma^m = (a_1 a_{1+r} a_{1+2r} \cdots)$. For instance, if $\sigma = (12345)$, then to find σ^{297}, write $297 = 295 + 2$ so that $\sigma^{297} = \sigma^2 = (13524)$.

The proof of Theorem 15(a) is simply the following canonical algorithm.

Algorithm 1 (Cycle Decomposition Algorithm) — To decompose any given permutation σ into cycles.
Stp.1. To start a new cycle, pick the smallest element a of $[n]$ not in any previous cycle. (Start the algorithm with $a = 1$.)
Stp.2. Trace $a = \sigma^0(a), \sigma(a), \sigma^2(a), \ldots$ until the first repetition when $\sigma^N(a) = a$. Add the cycle $(a, \sigma(a), \sigma^2(a), \cdots, \sigma^{N-1}(a))$ to your list.
Stp.3. Repeat Step 1 till all elements of $[n]$ are part of some cycle.
Stp.4. Remove all cycles of length 1. Compose the remaining cycles together in any order to get the cycle decomposition for σ.

Note that $\mathfrak{S}_n$ is a finite group, so Steps 2 and 3 will always terminate. In the cycle decomposition, any elements fixed by σ are simply omitted from the cycle notation. The decomposition is unique and the cycles are disjoint because the relation $a \sim_\sigma b \Leftrightarrow \exists i : a = \sigma^i(b)$ is an equivalence relation on $[n]$; the equivalence classes of $\sim_\sigma$ are called the ***orbits*** of σ.

Example 58

The cycle decomposition algorithm applied to the ξ above yields the cycles (134) and (25) so that $\xi = (134)(25)$. From this, it is obvious that

$$\xi^{100} = (134)^{100}(25)^{100} = (134) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 4 & 1 & 5 \end{pmatrix}.$$

In this case, the orbit of 1 is $\{1, 3, 4\}$ and the orbit of 2 is $\{2, 5\}$; there are only two orbits.

From this we immediately see two things:

Proposition 5
The order of a permutation is the least common multiple of the lengths of the cycles in its cycle decomposition.

Proposition 6
As soon as $n \geqslant 3$, the group $\mathfrak{S}_n$ is non abelian.

Proof. For $n \geqslant 3$, the cycles (12) and (13) don't commute. ■

We know that any permutation in $\mathfrak{S}_n$ can be written essentially uniquely as a product of disjoint cycles. By contrast, every permutation can be written in many ways as the product of nondisjoint cycles. However, there is something common among these decompositions: a "parity."

Definition 47. A 2-cycle is called a ***transposition***.

Note that, by definition, every transposition has order 2.

Proposition 7
Every $\sigma \in \mathfrak{S}_n$ can be written as a product of transpositions. Equivalently, $\mathfrak{S}_n = \langle T \rangle$, where $T = \{(ij) | 1 \leqslant i < j \leqslant n\}$. In general, this decomposition need not be unique.

Proof. This follows from Theorem 15(a) along with the simple observation that

$$(a_1 a_2 \cdots a_k) = (a_1 a_k)(a_1 a_{k-1}) \cdots (a_1 a_3)(a_1 a_2).$$

Notice that $(123) = (13)(12) = (12)(13)(12)(13) = (12)(23)$. ■

Note that while the transpositions in the decomposition above are different, the parity of the number of transpositions in the same. As we shall see, this is no coincidence.

4.5.3 Alternating Groups

Observe that the group $\mathfrak{S}_n$ "acts on" any set with n elements, i.e. for any set $A = \{a_1, \cdots, a_n\}$, we can define a map $\mathfrak{S}_n \times A \to A$ by sending $(\sigma, a_j) \mapsto a_{\sigma(j)}$. We denote this by $\sigma(a_j) := a_{\sigma(j)}$. Note that this action satisfies the properties $\mathrm{id}_{[n]}(a_j) = a_j$ and $\sigma(\tau(a_j)) = (\sigma \circ \tau)(a_j)$. (We will make the notion of group actions more precise later.) To make the notion of "parity" precise, we make a series of definitions.

Definition 48. Let $\mathbf{x} = (x_1, \cdots, x_n)$ be an n-tuple of independent variables x_j for $1 \leqslant j \leqslant n$.
(a) Define the ***Vandermonde polynomial*** of $\mathbf{x}$ to be $\Delta(\mathbf{x}) := \prod_{1 \leqslant i < j \leqslant n} (x_i - x_j)$.
(b) For any $\sigma \in \mathfrak{S}_n$, define $\sigma\mathbf{x} = \sigma(\mathbf{x}) := (\sigma(x_1), \cdots, \sigma(x_n)) = (x_{\sigma(1)}, \cdots, x_{\sigma(n)})$.

Notice that (a) for any $\sigma, \tau \in \mathfrak{S}_n$, $(\sigma \circ \tau)(\mathbf{x}) = \sigma(\tau\mathbf{x})$, and that (b) for any $\sigma \in \mathfrak{S}_n$, $\Delta(\sigma(\mathbf{x})) = \pm\Delta(\mathbf{x})$. Therefore, the following definition makes sense.

Definition 49. Given any $\sigma \in \mathfrak{S}_n$, define the ***sign*** of σ to be $\mathrm{sign}(\sigma) := \Delta(\sigma\mathbf{x})/\Delta(\mathbf{x})$. The sign of σ is sometimes also denoted as $(-1)^\sigma$.

Clearly, this is independent of the choice of $\mathbf{x}$, and depends only on σ.

Definition 50. A permutation $\sigma \in \mathfrak{S}_n$ is said to be ***even*** if $\mathrm{sign}(\sigma) = 1$ and ***odd*** if $\mathrm{sign}(\sigma) = -1$.

The following lemma is the key:

Lemma 8

For $n \geqslant 2$, the map $\text{sign} : \mathfrak{S}_n \to \{\pm 1\}$ is a surjective homomorphism.

Proof. For any $\tau, \sigma \in \mathfrak{S}_n$,

$$\text{sign}(\sigma \circ \tau) = \frac{\Delta(\sigma \circ \tau(\mathbf{x}))}{\Delta(\mathbf{x})} = \frac{\Delta(\sigma(\tau\mathbf{x}))}{\Delta(\tau\mathbf{x})}\frac{\Delta(\tau\mathbf{x})}{\Delta(\mathbf{x})} = \text{sign}(\sigma)\,\text{sign}(\tau).$$

Now notice that for $n \geqslant 2, \text{sign}(\text{id}) = 1$ and $\text{sign}(12) = -1$. ■

We now prove what we promised.

Proposition 8

Any transposition is an odd permutation. A permutation is odd iff it is the product of an odd number of transpositions, and it is even iff it is the product of an even number of transpositions. In particular, the parity of the number of transpositions in any decomposition of a permutation into transpositions is constant.

Proof. Notice that for any transposition (ij), if $\lambda = (1i)(2j)$, then $(ij) = \lambda(12)\lambda$. This means that $\text{sign}(ij) = \text{sign}(\lambda)\,\text{sign}(12)\,\text{sign}(\lambda) = \text{sign}(\lambda)^2\,\text{sign}(12) = -1$. Let σ be any transposition, and write σ as a product of m not-necessarily-disjoint transpositions; then $\text{sign}(\sigma) = (-1)^m$. ■

From this, and the observation made above in the proof of Proposition 7, we see that a k-cycle is odd iff k is even. Therefore, a permutation σ is odd iff the number of cycles of even length in its decomposition is odd. We are now ready to talk about alternating groups.

Definition 51. The ***alternating group on n letters***, denoted by $\mathfrak{A}_n$ or A_n, is the kernel of the sign homomorphism $\text{sign} : \mathfrak{S}_n \to \{\pm 1\}$.

In other words, the alternating group on n letters is simply the group of even permutations on n letters. It is clear that $\mathfrak{A}_n \trianglelefteq \mathfrak{S}_n$ and for $n \geqslant 2$, $\mathfrak{S}_n/\mathfrak{A}_n \cong \{\pm 1\}$ so that $|\mathfrak{A}_n| = n!/2$. This means that for small n, $\mathfrak{A}_n$ is already familiar to us. $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are both trivial, while $|\mathfrak{A}_3| = 3$ so that $\mathfrak{A}_3 \cong \mathbb{Z}/3\mathbb{Z}$. It is easy to see that for $n \geqslant 4, \mathfrak{A}_n$ is nonabelian.

In general, there is a trend that for small n, the group $\mathfrak{S}_n$ is the set of symmetries of a regular figure in some small-dimensional Euclidean space, and the subgroup $\mathfrak{A}_n$ is the set of *rotations* of the same figure in that space. For instance, for $n = 2$, this is true for a line segment in $\mathbb{R}^1$, for $n = 3$, this is true for a triangle in $\mathbb{R}^2$, for $n = 4$, this is true for a tetrahedron in $\mathbb{R}^3$. Let's show this for $n = 4$.

Correction. This is not true for $n = 5$ and the icosahedron in $\mathbb{R}^3$. While it is correct that the group of rotations of an icosahedron is $\mathfrak{A}_5$, the full symmetry group of the icosahedron is actually $\mathfrak{A}_5 \times Z_2$, which is NOT isomorphic to $\mathfrak{S}_5$, even though both have the same cardinality (120).

Theorem 16

The group $\mathfrak{A}_4$ is the group of rotations of a regular tetrahedron in $\mathbb{R}^3$.

Proof. Recall that a tetrahedron has 4 vertices, 6 edges, and 4 faces, each of which is an equilateral triangle. Label the vertices of the tetrahedron by 1, 2, 3, and 4. There are two types of axes of symmetry: one that joins a vertex to the center of the opposite face, and the other that joins the midpoints of opposite edges. The first one allows rotation by $2\pi/3$ and $4\pi/3$; for instance, about the axis through vertex 1, the rotations by $2\pi/3$ and $4\pi/3$ correspond to the permutations (234) and $(234)^2 = (243)$ of the vertices. Similarly, we get the permutations $(341), (314), (412), (421), (123), (132)$ by such rotations. The second kind (of which there are three) allows only rotation by π; for instance, rotation about π across the axis joining the midpoints of edges 12 and 34 gives rise to the permutation $(12)(34)$. Similarly, we get the permutations $(13)(24)$ and $(14)(23)$. In all, we have $2 \times 4 + 3 = 11$ nonindentity even permutations in $\mathfrak{S}_4$, so that along with the identity this gives us 12 even permutations. Since $|\mathfrak{A}_4| = \frac{1}{2} \cdot 4! = 12$, these must be all. ■

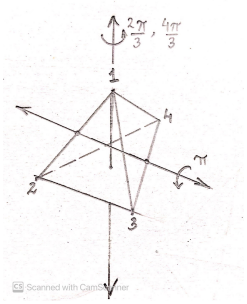


Figure 1: Rotations of a Tetrahedron

Note that any rotation of the first type composed with any rotation of the second type can be composed together to build the whole group. Therefore, the above discussion also tells us the nature of all the subgroups of $\mathfrak{A}_4$. They are precisely:

(a) Four 3-element subgroups generated by rotations of the first kind: $\langle(234)\rangle, \langle(341)\rangle, \langle(412)\rangle$ and $\langle(123)\rangle$.
(b) Three 2-element subgroups generated by rotations of the second kind: $\langle(12)(34)\rangle, \langle(13)(24)\rangle$ and $\langle(14)(23)\rangle$.
(c) The unique 4-element subgroup generated by any two of the rotations of the second kind:

$$\langle(12)(34), (13)(24)\rangle = \{e, (12)(34), (13)(24), (14)(23)\}.$$

It is easy to see that this last subgroup of order 4 is isomorphic to K_4, and that it is the only nontrivial proper normal subgroup of $\mathfrak{A}_4$. In particular, $\mathfrak{A}_4$ is not simple. However, this is the only exception.

Theorem 17

For $n = 1, 2, 3$ and $n \geqslant 5$, $\mathfrak{A}_n$ is a simple group.

We don't (yet) have the tools to prove this, but we will. Let's first look at something different: the notion of a group acting on a set.

5 Group Actions

One of the most fundamental notions and a unifying theme across mathematics is the notion of an algebraic structure–a group, ring, or field–acting on another kind of structure–a set, an abelian group, a topological space, etc. Usually when this happens, we are able to gain much understanding about both of these structures. Let's see how that works precisely.

5.1 Basic Definitions

Definition 52. A ***(left) group action*** of a group G on a set X is a map $G \times X \to X$ written $(g, x) \mapsto g \cdot x$ satisfying the following axioms:
(a) (Compatibility) $\forall g, h \in G, \forall x \in X : g \cdot (h \cdot x) = (gh) \cdot x$.
(b) (Identity) $\forall x \in X : e_G \cdot x = x$.
Given a group action $G \times X \to X$, we then say that the group G acts on the set X, and write $G \circlearrowright X$.

Notice that we could have similarly defined the notion of a right group action. Often, when there is no danger of confusion, we drop the central dot and denote a group action simply by gx. Sometimes, instead of $g \cdot x$, group actions are denoted by x^g; we will refrain from using this notation.

Observe that if a group G acts on a set X, then for each $g \in G$ we get a map $\rho_g : X \to X$ given by $x \mapsto g \cdot x$. By properties (a) and (b), this has a two sided inverse $\rho_{g^{-1}}$, so that $\rho_g \in \mathfrak{S}_X$. Further, the map $\rho : G \to \mathfrak{S}_X$ given by $g \mapsto \rho_g$ is a homomorphism. This homomorphism is called the ***permutation representation*** associated to the group action. Observe that this process is reversible: given a homomorphism $\rho : G \to \mathfrak{S}_X$, we get a natural action $G \times X \to X$ given by $(g, x) \mapsto \rho_g(x)$.

Note that when X is simply a set, this is the best we can do. However, if X has more structure than simply that of a set, we can look at group actions compatible with the structure of X. For example, if X is a group or ring or field, we can look at homomorphisms $\rho : G \to \mathrm{Aut}(X)$, if it's a topological space, we can look at homomorphisms $\rho : G \to \mathrm{Homeo}(X)$, etc.

Definition 53. Let $\rho : G \to \mathfrak{S}_X$ be the permutation representation corresponding to a group action $G \times X \to X$.
(a) The kernel $\ker \rho$ is called the ***kernel*** of the group action. It is the subgroup of G that fixes every element of X.
(b) The group action is said to be ***trivial*** if $\ker \rho = G$ and ***faithful*** if $\ker \rho = \{e\}$. The condition that a group action be trivial is that the all elements of the group induce the same (i.e. identity) permutation; the condition that a group action be faithful is that distinct elements of the group induce distinct permutations.

Let's now look at some examples.

Example 59

Given any group G and any set X, G trivially acts on the set X by taking $\rho : G \to \mathfrak{S}_X$ to be the null map.

Example 60

For any $n \in \mathbb{N}_0$, the group $\mathbb{R}^\times$ acts on the abelian group $(\mathbb{R}^n, +)$ by $s \cdot (v_1, \cdots, v_n) = (sv_1, \cdots, sv_n)$. More generally, if V is any vector space over the scalar field $\mathbb{F}$, then $\mathbb{F}^\times$ acts on V by, well, scaling.

Example 61

For any set X, the set $\mathfrak{S}_X$ acts on X by $\sigma \cdot x = \sigma(x)$. The associated permuation representation is the identity map $\rho = \mathrm{id}_{\mathfrak{S}_X} : \mathfrak{S}_X \to \mathfrak{S}_X$; it is most certainly faithful. Also, $\mathfrak{S}_n$ acts on any set with n elements: if $X = \{x_1, \cdots, x_n\}$, then $\mathfrak{S}_n$ acts on X by $\sigma(x_i) = x_{\sigma(i)}$.

Example 62

For any $n \in \mathbb{N}$, $\mathfrak{S}_n$ acts on the ring of polynomials $\mathbb{R}[x_1, \cdots, x_n]$ in such a way that is compatible with their addition and multiplication. In other words, $\forall \sigma \in \mathfrak{S}_n, \forall p, q \in \mathbb{R}[x_1, \cdots, x_n] : \sigma(p+q) = \sigma(p) + \sigma(q)$ and $\sigma(pq) = \sigma(p)\sigma(q)$.

Example 63

The group $\mathfrak{A}_4$ acts on the set of vertices, or the set of edges, or the set of faces of a regular tetrahedron. The group $\mathfrak{S}_4$ acts on the set of vertices, or the set of edges, or the set of faces, or the set of diagonals etc., of a cube.

Note that while the action of $\mathfrak{S}_4$ on the set of diagonals of a cube is faithful (in fact, the corresponding permutation representation is an isomorphism), the action of $\mathfrak{S}_4$ on the set of 3 opposite pairs of faces of a cube is not faithful. In general, a a faithful action of a group G on a set X embeds $\rho : G \hookrightarrow \mathfrak{S}_X$.

Example 64

Let G be a group and set $X = G$. Then G acts on X by left-multiplication (also called left-translation). More precisely, consider the action $G \times X \to X$ given by $g : x \mapsto gx$. This is called the ***left regular action*** of a group G on itself. The cancellation property shows that this action is faithful.

From this we immediately get:

Theorem 18 (Cayley's Theorem)

Every group is isomorphic to a subgroup of some symmetric group. More precisely, the permutation representation corresponding to the left regular action of G on itself is an embedding $\rho : G \hookrightarrow S_{|G|}$.

Historically, finite groups were not studied axiomatically the way we are doing, but rather as subgroups of the symmetric groups. Subgroups of the symmetric group are called ***permutation groups***. Cayley's Theorem tells us that the approaches are equivalent. Another thing I want to point out is that given any group action, we can work our cycle decompositions in this case too. If a group G acts on a finite set X, then we get a map $\rho : G \to \mathfrak{S}_X$ and $\mathfrak{S}_X$ finite, so we may perform the same cycle decomposition on ρ_g as before, and this can be done even without choosing a bijection $\theta : [n] \to X$. Note that this does not require G to be finite.

5.2 Orbit-Stabilizer Theorem, Not-Burnside's Lemma, Polyá Enumeration

One of the most fundamental theorem in the theory of group actions is the theorem connecting the size of the orbit of an element to the size of its stabilizer.

Definition 54. Let $G \curvearrowright X$.

(a) The relation $\sim$ on X given by $x \sim y \Leftrightarrow \exists g \in G : x = gy$ is an equivalence relation. The equivalence class of any element $x \in X$ is called the ***orbit*** of x, and is denoted by Gx or $\mathcal{O}_x$. The set of equivalence classes is called the set of orbits of the action of G on X, and is denoted by X/G.

(b) For any $x \in X$, the subgroup $G_x := \{g \in G : gx = x\} \leqslant G$ is called the ***stabilizer*** of x. This is sometimes denoted by $\mathrm{stab}(x)$.

Because X/G is a partition of X, the following lemma is clear.

Lemma 9

If $G \curvearrowright X$, then $X = \coprod_{\mathcal{O} \in X/G} \mathcal{O}$. If X is finite, then $|X| = \sum_{\mathcal{O} \in X/G} |\mathcal{O}|$.

Example 65

Let $\mathbb{R}^\times$ act on $\mathbb{R}$ by multiplication. Then $\mathbb{R}/\mathbb{R}^\times = \{\{0\}, \mathbb{R}^\times\}$, i.e. there are two orbits: the orbit of zero is simply $\{0\}$ and the orbit of any nonzero element is $\mathbb{R}^\times$. The stabilizer

$$\mathbb{R}^\times_x = \begin{cases} \mathbb{R}^\times, & \text{if } x = 0, \\ \{1\}, & \text{if } x \neq 0. \end{cases}$$

Definition 55. A group action is called transitive if there is only one orbit. In other words, a group action of group G on set X is transitive if $\forall x, y \in X : \exists g \in G : x = gy$.

Example 66

Consider the action of $\mathfrak{A}_4$ on the set V of vertices of a tetrahedron. As usual, label the vertices of the tetrahedron by 1, 2, 3, and 4. This is clearly transitive, so that the orbit of any vertex is the whole set V. The stabilizer of any vertex is the 3-element subgroup of $\mathfrak{A}_4$ generated by the rotation through the axis of symmetry through that vertex. For instance, $\operatorname{stab}(1) = \langle(234)\rangle$. Observe that in this case, for any vertex v, $|\mathcal{O}_v| \cdot |\operatorname{stab}(v)| = 4 \times 3 = 12 = |\mathfrak{A}_4|$. This is not a coincidence.

From these examples we see that the bigger the orbit of an element, the smaller its stabilizer. Let's make this precise.

Theorem 19 (Orbit-Stabilizer Theorem)

If $G \curvearrowright X$, then for any $x \in X$ there is a bijection $G/G_x \xrightarrow{\sim} Gx$. In particular, if G is finite, then for any element $x \in X$: $|Gx| \cdot |G_x| = |G|$.

Proof. For fixed $x \in X$, consider the map $\varphi_x : G/G_x \to Gx$ by $gG_x \mapsto gx$. This is well-defined and injective because $gG_x = g'G_x \Leftrightarrow g^{-1}g' \in G_x \Leftrightarrow (g^{-1}g')x = x \Leftrightarrow gx = g'x$. It is also surjective by definition, so that it is a bijection. If G is finite, then $|G/G_x| = |G|/|G_x| = |Gx|$. ∎

Example 67

The above is precisely the statement of Lagrange's Theorem if you take consider the natural action by left multiplication of G on the set $X = G/H$ of left cosets of a subgroup $H \leqslant G$.

From this, we prove the lemma that is usually (but incorrectly) attributed to mathematician William Burnside, but which is originally due to Cauchy and Frobenius. First, we need some notation.

Definition 56. Let $G \curvearrowright X$. For a particular $g \in G$, let $X^g := \{x \in X : gx = x\}$. This is the set of elements of X fixed by G; in other words $X^g = \{x \in X : g \in G_x\}$.

Lemma 10 (The Lemma that is not Burnside's/Cauchy-Frobenius Formula)

Let a finite group G act on a finite set X. Then

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

In other words, the number of orbits is the average number of elements fixed by the elements of G.

Proof. Notice first that $\sum_{g\in G}|X^g| = |\{(g,x)\in G\times X : g\cdot x = x\}| = \sum_{x\in X}|G_x|$. Now, using the Orbit-Stabilizer Theorem, we get that

$$\sum_{x\in X}|G_x| = |G|\sum_{x\in X}\frac{1}{|Gx|} = |G|\sum_{\mathcal{O}\in X/G}\sum_{x\in\mathcal{O}}\frac{1}{|\mathcal{O}|} = |G|\sum_{\mathcal{O}\in X/G}1 = |G|\cdot|X/G|.$$

■

This innocuous-looking statement has beautiful combinatorial consequences.

Example 68

In how many ways can we color the vertices of a regular tetrahedron with the colors red, green, and blue? (Note that two colorings are the same if there is a rotation which gets from one to another.)

Solution. 15. Consider the set X of all possible colorings; this is just the set of functions $\{1,2,3,4\}\to\{R,G,B\}$, so that it has size $|X| = 3^4 = 81$. The group $\mathfrak{A}_4$ acts on X, and we are interested in the number of orbits $|X/\mathfrak{A}_4|$. By Lemma 10, $|X/\mathfrak{A}_4| = \frac{1}{12}\sum_{\sigma\in\mathfrak{A}_4}|X^\sigma|$.

(a) The identity $e\in\mathfrak{A}_4$ fixes everything, so that $|X|^e = |X| = 81$.
(b) If $\sigma = (1)(234)$, then to count the number of colorings fixed by σ, observe that it fixes vertex 1, so we can color it arbitrarily in 3 ways, and it cyclically permutes vertices 2, 3, 4, which must therefore be the same color, which can be chosen independently in 3 ways, so that $|X^{(234)}| = 3\times 3 = 9$. Similarly, for any rotation σ of the first kind, $|X^\sigma| = 9$.
(c) If $\sigma = (12)(34)$, then σ switches vertices 1 and 2 and switches vertices 3 and 4. Therefore, 1 and 2 must be the same color, chosen in 3 ways, and 3 and 4 must be the same color, chosen in another 3 ways. This means $|X^{(12)(34)}| = 3\times 3 = 9$.

Putting this all together, we see that $|X/\mathfrak{A}_4| = \frac{1}{12}(1\times 81 + 8\times 9 + 3\times 9) = 15$, so there a 15 inequivalent colorings of the vertices of a tetrahedron by three colors. These are nothing but $\{R^4, R^3G, R^3B, R^2G^2, R^2GB, R^2B^2, RG^3, RG^2B, RGB^2, RB^3, G^4, G^3B, G^2B^2, GB^3, B^4\}$.

From the previous example, the following theorem is clear.

Theorem 20 (Unweighted Polyá Enumeration)

Let G be a finite group and X, C finite sets and let $|X| = n$. Observe that if $G\curvearrowright X$, then $G\curvearrowright C^X$ simply by $g\cdot f = f\circ\rho_{g^{-1}}$. For each $g\in G$, $\rho_g\in\mathfrak{S}_X\cong\mathfrak{S}_n$, so that we may define $c(g)$ to be the number of disjoint cycles in the cycle decomposition of ρ_g (including the 1-element cycles, i.e. fixed points). Then:

$$|C^X/G| = \frac{1}{|G|}\sum_{g\in G}|C|^{c(g)}$$

Intuitively, a map $f: X\to C$ corresponds exactly to a coloring, and $|C|^{c(g)}$ measures precisely the number of colorings invariant under $g\in G$. Observe that, while this was true of Example 68, in general, $c(g)$, which is the number of cycles in ρ_g, need NOT be the number of cycles in g itself. This is illustrated by the following example:

Example 69

In how many ways can we color the faces of a cube using n colors?

Solution. The number of distinct (rotationally inequivalent) colorings is $\frac{1}{24}(n^6+3n^4+12n^3+8n^2)$. To see this, take $X=[6]$ and label the faces from 1 through 6 by F1-R2-Ba3-L4-T5-Bo6, take C to be the set of n colors, and G to be the group of rotational symmetries of the cube (which we know is isomorphic to $\mathfrak{S}_4$, but that information is not needed for this problem). By the Polyá Enumeration Theorem,

$$|C^X/G| = \frac{1}{24}\sum_{\sigma\in G} n^{c(\sigma)}.$$

Enumerating as before,

(a) If $\sigma = e$, then $\rho_e = (1)(2)(3)(4)(5)(6)$, so that $c(e)=6$.
(b) For the $3\times 2 = 6$ rotations by $\pi/2$ and $3\pi/2$ about the 3 short diagonals: for instance, if the axis is through $T-Bo$ and rotation counterclockwise, then $\rho_\sigma = (1234)(5)(6)$. This gives us $c(\sigma)=3$.
(c) For the $3\times 1 = 3$ rotations by π about the short diagonals: for the same axis, we get $\rho_\sigma = (13)(24)(5)(6)$. This gives us $c(\sigma)=4$.
(d) For the $4\times 2 = 8$ rotations by $2\pi/3$ and $4\pi/3$ about the long diagonals: for instance, about one of them, we get $\rho_\sigma = (126)(453)$. In this case, $c(\sigma)=2$.
(e) For the $6\times 1 = 6$ rotations by π about the medium diagonals: for instance, about one of them, we get $\rho_\sigma = (13)(26)(45)$. In this case, $c(\sigma)=3$.

Putting all of this together, we get that the number of colorings is

$$\frac{1\times n^6 + 6\times n^3 + 3\times n^4 + 8\times n^2 + 6\times n^3}{24} = \frac{1}{24}(n^6+3n^4+12n^3+8n^2)$$

as needed.

As you can see, this is a pretty powerful tool! For instance, we can now immediately say that then number of cube face colorings with 25 colors is 10229375, without listing out a single one! Similar techniques can be applied to necklaces (with X being the set of beads and $G=Z_{|X|}$) and to bracelets (with X being the set of beads and $G=D_{2|X|}$) to obtain coloring formulae. One of these is on your HW. There is another, more general or weighted Polyá Enumeration formula, which is slightly more complicated, but significantly more powerful. We only mention the statement.

Theorem 21 (Weighted Polyá Enumeration/Redfield-Polyá Theorem)

Let G be a finite group, X a finite set with $|X|=n$, but let the set C of colors be possibly infinite. Suppose we have a function $w: C\to\mathbb{N}_0$ assigning to each color a weight. Let $f(t)=\sum_{r=0}^{\infty}|w^{-1}(r)|t^r$ be the generating function for the number of colors of a given weight. Define the ***cycle index***

$$Z_G(x_1,\cdots,x_n) = \frac{1}{|G|}\sum_{g\in G}\prod_{k=1}^{n} x_j^{c_k(g)},$$

where $c_k(g)$ is the number of k-cycles in ρ_g. If we define the weight of a coloring $\phi: X\to C$ by $\tilde{w}(\phi)=\sum_{x\in X} w(\phi(x))$, then the generating function for the number of colored arrangements by weight is:

$$\sum_{r=0}^{\infty}|\tilde{w}^{-1}(r)|t^r = Z_G(f(t),f(t^2),\cdots,f(t^n)).$$

We can recover the unweighted version by letting all colors have weight zero, so that $f(t)=n$. There is an even more general theorem obtained by replacing t everywhere with $\mathbf{t}=(t_1,t_2,\cdots)$ where the weight of each color is an ordered tuple of nonnegative integers. Needless to say, we will not attempt to prove any of these here; nonetheless, the core idea in the proof is the same principle as we've seen before–by counting the number of fixed points. This formula has applications to to graph theory and combinatorics, chemistry (counting the number of distinct acyclic molecules), computer science (counting the number of rooted ternary trees), etc.

5.3 Rotation Groups and Platonic Solids

In this section, we use group theory to classify all Platonic solids in $\mathbb{R}^3$. First we review some definitions.

Definition 57. In $\mathbb{R}^3$, a ***Platonic solid*** is a regular convex polyhedron. In other words, it is a polyhedron with congruent regular polygonal faces with the same number of faces meeting at each vertex.

Example 70

The tetrahedron, the cube, the octahedron, the dodecahedron and the icosahedron are all Platonic solids. Note that the tetrahedron is self-dual, the cube and the octahedron form a dual pair, and the dodecahedron and the icosahedron form a dual pair.

Platonic solids were studied extensively since antiquity, because they were considered the epitomes of symmetry and geometrical beauty. They are named after the Greek philosopher Plato, who associated them with the classical elements (earth, air, water, fire), and a new element he called the "aether." He also remarked the dodecahedron was what the gods used to "arrange constellations in heaven."

We will show that these are, in fact, all. Any proof of the classification of all Platonic solids involves two steps:

(a) Showing that there can't be any more than those mentioned above.
(b) Showing that, in fact, the above solids *do* exist.

The second task is mechanical, and can be performed by explicitly writing out coordinates. The first is more subtle. Euclid's *Elements* contains an elementary proof of the first part. We will use group theory to show that same.

Definition 58. The following should be familiar from linear algebra.

(a) Given any vectors $v, w \in \mathbb{R}^n$ (written as column vectors), their ***dot product*** is given by $v \cdot w = v^\top w$.
(b) For any $v \in \mathbb{R}^n$, the ***length*** of v is given by $\|v\| = \sqrt{v \cdot v}$. For any $v \in \mathbb{R}^n, \|v\| \geqslant 0$ with equality iff $v = 0$.

The Cauchy-Schwarz Inequality tells us that for any $v, w \in \mathbb{R}^n, |v \cdot w|^2 \leqslant \|v\|^2 \|w\|^2$. Therefore, the following definition makes sense:

(c) For $v, w \neq 0 \in \mathbb{R}^n$, the ***angle*** between v and w is the $\theta = \cos^{-1} \dfrac{v \cdot w}{\|v\| \cdot \|w\|} \in [0, \pi]$.
(d) A linear map $Q \in \mathrm{GL}_n(\mathbb{R})$ is ***orthogonal*** if $QQ^\top = Q^\top Q = I_n$.

An orthogonal linear map preserves dot products because $Qv \cdot Qw = (Qv)^\top Qw = v^\top Q^\top Qw = v^\top Iw = v^\top w$. In particular, it preserves lenghts and angles, so that it is an isometry of $\mathbb{R}^n$ that preserves the origin. From $QQ^\top = I$, it is clear that $|Q|^2 = 1$ so that $|Q| = \pm 1$. An orthogonal linear map preserves orientations if $|Q| = 1$, otherwise it reverses orientations.

Definition 59. Let $n \geqslant 1$.

(a) The ***orthogonal group*** in dimension n over $\mathbb{R}$ is $\mathrm{O}(n, \mathbb{R}) = \{Q \in \mathrm{GL}_n(\mathbb{R}) : QQ^\top = I_n\} \leqslant \mathrm{GL}_n(\mathbb{R})$.
(b) The ***special orthogonal group*** in dimension n over $\mathbb{R}$ is $\mathrm{SO}(n, \mathbb{R}) = \mathrm{O}(n, \mathbb{R}) \cap \mathrm{SL}_n(\mathbb{R})$. It is the group of orientation-preserving isometries, i.e. rotations, of n-dimensional Euclidean space.

Example 71

The group $\mathrm{SO}(2, \mathbb{R}) \cong \mathbb{S}^1$. It is a fun easy exercise to prove that the only finite subgroups of $\mathrm{SO}(2, \mathbb{R})$ are the Z_n for $n \geqslant 1$.

Observe that a unit vector $v \in \mathbb{R}^n$ is an eigenvector of a rotation $Q \in \mathrm{SO}(n, \mathbb{R})$ with eigenvalue 1 iff $Qv = v$. This means that Q fixes v, i.e. v is a unit vector along an ***axis*** of the rotation Q. In this case,

v is called a ***pole*** of Q. Conversely, given a pole of rotation, it is clear that we get an eigenvector with eigenvalue 1.

Example 72
If v is a pole of Q, then so is $-v$.

Example 73
Every unit vector v is a pole of I_n.

To classify all Platonic solids, it suffices to classify all finite subgroups of $\mathrm{SO}(3,\mathbb{R})$. For that we will need the following theorem:

Theorem 22 (Euler Axis Theorem)
If n is odd, then every element of $Q \in \mathrm{SO}(n,\mathbb{R})$ has a pole. In other words, in odd-dimensional Euclidean space, every rotation must have an axis of rotation.

The more traditional language for this, and the case that Euler proved, is "in 3-dimensional space, every displacement of a rigid body that leaves at least one point fixed must be a rotation about some axis passing through the body."

Observe that this is not the case for even dimensions. For instance, the rotation $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SO}(2,\mathbb{R})$ has no axis in $\mathbb{R}^2$. Essentially, any proof of this theorem boils down to the fact that a polynomial of odd degree over the reals must have a real root. We give a proof of which the connection to the above statement needs some thought.

Proof. It suffices to show that for odd n, if $Q \in \mathrm{SO}(n,\mathbb{R})$, then $|Q-I| = 0$. But that is simple:

$$|Q-I| = |(Q-I)^\top| = |Q^\top - I| = |Q^{-1} - Q^{-1}Q| = |Q^{-1}(I-Q)| = |Q^{-1}| \cdot |I-Q| = (-1)^n |Q-I|.$$

■

The particular case we are interested in is $n = 3$.

Corollary 22.1 — Every element of $Q \in \mathrm{SO}(3,\mathbb{R})$ has a pole. Every element of Q is, upto change of basis, given by $\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ for some $\theta \in [0, 2\pi)$. Additionally, every nonidentity element has exactly two antipodal poles.

How is that relevant? Well, here's the promised proof:

Theorem 23
The only finite subgroups of $\mathrm{SO}(3,\mathbb{R})$ are the cyclic groups Z_n, the dihedral groups D_{2n}, $\mathfrak{A}_4, \mathfrak{S}_4$ and $\mathfrak{A}_5$.

Proof. Let $G \leqslant \mathrm{SO}(3,\mathbb{R})$ be finite, and let $|G| = N$. Let $X = \{p \in \mathbb{R}^3 : \|p\| = 1 \wedge \exists g \neq e \in G : gp = p\}$. The key observation is that $G \circlearrowright X$: if p is a pole of g, then hp is a pole of the conjugate hgh^{-1}. For $p \in X$, let $r_p := |G_p|$; by definition, $r_p \geqslant 2$. Consider the incidence correspondence $\Sigma = \{(g,p) \in (G \setminus \{e\}) \times X : gp = p\}$. Consider $\pi_1 : \Sigma \to G \setminus \{e\}$; as we've observed, this map is exactly $2:1$, so that $|\Sigma| = 2N - 2$. But now, by $\pi_2 : \Sigma \to X$, the same is simply $|\Sigma| = \sum_{p \in X}(r_p - 1)$. By the Orbit-Stabilizer Theorem, it is clear that r_p depends only on $\mathcal{O}_p$. Therefore, say $X/G = \{\mathcal{O}_1, \cdots, \mathcal{O}_k\}$; let $n_i := |\mathcal{O}_i|$ and let $r_i := r_p$ for any $p \in \mathcal{O}_i$. Then $|\Sigma| = \sum_{p \in X}(r_p - 1) = \sum_{i=1}^{k} n_i(r_i - 1)$. But, again by

the Orbit-Stabilizer Theorem, $n_i r_i = N$. This means that $2N - 2 = \sum_{i=1}^{k} N\left(1 - \frac{1}{r_i}\right)$, so that

$$2 - \frac{2}{N} = \sum_{i=1}^{k}\left(1 - \frac{1}{r_i}\right).$$

At this point, the problem is almost solved! This equation is the key. Observe that the number on the left is < 2 while each summand on the right is $\geqslant 1/2$, so that this immediately tells us that $k \leqslant 3$. In fact, the case $k = 1$ is not possible either, because the number on the left is $\geqslant 1$, while a single number on the right would be < 1. Therefore, we are left with only two cases:

(a) Case $k = 2$. This gives us $\frac{2}{N} = \frac{1}{r_1} + \frac{1}{r_2}$. But now, $r_1, r_2 \leqslant N$, so the only possibility is $r_1 = r_2 = N$, so that $n_1 = n_2 = 1$. This tells us that there are exactly two poles, each stabilized by the whole group, and they can't be carried over to one another; this means that the entire group is a subgroup of rotations around one fixed axis, which is a copy of $\mathrm{SO}(1, \mathbb{R}) = \mathbb{S}^1$. Any finite subgroup of this is a cyclic group. Therefore, in this case, $G \cong Z_n$ for some $n \geqslant 1$.

(b) Case $k = 3$. This gives us the equation $\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} = 1 + \frac{2}{N}$. WLOG, let $r_1 \leqslant r_2 \leqslant r_3$. Observe that if $r_1 \geqslant 3$, then the LHS $\leqslant 1$, whereas the RHS > 1. Therefore, $r_1 = 2$.

1. Subcase $r_2 = 2$. This implies $N = 2r_3$ so that $n_3 = 2$. Observe that this means that $|\mathcal{O}_3| = 2$; but then the two poles in $\mathcal{O}_3$ must be antipodal to one another, say p and $-p$. Now $|G/G_p| = 2$, so that exactly half of the group of the group fixes p and $-p$, and the other half of the group switches them. This means that G_p must be a subgroup of rotations about one axis. Since $|G_p| = N/2$, this tells us precisely that $G_p \cong Z_{N/2}$. In this case, it is not hard to see that $G \cong D_N$. (What poles are contained in the first two orbits?)
2. Subcase $r_2 \geqslant 3$. In this case, if $r_2 > 3$, then LHS $\leqslant 1$; this means that $r_2 = 3$. This gives us the equation $\frac{1}{r_3} = \frac{1}{6} + \frac{2}{N}$, which tells us that $3 \leqslant r_3 \leqslant 5$. Therefore, three possibilities remain:
   i. $r_i = 2, 3, 3$, $n_i = 6, 4, 4$ and $N = 12$
   ii. $r_i = 2, 3, 4$, $n_i = 12, 8, 6$ and $N = 24$.
   iii. $r_i = 2, 3, 5, n_i = 30, 20, 12$ and $N = 60$.

   These are all the possibilities, and it is not hard to see that in fact all of these possibilities do arise, as simply $\mathfrak{A}_4, \mathfrak{S}_4$ and $\mathfrak{A}_5$ respectively. Observe that the n_i are the number of edges, faces, and vertices OR the number of edges, vertices and faces. These give us the three dual pairs: the tetrahedron, the cube-octahedron pair, and the icosahedron-dodecahedron pair.

■

Of course, the last part of this proof is informal, and does not actually give a construction for these groups inside of $\mathrm{SO}(3, \mathbb{R})$. To complete the proof, one has to prove that these possibilities indeed arise; but that is not hard to do once we write down coordinates for the vertices of each of the above polyhedra, which is left as an exercise for the reader. From the above, we have shown that:

Theorem 24
The only Platonic solids are the ones we already know.

5.4 Groups Acting on Themselves by Conjugation–The Class Equation

One of the most important group actions that tells us a lot about the group structure is conjugation. It is helpful to think of conjugation always as a "change of variables" or "change of labelling."

Definition 60. Any group $G \circlearrowright G$ by $(g, h) \mapsto ghg^{-1}$. In fact, the corresponding permutation representation is the map $\mathrm{conj} : G \to \mathrm{Aut}(G)$.

In general, this action is not faithful, nor is it transitive.

Example 74
The kernel of the above action is $Z(G)$. The orbit of any element of $Z(G)$ consists only of itself.

Definition 61. Let $G \curvearrowright G$ by conjugation.
(a) The orbits of G acting on itself are called ***conjugacy classes***.
(b) The stabilizer of any element $g \in G$ is called the ***centralizer*** of g in G, and is written as either $C_G(g)$ or $Z(g)$.

Observe that the centralizer of any element in the group is the subgroup of elements that commute with it. The following example is clear:

Example 75
For $g \in G, Z(g) = G \Leftrightarrow g \in Z(G)$. Similarly, $Z(G) = \bigcap_{g \in G} Z(g)$.

Proposition 9
Any two elements in the same conjugacy class have the same order.

Proof. Let x, y s.t. $x = gyg^{-1}$. The claim follows because $x^k = (gyg^{-1})^k = gy^k g^{-1}$. ■

Conjugation also induces an action $G \curvearrowright \wp(G)$. Similarly, it induces an action on the set $\mathrm{Sub}(G)$ of all subgroups of G.

Proposition 10
For any $g \in G$ and $H \leqslant G$, $\mathrm{conj}_g : H \to gHg^{-1}$ is an isomorphism. In particular, $H \cong gHg^{-1}$.

Definition 62. Let $G \curvearrowright \wp(G)$ by conjugation.
(a) Two subsets $S, T \in \wp(G)$ are said to be ***conjugate*** if they lie in the same orbit in $\wp(G)/G$, i.e. if $\exists g \in G : gSg^{-1} = T$.
(b) For any $S \subseteq G$, the stabilizer of S under the action $G \curvearrowright \wp(G)$ by conjugation is called the ***normalizer*** of S in G, and is written $N_G(S)$. In other words, $N_G(S) = \{g \in G : gSg^{-1} = S\}$.
(c) Now observe that $N_G(S) \curvearrowright S$ by conjugation. The kernel of this action is called the ***centralizer*** of S in G, and is denoted by $C_G(S)$ or $Z(S)$. In other words, $Z(S) = \{g \in G : \forall s \in S : gsg^{-1} = s\}$. From the definition, it is clear that $Z(S) \leqslant N_G(S)$.

It is clear that for any $H \leqslant G, H \trianglelefteq N_G(H)$, and that $H \trianglelefteq G \Leftrightarrow N_G(H) = G$. Let's get to one of the easy but striking consquences of this definition.

Theorem 25 (The Class Equation)
Let G be a finite group, and let $\mathcal{C}_1, \ldots, \mathcal{C}_r$ be distinct conjugacy classes not contained in the center $Z(G)$. Then

$$|G| = |Z(G)| + \sum_{i=1}^{r} |\mathcal{C}_i|.$$

Further, $\forall i \in [r] : |\mathcal{C}_i| \geqslant 2$ and $|\mathcal{C}_i|$ divides $|G|$.

Proof. The conjugacy classes are precisely the one-element $\{z_j\}$ for $z_j \in Z(G)$ and $\mathcal{C}_1, \cdots, \mathcal{C}_r$. The proof follows from Lemma 9 and the Orbit-Stabilizer Theorem. ■

Note that the class equation is also written sometimes as $|G| = \underbrace{1 + 1 + \cdots + 1}_{Z(G) \text{ times}} + \sum_{i=1}^{r} |\mathcal{C}_i|$.

Example 76

The class equation for an abelian group gives us no new information because $Z(G), r = 0$, and each conjugacy class has size 1.

Example 77

The conjugacy classes in $\mathfrak{S}_3$ are $\{()\}, \{(12), (13), (12)\}$ and $\{(123), (132)\}$. The class equation tells us that $6 = 1 + 3 + 2$.

Example 78

The conjugacy classes in $\mathfrak{A}_4$ are $\{()\}, \{(12)(34), (13)(24), (14)(23)\}, \{(123), (243), (341), (421)\}$ and $\{(132), (234), (314), (412)\}$. The class equation tells us that $12 = 1 + 3 + 4 + 4$. This also shows that the only normal subgroup of $\mathfrak{A}_4$ is $\langle (12)(34), (13)(24) \rangle \cong K_4$.

The key idea behind the applications of the class equation is that if $|G|$ is not divisible by a lot of primes, then we can get a lot of information about the structure of the conjugacy classes by the size restrictions. One example of the above is:

Proposition 11

Let G be a finite p-group, i.e. let $|G| = p^n$ for some prime p and $n \geqslant 1$. Then
(a) The center $Z(G)$ is nontrivial.
(b) G contains an element (equivalently subgroup) of order p.
(c) G contains a subgroup of order p^m for every $0 \leqslant m \leqslant n$.

Proof. (a) By the class equation $p^n = |G| = |Z(G)| + \sum_{i=1}^{r} |\mathcal{C}_i|$. Now $|\mathcal{C}_i| \geqslant 2$ and $|\mathcal{C}_i| \mid p^n$ means that $\forall i : p \mid |\mathcal{C}_i|$. From this, we see that $p \mid p^n - \sum_{i=1}^{r} |\mathcal{C}_i| = |Z(G)|$. But now, $|Z(G)| \geqslant 1$, so that $|Z(G)| \geqslant p$.

(b) Let $g \neq e \in G$. Then $|g| \mid p^n$ so that $|g| = p^j$ for some $1 \leqslant j \leqslant n$. Then $g^{p^{j-1}} \in G$ has order p.

(c) We proceed by induction on n. The case $n = 1$ is easily verified. Suppose $|G|$ is a group with $|G| = p^n$ for some $n \geqslant 2$, and we have verified the proposition for $n - 1$. The case $m = 0$ is trivial, so assume that $1 \leqslant m \leqslant n$. Consider $Z(G)$: by part (a), it is nontrivial, so it is itself a finite p-group; by part (b), it has a subgroup $N \leqslant Z(G)$ of order p. Since $N \leqslant Z(G), N \trianglelefteq G$. The quotient $|G/N|$ has order p^{n-1} and $0 \leqslant m - 1 \leqslant n - 1$, we apply the inductive hypothesis to find a subgroup $\bar{H}$ of G/N of order p^{m-1}. Then the complete preimage H of $\bar{H}$ under the natural projection $\pi : G \to G/N$ is a subgroup of G of order p^m.

■

This already has nice consequences:

Corollary 25.1 — There are only two (isomorphism classes) of groups of order p^2: namely, Z_{p^2} and $Z_p \times Z_p$. In particular, every group of order p^2 is abelian.

Note how this generalizes our verification that the only groups of order 4 are Z_4 and K_4 to arbitrary prime squares.

Proof. Let G be a group with $|G| = p^2$. Since $|Z(G)| \neq 1, |Z(G)| \in \{p, p^2\}$ so that $|G/Z(G)| \in \{p, 1\}$. This means that $G/Z(G)$ is cyclic, and so by a HW problem, trivial. This means that G is abelian. If G has an element of order p^2, then $G \cong Z_{p^2}$. Hence assume that all nonidentity elements of G have order p. Let $x \neq e \in G$ and $y \in G \setminus \langle x \rangle$. Then both x and y have order p, and so $\langle x \rangle \times \langle y \rangle \cong Z_p \times Z_p$. Observe that $\langle x \rangle \cap \langle y \rangle = \{e_G\}$, so that the map $\langle x \rangle \times \langle y \rangle \to G$ given by $(x^a, y^b) \mapsto x^a y^b$ is injective. Since both groups are finite of order p^2, it must be bijective and so an isomorphism. ■

This is an excellent resource for more on conjugacy classes.

5.5 Some More Groups of Small Order

We are now ready to classify all groups of order 8.

Theorem 26

There are five groups of order 8: $Z_8, Z_4 \times Z_2, Z_2^3, D_8$ and Q.

Proof. Let $|G| = 8$. Let $g \in G \setminus \{e\}$; then by Lagrange, $|g| \in \{2, 4, 8\}$. If $|g| = 8$, then $G \cong Z_8$. Hence assume that all elements of G have order 2 or 4. Again, if every element had order 2, then by a HW problem, $G \cong Z_2^3$. If $\exists a \in G$ with $|a| = 4$, then $H = \langle a \rangle \leqslant G$ has index 2, so that $H \trianglelefteq G$. Let $b \in G \setminus H$. Then $bab^{-1} \in H$; further bab^{-1} has the same order as a. This leaves us with two possibilities:

(a) If $bab^{-1} = a$, then $[a, b] = \{e\}$. This means that a and b commute. Then b must have order 2, which gives us $G \cong \langle a, b | a^4, b^2, [a, b] \rangle \cong Z_4 \times Z_2$.

(b) The only case left is $bab^{-1} = a^{-1}$. Now $|b| \in \{2, 4\}$. Accordingly we have the two presentations:

1. If $|b| = 2$, then $G = \langle a, b | a^4, b^2, (ba)^2 \rangle \cong D_8$.
2. If $|b| = 4$, then $G = \langle a, b | a^4, b^4, bab^{-1}a \rangle$. It is easy to see in this case that the map $a \mapsto i, b \mapsto j$ and $ab \mapsto k$ is an isomorphism of G with Q_8.

■

Also, from the preceding section, the only two groups of order 9 are Z_9 and $Z_3 \times Z_3$. Therefore, we have succeed in classifying all groups of order $\leqslant 9$.

5.6 Conjugacy in $\mathfrak{S}_n$ and the Simplicity of $\mathfrak{A}_5$

We now see what it means to be conjugate in $\mathfrak{S}_n$. By applying Claim to a permutation, a bijection $\theta : \Omega \to \Omega$, we get that for any set Ω, $\text{conj}_\theta : \Omega \to \Omega$ is an automorphism. We already knew that! But notice what this tells us: it tells us the conjugation by any element of the symmetric group is just a relabelling of elements. The following makes this precise:

Lemma 11

Let $\sigma, \tau \in \mathfrak{S}_n$, and suppose σ has cycle decomposition

$$(a_1, a_2, \cdots, a_{k_1})(b_1, b_2, \cdots, b_{k_2}) \cdots .$$

Then the conjugate $\tau\sigma\tau^{-1}$ has the cycle decomposition

$$(\tau(a_1), \tau(a_2), \cdots, \tau(a_{k_1})(\tau(b_1), \tau(b_2), \cdots, \tau(b_{k_2}) \cdots .$$

In other words, it is obtained by replacing every entry i in the cycle composition by $\tau(i)$.

Proof. Observe that $\sigma(i) = j \Leftrightarrow \tau\sigma\tau^{-1}(\tau(i)) = \tau(j)$. ■

The notion of the above invariant cycle structure is made precise by the following definition:

Definition 63. If $\sigma \in \mathfrak{S}_n$ is the product of disjoint cycles (inlcluding length 1) of length $n_1, \cdots, n_r$, then this multiset $\lambda = \{n_1, \cdots, n_r\}$ is called the ***cycle type*** of σ.

Then the above lemma can be summarized by saying any two elements in the same conjugacy class have the same cycle type. Conversely, it is clear that any two elements having the same cycle type are in the same class. Since each cycle type λ is a partition of n, i.e. $\lambda \dashv n$, we get the following really beautiful theorem:

Theorem 27

The conjugacy classes in $\mathfrak{S}_n$ are in bijection with the partitions of n. In particular, the number of conjugacy classes in $\mathfrak{S}_n$ is $p(n)$, the number of partitions of n.

Example 79

If $\sigma_1 = (1)(35)(2476)$ and $\sigma_2 = (2)(57)(3164)$, then we choosing $\tau = (1235764)$ gives us $\tau\sigma_1\tau^{-1} = \sigma_2$.

Example 80

For $n = 4$, the following gives a table for the partitions of 4 and a representative of each conjugacy class having the given cycle type:

$\lambda \dashv 4$	Representative of Conjugacy Class
$\{1,1,1,1\}$	$()$
$\{1,1,2\}$	(12)
$\{1,3\}$	(123)
$\{2,2\}$	$(12)(34)$
$\{4\}$	(1234)

This tells us that the class equation for $\mathfrak{S}_4$ is $24 = 1 + 6 + 8 + 3 + 6$, something we already knew.

Suppose the partition λ consists of b_1 1's, b_2 2's, etc., so that $n = b_1 + 2b_2 + \cdots + nb_n$. Then what is the size of the conjugacy class corresponding to this partition? Well, the number of ways of dividing a set of n elements into labelled subsets of sizes $c_1, \cdots, c_k$ is the multinomial coefficient

$$\binom{n}{c_1; \cdots; c_k} = \frac{n!}{c_1! c_2! \cdots c_k!}.$$

If the subsets are unlabelled, you also have to divide by the number of ways to permute the subsets of the same size. This gives you:

$$\frac{n!}{(1!)^{b_1}(2!)^{b_2}\cdots(n!)^{b_n}} \cdot \frac{1}{b_1! b_2! \cdots b_n!} = \frac{n!}{\prod_{i=1}^n (i!)^{b_i}(b_i!)}.$$

Now within each selected cycle of length i, we can cyclically permute the elements inside in $(i-1)!$ ways, so that gives a factor of $(i-1)!^{b_i}$ in the numerator. The end result is that the conjugacy class has size:

$$\frac{n!}{\prod_{i=1}^n (i!)^{b_i}(b_i!)} \cdot \prod_{i=1}^n (i-1)!^{b_i} = \frac{n!}{\prod_{i=1}^n i^{b_i} \cdot b_i!}.$$

Therefore, the class equation for $\mathfrak{S}_n$ in general reads:

$$n! = \sum_b \frac{n!}{\prod_{i=1}^n i^{b_i} \cdot b_i!},$$

where the sum runs over all tuples $b = (b_1, \cdots, b_n)$ of nonnegative integers s.t. $\sum_{i=1}^n i b_i = n$. This gives a nice way to index partitions that will be useful later: for $b = (b_1, \cdots, b_n)$ a tuple of nonnegative integers s.t. $\sum_{i=1}^n i b_i = n$, denote the partition $\{\underbrace{1, \cdots, 1}_{b_1 \text{ times}}, \ldots, \underbrace{n, \cdots, n}_{b_n \text{ times}}\}$ by $\lambda_b \dashv n$ and the conjugacy class in $\mathfrak{S}_n$ by $\mathcal{C}_b$.

It would be tempting to assume that the same statement (i.e. conjugate iff same cycle type) is true for $\mathfrak{A}_n$ too; however, that is NOT the case. The reason for that is that there may be elements of $\mathfrak{A}_n$ having the same cycle type may be related by conjugation by elements of $\mathfrak{S}_n \setminus \mathfrak{A}_n$ but not $\mathfrak{A}_n$; in that case, said elements may not be conjugate in $\mathfrak{A}_n$. This does indeed happen. The following section is devoted to making sense of when exactly that happens. First we prove a general fact:

Lemma 12

If group $G \curvearrowright X$ transitively, and $H \leqslant G$ with $|G/H| = 2$, then the induced action $H \curvearrowright X$ has either one or two orbits.

Proof. First note the induced action corresponds the permutation representation $\rho \circ \iota : H \to \mathfrak{S}_X$ where $\iota : H \hookrightarrow G$ and $\rho : G \to \mathfrak{S}_X$. Now let $a \in G \setminus H$; then $G = H \cup Ha$. This means that for fixed $x \in X : S = Gx = Hx \cup Hax$. If $Hx = Hax$, then $X = Hx$ so that there is only on orbit; this happens iff $ex \in Hax \Leftrightarrow \exists b \notin H : bx = x$. If this is not the case, then there are two orbits, that of x and that of ax. ■

From the above proof, it is clear that there is one orbit iff $\exists x \in X : \exists b \notin H : bx = x \Leftrightarrow \forall x \in X : \exists b \notin H : bx = x$. One direction of the iff is clear; for the other direction, let $x \in X$ and $b \notin H$ be given s.t. $bx = x$, and let x' be any other element. By transitivity of G, $\exists g \in G : x' = gx$. Then $gbg^{-1} \notin H : gbg^{-1}x' = x'$.

Proposition 12

If $\mathcal{C} \subseteq \mathfrak{S}_n$ is a conjugacy class, then there are three possibilities:

(a) $\mathcal{C} \cap \mathfrak{A}_n = \varnothing$. This happens iff $\mathcal{C}$ contains odd permutations.

(b) $\mathcal{C} \subseteq \mathfrak{A}_n$ is a conjugacy class in $\mathfrak{A}_n$. This happens iff some element of $\mathcal{C}$ commutes with an odd permutation.

(c) $\mathcal{C} \subseteq \mathfrak{A}_n$ is the union of two conjugacy classes of $\mathfrak{A}_n$. This happens otherwise, and in this case both the conjugacy classes that $\mathcal{C}$ breaks into have the same size.

Proof. Apply the above lemma to $G = \mathfrak{S}_n, H = \mathfrak{A}_n, X = \mathcal{C}$. The conjugacy class remains stable iff $\exists \sigma \in \mathcal{C}, \tau \in \mathfrak{S}_n \setminus \mathfrak{A}_n$ s.t. $\tau\sigma\tau^{-1} = \sigma$. For the last claim, assume that $\sigma \in \mathcal{C}$ and $\tau \in \mathfrak{S}_n \setminus \mathfrak{A}_n$ then the two orbits are the orbits of σ and $\tau\sigma\tau^{-1}$; but now conj_τ is an bijection between them, so that they have the same size. ■

The following gives a handy criterion of determining whether the elements of a conjugacy class commute with odd permutations:

Proposition 13

Given a tuple $b = (b_1, \cdots, b_n)$, the conjugacy class $\mathcal{C}_b$ remains stable iff EITHER it contains a cycle of even length OR it contains two cycles of the same length. Equivalently, a conjugacy class splits iff $\forall i : b_{2i} = 0$ and $b_{2i+1} \leqslant 1$.

Proof. If $\sigma \in \mathcal{C}_b$ contains a cycle c of even length, then σ commutes with $c \in \mathfrak{S}_n \setminus \mathfrak{A}_n$. If it contains two distinct cycles $(a_1 a_2 \cdots a_\ell)$ and $(b_1 b_2 \cdots b_\ell)$ of odd lenght ℓ, then it commutes with $(a_1 b_1)(a_2 b_2) \cdots (a_\ell b_\ell) \in \mathfrak{S}_n \setminus \mathfrak{A}_n$. Conversely, assume that $\sigma = c_1 c_2 \cdots c_s$ is a product of distinct odd cycles. Suppose $\tau \in \mathfrak{S}_n$ commutes with σ; then τ must fix each of the cycles. This means that τ has the form $\tau = c_1^{a_1} c_2^{a_2} \cdots c_s^{a_s}$ for $a_i \in \mathbb{Z}$; this shows that $\tau \in \mathfrak{A}_n$. ■

Example 81

The following gives the structure of conjugacy classes for $n = 5$.

b	λ	Representative of $\mathcal{C}_b$	$\mathfrak{S}_5$	$\mathfrak{A}_5$
$(5,0,0,0,0)$	$\{1,1,1,1,1\}$	$()$	1	1
$(3,1,0,0,0)$	$\{1,1,1,2\}$	(12)	10	–
$(2,0,1,0,0)$	$\{1,1,3\}$	(123)	20	20
$(1,0,0,1,0)$	$\{1,4\}$	(1234)	30	–
$(0,0,0,0,1)$	$\{5\}$	(12345)	24	$12+12$
$(1,2,0,0,0)$	$\{1,2,2\}$	$(12)(34)$	15	15
$(0,1,1,0,0)$	$\{2,3\}$	$(12)(345)$	20	–

From this, we are ready to prove what we wanted.

Theorem 28

$\mathfrak{A}_5$ is simple.

Proof. From the above example, the class equation for $\mathfrak{A}_5$ is $60 = 1 + 20 + 12 + 12 + 15$. Now observe that any $N \trianglelefteq \mathfrak{A}_5$ must be a union of conjugacy classes that also happens to be a subgroup, but in the above sum, there is no subsum containing 1 that is a divisor of 60. ■

One proof of the general case $n \geqslant 5$ (that can be found in DF) proceeds by induction, and the above theorem is crucial to establishing the base case.

5.7 Sylow Theorems

The Sylow Theorems, named after Norwegian mathematician Peter Ludwig Sylow who proved them in 1873, provide a huge deal of information about the structure of subgroups of a given finite group. Let's first go ahead and state them.

Theorem 29 (Sylow Theorems)

Let G be a finite group and p be a prime.

1. If $|G| = p^e m$ for some $p \nmid m$. Then G contains a subgroup of order p^e. Such a subgroup is called a Sylow p-subgroup of G, and the set of all Sylow p-subgroups of G is denoted by $\mathrm{Syl}_p(G)$.
2. If $K \in \mathrm{Syl}_p(G)$ and H is any p-subgroup of G, then H is contained in some conjugate of K, i.e. $\exists g \in G : H \subseteq gKg^{-1}$. In particular, any two Sylow p-subgroups of G are conjugate in G.
3. If $n_p := |\mathrm{Syl}_p(G)|$, then $n_p \mid m$ and $n_p \equiv 1 \pmod{p}$.

The whole power of the Sylow Theorems rests on the fact that the first one tells us that we can find a Sylow p-subgroup K of G, the second one tells us that $K \trianglelefteq G$ iff $n_p = 1$, and the third one imposes powerful number-theoretic restrictions on n_p. These restrictions sometimes allow us to conclude that $n_p = 1$, and then lo! we have found a normal subgroup of G. Otherwise, we may try to interplay the different Sylow p-subgroups against each other for values primes p, and sometimes that provides very valuable information too.

Example 82

If $p \nmid |G|$, then $\mathrm{Syl}_p(G) = \{\{e\}\}$, and all the statements are trivially true. If $|G| = p^e$, then $\mathrm{Syl}_p(G) = \{G\}$ and all the statements are trivially true.

Example 83

Let's verify the theorems for $G = \mathfrak{A}_4$ and $p = 2$.

1. $|G| = 12 = 2^2 3$ so that $e = 2, m = 3$. We know that $\mathfrak{A}_4$ has a unique subgroup of order 4, namely $H = \langle (12)(34), (13)(24) \rangle \cong K_4$. Therefore, $\mathrm{Syl}_2(\mathfrak{A}_4) = \{H\}$.
2. It is clear that any 2-subgroup of $\mathfrak{A}_4$ is contained in a conjugate (in fact, the conjugate by e) of the above subgroup H. In fact, $H \trianglelefteq \mathfrak{A}_4$.
3. Here $n_2 = 1$, $n_2 \mid 3$ and $n_2 \equiv 1 \pmod{2}$.

How about for $G = \mathfrak{A}_4$ and $p = 3$?

1. $12 = 3 \cdot 4$ so that $e = 1, m = 4$. In this case, $\mathrm{Syl}_3(\mathfrak{A}_4) = \{\langle (123) \rangle, \langle (234) \rangle, \langle (341) \rangle, \langle (412) \rangle\}$.
2. Since $e = 1$, there are no nontrivial $3-$subgroups other than the Sylow 3-subgroups.
3. Here $n_2 = 4, n_2 \mid 4$ and $n_2 \equiv 1 \pmod{3}$.

Example 84

Suppose $G = \mathfrak{S}_3$ and $p = 3$. Then $n_3 \mid 2 \wedge n_3 \equiv 1 \pmod{3} \Rightarrow n_3 = 1$; therefore, there is a unique Sylow 3-subgroup of $\mathfrak{S}_3$, which must be normal in $\mathfrak{S}_3$. Therefore, $\mathrm{Syl}_3(\mathfrak{S}_3) = \{\mathfrak{A}_3\}$.

Example 85

Let's see if starting with the Sylow Theorems, we can get some information about group structure. Suppose G is a group of order 15. Then $n_3|5 \wedge n_3 \equiv 1 \pmod 3 \Rightarrow n_3 = 1$. Similarly, $n_5|3$ and $n_5 \equiv 1 \pmod 5 \Rightarrow n_5 = 1$. Therefore, $\exists! H \in \mathrm{Syl}_3(G)$ and $\exists! K \in \mathrm{Syl}_5(G)$; we will see in the next section on recognizing direct products that this is sufficient to conclude that $G \cong H \times K \cong Z_{15}$.

The proof of the Sylow Theorems needs two key lemmas:

Lemma 13

If $n = p^e m$ for some $p \nmid m$, then $p \nmid \binom{n}{p^e}$.

Proof. The case $e = 0$ is clear. For $e \geqslant 1$, note that

$$\binom{n}{p^e} = \prod_{k=0}^{p^e-1} \frac{n-k}{p^e-k}.$$

For $k = 0, n/p^e = m$, which is not divisible by p. For $1 \leqslant k \leqslant p^e - 1$, if $k = p^f l$ for $p \nmid f$, then $f < e$; so that $n - k = p^f(p^{e-f}m - l)$ and $p^e - k = p^f(p^{e-f} - l)$ and since $e - f \geqslant 1$, $p \nmid p^{e-f}m - l, p^{e-f} - l$; therefore each term in product is indivisible by p. ■

Lemma 14

Note that $G \curvearrowright \wp(G)$ by left-multiplication. If $U \subseteq G$ and $H = \mathrm{stab}(U)$, then $|H|$ divides $|U|$.

Proof. Note that H acts faithfully on U by left-multiplication, and for $h \in H, U^h = \begin{cases} U, & \text{if } h = e, \\ \varnothing, & \text{otherwise.} \end{cases}$ This means that $U = \coprod_{\mathcal{O} \in U/H} \mathcal{O}$ where for each orbit, $|\mathcal{O}| = |H|$. Therefore, $|U| = |U/H| \cdot |H|$. ■

We are now in a position to prove the Sylow Theorems.

Main Proof. If $e = 0$, then all theorems are trivial. Hence assume $e \geqslant 1$.

1. Let $\mathcal{S} = \{U \subseteq G : |U| = p^e\} \subseteq \wp(G)$ be the set of subsets of G of order p^e. Then $G \curvearrowright \mathcal{S}$ by left-multiplication. Now by Lemmas 9 and 13, $p \nmid \binom{n}{p^e} = \sum_{\mathcal{O} \in \mathcal{S}/G} |\mathcal{O}|$; therefore, $\exists U \in \mathcal{S}$ s.t. $p \nmid |\mathcal{O}_U|$. Let $H = \mathrm{stab}(U)$; by Lemma 14, $|H| \mid |U| = p^e$. By the Orbit-Stabilizer Theorem, $p^e \mid p^e m = |G| = |H| \cdot |\mathcal{O}_U|$, but since $p \nmid |\mathcal{O}_U|$, this means that $p^e \mid |H|$. This shows that $|H| = p^e$.
2. If $H = \{e\}$, the claim is trivial, hence assume $|H| \geqslant p$. Note that $G \curvearrowright G/K$ by left-multiplication and $\mathrm{stab}(eK) = K$. Restricting the action to $H \subseteq G \curvearrowright G/K$ gives us an action of a p-group H on a set G/K with $p \nmid |G/K|$; by HW5 Q3, $\exists gK \in G/K$ s.t. all of H fixes gK, i.e. $H \subseteq \mathrm{stab}(gK) = g\,\mathrm{stab}(eK)g^{-1} = gKg^{-1}$.
3. Theorem 2 tells us that $G \curvearrowright \mathrm{Syl}_p(G)$ transitively by conjugation. For fixed $K \in \mathrm{Syl}_p(G), \mathrm{stab}(K) = N_G(K) \geqslant K$. so that $p^e = |K| \mid N_G(K)$. By the Orbit-Stabilizer Theorem, $p^e m = |G| = |\mathcal{O}_K| \cdot |\mathrm{stab}(K)| = n_p \cdot |N_G(K)|$, so that $n_p \mid m$.
   Next, $K \curvearrowright \mathrm{Syl}_p(G)$ by conjugation; first we show that $\mathrm{Syl}_p(G)^K = \bigcap_{k \in K} \mathrm{Syl}_p(G)^k = \{K\}$: suppose that $H \in \mathrm{Syl}_p(G)$ is fixed under conjugation by every element of K, i.e. $K \leqslant N_G(H)$. But then $H, K \in \mathrm{Syl}_p(N_G(H))$ with $H \trianglelefteq N_G(H)$ so that Theorem 2 tells us that $H = K$. Finally, Lemma 9 tells us that
   $$n_p = |\mathrm{Syl}_p(G)| = \sum_{\mathcal{O} \in \mathrm{Syl}_p(G)/K} |\mathcal{O}| = 1 + \sum_{\substack{\mathcal{O} \in \mathrm{Syl}_p(G)/K \\ \mathcal{O} \neq \{K\}}} |\mathcal{O}|,$$
   where for the orbits $\mathcal{O}$ other than $\{K\}$, $2 \leqslant |\mathcal{O}| \nmid |K| = p^e$, so that $p \mid |\mathcal{O}|$. Therefore, reducing mod p, we get the $n_p \equiv 1 \pmod p$.

■

5.8 Recognizing Direct and Semidirect Products

Now that we have Sylow Theorems, we need some ways to recognize the structure of a group from the structure of its subgroups. The following theorem is a step towards that:

Theorem 30 (Recognizing Direct Products)
Suppose G is a finite group. Suppose we know subgroups $H, K \leqslant G$ s.t.
(a) $H, K \trianglelefteq G$.
(b) $H \cap K = \{e\}$.
(c) $|H| \times |K| = |G|$.
Then $G \cong H \times K$. Note in particular that the condition (b) can be replaced by $(|H|, |K|) = 1$.

Proof. Let $h \in H, k \in K$. Consider the commutator $[h,k] = hkh^{-1}k^{-1}$. Now $H \trianglelefteq G \Rightarrow [h,k] = h(kh^{-1}k^{-1}) \in H$, and $K \trianglelefteq G \Rightarrow [h,k] = (hkh^{-1})k^{-1} \in K$, so that $[h,k] \in H \cap K = \{e\}$. Therefore, every element of H commutes with every element of K. Consider the map $\varphi : H \times K \to G$ by $(h,k) \mapsto hk$. A priori this is only a set map, but because of the observation above, $\varphi((h_1,k_1)\cdot(h_2,k_2)) = h_1(h_2k_1)k_2 = h_1(k_1h_2)k_2 = \varphi(h_1,k_1)\cdot\varphi(h_2,k_2)$, so that this is in fact a homomorphism. By hypothesis (b), this map is injective; since G and $H \times K$ are finite sets of the same cardinality, this is the required isomorphism. The last statement is true because if $g \in H \cap K$, then $|g| \mid |H| \wedge |g| \mid |K| \Rightarrow |g| \mid (|H|, |K|) = 1$. ∎

Example 86
This completes the proof started in Example 85 that there is only one group of order 15.

Let's look at another example that will help us get intuition for what comes next.

Example 87
Suppose G is a group of order 21. Then $n_3 \mid 7 \wedge n_3 \equiv 1 \pmod 3 \Rightarrow n_3 \in \{1, 7\}$, and $n_7 \mid 3 \wedge n_7 \equiv 1 \pmod 7 \Rightarrow n_7 = 1$.
(a) If $n_3 = 1$, then as before, $G \cong Z_3 \times Z_7$.
(b) The case $n_3 = 7$ is trickier. Note that in this case there are 7 Sylow 3-subgroups (all of which be isomorphic to Z_3) and a unique Sylow 7-subgroup, say H, which is normal and isomorphic to Z_7. Observe that any of the Sylow 3-subgroups must intersect only in the identity, and any of the Sylow 3-subgroups interesects H only in the identity, so that tells us that these must contain exactly $2 \times 7 + 6 \times 1 = 20$ nonidentity elements; so these must be all.
Now let $H = \langle a \rangle$ and $b \in G \setminus H$; by the above discussion, $|b| = 3$. Since $H \trianglelefteq G, bab^{-1} \in H$, say $bab^{-1} = a^j$. Then $a = b^3ab^{-3} = a^{j^3}$ so that $j^3 \equiv 1 \pmod 7$ so that $j \in \{1, 2, 4\}$.

1. Suppose $j = 1$; then a and b commute. This tells us that $G \cong \langle a, b | a^7, b^3, [a,b] \rangle \cong Z_3 \times Z_7$ for which $n_3 = 1$, which is not possible.
2. Suppose $j = 4$; then by choosing $c = b^2$ instead of b would give us that $cac^{-1} = a^{4^2} = a^2$. This means that WLOG we can assume that $j = 2$. This gives us the group presentation $\langle a, b | a^7 = b^3 = e, bab^{-1} = a^2 \rangle$.

It is not hard to see that the subgroup of $\mathfrak{S}_7$ generated by $a = (1234567)$ and $y = (235)(476)$ has exactly this presentation, so that in fact, there is a unique nonabelian group of order 21.

An exactly analogous argument can be used to show that we are essentially done classifying groups of order pq. We record the statement here, and the proof is left as an exercise to the reader. (It can also be found in DF.)

Theorem 31 (Classification of Groups of Order pq)

Let G be a group with $|G| = pq$ for primes $p < q$. Then:
(a) If $q \not\equiv 1 \pmod{p}$, then G can only be $Z_p \times Z_q$
(b) If $q \equiv 1 \pmod{p}$, then either $G \cong Z_p \times Z_q$ OR G is the unique nonabelian group of order pq. Further, G can be realized a subgroup of $\mathfrak{S}_q$.

Essentially, what happened in the above example is that we found a subgroup $H \trianglelefteq G$, and a subgroup $K \leqslant G$ not necessarily normal, but with an action of $K \curvearrowright H$ by conjugation. This leads us naturally to the definition of semidirect products.

Proposition 14

Suppose H and K are groups, and let $\varphi : K \to \mathrm{Aut}(H)$ be any homomorphism. Let $G = \{(h, k) : h \in H, k \in K\}$ with the following law of composition:

$$(h_1, k_1) \cdot (h_2, k_2) = (h_1 \, \varphi_{k_1}(h_2), k_1 k_2).$$

Then this multiplication makes G into a group of order $|H| \times |K|$, and the subgroups $\{(h, e_K)\}$ and $\{(e_H, k)\}$ are copies of H and K as subgroups of G. Further,
(a) $H \trianglelefteq G$.
(b) $H \cap K = \{e\}$.
(c) $\forall h \in H, k \in K : khk^{-1} = \varphi_k(h)$.

Proof Sketch. It is straightforward to verify that the given construction makes a group with identity (e_G, e_H) and inverse $(h, k)^{-1} = (\varphi_{k^{-1}}(h^{-1}), k^{-1})$. The other propositions are similarly verified. The reader is encouraged to fill in the details. ■

Definition 64. In the notation of the above proposition, G is called the ***semidirect product*** of H and K w.r.t φ, and is denoted by $H \rtimes_\varphi K$. (If there is no danger of confusion, or if all such semidirect products for nontrivial φ are isomorphic, then we denote it simply by $H \rtimes K$.)

The notation is there to remind us that $H \trianglelefteq H \rtimes K$, but it is not necessarily true that $K \trianglelefteq H \rtimes K$.

Example 88

Taking φ to be the null map recovers the direct product, i.e. $H \rtimes_e K = H \times K$, and in this case $K \trianglelefteq H \rtimes_e K$ too.

Example 89

Taking $H = Z_n = \langle r \rangle$, $K = Z_2 = \langle s \rangle$ and $\varphi : K \to \mathrm{Aut}(H)$ by $\varphi_s = \mathrm{inv}$ recovers the dihedral group, i.e. $D_{2n} \cong Z_n \rtimes_\varphi Z_2$. In general, if H is any abelian group and $K = Z_2 = \langle s \rangle$, the $\varphi : K \to \mathrm{Aut}(H)$ by $\varphi_s = \mathrm{inv}$ creates a semidirect product $H \rtimes Z_2$. The group $Z_\infty \rtimes_\varphi Z_2$ is sometimes denoted by D_∞.

Example 90

Taking $H = Z_7 = \langle a \rangle, K = Z_3 = \langle b \rangle$ and $\varphi : K \to \mathrm{Aut}(H)$ by $\varphi_b(a) = a^2$ recovers the (unique) nonabelian group of order 21, i.e. it is $Z_7 \rtimes Z_3$. Similarly, the (unique) nonabelian group of order pq for $p \mid q-1$ can be denoted by $Z_q \rtimes Z_p$. (Uniqueness can be shown by realizing that homomorphisms $Z_p \to \mathrm{Aut}(Z_q)$ are equivalent in an appropriate sense.)

As before, we have a recognition theorem.

Theorem 32 (Recognizing Semidirect Products)

Suppose G is a finite group with subgroups H and K such that:
(a) $H \trianglelefteq G$.
(b) $H \cap K = \{e\}$.
(c) $|H| \times |K| = |G|$.
Let $\varphi : K \to \mathrm{Aut}(H)$ be the homomorphism $k \mapsto \mathrm{conj}_k$. Then $G \cong H \rtimes_\varphi K$.

Proof. As before, consider the map $\Phi : H \rtimes_\varphi K \to G$ given by $(h, k) \mapsto hk$. This is a homomorphism because:

$$\Phi((h_1, k_1)\cdot(h_2, k_2)) = \Phi(h_1\varphi_{k_1}(h_2), k_1k_2) = h_1\varphi_{k_1}(h_2)k_1k_2 = h_1k_1h_2k_1^{-1}k_1k_2 = h_1k_1h_2k_2 = \Phi(h_1, k_1)\cdot\Phi(h_2, k_2).$$

By condition (b), this map is injective, so that by condition (c) it is an isomorphism. ■

5.9 A Few More Groups of Small Order (Final)

Let's now use the above theory to classify some more groups!

Theorem 33 (Classification of Groups of order $2p$)

Let G be a group with $|G| = 2p$ for prime p. Then either $G \cong Z_{2p}$ or $G \cong D_{2p}$.

Note that this is a special case of Theorem 31, but since we didn't prove that, let's at least prove this.

Proof. We've already done the case $p = 2$; hence assume that $p \geqslant 3$. Observe that $n_p \equiv 1 \pmod p \wedge n_p \mid 2 \Rightarrow n_p = 1$. Let H be the unique Sylow p-subgroup of G, and let K be any Sylow 2-subgroup. Then H and K satisfy the conditions of Theorem 32, so this tells us that $G \cong Z_p \rtimes_\varphi Z_2$ for some homomorphism $\varphi : Z_2 \to \mathrm{Aut}(Z_p)$. If $Z_2 = \langle s\rangle$ and $Z_p = \langle r\rangle$, then $\varphi : s \mapsto (r \mapsto r^j)$ for some $1 \leqslant j \leqslant p-1$. For φ_s to have order 2, we must have $r^{j^2} = \varphi_s^2(r) = \mathrm{id}_H(r) = r$ so that $j^2 \equiv 1 \pmod p \Rightarrow j \equiv \pm 1 \pmod p$. If $j = 1$, we recover $G \cong Z_p \times Z_2 \cong Z_{2p}$; if $j = -1$ we recover $G \cong D_{2p}$. ■

This means that we are done classifying groups of order $\leqslant 11$. Let's now do 12.

Theorem 34

There are five groups of order 12: these are $Z_4 \times Z_3, K_4 \times Z_3, \mathfrak{A}_4, D_{12}$ and a unique nontrivial $Z_3 \rtimes Z_4$.

Proof. Let G be a group s.t. $|G| = 12$. Then $n_2|3 \wedge n_2 \equiv 1 \pmod 2 \Rightarrow n_2 \in \{1, 3\}$, and $n_3|4 \wedge n_3 \equiv 1 \pmod 3 \Rightarrow n_3 \in \{1, 4\}$.

(a) If $n_2 = n_3 = 1$, let H be the unique Sylow 2-subgroup and K be the unique Sylow 3-subgroup. Then by Theorem 30, $G \cong H \times K$. Since we have two choices for H, namely $H = Z_4$ and $H = K_4$, and one choice for K, namely $K = Z_3$, this gives us the two (abelian) possibilities $G \cong Z_4 \times Z_3$ and $G \cong K_4 \times Z_3 \cong Z_2 \times Z_2 \times Z_3 \cong Z_2 \times Z_6$.

(b) Now suppose that $n_2 = 1$ and $n_3 = 4$. Let $\mathrm{Syl}_3(G) = \{K_1, \cdots, K_4\}$. By Sylow Theorem 2, $G \curvearrowright \mathrm{Syl}_3(G)$ transitively by conjugation, and this gives a permutation representation $\rho : G \to \mathfrak{S}_4$. We show that G maps isomorphically to $\mathfrak{A}_4$. Now for $i \in [4]$, by the Orbit-Stabilizer Theorem, the stabilizer $\mathrm{stab}(K_i) = N_G(K_i)$ has order 3, but it contains at least K_i, so that $N_G(K_i) = K_i$. Therefore, $\ker\rho = \bigcap_{i=1}^4 \mathrm{stab}(K_i) = \{e\}$, so that the action is faithful, and $\rho : G \hookrightarrow \mathfrak{S}_4$. Then $\mathrm{im}\,\rho \trianglelefteq \mathfrak{S}_4$ is a subgroup of order 12 and index 2; but from the class equation of $\mathfrak{S}_4$, which is $24 = 1 + 6 + 8 + 3 + 6$, the only subgroup of $\mathfrak{S}_4$ of order 12 is $\mathfrak{A}_4$. Therefore, in this case, $G \cong \mathfrak{A}_4$.

(c) We show that the case $n_2 = 3, n_3 = 4$ is not possible. Suppose that there were 4 Sylow 3-subgroups, then their pairwise intersections must be trivial, this gives us $4 \times 2 = 8$ nonidentity elements, leaving the identity and three other nonidentity elements, which must then form the unique Sylow 2-subgroup (because the intersection of any Slyow 2-subgroup and Sylow 3-subgroup must be trivial.)

(d) The only case left is $n_2 = 3, n_3 = 1$. In this case, if H is the unique Sylow 3-subgroup (so $H \cong Z_3$), and K is any Sylow 2-subgroup, then Theorem 32 tells us that $G \cong Z_3 \rtimes_\varphi K$ for some homomorphism $\varphi : K \to \mathrm{Aut}(Z_3)$. Observe that we've seen before that $\mathrm{Aut}(Z_3) \cong Z_2$: if $Z_3 = \{e, b, b^2\}$, then the only nonidentity automorphism of Z_3 swaps b and b^2. There are only two possibilities for K: $K = Z_4$ and $K = K_4$. Let's treat these separately.

1. If $K = Z_4 = \langle a \rangle$, then we want to classify all homomorphisms $Z_4 \to Z_2$; but it is clear that there only two such homomorphisms, the null map corresponding to $a \mapsto e$ and the map that sends a to the nonidentity element. The null map gives us the direct product $Z_3 \times Z_4$ for which $n_2 = 1$, so that is not possible. The only other map left is the map $Z_4 \to \mathrm{Aut}(Z_3)$ given by $\varphi_a = \mathrm{inv}$. This gives us the unique nontrivial semidirect product $Z_3 \rtimes Z_4$ with presentation $\langle a, b | a^4 = b^3 = e, aba^{-1} = b^{-1} \rangle$. (We still need to show that there is no collapsing and that such a group exists, but it is easy to see that $a = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} e^{2\pi\imath/3} & 0 \\ 0 & e^{4\pi\imath/3} \end{pmatrix}$ generate a subgroup of $\mathrm{SL}_2(\mathbb{C})$ with exactly these relations, so that such a group indeed exists.)
2. Suppose $K = K_4$, and we want to classify all nontrivial maps $K \to \mathrm{Aut}(Z_3)$. Observe that in this case, the stabilizer of b for the action $K \curvearrowright \{b, b^2\}$ must have size 2, so that there is $u \neq e \in K : ubu = b$, and also $v \in K : vbv = b^2$. Since K is abelian, $uv = vu$, so this gives a presentation $Z_3 \rtimes K_4 = \langle u, v, b | u^2 = v^2 = b^3 = e, uv = vu, ub = bu, vb = b^2 v \rangle$. This is sufficient to determine the group table completely, so there is at most one such isomorphism class. But the group D_{12} is missing from our discussion so far (i.e. it's not isomorphic to any of the ones we've found above), and we know it's a group of order 12, so that this must be it. (Exercise: find out the isomorphism explicitly.)

■

With this, we are done classifying groups of order $\leqslant 15$. The following table summarizes our findings: for $n \in \mathbb{N}$, let $g(n)$ denote the number of (isomorphism classes of) groups of order n.

n	$g(n)$	Classes
1	1	Z_1
2	1	Z_2
3	1	Z_3
4	2	Z_4, Z_2^2
5	1	Z_5
6	2	$Z_6, \mathfrak{S}_3$
7	1	Z_7
8	5	$Z_8, Z_4 \times Z_2, Z_2^3, D_8, Q$
9	2	Z_9, Z_3^2
10	2	Z_{10}, D_{10}
11	1	Z_{11}
12	5	$Z_{12}, K_4 \times Z_3, \mathfrak{A}_4, D_{12}, Z_3 \rtimes Z_4$
13	1	Z_{13}
14	2	Z_{14}, D_{14}
15	1	Z_{15}

Beyond this, there are already 14 groups of order 16. Needless to say, classifying them is beyond the scope of this course. In general, as you've observed, it becomes harder and harder to classify groups as their sizes increase. Nonetheless, it is a tremendous feat of modern mathematics that we have classified all finite *simple* groups, and this classification was completed in 2012, with the help of powerful computer machinery.

To conclude, we mention only a result that (probably) confirms your intuition about some *abelian* groups.

Theorem 35 (Fundamental Theorem of Finitely Generated Abelian Groups)

If G is a finitely generated abelian group, then G can be written as $G = \mathbb{Z}^r \times \prod_{i=1}^{k} (\mathbb{Z}/p_i^{a_i}\mathbb{Z})$ for some primes p_i and exponents $a_i \geqslant 1$. Moreover, this decomposition into ***elementary divisors*** is unique upto the order of factors. The part $\mathbb{Z}^r$ is called the ***free part*** of G, r is called the ***rank*** of G, and the remaining part is called the ***torsion subgroup*** of G, and is denoted by G_{tors}.

(Note that there is another form of this theorem that deals with the ***invariant factor decomposition***.) In particular, G is finite $\Leftrightarrow G = G_{\text{tors}} \Leftrightarrow r = 0$. From this, it is clear that the isomorphism classes of abelian groups of order $n = \prod_{i=1}^{k} p_i^{e_i}$ and in bijective correspondence with the cartesian product of all partitions of the e_i's. In particular, there are $\prod_{i=1}^{k} p(e_i)$ such groups, where p denotes the partition function.

The proof of this theorem is (just) beyond the reach of this course, and is usually part of the content of a second course in abstract algebra. That concludes our journey of group classifications.

6 A First Encounter with Category Theory

Now that we've talked a lot about groups, let's briefly talk about some other structures from abstract algebra.

6.1 Rings and Fields

Definition 65. A ***ring*** is an ordered triple $(R, +, \times)$ where R is a set and $+, \times : R \times R \to R$ are binary operations satisfying the following axioms:
(a) (Additive Structure) $(R, +)$ is an abelian group.
(b) (Multiplicative Structure) $(R, \times)$ is a semigroup, i.e. $\times$ is associative.
(c) (Distributive Laws) $\forall a, b, c \in R : a \times (b + c) = a \times b + a \times c$ and $(a + b) \times c = a \times c + b \times c$.
Further,
(d) If $\times$ is commutative, the R is said to be a ***commutative ring.***
(e) If $(R, \times)$ is a monoid, i.e. it has an identity (usually denoted by 1), then R is said to be ***unitary ring***, or a ***ring with unit***.

We shall usually simply write ab in stead of $a \times b$; 0 is taken to be the identity of $(R, +)$ and $-a$ the additive inverse. The conditions all seem fairly natural and general, except possibly the requirement that $(R, +)$ be an *abelian* group: that is actually forced upon us by distributivity, i.e. even if we assume R to be unitary but $(R, +)$ not necessarily abelian, the distributive laws tell us that

$$a + a + b + b = (1 + 1)a + (1 + 1)b = (1 + 1)(a + b) = 1(a + b) + 1(a + b) = a + b + a + b,$$

from which $a + b = b + a$.

Example 91
The ***zero ring*** $\{0\}$ is the unique ring of size 1. Only in this ring, $1 = 0$.

Example 92
The prototypical example of a ring is the ring of ***Zahlen***, i.e. integers, $\mathbb{Z}$.

Example 93
For any $n \in \mathbb{N}_0$, we have the ring $\mathbb{Z}/n\mathbb{Z}$. This is an example of a more general construction called a ***quotient ring***.

Example 94
The ring of even integers, $2\mathbb{Z}$, is a commutative nonunitary ring.

Example 95
For $n \in \mathbb{N}$, the ring of $n \times n$ matrices with real (resp. complex) entries is denoted by $\mathcal{M}_n(\mathbb{R})$ (resp. $\mathcal{M}_n(\mathbb{C})$). For $n \geqslant 2$, this is a noncommutative unitary ring.

Some authors like to reserve the word *ring* for commutative unitary rings (or cr1ngs), and specifiy explicitly if their rings are noncommutative or unitary; we will not follow that convention here.

The following are elementary consequences of the definitions:

Proposition 15

Let R be a ring. Then for all $a, b \in R$:
(a) $0a = a0 = 0$.
(b) $(-a)b = a(-b) = -(ab)$.
(c) $(-a)(-b) = ab$.
(d) If $1 \in R$, then it is unique and $-a = (-1)a$.

Proof. These follow immediately from the axioms. Details can be found in Proposition 1 of DF 7.1. ■

We make two basic definitions:

Definition 66. If R and S are rings and $R \subseteq S$, then R is called a ***subring*** of S.

We have the usual handy characterization: to check that R is a subring of S, it suffices to check that it is nonempty and closed unders subtraction and multiplication. Some authors also require that $1_S \in R$; that will usually be clear from context.

Definition 67. An element $u \in R$ is said to be a ***unit*** if $\exists u' \in R : uu' = 1$. The set of units of R form a group under multiplication, and this is denoted by $R^\times$.

Example 96

$\mathbb{Z}^\times = \{\pm 1\} \cong Z_2$.

Example 97

$(\mathbb{Z}/n\mathbb{Z})^\times = \{\bar{a} : (a, n) = 1\}$.

Example 98

The ring of ***Gaussian integers***, defined to be $\mathbb{Z}[\imath] := \{a + b\imath : a, b \in \mathbb{Z}\}$ is a subring of $\mathbb{C}$. It is not hard to show that $\mathbb{Z}[\imath]^\times = \{\pm 1, \pm\imath\}$.

Example 99

Consider the set R of functions $f : [0, 1] \to \mathbb{R}$. This is ring under pointwise addition and multiplication, i.e. $(f + g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$. Now $f \in R^\times \Leftrightarrow f([0, 1]) \not\ni 0$. The subset of continuous functions $f : [0, 1] \to \mathbb{R}$ forms a subring of R.

Example 100

For any ring nontrivial R, we may create the ***polynomial ring*** $R[x]$. Then R is a subring of $R[x]$. Similarly, we may create the polynomial ring in finitely many varibales $R[x_1, \cdots, x_n]$; then we have a chain of proper subrings $\{0\} \subsetneq R \subsetneq R[x_1] \subsetneq \cdots \subsetneq R[x_1, \cdots, x_n]$.

Example 101

If R is a ring, then we denote by $R[[x]]$ the ring of ***formal power series*** in x with coefficients in R. In other words, $R[[x]] := \{\sum_{n=0}^\infty a_n x^n : a_n \in R\}$ with coefficient-wise addition and usual multiplication $\left(\sum_{n=0}^\infty a_n x^n\right)\left(\sum_{n=0}^\infty b_n x^n\right) = \sum_{n=0}^\infty c_n x^n$ where $c_n = \sum_{r=0}^n a_r b_{n-r}$. (Note that we're only taking ***finite sums***, so that this is well-defined irrespective of convergence issues.)

Example 102

Let p be a fixed prime. The set $\mathbb{Z}_{(p)} = \{a/b \in \mathbb{Q} : (a,b) = 1 \wedge p \nmid b\}$ is called the ***localization*** of $\mathbb{Z}$ at (p). It is a subring of $\mathbb{Q}$. Note that $\mathbb{Z}_{(p)}^\times = \{a/b \in \mathbb{Q} : (a,b) = 1 \wedge p \nmid b \wedge p \nmid a\}$. For example, $2/3 \in \mathbb{Z}_{(5)}^\times$.

Definition 68. Let R be a cr1ng. Then define $\mathrm{GL}_n(R) := \mathcal{M}_n(R)^\times$; in other words, $\mathrm{GL}_n(R)$ is the set of $n \times n$ matrices A with entries in R s.t. $\det A \in R^\times$. Similary, define $\mathrm{SL}_n(R)$ to be the set of $n \times n$ matrices A with entries in R s.t. $\det A = 1$.

Example 103

Observe that with our previous definitions, $\mathrm{GL}_n(\mathbb{Z}) = \mathrm{SL}_n^{\pm}(\mathbb{Z})$.

Observe that for $R = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, we have $R^\times = R \setminus \{0\}$, i.e. all nonzero elements are invertible. Such rings are given a special name:

Definition 69. A commutative unitary ring R with $1 \neq 0$ s.t. $R^\times = R \setminus \{0\}$ is called a ***field***.

We immediately get the following usual properties:

Proposition 16

If F is a field, then:
(a) If $x \neq 0$, then $xy = xz \Rightarrow y = z$.
(b) If $x, y \neq 0$, then $xy \neq 0$.

Can you guess what a ***subfield*** is?

Example 104

We have a tower of subfields $\mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}$.

Example 105

For any field F, the group $\mathrm{GL}_1(F) \cong F^\times$.

Example 106

Given any field F, we can form the ***field of rational functions*** in variable x with coefficients in F. In other words, $F(x) := \{p(x)/q(x) : p(x), q(x) \in F[x], q \not\equiv 0\}$.

Example 107

Let D be a rational number that is not a perfect square in $\mathbb{Q}$. Then the set $\mathbb{Q}[\sqrt{D}] := \{a + b\sqrt{D} : a, b \in \mathbb{Q}\}$ is a ***quadratic subfield*** of $\mathbb{C}$. If $D \geqslant 0$, then it is a subfield of $\mathbb{R}$. These are special examples of ***number fields***, objects of study in algebraic number theory.

Example 108

Observe that for any prime p, the ring $\mathbb{Z}/p\mathbb{Z}$ is actually a field! This is called the ***finite field*** of order p, and is denoted by $\mathbb{F}_p$.

We now talk about relationships and maps between rings:

Definition 70. Let R and S be rings. A map $\varphi : R \to S$ is called a ***ring homomorphism*** if:
(a) (Additive Structure) $\varphi : (R, +) \to (S, +)$ is a group homorphism, i.e. $\forall a, b \in R : \varphi(a + b) = \varphi(a) + \varphi(b)$.
(b) (Multiplicative Structure) $\varphi : (R, \times) \to (S, \times)$ is a homomorphism of semigroups, i.e. $\forall a, b \in R : \varphi(ab) = \varphi(a)\varphi(b)$.
Note that some authors also require that $\varphi(1_R) = 1_S$. The set of ring homomorphism $R \to S$ is denoted by $\mathrm{Hom}(R, S)$.

Similarly, if E, F are fields, then a map $\varphi : E \to F$ is called a ***field homomorphism*** if it respects both the additive structure and multiplicative structure of E and F. Some authors require also that $\varphi(1_E) = 1_F$ to avoid trivial homomorphisms like $\varphi \equiv 0$. We have the following usual definitions:

Definition 71. Let $\varphi : R \to S$ be a ring homomorphism. The ***kernel*** of φ is the fiber over 0_S, i.e., it is the kernel of the additive homomorphism $\varphi : (R, +) \to (S, +)$.

Definition 72. A ring homomorphism $\varphi : R \to S$ is said to be an ***isomorphism*** if there is a ring homomorphism $\psi : S \to R$ s.t. $\psi \circ \varphi = \mathrm{id}_R$ and $\varphi \circ \psi = \mathrm{id}_S$. In this case, we write $R \cong S$.

Example 109

The reduction-mod-n map $\varphi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ is a ring homomorphism.

Example 110

Let R be a ring. For fixed $\alpha \in R$, consider the map $\cdot\big|_\alpha : R[x] \to R$ given by $p(x) \mapsto p(\alpha)$. This is a ring homomorphism. It is surjective by definition, but it is not, in general, injective.

The following is the analog of the corresponding theorem for groups:

Theorem 36

Let R and S be rings, and let $\varphi : R \to S$ be a homomorphism.
(a) The image $\varphi(R)$ is a subring of S.
(b) The kernel $\ker \varphi$ is closed under addition and multiplication by elements of R.

Nonempty subsets $I \subseteq R$ that satisfy $a, b \in I \Rightarrow a - b \in I$ and $r \in R, a \in I \Rightarrow ra \in I$ are called ***ideals*** of R. Part (b) of the above theorem says that $\ker \varphi$ is an ideal of R. These are the analogs of normal subgroups.

We also have the following definitions:

Definition 73. Let R be a ring. Then a ring homomorphism $\varphi : R \to R$ is called an ***endomorphism*** of R. The set of endomorphisms of R forms a ring under pointwise addition and composition, denoted by $\mathrm{End}(R)$, and called the ***endomorphism ring*** of R. Its group of units is called the ***automorphism group*** of R, i.e. $\mathrm{Aut}(R) := \mathrm{End}(R)^\times$.

6.2 Modules and Vector Spaces

We are now ready to make precise the notions of ring and field actions.

Definition 74. Let R be a ring. A ***left R-module*** is an abelian group $(M, +)$ along with a left-action of R, i.e. a map $\times : R \times M \to M$ denoted by $(r, m) \mapsto rm$ that satisfies the following axioms:
(a) (Compatibility with Multiplication) $\forall r, s \in R, m \in M : r(sm) = (rs)m$.
(b) (Distributive Laws) $\forall r, s \in R$ and $m \in M : (r + s)m = rm + sm$ AND $\forall r \in R, m, n \in M : r(m + n) = rm + rn$.
Further, if $1 \in R$, then we require that $\forall m \in M : 1m = m$.

The last condition is to avoid pathologies like $rm = 0$ for all r, m. Note that again the distributive laws tell us immediately that $0_R m = 0_M$ for every $m \in M$.

We also have the natural definition:

Definition 75. Let R be a ring and M a left R-module. A ***left R-submodule*** of M is a subgroup $(N, +) \leqslant (M, +)$ s.t. $\times|_{R \times N} : R \times N \to N$, i.e. $\forall r \in R, n \in N : rn \in N$.

Note that we may similarly define right R-modules. Note that for commutative rings R, we may define simply R-modules. Unless otherwise specified, we take R-modules to mean left R-modules.

Example 111

Let M be any R-module. Then M has two obvious submodules: $\{0\}$ and M itself. The former is called the ***trivial submodule*** of M.

Example 112

Let R be any ring. Then R can be considered a module over itself. The submodules of R as a module over itself are precisely the ideals $I \subseteq R$.

Example 113

Let R be a cr1ng, and $n \in \mathbb{N}$. Define R^n to be the set of n-tuples of elements of R. Then R^n is an R-module, called the ***free module of rank n over R***.

We now are ready to talk about relationships between R-modules.

Definition 76. Let R be a ring and M, N be R-modules. A map $\varphi : M \to N$ is called an ***R-module homomorphism*** (or simply ***R-linear***) if it respects their R-module structures, i.e.:
(a) (Compatibility with Additive Structure) $\varphi : (M, +) \to (N, +)$ is a homomorphism of abelian groups, i.e. $\forall m, m' \in M : \varphi(m + m') = \varphi(m) + \varphi(m')$.
(b) (Compatibility with Action of R) $\forall r \in R, m \in M : \varphi(rm) = r\varphi(m)$.
We denote the set of all R-module homomorphisms $M \to N$ by $\mathrm{Hom}_R(M, N)$.

Note that something special has happened here:

Proposition 17

Let R be a cr1ng, and let M, N be R-modules. Then the set $\mathrm{Hom}_R(M, N)$ is an abelian group under pointwise addition, and under the action $(r\varphi)(m) = r\varphi(m)$ forms an R-module itself.

Definition 77. With addition as above and multiplication by function composition, $\mathrm{End}_R(M) := \mathrm{Hom}_R(M, M)$ is a ring itself, and is called the ***endomorphism ring*** of M. The group $\mathrm{Aut}_R(M) := \mathrm{End}_R(M)^\times$ is called the ***automorphism group*** of M.

We now look at some special kinds of modules.

Example 114

The set of all continuous functions $f : [0, 1] \to \mathbb{C}$ forms a $\mathbb{C}$-module.

In fact, you've seen $\mathbb{C}$-modules before.

Definition 78. If F is a field, then an F-module V is called a ***vector space*** over F. If V, W are vector spaces over F, then the elements of $\mathrm{Hom}_F(V, W)$ are called ***linear maps*** or ***linear transformations*** of vector spaces.

Example 115

For any field F, the vector space F^n is said to be the ***n-dimensional vector space over*** F. (It is unique upto isomorphism.)

Example 116

For any field F, the space $V = F[x]$ is an F-vector space.

Example 117

The set of continuous (resp. differentiable) functions $f : \mathbb{R} \to \mathbb{R}$ is an $\mathbb{R}$-vector space.

Example 118

The collections of solutions of a linear, homogenous, constant coefficient differential equation (e.g. $y'' - 2y' + 2y = 0$) form an $\mathbb{R}$ (resp. $\mathbb{C}$) vector space.

Example 119

Let L/K be a field extension (i.e. $K \subseteq L$ is a subfield). Then L is a K-vector space. Then the group $\mathrm{Aut}_K(L)$ is the group of K-linear automorphisms of L, i.e. the automorphisms $\sigma : L \to L$ s.t. $\sigma|_K = \mathrm{id}_K$. This group is also denoted by $\mathrm{Aut}(L/K)$. If the extension L/K is ***Galois*** (which we don't have the tools to define here), then the group $\mathrm{Aut}(L/K)$ is denoted by $\mathrm{Gal}(L/K)$, and is called the ***Galois group*** of L/K.

Example 120

Given a vector space V over a field F, the group $\mathrm{Aut}_F(V)$ is denoted simply by $\mathrm{GL}(V)$, and is called the ***general linear group*** of V. Can you see why $\mathrm{GL}_n(F) = \mathrm{GL}(F^n)$ and also what the definition of $\mathrm{SL}(V)$ might be?

6.3 Introduction to Categories

6.3.1 Basic Definitions

We've seen a lot of instances where we have a collection of "things" and certain "maps" between those things. Seeing collections of this sort pop up all over math, and considering relationshipships between these collections, algebraic topologists Eilenberg and Mac Lane introduced the notion of categories in 1942-45. It is a fundamental language for many areas of modern mathematical research, e.g. in algebraic geometry.

Definition 79. A ***category*** $\mathscr{C}$ consists of the following information:
(a) A collection $\mathrm{Ob}(\mathscr{C})$ of ***objects*** in $\mathscr{C}$.
(b) For each pair of objects $A, B \in \mathrm{Ob}(\mathscr{C})$, a collection $\mathrm{Mor}_{\mathscr{C}}(A, B)$ of ***morphisms*** or ***arrows*** between A and B. These are written $f : A \to B$.
(c) For every ordered triple of objects $A, B, C \in \mathrm{Ob}(\mathscr{C})$, a ***law of composition*** of morphisms:

$$\mathrm{Mor}_{\mathscr{C}}(B, C) \times \mathrm{Mor}_{\mathscr{C}}(A, B) \to \mathrm{Mor}_{\mathscr{C}}(A, C)$$

denoted by $(g, f) \mapsto g \circ f$.
The objects and morphisms are required to satisfy the following axioms:
(a) The collections $\mathrm{Mor}_{\mathscr{C}}(A, B)$ and $\mathrm{Mor}_{\mathscr{C}}(C, D)$ are disjoint unless $A = C$ and $B = D$.
(b) (Associativity of Composition) For every ordered quadruple of objects A, B, C, D and triple of morphisms $(h, g, f) \in \mathrm{Mor}_{\mathscr{C}}(C, D) \times \mathrm{Mor}_{\mathscr{C}}(B, C) \times \mathrm{Mor}_{\mathscr{C}}(A, B)$ we have $(h \circ g) \circ f = h \circ (g \circ f)$.
(c) (Existence of Identity) $\forall A \in \mathrm{Ob}(\mathscr{C})$, $\exists\, \mathrm{id}_A \in \mathrm{Mor}_{\mathscr{C}}(A, A) : \forall B \in \mathrm{Ob}(\mathscr{C})$ we have that $f \circ \mathrm{id}_A = f$ for every $f \in \mathrm{Mor}_{\mathscr{C}}(A, B)$ and $\mathrm{id}_A \circ g = g$ for every $g \in \mathrm{Mor}_{\mathscr{C}}(B, A)$.

Note that:
(a) It is an easy exercise that any identity morphism must be unique.
(b) $\mathrm{Mor}_{\mathscr{C}}(A, B)$ is sometimes also denoted by $\mathrm{Hom}_{\mathscr{C}}(A, B)$, but we will avoid that notation because nowadays $\mathrm{Hom}_{\mathscr{C}}(A, B)$ is usually reserved for ***abelian categories***.
(c) In writing $\mathrm{Mor}_{\mathscr{C}}(A, B)$, the subscript $\mathscr{C}$ is usually dropped if the category is clear from context.
(d) We are necessarily vague about "collections" of objects in stead of sets. This is to circumvent foundational issues, e.g. to avoid talking about "sets of sets" and running into Russell's paradox etc. This is because this is not a graduate course on advanced mathematical logic.

The following are examples of categories you have seen before:
(a) (Set): sets and functions.
(b) (Grp): groups and group homomorphisms.
(c) (Ab): abelian groups and group homomorphisms.
(d) (Rng): rings are ring homomorphisms.
(e) (CRng): commutative rings and ring homomorphisms.
(f) ($\mathsf{CR1ng}$): commutative unitary rings and ring homomorphisms.
(g) (Mod_R): R-modules and R-module homomorphisms; e.g. ($\mathsf{Mod}_{\mathbb{Z}}$), ($\mathsf{Mod}_{F[x]}$), etc.
(h) (Vec_F): F-vector spaces and F-linear maps; e.g. ($\mathsf{Vec}_{\mathbb{R}}$), ($\mathsf{Vec}_{\mathbb{C}}$), ($\mathsf{Vec}_{\mathbb{F}_p}$), etc.
(i) (FDVec_F): finite dimensional F-vector spaces and F-linear maps; e.g. ($\mathsf{FDVec}_{\mathbb{R}}$), ($\mathsf{FDVec}_{\mathbb{F}_p}$), etc.

The following are examples of categories that you may not have seen so far, and but you will see at least once in your mathematics courses at college.
(a) (Top): Topological spaces and continuous maps.
(b) (Top_*): Pointed topological spaces (X, p) with continuous maps respecting the distinguished point, i.e. $f : (X, p) \to (Y, q)$ s.t. $f(p) = q$.
(c) (TopGrp): Topological groups and continous homomorphisms.
(d) (Man): Topological manifolds and continuous maps.
(e) (SMan): Smooth manifolds and smooth maps.
(f) (Lie): Lie groups and smooth homomorphisms.
(g) (CW): CW complexes and continuous maps.
(h) (Var_F): Varieties over the field F and morphisms of varieties.
(i) (AbVar_F): Abelian varieties over the field F with morphisms of varieties that are also abelian group homomorphisms.
(j) (Ell_F): Elliptic curves over the field F and isogenies.

The following are examples of examples of objects that you may have seen before, but may not have thought of as categories:

(a) Let $(I, \preceq)$ be a partially ordered set. Then we can form a category (I) whose objects are the elements of I and s.t. for $x, y \in I$ there is a unique morphism $x \to y$ if $x \preceq y$ and none otherwise. (See why this is a category?) Similarly, we can form the "dual" category with a single morphism $x \to y$ if $x \succeq y$ and none otherwise.

(b) Let X be any set. Then we can form a category of the poset $(\wp(X), \subseteq)$.

(c) Let X be a topological space. Then we can form a category of the poset $(\mathscr{T}(X), \subseteq)$.

The last three categories are example of ***small categories***, which are categories $\mathscr{C}$ s.t. $\mathrm{Ob}(\mathscr{C})$ is a set. We can also define ***locally small categories***, which are categories $\mathscr{C}$ s.t. for each $A, B \in \mathrm{Ob}(\mathscr{C})$, $\mathrm{Mor}_{\mathscr{C}}(A, B)$ is a set. Can you identify which of the above categories are small or locally small?

We can similarly define notions of subcategories:

Definition 80. A ***subcategory*** $\mathscr{A}$ of a category $\mathscr{B}$ is a category whose objects are some of the objects of $\mathscr{B}$ and whose morphisms are some of the morphisms of $\mathscr{B}$ s.t. they include the identity morphism for each of the objects and are closed under composition.

A subcategory $\mathscr{A}$ of $\mathscr{B}$ is called ***full*** if it contains all of $\mathrm{Mor}_{\mathscr{B}}(A, B)$ for each $A, B \in \mathrm{Ob}(\mathscr{A}) \subseteq \mathrm{Ob}(\mathscr{B})$.

Example 121

(Ab) is a full subcategory of (Grp).

Example 122

We have a chain of subcategories ($\mathsf{CR1ng}$) in (CRng) in (Rng). Whether these subcategories are full depends on your definition of ring homomorphisms.

Example 123

(FDVec_F) is a full subcategory of (Vec_F) for each field F.

Example 124

(SMan) is a subcategory of (Man), but it is not full.

Now that we have notions of *morphisms*, we better talk about *isomorphisms*, *endomorphisms*, *automorphisms*, etc.

Definition 81. Let $\mathscr{C}$ be any category.

(a) Let $A, B \in \mathrm{Ob}(\mathscr{C})$. A morphism $f : A \to B$ is called an ***isomorphism*** if there is a–necessarily unique–morphism $g : B \to A$ s.t. $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$. In this case, we write $A \cong B$.

(b) For any object $A \in \mathrm{Ob}(\mathscr{C})$, any morphism $f : A \to A$ is called an ***endomorphism*** of A. The collection of endomorphisms is $\mathrm{End}_{\mathscr{C}}(A) := \mathrm{Mor}_{\mathscr{C}}(A, A)$.

(c) An endomorphism which is also an isomorphism is called an ***automorphism***. The collection of automorphisms from an object $A \in \mathrm{Ob}(\mathscr{C})$ to itself, denoted by $\mathrm{Aut}_{\mathscr{C}}(A)$ or simply $\mathrm{Aut}(A)$.

If there is a unique isomorphism $f : A \to B$, then we call it a ***canonical isomorphism***, and we sometimes write $A = B$. Note that if $\mathscr{C}$ is a locally small category, then in fact each $\mathrm{Aut}(A)$ has the natural structure of a group.

6.3.2 Functors

Now that we have morphisms within a category, we'd now like to talk about relations between categories. This is where we actually see how powerful category theory is.

Definition 82. Let $\mathscr{C}$ and $\mathscr{D}$ be categories. A ***covariant functor*** $\mathscr{F} : \mathscr{C} \to \mathscr{D}$ is the following information:

(a) A map of collections $\mathscr{F} : \mathrm{Ob}(\mathscr{C}) \to \mathrm{Ob}(\mathscr{D})$, i.e. to each object $A \in \mathrm{Ob}(\mathscr{C})$, it assigns an object $\mathscr{F}(A) \in \mathrm{Ob}(\mathscr{D})$.

(b) For each $A, B \in \mathrm{Ob}(\mathscr{C})$, a map of collections $\mathscr{F} : \mathrm{Mor}_{\mathscr{C}}(A, B) \to \mathrm{Mor}_{\mathscr{D}}(\mathscr{F}(A), \mathscr{F}(B))$, i.e. to each morphism $f : A \to B$, it assigns a morphism $\mathscr{F}(f) : \mathscr{F}(A) \to \mathscr{F}(B)$.

These mappings are required to respect composition and identites:

(a) If $A \xrightarrow{f} B \xrightarrow{g} C$, then $\mathscr{F}(A) \xrightarrow{\mathscr{F}(f)} \mathscr{F}(B) \xrightarrow{\mathscr{F}(g)} \mathscr{F}(C)$ s.t. $\mathscr{F}(g \circ f) = \mathscr{F}(g) \circ \mathscr{F}(f)$.

(b) For each $A \in \mathrm{Ob}(\mathscr{C})$, $\mathscr{F}(\mathrm{id}_A) = \mathrm{id}_{\mathscr{F}(A)}$.

In many cases, when the covariant functor is understood, it is common to denote $\mathscr{F}(f)$ by f_*, and to call it the ***pushforward*** of f.

Example 125

For any category $\mathscr{C}$, identity functor $\mathrm{id}_{\mathscr{C}}$ is a covariant functor $\mathrm{id}_{\mathscr{C}} : \mathscr{C} \to \mathscr{C}$.

Example 126

For any subcategory $\mathscr{A}$ of $\mathscr{B}$, we have an inclusion functor $\iota : \mathscr{A} \to \mathscr{B}$.

Example 127

For $\mathscr{C} = (\mathsf{Grp}), (\mathsf{Top})$, etc., we have the ***forgetful functor*** $\mathscr{F} : \mathscr{C} \to (\mathsf{Set})$ that remembers only the set-theoretic structure of the underlying objects.

Example 128

The ***abelianizing functor*** $\mathrm{ab} : (\mathsf{Grp}) \to (\mathsf{Ab})$ associates to each group G its abelianization G^{ab}, and to each homomorphism $\varphi : G \to H$ associates the corresponding induced homomorphism $\varphi^{\mathrm{ab}} : G^{\mathrm{ab}} \to H^{\mathrm{ab}}$ defined as follows: let π_H denote the quotient $\pi_H : H \to H/\langle [H, H] \rangle =: H^{\mathrm{ab}}$, and consider the composition $\pi_H \circ \varphi : G \to H^{\mathrm{ab}}$; by the universal property of abelianization, this induces a unique homomorphism from $G^{\mathrm{ab}} \to H^{\mathrm{ab}}$, call this φ^{ab}. In other words, this is defined so that $\varphi^{\mathrm{ab}} \circ \pi_G = \pi_H \circ \varphi$, i.e. so that the following diagram commutes:

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & H \\ \downarrow{\scriptstyle \pi_G} & & \downarrow{\scriptstyle \pi_H} \\ G^{\mathrm{ab}} & \xrightarrow{\varphi^{\mathrm{ab}}} & H^{\mathrm{ab}} \end{array}$$

Example 129

(This example was what motivated Eilenberg and Mac Lane to come up with categories.) The ***fundamental group functor*** $\pi_1 : (\mathsf{Top}_*) \to (\mathsf{Grp})$ assigns to each pointed topological space (X, p) a group $\pi_1(X, p)$ and to each continuous map $f : (X, p) \to (Y, q)$ the induced pushforward $f_* : \pi_1(X, p) \to \pi_1(Y, q)$ given by $f_*([\gamma]) = [f \circ \gamma]$.

Sometimes, it is necessary to deal with functors that take morphisms to morphisms in the opposite direction. These are naturally called ***contravariant functors***.

Definition 83. Let $\mathscr{C}$ and $\mathscr{D}$ be categories. A ***contravariant functor*** $\mathscr{F} : \mathscr{C} \to \mathscr{D}$ is the following information:

(a) A map of collections $\mathscr{F} : \mathrm{Ob}(\mathscr{C}) \to \mathrm{Ob}(\mathscr{D})$.

(b) For each $A, B \in \mathrm{Ob}(\mathscr{C})$, a map of collections $\mathscr{F} : \mathrm{Mor}_{\mathscr{C}}(A, B) \to \mathrm{Mor}_{\mathscr{D}}(\mathscr{F}(B), \mathscr{F}(A))$, i.e. to each morphism $f : A \to B$, it assigns a moprhism $\mathscr{F}(f) : \mathscr{F}(B) \to \mathscr{F}(A)$.

These mappings are required to respect composition and identities:

(a) If $A \xrightarrow{f} B \xrightarrow{g} C$, then $\mathscr{F}(A) \xleftarrow{\mathscr{F}(f)} \mathscr{F}(B) \xleftarrow{\mathscr{F}(g)} \mathscr{F}(C)$ s.t. $\mathscr{F}(g \circ f) = \mathscr{F}(f) \circ \mathscr{F}(g)$.

(b) For each $A \in \mathrm{Ob}(\mathscr{C})$, $\mathscr{F}(\mathrm{id}_A) = \mathrm{id}_{\mathscr{F}(A)}$.

Again, when the contravariant functor is understood, it is common to denote $\mathscr{F}(f)$ by f^*, and to call it the ***pullback*** of f.

Example 130

Perhaps the most familiar example of this is the ***dual space functor***: for any field F, the functor $\vee : (\mathsf{Vec}_F) \to (\mathsf{Vec}_F)$ assigns to each vector space V/F its dual $V^\vee := \mathrm{Hom}_F(V, F^1)$, i.e. the vector space of linear maps from V to F, and to each linear map $\varphi : V \to W$ its transpose $\varphi^\vee = \varphi^\top : W^\vee \to V^\vee$ given by precomposition: $l \mapsto l \circ \varphi$, i.e. such that the following diagram commutes:

$$\begin{array}{ccccc} V & \xrightarrow{\varphi} & W & \xrightarrow{l} & F \\ & \searrow & & \nearrow_{\varphi^\vee(l)} & \end{array}$$

Example 131

The functor $\mathscr{C}(-, \mathbb{R}) : (\mathsf{Top}) \to (\mathsf{CR1ng})$ that assigns to a topological space X the ring of continuous functions $f : X \to \mathbb{R}$, and to each continuous map $\varphi : X \to Y$ the precomposition $\varphi^* : \mathscr{C}(Y, \mathbb{R}) \to \mathscr{C}(X, \mathbb{R})$ by $[g : Y \to \mathbb{R}] \mapsto [\varphi^*(g) = g \circ \varphi : X \to \mathbb{R}]$.

Example 132

Observe that if A and B are abelian groups, then $\mathrm{Hom}(A, B)$ can be given the structure of an abelian group by pointwise composition, i.e. if $\xi, \psi \in \mathrm{Hom}(A, B)$, then define $\xi\psi$ by $(\xi\psi)(a) := \xi(a)\psi(a)$ (or if written additively by $(\xi + \psi)(a) := \xi(a) + \psi(a)$). Now suppose that Z is a fixed abelian group. Then we may define a contravariant functor $\mathrm{Hom}(-, Z) : (\mathsf{Ab}) \to (\mathsf{Ab})$ by $A \mapsto \mathrm{Hom}(A, Z)$ and $\varphi : A \to B$ going to $\varphi^* : \mathrm{Hom}(B, Z) \to \mathrm{Hom}(A, Z)$ taking $[\xi : B \to Z] \mapsto [\xi \circ \varphi : A \to Z]$. When $Z = \mathbb{S}^1$, we call the functor $\mathrm{Hom}(-, \mathbb{S}^1)$ the ***dual functor*** or ***character functor*** in the category of abelian groups. (The reason for this nomenclature will become evident when you study representation theory.)

The next objects of study in category theory are ***natural transformations*** of functors, which are maps between functors, when two functors are isomorphic, etc. This allows us to the study the ***equivalence of categories***, i.e. when two categories carry essentially the same information. We will not be able to get into the specifics here. However, you will show on your homework that the categories (Ab) and $(\mathsf{Mod}_{\mathbb{Z}})$ are equivalent.

6.4 Universal Constructions

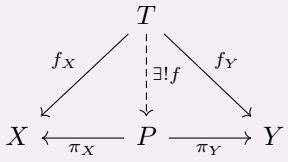
(Disclaimer: Some of the following section is adapted from Prof. Ravi Vakil's *The Rising Sea*.)

One of the most beautiful aspects of category theory is the notion of the universal constructions. These seem artifical and pretty hard to understand, but once you become familiar with these you will realize that these sometimes give us the most beautiful and elegant proofs of theorems. There's a way to formalize the notion of ***universal objects*** and ***universal arrows***, but we'll instead talk about them a bit informally.

6.4.1 Products

As a motivating question, consider the product of two sets X and Y. One way to define this is using ordered pairs: you define $X \times Y$ to be the set of all ordered pairs (x, y) for $x \in X$ and $y \in Y$. Now suppose you met some aliens, and they described to you a construction, where given two sets X and Y you look at the set of all elements of the form $\frac{x}{y}$ for $x \in X$ and $y \in Y$. Ha! You recognize that this is the same thing as what you've been calling the "Cartesian product." Better yet, there's a bijection between them that preserves their essential structures–that they carry information about pairs of elements of X and Y. Therefore, a better way to capture the essential information carried by the product is not to define it in terms of ordered pairs, but to define it in a way that captures the notion you want it to embody:

Definition 84. Given two sets X and Y, their ***product*** P is a set with functions $\pi_X : P \to X$ and $\pi_Y : P \to Y$ s.t. given any *other set* T with functions $f_X : T \to X$ and $f_Y : T \to Y$, there is a unique function $f : T \to P$ s.t. $f_X = \pi_X \circ f$ and $f_Y = \pi_Y \circ f$. In other words, $\exists! f$ s.t. the following diagram commutes:

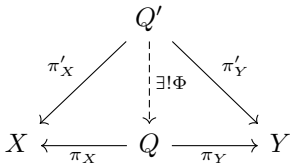


Clearly, both our definition and the aliens' definition satisfy this property. The great thing is: this property is enough to tell us that any two objects satisfying it must be "the same."

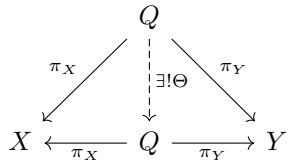
Theorem 37

A product of two sets X and Y, if it exists, is unique upto a unique isomorphism that preserves projections, so that we may call it *the* product and denote it unambiguosly by $X \times Y$.

Proof. Suppose (Q, π_X, π_Y) and (Q', π'_X, π'_Y) satisfy the universal property for the product of sets X and Y. Substituting $P = Q$ and $T = Q'$ in the above definition, we get that $\exists! \Phi : Q' \to Q$ s.t. $\pi'_X = \pi_X \circ \Phi$ and $\pi'_Y = \pi_Y \circ \Phi$, i.e. s.t. the following commutes:



Now, using the same diagram but with $P = Q'$ and $T = Q$ tells us that $\exists! \Psi : Q \to Q'$ s.t. $\pi_X = \pi'_X \circ \Psi$ and $\pi_Y = \pi'_Y \circ \Psi$. Observe that the map $\Phi \circ \Psi : Q \to Q$ is s.t. $\pi_X = \pi_X \circ \Phi \circ \Psi$ and $\pi_Y = \pi_Y \circ \Phi \circ \Psi$. Now apply the diagram to $P = Q$ and $T = Q$.

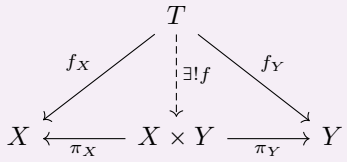


This tells us that $\exists! \Theta : Q \to Q$ s.t. $\pi_X = \pi_X \circ \Theta$ and $\pi_Y = \pi_Y \circ \Theta$. But now, I can give you two such Θ! Both of $\Theta = \mathrm{id}_Q$ and $\Theta = \Phi \circ \Psi$ work. By uniqueness, this means that $\mathrm{id}_Q = \Theta = \Phi \circ \Psi$. The analogous argument with $P = T = Q'$ tells us that $\mathrm{id}_{Q'} = \Psi \circ \Phi$. Therefore, Φ and Ψ are the required unique isomorphisms preserving the projections. ∎

It is clear that this theorem, and indeed the definition, only tell us that *if* such an object were to exist, *then* it must be essentially unique–we cannot use this to conclude that such a construction exists! Indeed, to show such a thing exists is when we have to resort to talking about ordered pairs (x, y) or

when the aliens have to resort to talking about their stacks $\frac{x}{y}$. Of course, everything we've done here generalizes to arbitrary categories.

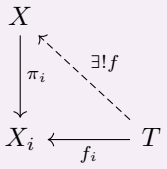
Definition 85. Let $\mathscr{C}$ be any category, and let $X, Y \in \mathrm{Ob}(\mathscr{C})$ be any objects. Then the ***product*** $X \times Y$ is an object of $\mathscr{C}$ along with morphisms $\pi_X : X \times Y \to X$ and $\pi_B : A \times B \to B$ s.t. given any other object $T \in \mathrm{Ob}(\mathscr{C})$ and morphisms $f_X : T \to X$ and $f_Y : T \to Y$, there is a unique morphism $f : T \to X \times Y$ s.t. $f_X = \pi_X \circ f$ and $f_Y = \pi_Y \circ f$. In other words, $\exists! f$ s.t. the following diagram commutes:



By the standard universal property argument, such an object, if it exists, must be unique upto unique isomorphism preserving projections, so that we may call it *the* product.

In fact, what's stopping us from looking at products of more than two things?

Definition 86. Let $\mathscr{C}$ be any category, and let $\{X_i\}_{i \in I}$ be any collection of objects of $\mathscr{C}$ indexed by the set I. The ***product*** $X = \prod_{i \in I} X_i$ is an object of $\mathscr{C}$ with morphisms $\pi_i : X \to X_i$ for each $i \in I$ s.t. given any other object $T \in \mathrm{Ob}(\mathscr{C})$ and morphisms $f_i : T \to X_i$ for each $i \in I$, there is a unique morphism $f : T \to X$ s.t. $f_i = \pi_i \circ f$ for each $i \in I$. In other words, $\exists! f$ s.t. the following diagram commutes for each $i \in I$:



By the standard universal property argument, such an object, it if exists, must be unique upto unique isomorphism preserving projections, so that we may call it *the* product.

Note that the phrase "preserving projections" is essential, e.g. there are many isomorphisms $\varphi : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ (e.g. the involution swapping the factors is one), but there's only one isomorphism that preserves the projections onto the first and second factors.

Example 133

Definition 11 (and a little verification) tells us that the product exists in the category (Set).

Example 134

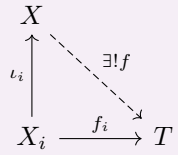
Theorem 9 tells us that the product exists in the category (Grp). From this it is also clear that the same construction works in the category (Ab).

A word of caution: not all categories admit products. However, these are significantly hard to come by (one example would be $(\mathsf{Ell}_{\mathbb{C}})$). In fact, the product exists in almost all of the categories you've met so far, though it might look different than you imagine. Here's a fun exercise to test your understanding. Suppose X is any set. Consider the category of the poset $(\wp(X), \subseteq)$, i.e. the category whose objects are subsets $U \subseteq X$, with a single morphism $U \to V$ if $U \subseteq V$ and none otherwise. What is the product in this category? (Hint: Your answer should be one word.)

6.4.2 Coproducts

Essentially for every construction is category theory, there is a "dual" or an opposite construct, which consists of the same definitions, except with all the arrows flipped. Here we have what is called the ***coproduct***.

Definition 87. Let $\mathscr{C}$ be any category, and let $\{X_i\}_{i\in I}$ be any collection of objects of $\mathscr{C}$ indexed by the set I. The ***coproduct*** $X = \coprod_{i\in I} X_i$ is an object of $\mathscr{C}$ with morphisms $\iota_i : X_i \to X$ for each $i \in I$ s.t. given any other object $T \in \mathrm{Ob}(\mathscr{C})$ and morphisms $f_i : X_i \to T$ for each i, there is a unique morphisms $f : X \to T$ s.t. $f = f_i \circ \iota_i$ for each $i \in I$. In other words, $\exists! f$ s.t. the following diagram commutes for each $i \in I$:



By the standard universal property argument, such an object, if it exists, must be unique upto unique isomorphism preserving inclusions, so that we may call it *the* coproduct.

Example 135

Theorem 10 tells us that the disjoint union is the coproduct in the category of sets.

Example 136

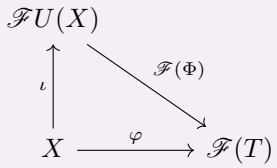
Theorem 11 tells us that the free product is the coproduct in the category of sets.

Again, the coproduct might look a bit different than these: if we return to our category $(\wp(X), \subseteq)$, what is the coproduct here? (Hint: It's the *other* one word.)

6.4.3 Universal Arrows

We are now ready to talk about slightly more general universal constructions.

Definition 88. Let $\mathscr{C}$ and $\mathscr{D}$ be categories, and let $\mathscr{F} : \mathscr{C} \to \mathscr{D}$ be a covariant functor. Let $X \in \mathrm{Ob}(\mathscr{D})$. A ***universal arrow*** from X to $\mathscr{F}$ is a pair $(U(X), \iota)$ where $U(X) \in \mathrm{Ob}(\mathscr{C})$ and $\iota \in \mathrm{Mor}_{\mathscr{D}}(X, \mathscr{F}U(X))$ s.t. for any $T \in \mathrm{Ob}(\mathscr{C})$ and $\varphi \in \mathrm{Mor}_{\mathscr{D}}(X, \mathscr{F}(T))$, $\exists! \Phi \in \mathrm{Mor}_{\mathscr{C}}(U(X), T)$ s.t. $\varphi = \mathscr{F}(\Phi) \circ \iota$, i.e. such that the following diagram commutes:



The functor $\mathscr{F}$ is then called ***right-adjoint***. This fits into the more general idea of adjoint functors, which you'll learn when you take an actual course that talks about category theory. As before, we first prove that this universal object $U(X)$ is essentially unique.

Theorem 38

For any fixed $X \in \mathrm{Ob}(\mathscr{C})$, there is a unique $U(X)$ upto isomorphism preserving ι.

Proof Sketch. Suppose $(U(X), \iota)$ and $(U'(X), \iota')$ are two such universal arrows. Then we get that $\exists! \Phi \in \mathrm{Mor}_{\mathscr{C}}(U(X), U'(X))$ s.t. $\mathscr{F}(\Phi)$ takes ι to ι' and $\exists! \Psi \in \mathrm{Mor}_{\mathscr{C}}(U'(X), U(X))$ s.t. $\mathscr{F}(\Psi)$ takes ι' to ι. Applying the universal construction one more time tells us that $\Psi \circ \Phi = \mathrm{id}_{U(X)}$ and $\Phi \circ \Psi = \mathrm{id}_{U'(X)}$. ■

The most common example of this kind is the kind you get when you have relatively concrete categories.

Definition 89. Let $\mathscr{C}$ and $\mathscr{D}$ be categories, and let $\mathscr{F} : \mathscr{C} \to \mathscr{D}$ be a covariant functor. Then the pair $(\mathscr{C}, \mathscr{F})$ is called a ***relatively concrete category over the base category*** $\mathscr{D}$ if the functor $\mathscr{F}$ is ***faithful***, i.e. if for each $A, B \in \mathrm{Ob}(\mathscr{C})$, the map $\mathscr{F} : \mathrm{Mor}_{\mathscr{C}}(A, B) \to \mathrm{Mor}_{\mathscr{D}}(\mathscr{F}(A), \mathscr{F}(B))$ is injective.
A relatively concrete category over $\mathscr{D} = (\mathsf{Set})$ is said to be an ***absolutely concrete category***, or simply a ***concrete category***.

Example 137

Take $\mathscr{C} = (\mathsf{Ab})$, $\mathscr{D} = (\mathsf{Grp})$, and $\mathscr{F} = \iota : (\mathsf{Ab}) \to (\mathsf{Grp})$ the inclusion functor. Then $((\mathsf{Ab}), \iota)$ is a concrete category over (Grp). For $G \in \mathrm{Ob}(\mathsf{Grp})$, the pair (G^{ab}, π_G) (where $\pi_G : G \to G/\langle [G, G] \rangle = G^{\mathrm{ab}}$ is the natural projection) is a universal arrow from G to ι. This tells us that the abelianization G^{ab} must be unique upto unique isomorphism.

Example 138

(Free Objects) Let $(\mathscr{C}, \mathscr{F})$ be a concrete category. (Usually we take $\mathscr{C}$ to consist of some kind of "decorated" sets, i.e. groups, abelian groups, rings, modules, vector spaces, etc.) For a set S, a ***free object on*** S ***in*** $\mathscr{C}$ is a universal arrow from S to the forgetful functor $\mathscr{F} : \mathscr{C} \to (\mathsf{Set})$. For example, we've see free objects on sets in the $\mathscr{C} = (\mathsf{Grp})$ in Theorem 12 in $\mathscr{C} = (\mathsf{Ab})$ in Definition 44.

Example 139

If you've studied topology (or perhaps if you're returning to these notes after having studied topology), can you guess what the free objects on sets in $\mathscr{C} = (\mathsf{Top})$ are?

Therefore, most of our constructions from Chapter II are actually special cases of universal constructions! The beauty of category theory is that we now needn't prove uniqueness for each separately.

That's about how far we will go with category theory. If you want to learn more, Steve Awodey's *Category Theory* is a great starting point. The standard, but much more difficult, resource is *Categories for the Working Mathematician* by Mac Lane himself.

6.5 What Next?

There are lot of directions one can go to that are natural successor courses. The following are suggestive examples:

(a) Linear Algebra: Linear algebra is the study of vector spaces and linear maps between them, especially as matrices. It has a wide range of applications in physics, computer science, data science, economics, etc.

(b) Representation Theory: This branch of mathematics tries to study groups via their actions. A ***representation*** of a group G is a homomorphism $\rho : G \to \mathrm{GL}(V)$, or equivalently, a linear action $G \circlearrowright V$. The representations of a group (or continuous representations of topological groups) often contain a lot of information about both G and V.

(c) Elementary Number Theory from an Algebraic Perspective: Some of the most basic questions are the questions involving the integers $\mathbb{Z}$. These are part of study of a branch of mathematics called

number theory. While number theory has been around for millenia, a fresh perspective on number theory using abstract algebra can often help simplify many concepts and proofs.

(d) Abstract Algebra: Another natural successor course would be a second course in abstract algebra, dealing with rings, fields, modules, field extensions, Galois Theory, etc. From that, you can branch off to:

   (1) Algebraic Number Theory: This branch of number theory uses advanced algebra to study properties *number fields*, which are subextensions $\mathbb{C}/K/\mathbb{Q}$ s.t. $\dim_{\mathbb{Q}}(K)$ is finite. These provide really beautiful insights into the working of integers. From here, you can go on and study category theory.

   (2) Algebraic Geometry: This branch of mathematics studies objects known as *varieties*, which are generalizations of curves to higher dimensions.

   (3) Algebraic Topology: This branch of topology studies topological spaces through their invariants like their homotopy groups, homology and cohomology groups, etc.

(e) Algebraic Combinatorics/Discrete Mathematics: This branch of mathematics tries to understand how to *count* various things.[2] We've seen some applications of group theory to combinatorics including Burnside's Lemma and Polyá enumeration. Another object of study in algebraic combinatorics is the representation theory of the symmetric group $\mathfrak{S}_n$.

(f) Mathematical Logic and/or Mathematical Philosophy: This branch of mathematics is a formal treatment of the foundations of mathematics. This can then lead you to works by Russell, Wittgenstein, Gödel, and many more.

(g) Algorithms/Coding Theory/Complexity Theory: This branch is at the intersection of theoretical computer science and mathematics. It studies how we can efficiently use computers to solve mathematics problems.

If instead you'd like a more well-rounded college mathematics education, I would also recommend checking out:

(a) Real and Complex Analysis: This branch of mathematics rigorously treats topics like differentiation, integration, power series, etc. This can then lead to:

   (1) Analytic Number Theory: This branch of number theory uses complex analysis to study the properties of numbers. A famous example of an object of study here is the Riemann zeta function $\zeta(s)$.

(b) Topology: This branch of mathematics studies topological spaces and continuous maps between them. This can then naturally lead to differential topology, the study of *smooth manifolds* and maps between them. This can further lead to the study of Lie Groups.

Of course, this is not meant to be a comprehensive survey of the branches of mathematics, and, as you can guess, some of these paths are not trees. For instance:

(a) You need both topology and algebra to study algebraic topology.

(b) You need both differential topology and algebraic topology to understand the beautiful connection between them (as outlined, for example, in Bott and Tu's *Differential Forms in Algebraic Topology*).

(c) Dirichlet's theorem on the infinitude of primes in arithmetic progressions uses tools from both algebraic and analytic number theory.

(d) You need algebraic number theory, algebraic geometry, and analytic number theory to study the various aspects of the theory elliptic curves.

Further, there are many interconnections across the various branches mentioned here. For example, Andrew Wiles's proof of Fermat's Last Theorem relies on heavy machinery from representation theory, algebraic number theory, algebraic geometry, analytic number theory, the theory of elliptic curves, and much more.

In conclusion, math is amazing! Go ahead and study as much as you can!

[2] This is a tremendous oversimplification, but let's roll with it for now.