SETS, GROUPS, AND GEOMETRY
Spring 2025

THOMAS BRAZELTON

Abstract. Course notes for MATH101: Sets, groups and geometry, taught at Harvard in Spring 2025.

0. Sets

A *set* is a collection of things, and these things are called elements. We won't give a formal definition of a set, since this gets us too deep into mathematical logic, so we'll kind of take a set as a given and build mathematics on top of it.

We denote by $\{1, 2, 3\}$ the set whose elements are the numbers 1, 2, and 3. These curly braces are used to list the elements of a set.

Example 0.1. The set

$$S = \{a, b, c, d\}$$

is a set consisting of four elements, which are *letters* a, b, c, and d.

Note 0.2. Elements are not allowed to be repeated! For instance, $\{a, b, a, c, d\}$ is not a valid set.[1]

Notation 0.3. We use the symbol $\in$ to denote if an element is in a set. So if $T = \{0, 4, 1, 6\}$, we might write

$$1 \in T$$

to mean that 1 is an element of T. We will write $\notin$ to say something is **not** an element of a set. So for instance

$$2 \notin T.$$

Example 0.4. We denote by $\mathbb{N}$ the set of all *natural numbers*, meaning counting numbers including zero:

$$\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}.$$

We denote by $\mathbb{Z}$ the set of all *integers*[2] meaning all positive and negative counting numbers:

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

We denote by $\mathbb{Q}$ the set of all *rational numbers*, meaning numbers of the form $\frac{p}{q}$ where p and q are integers, and $q \neq 0$.

Example 0.5. We don't just need to have numbers and letters be elements of sets. We can really let *anything* be an element in a set. For instance

$$S = \{\bigcirc, \triangle, \square\}.$$

[1] This is a convention that we're not allowing for repeated elements. We can build a different type of set theory where you *can* have repeated elements in sets, these are called *multisets*. The math that you build with these becomes a lot more complicated though.

[2] This letter comes from the German *Zahlen*, meaning "numbers."

We can also have *sets* being elements of sets. For instance we can take

$$B = \{\mathbb{N}, \mathbb{Z}, 3, \{4\}\}.$$

This is a set with four elements – the set of natural numbers, the set of integers, the number 3, and the set with one element which is the number 4. This might feel weird but we'll get used to it soon enough.

In the above examples, we didn't list out every element of a set when we wrote it, instead we did a . . . when the pattern became clear. For instance what is the following set:

$$A = \{0, 3, 6, 9, 12, 15, \ldots\}.$$

It is the set of all multiples of three! Instead of listing it out, we might *build it*, meaning give a rule for elements to be a part of it. This is done using set builder notation:

$$A = \{3n \colon n \in \mathbb{N}\}.$$

This means A is the set of all numbers of the form $3n$ where n is an element of $\mathbb{N}$.[3]

A special set is the *empty set*, which has no elements. We could write it as $\{\}$ if we wanted, but we use special notation for it, namely $\varnothing$.

0.1. **Cardinality.** If A is a set, we denote by $|A|$ the *cardinality* of the set, roughly meaning its size. It is the number of elements in the set, possibly infinite.

Example 0.6. The cardinality of some sets we've discussed are:

$$\begin{aligned} |\{a, b, c, d\}| &= 4 \\ |\mathbb{N}| &= \infty \\ |\mathbb{Z}| &= \infty \\ |\{\bigcirc, \triangle, \square\}| &= 3 \\ |\{\mathbb{N}, \mathbb{Z}, 3, \{4\}\}| &= 4 \\ |\varnothing| &= 0. \end{aligned}$$

0.2. **Subsets.** Note that every element in $\mathbb{N}$ is an element of $\mathbb{Z}$. When this happens, we write $\subseteq$, and we say one set is a *subset* of the other.

Definition 0.7. Given two sets A and B, we write $A \subseteq B$ if $x \in A$ implies that $x \in B$. In words, every element in A is also an element in B. We write $A \subsetneq B$ if A is *not* a subset of B.

Example 0.8. We have that $\mathbb{N} \subseteq \mathbb{Z}$.

Question 0.9. Given two sets A and B, how would you argue that A is *not* a subset of B?

You just have to find some element in A that is not in B.

Example 0.10. To argue that $A = \{3, 6, 8, 1\}$ is not a subset of $B = \{2, 6, 8, 1, 5\}$, we see that $3 \in A$ but $3 \notin B$. Therefore $A \subsetneq B$.

Example 0.11. Let $A = \{1, 2, 3\}$. Is it true that $\varnothing \subseteq A$?

Yes! The condition that $\varnothing \subseteq A$ means that for every $x \in \varnothing$ we have that $x \in A$. Since $\varnothing$ has no elements, this is true.[4] In fact $\varnothing \subseteq S$ for *any* set S.

[3] People who know a little CS, we might think about this as an infinite for loop (for all $n \in \mathbb{N}$, add $3 \cdot n$ to the set we're building, and let A be the resulting output). Obviously this wouldn't terminate on a computer, but we're mathematicians so we can let things happen infinitely many times and keep moving!

[4] We refer to statements like this as *vacuously true* – they're true because no elements exist to check the conditions on. For example I might say "every number which is both even and odd is equal to 7." This is a true statement, not because 7 is both even and odd, but because no numbers are both even and odd.

0.3. **Set equality.**

Question 0.12. What does it mean for two sets to be equal?

Example 0.13. We claim that $\{4, 1, 0\} = \{0, 1, 4\}$.

Answer 0.14. Two sets A and B are equal if they have the same elements. Phrased differently, $x \in A$ implies $x \in B$ and $x \in B$ implies $x \in A$. That is, $A \subseteq B$ and $B \subseteq A$.

0.4. **Operations with sets.** Given two sets A and B we denote by $A \cup B$ their *union*, meaning the set of all elements in A or in B.

$$A \cup B = \{x \colon x \in A \text{ or } x \in B\}.$$

Example 0.15. We have that

$$\{1, 2, 3\} \cup \{4, 5, 6\} = \{1, 2, 3, 4, 5, 6\}.$$

Note we don't allow repeats, so

$$\{1, 2, 3\} \cup \{3, 4\} = \{1, 2, 3, 4\}.$$

Given two sets A and B, we denote by $A \cap B$ their *intersection*, meaning the set of all elements in both A **and** B:

$$A \cap B = \{x \colon x \in A \text{ and } x \in B\}.$$

Example 0.16. We have

$$\{1, 2, 3, 4\} \cap \{3, 4, 5, 6\} = \{3, 4\}.$$

Question 0.17. What is

$$\{1, 2, 3\} \cap \{4, 5, 6\}?$$

It is the empty set! There are no elements in both sets.

Finally we denote by $A - B$ their *difference*, meaning

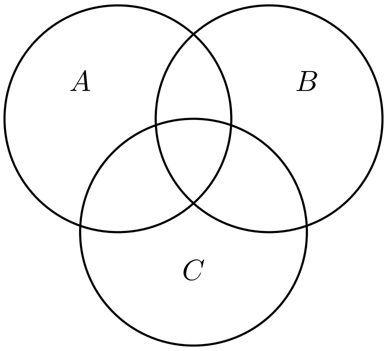
$$A - B = \{x \colon x \in A \text{ and } x \notin B\}.$$

For instance

$$\{1, 2, 3\} - \{3, 4, 5\} = \{1, 2\}.$$

Note that difference depends on the order of sets! We always have that $A \cup B = B \cup A$ and $A \cap B = B \cap A$, but $A - B$ and $B - A$ might be different sets.

Venn diagrams are a great way to visualize sets and their overlaps:



0.5. **Power sets.** Given a set A, we denote by

$$\mathcal{P}(A) := \{X \colon X \subseteq A\}$$

the *power set* of A, meaning the set of all subsets of A.

Question 0.18. What is the power set of $\{1, 2\}$?

It is the set

$$\mathcal{P}(\{1, 2\}) = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}.$$

Don't forget that $\varnothing \subseteq S$ and $S \subseteq S$ for every set S.

Question 0.19. If S has cardinality n, what do you think the cardinality of the power set $\mathcal{P}(S)$ is? Think about this.

0.6. **The real numbers.** We denote by $\mathbb{R}$ the set of *real numbers.* These are numbers we think about as lying on the number line, but need not be rational. For instance $\pi \in \mathbb{R}$ but $\pi \notin \mathbb{Q}$.[5] It's not super easy to define $\mathbb{R}$ formally, so we'll come back to this later in the class.

We define *intervals* to be subsets of $\mathbb{R}$. You may have seen the notation $[0, 1]$ before. This refers to the *closed interval* between zero and one. Explicitly in terms of set builder notation, we would write:

$$[0, 1] = \{x \in \mathbb{R} \colon 0 \leq x \text{ and } x \leq 1\}.$$

We also have open intervals, denoted by (a, b). For instance

$$(2, 3) := \{x \in \mathbb{R} \colon 2 < x \text{ and } x < 3\}.$$

0.7. **Cartesian products.**

Definition 0.20. An *ordered pair* is a tuple of two things (x, y).

Definition 0.21. Given two sets A and B, we define their *(Cartesian) product* denoted $A \times B$ by

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\}.$$

Example 0.22. If $A = \{x, y, z\}$ and $B = \{1, 2\}$ then

$$A \times B = \{(x, 1), (x, 2), (y, 1), (y, 2), (z, 1), (z, 2)\}.$$

Example 0.23. When we graph things on the xy-plane, we are thinking about a *subset* of $\mathbb{R}^2$

$$\mathbb{R}^2 = \{(x, y) : x \in \mathbb{R} \text{ and } y \in \mathbb{R}\}.$$

In particular if $y = f(x)$ is a function, we could graph the subset

$$\left\{(x, y) \in \mathbb{R}^2 : y = f(x)\right\}.$$

This is most of what we do in high school algebra - studying subsets of $\mathbb{R}^2$ of this form.

Observe: For any two finite sets A and B, we have that

$$|A \times B| = |A| \cdot |B|.$$

We can iterate multiplying sets, for instance if $A_1, A_2, \ldots, A_n$ are all sets, then we denote by

$$A_1 \times \cdots \times A_n = \{(x_1, \ldots, x_n) : x_i \in A_i \text{ for each } i = 1, \ldots, n\}.$$

We might use shorthand notation for this:

$$\prod_{i=1}^{n} A_i = A_1 \times \cdots \times A_n.$$

[5] This is not super easy to prove, but we'll see examples later of irrational numbers.

This type of notation is common also for unions and intersections of more than two sets:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup \cdots \cup A_n$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap \cdots \cap A_n.$$

0.8. **Index sets.** If we have sets $A_1, A_2, \ldots, A_n$, another way to phrase this is that we have sets *indexed* over the set $I = \{1, 2, \ldots, n\}$. In other words for each $i \in I$ we have a set A_i. In that way we can rewrite the operations above as

$$\prod_{i \in I} A_i, \qquad \bigcup_{i \in I} A_i, \qquad \bigcap_{i \in I} A_i.$$

From this perspective it's not really important that I was a set of natural numbers. We can have sets A_i indexed over *any* index set I.

0.9. **Complements.** If $B \subseteq A$ is a subset, we denote by B^c the *complement* of B in A, meaning everything that is in A and not in B:

$$B^c = \{x \in A : x \notin B\}.$$

Note that Hammack writes this as $\bar{B}$.

1. Axiomatic rules for sets

We've mentioned that it's hard to define sets, but that they satisfy certain rules. We'll lay these out now. These rules were developed by Zermelo and Fraenkel in the first few decades of the 20th century, building on work in formal logic and set theory in the 19th century. We call these axioms **ZF** after Zermelo and Fraenkel, and there are 8 axioms in total.

Definition 1.1. A set X is a *pure set* or a *hereditary set* if all of its elements are themselves sets, and all of the elements of those sets are sets, and so on.

Example 1.2. The empty set $\varnothing$ is vacuously a pure set. The set $\{\varnothing\}$ or $\{\varnothing, \{\varnothing\}\}$, are also pure, for instance.

Pure set theory: Let's treat this like a game, and temporarily forget everything we're allowed to do with sets. Our pieces are pure sets, and here are the rules.

(1) given any two sets A and B, you are allowed to ask if they are equal, and the answer is either true or false.[6]
(2) given any two sets A and B you're allowed to ask if $A \in B$, and the answer is either true or false.
(3) you're allowed to use as many variables as you want to represent sets
(4) you're allowed to negate any statement and ask if it is true or false (i.e. is it true that $A \neq B$)
(5) you can make "for all" and "implies"statements, like "for all $X \in A$" this "implies" that $X \in B$ (meaning $A \subseteq B$)
(6) you can make "there exists" statements like "there exists $x \in A$ so that $X \notin B$" (meaning $A \not\subseteq B$).

[6] Just like anything in math, we could ask what happens if we remove some of the basic building blocks. What happens if we let statements like $A \in B$ admit another truth value - not true or false but something else? What if, for instance, the *truth* of a statement is a number in the interval $[0, 1]$ where 0 is absolutely false and 1 is absolutely true, but we can have intermediate stages? These kinds of questions lead us to something called *fuzzy logic*, a fascinating detour we sadly won't explore in this class.

On top of these ground rules we're going to have some *axioms*. An *axiom* is like a mathematical rule. They are some base facts that you take for granted, and build mathematics off of. By no means are the axioms we're laying out here the only axioms you could build mathematics off of, and we're not even necessarily saying they're "true." They just end up leading to a convenient formulation of a lot of things we want to do in math.

Note 1.3. The numbering here is not a standard thing, I'm just using it to keep track of stuff easier.

ZF1: *(Axiom of extensionality)* Two sets are equal if they have the same elements.

ZF2: *(Axiom of union)* Unions of sets exist.[7]

ZF3: *(Axiom of power set)* Power sets exist – if A is a set then $\mathcal{P}(A)$ is a valid set.

ZF4: *(Axiom of pairing)* If A and B are sets, then the set $\{A, B\}$ exists.

Corollary 1.4. If A is a set then $\{A\}$ is a set.

Proof. Since A is a set, we can apply the axiom of pairing to A and itself to form the set $\{A, A\}$. Since sets can't have repeated elements, this set $\{A, A\}$ guaranteed by the axiom of pairing only has *one* element, so we abbreviate it $\{A\}$. □

ZF5: *(Axiom of regularity)* If S is a nonempty set, then it contains an element $T \in S$ so that T and S are disjoint sets (have no elements in common).

This is maybe nonintuitive but it has some important applications.

Corollary 1.5. No set can contain itself as an element.

Proof. Let A be any set, and consider the set $S = \{A\}$. By **ZF5**, S contains an element that is disjoint from itself, and since S only has one element, this implies that S is disjoint from A. In other words A and $\{A\}$ have no elements in common, so in particular $A \notin A$. □

ZF6: *(Axiom schema of specification)* You can build sets with set builder notation.[8]

Explicitly, **ZF6** says that the following type of set building is allowed:[9]

$$\{x \in A : \text{something about } x \text{ is true}\}.$$

But this type of set building is not allowed:

$$\{x : \text{something about } x \text{ is true}\}.$$

Why can't we let the latter exist?

Russell's paradox: Suppose we're allowed to build sets of that form, and we take

$$S = \{x : x \notin x\}.$$

We've already seen that no set can contain itself, so $x \notin x$ for every set x. In particular S contains *every set*. But S itself is a set, which means $S \in S$. But also $S \notin S$. These can't both be true, so we've broken math!

[7] The precise statement is if A is a pure set, there exists a set $\cup_{B \in A} B$ which is a union of all the elements of A (the most precise statements says there is a set *containing* $\cup_{B \in A} B$, and we can shorten this to $\cup_{B \in A} B$ using the axiom of pairing). For CS people, this is an axiomatization of the process of *flattening* a set or a list.

[8] We're being vague here – **ZF6** tells you more concretely *what kinds of formulas* you're allowed to use in set builder notation, but let's treat this as a black box for the time being.

[9] We're being intentionally vague with this "something about x." The precise things that are allowed to be here are what are called *first order formulas*. We'll get into these more next week.

It's generally advisable not to break math, so we exclude sets built like this. The point is not whether $S \in S$ or whether $S \notin S$, the point is such a set S *cannot be allowed to exist* if we want a logically consistent framework of math.

Barber's paradox (a common application of Russell's paradox): A barber cuts everyone's hair who doesn't cut their own hair. Does the barber cut their own hair?

ZF7: *(Axiom schema of replacement)* The domains of functions are sets (roughly speaking).

ZF8: *(Axiom of infinity)* There exists a set with infinitely many elements.

There is a 9th mysterious axiom, called the *axiom of choice*. This isn't one of the ZF axioms, so when we use it we often refer to **ZFC** which is ZF + Choice. We won't go into this as much in this class, but it will become super important later in proof-based mathematics.

2. Logic

A *statement* is any mathematical sentence that can definitively evaluated as true or false.

Here are some examples of statements:

(1) It is Monday today
(2) The number 2 is even
(3) The number 2 is not even
(4) There exists a finite subset of X.
(5) Every natural number is divisible by a prime number
(6) Every subset of an infinite set is infinite.

We can evaluate each of these as true or false.

Let P be a mathematical statement. Then we can assign it a *truth value* meaning an element of the set $\{T, F\}$ where T stands for true and F stands for false.

We can *negate* mathematical statements, which swaps the truth value of the statement. We denote this new statement by $\neg P$ (Hammack writes $\sim P$)

P	$\neg P$
It is Monday today	It is not Monday today
The number 2 is even	The number 2 is not even
The number 2 is not even	The number 2 is even

Pause – what happened here? Let P be "the number 2 is even." Then we just said

$$\neg\neg P \text{ is the same statement as } P.$$

This is called *double negation elimination*.[10] It's an admissible rule in our logical framework that we can cancel two negation symbols when they appear right next to each other.

Let's keep negating:

P	$\neg P$
There exists a finite subset of X	There does not exist a finite subset of X
or	For every subset of X, it is not finite.

Interesting – when we negate a "there exists" statement, we get a "for every" statement.

Let's keep negating:

[10] There exist frameworks of logic that *explicitly reject this*, but classical logic accepts it and so will we in this class.

P	$\neg P$
Every natural number is divisible by a prime number	Not every natural number is divisible by a prime number
or	There exists a natural number which is not divisible by a prime number
or more nicely written:	There exists a natural number which is not divisible by *any* prime number

Same deal – negating an "every" statement gets us a "there exists" statement. Finally:

P	$\neg P$
Every subset of an infinite set is infinite	Not every subset of an infinite set is infinite.
or	There exists a finite subset of an infinite set.

2.1. **"And" and "or".** We can combine statements with the words "and" and "or" to get new statements.

Notation 2.1. Given two statements P and Q we write $P \wedge Q$ to mean "P *and* Q."

Question: How does the truth of $P \wedge Q$ depend on the truth of P and the truth of Q?

Let's consider an example:

$$P = \text{“it is Monday”}$$
$$Q = \text{“it is raining”}.$$

Then

$$P \wedge Q = \text{“it is Monday } \textbf{and} \text{ it is raining”}$$

Let's consider the four possibilities for P being true and false and Q being true and false.

it is Monday	it is raining	$\Rightarrow$	$P \wedge Q$ is true
it is Monday	it is not raining	$\Rightarrow$	$P \wedge Q$ is false
it is not Monday	it is raining	$\Rightarrow$	$P \wedge Q$ is false
it is not Monday	it is not raining	$\Rightarrow$	$P \wedge Q$ is false.

We can encode this more concisely in a *truth table*:

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F.

Definition 2.2. If P and Q are statements, we denote by $P \vee Q$ the new statement "P **or** Q."

Note that $P \vee Q$ will be true if at least one of P and Q are true.

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

2.2. **Conditional statements.**

Definition 2.3. Given two statements P and Q, we write $P \Rightarrow Q$ to mean "P **implies** Q." In other words, whenever P is true, this implies Q must be true as well. We might also phrase this as "if P then Q."

This one is a little weirder. It's still a statement, so we can input truth values for P and Q and get out a truth value for $P \Rightarrow Q$.

Hammack talks about this as a "promise." Explicitly, the statement $P \Rightarrow Q$ is the *promise* that any time P is true, then Q will also be true. The truth value of $P \Rightarrow Q$ refers to whether or not the promise was broken.

Example 2.4. Let P be the statement "you pass the exam" and Q be the statement "you pass the class." The statement $P \Rightarrow Q$ could be a promise the professor makes to the students: "**if** you pass the exam **then** you pass the class."

In this case truth values of P and Q could be different scenarios that could play out:

you passed the exam	you passed the class	$\Rightarrow$	cool!
you passed the exam	you didn't pass the class	$\Rightarrow$	the promise was broken
you didn't pass the exam	you passed the class	$\Rightarrow$	cool, the promise wasn't broken
you didn't pass the exam	you didn't pass the class	$\Rightarrow$	the promise wasn't broken

We can write this as a truth table

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Definition 2.5. The *converse* of a statement $P \Rightarrow Q$ is the statement $Q \Rightarrow P$. These are *not equivalent statements.*

2.3. **If and only if.**

Definition 2.6. We write $P \Leftrightarrow Q$ to mean P **if and only if** Q. It is a shorthand for $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$. It is a promise that whenever P is true then Q will be true **and** whenever Q is true then P will be true.

We can encode this in a truth table

P	Q	$P \Leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

Roughly speaking, $P \Leftrightarrow Q$ means that P and Q are the same statement — the truth of one is equivalent to the truth of the other.

2.4. **Quantifiers.**

Definition 2.7. Let X be a set, and let $P(x)$ be a statement (not necessarily true or false) that can be made about any element $x \in X$.

Example 2.8. Let $X = \mathbb{N}$, then $P(x)$ could be any statement you can make about natural numbers, for example "x is even" or "x is prime."

How would you say $P(x)$ is true *for every* $x \in X$? We could try to list out the elements of X in some order, i.e.

$$X = \{x_1, x_2, x_3, \ldots\},$$

and then we could write

$$P(x_1) \wedge P(x_2) \wedge P(x_3) \wedge \cdots$$

This is cumbersome notation, and as we'll soon see we sometimes can't even order the elements in X like that. So we want a shorthand.

Definition 2.9. The symbol $\forall$ means "for all." We use it in the following way:

$$\forall x \in X,\ P(x)$$

this means "for all $x \in X$, $P(x)$ is true."

Similarly, we could make the statement "there exists an $x \in X$ for which $P(x)$ is true." This means either $P(x_1)$ is true, or $P(x_2)$ is true, or $P(x_3)$ is true,... so we could write this as

$$P(x_1) \vee P(x_2) \vee P(x_3) \vee \cdots$$

Again we bump into the same issue that this is cumbersome notation.

Definition 2.10. The symbol $\exists$ means "there exists." In other words

$$\exists x,\ P(x)$$

means "there exists an $x \in X$ for which $P(x)$ is true."

Terminology 2.11. The symbols $\forall$ and $\exists$ are called *quantifiers.*

Example 2.12. We've seen these in calculus — given a function $f \colon \mathbb{R} \to \mathbb{R}$, the statement "$f$ is *continuous* at x_0" is shorthand for the statement

$$\forall \epsilon > 0 \exists \delta > 0 \left(|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon\right).$$

3. On implication and proof

Given statements $P \Rightarrow Q$, it is either a true or false statement, and this doesn't depend on the truth of P or Q. As an example let $P(n)$ be "n is divisible by 10" and let $Q(n)$ be "n is even." Then we claim

$$P(n) \Rightarrow Q(n)$$

is a true statement. It's totally fine that P can be false (look at $P(13)$). What matters is the implication.

Question 3.1. Suppose we know P is true and $P \Rightarrow Q$ is true. What can we say about Q?

Answer 3.2. We can deduce that Q must be true!

This type of reasoning is called *deduction* or *logical inference.* We usually write it as

$$\begin{array}{c} \text{statement} \\ \text{statement} \\ \vdots \\ \text{statement} \\ \hline \text{conclusion} \end{array}$$

The logical rule we just outlined is called *modus ponens*

$$\begin{array}{c} P \Rightarrow Q \\ P \\ \hline Q \end{array}$$

We have two more important rules of deduction.

Definition 3.3. *Modus tollens* is the deduction

$$\begin{array}{c} P \Rightarrow Q \\ \neg Q \\ \hline \neg P \end{array}$$

If P implies Q, and Q is false, then P cannot be true.

Definition 3.4. *Elimination* is the deduction

$$\begin{array}{c} P \vee Q \\ \neg P \\ \hline Q \end{array}$$

If P or Q is true, and P is false, then Q must be true.

Definition 3.5. A *theorem* is a mathematical statement that is true, and a *proof* is a line of deduction that demonstrates its truth from other statements that are known to be true. A *lemma* is a mathematical statement proved on the way to proving a theorem. A *corollary* is a result that follows from the statement of a theorem. A *proposition* is another word for a mathematical statement that is asserted to be true, often a smaller or more obvious result than a lemma or theorem.[11]

How do we prove a mathematical statement? There are lots of different ways to prove something, and it depends upon how the statement is phrased. We'll go through some proof styles and give examples.

3.1. **Direct proof.** Suppose we want to prove a proposition of the following form.

Proposition 3.6. If P then Q.

This doesn't mean that either P or Q are necessarily true, it just means that P will *imply* Q (symbolically, $P \Rightarrow Q$).

A valid line of reasoning here is called *direct proof.* We suppose that P is true, we carry out some logical inference, and we arrive at the conclusion that Q is true.

Proof. Suppose P. Then … therefore Q. □

Definition 3.7. We say a number $n \in \mathbb{N}$ is *odd* if $n = 2k + 1$ for some $k \in \mathbb{N}$. We say $n \in \mathbb{N}$ is *even* if it is of the form $n = 2k$ for some $k \in \mathbb{Z}$.

Proposition 3.8. If x is odd then x^2 is odd.

We can start by filling out the start and bottom of the proof:

[11] The lines between "proposition"/"lemma"/"theorem" are blurry and often very subjective.

Proof. Suppose x is odd.

...

Therefore x^2 is odd. □

We should *use the definition.*

Proof. Suppose x is odd. Then $x = 2a + 1$ for some integer a, by definition of an odd number.

...

Then $x^2 = 2b + 1$. Therefore x^2 is odd. □

Now we need some line of reasoning to fill in the gaps. Here it comes from expanding x^2 in terms of a, and finding the right expression for b:

Proof. Suppose x is odd. Then $x = 2a + 1$ for some integer a, by definition of an odd number. We have that

$$x^2 = (2a + 1)^2 = 4a^2 + 4a + 1 = 2(2a^2 + 2a) + 1.$$

Let $b = 2a^2 + 2a$. Then $x^2 = 2b + 1$. Therefore x^2 is odd. □

3.2. **Proof by case.** Suppose I want to prove something is true for all elements in a set X. It might be easier to break X into two smaller sets $X = X_1 \cup X_2$, and do two proofs — prove the statement for elements in X_1 and the statements for elements in X_2.

For example, when proving things about natural numbers $n \in \mathbb{N}$, it can occasionally be helpful to break into two cases: when n is even and when n is odd.

Proposition 3.9. For any $n \in \mathbb{N}$, the number $n^2 + 3n + 1$ is odd.

Proof. First suppose n is even. Then $n = 2k$ for some $k \in \mathbb{N}$. Then we have that

$$n^2 + 3n + 1 = (2k)^2 + 3(2k) + 1 = 6k^2 + 6k + 1 = 2(3k^2 + 3k) + 1.$$

Letting $b = 3k^2 + 3k$, we have that $n^2 + 3n + 1 = 2b + 1$, so $n^2 + 3n + 1$ is odd.

Suppose n is odd. Then $n = 2k + 1$ for some $k \in \mathbb{Z}$. Then we have that

$$\begin{aligned} n^2 + 3n + 1 &= (2k + 1)^2 + 3(2k + 1) + 1 \\ &= (4k^2 + 4k + 1) + (6k + 3) + 1 \\ &= 4k^2 + 10k + 5 \\ &= 2(2k^2 + 5k + 2) + 1. \end{aligned}$$

Letting $b = 2k^2 + 5k + 2$, we have that $n^2 + 3n + 1 = 2b + 1$, hence $n^2 + 3n + 1$ is odd. □

4. Contrapositive proof

Suppose we want to argue that P implies Q. As we have seen, in order to prove that $P \Rightarrow Q$ directly, we suppose P as a hypothesis, we carry out some logical deductions, and then we arrive at Q. Recall the truth table for $P \Rightarrow Q$:

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

We can compare this with the truth table for $(\neg Q) \Rightarrow (\neg P)$:

P	Q	$(\neg Q) \Rightarrow (\neg P)$
T	T	T
T	F	F
F	T	T
F	F	T

They are the same! What this means is that a proof that $P \Rightarrow Q$ is a valid proof that $(\neg Q) \Rightarrow (\neg P)$ and vice versa. This gives us a new style of proof we can try to carry out.

Goal: Prove that $P \Rightarrow Q$.

Strategy 1 (direct proof): Assume P as a hypothesis, carry out some reasoning, and show that you arrive at Q.

Strategy 2 (contrapositive proof): Assume $\neg Q$ as a hypothesis, carry out some reasoning, and show that you arrive at $\neg P$.

Let's see two examples.

Proposition 4.1. Let $x \in \mathbb{Z}$. If $9x + 5$ is even, then x is odd.

Here P is the statement "$9x + 5$ is even" while Q is the statement "x is odd."

Direct proof. Let's prove $P \Rightarrow Q$. Suppose $9x + 5$ is even. Then $9x + 5 = 2k$ for some integer $k \in \mathbb{Z}$. Subtracting $8x$ from each side, we get

$$x = 2k - 8x + 5.$$

Letting $b = k - 4x + 2$, we have that

$$x = 2b + 1,$$

hence x is odd. □

Contrapositive proof. Let's prove $\neg Q \Rightarrow \neg P$. Suppose x is not odd (meaning x is even). Then $x = 2a$ for some $a \in \mathbb{Z}$. Then

$$9x + 5 = 18a + 5 = 2(9a + 2) + 1.$$

Therefore $9x + 5$ is odd. □

Which proof do we prefer? They are both completely valid, but the second one seems a little cleaner. This is because it's easy to take info about x and turn it into info about $9x + 5$, but it's more cumbersome to go the other way around.

Proposition 4.2. Let $x \in \mathbb{Z}$. If $x^2 + 4x + 3$ is even, then x is odd.

Direct proof. Suppose $x^2 + 4x + 3$ is even, so

$$x^2 + 4x + 3 = 2k$$

for some k. Then... bleh. □

Contrapositive. Suppose x is even, then $x = 2n$ for some $n \in \mathbb{Z}$. Then

$$x^2 + 4x + 3 = (2n)^2 + 4(2n) + 3 = 4n^2 + 8n + 3 = 2(2n^2 + 4n + 2) + 1.$$

Letting $c = 2n^2 + 4n + 2$, we have that $x^2 + 4x + 3 = 2c + 1$, so $x^2 + 4x + 3$ is odd. □

Proposition 4.3. Let $x, y \in \mathbb{Z}$. If $3 \mid (xy)$ then $3 \mid x$ or $3 \mid y$.

What is the contrapositive of this statement?

- P is the statement $3 \mid (xy)$
- Q is the statement "$3 \mid x$ **or** $3 \mid y$"

The contrapositive is $(\neg Q) \Rightarrow (\neg P)$. The negation of Q is "$3 \nmid x$ *and* $3 \nmid y$."

4.1. **Beware the fallacy of the converse.** Consider the statement $P \Rightarrow Q$. We've defined

- its *converse*, which is $Q \Rightarrow P$
- its *contrapositive*, which is $\neg Q \Rightarrow \neg P$.

The statement $P \Rightarrow Q$ is *equivalent* to its contrapositive. That's why we can prove $P \Rightarrow Q$ by proving $(\neg Q) \Rightarrow (\neg P)$. Don't get the contrapositive and converse mixed up though.

Fallacy of the converse: We know $P \Rightarrow Q$ is true. Suppose Q, then we can conclude P.

5. Contradiction, if and only if

Suppose we want to prove a direct statement $P \Rightarrow Q$. We can look at the truth table, and we see there's only one way that $P \Rightarrow Q$ could fail to hold, namely if P is true and Q is false:

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

So let's show this possibility *can never occur*. How would we do that without a direct proof?

Suppose someone walks up to you and says "I have this mathematical proposition P, and if P is true, then it logically follows that 2 is odd." You might say, "good for you, but whatever P is, it can't be true because 2 isn't odd, so I'm going to conclude that P is false." That's the point of contradiction! We're going to show something is false by using it to arrive at some kind of contradiction.

Definition 5.1. To prove $P \Rightarrow Q$ *by contradiction*, we assume P *and* we assume $\neg Q$, and then we carry out some formal deduction to arrive at a contradiction, i..e we show some other statement R and its converse $\neg R$ are both true.

Proposition 5.2. If $x + y$ is odd, then x is odd or y is odd.

Proof by contradiction. Suppose $x + y$ is odd, and suppose towards a contradiction that both x and y are even. Then $x + y$ is even, contradicting that $x + y$ is odd. □

We assumed P and $\neg Q$, and we arrived at $\neg P$. Since both P and $\neg P$ can't be true, we must have that $P \wedge \neg Q$ is false. Hence $P \Rightarrow Q$ is true.

Definition 5.3. A natural number $n \geq 2$ is *prime* if it is its only divisors are 1 and itself.

Remark 5.4. The number 1 isn't prime by convention.

Every integer factors *uniquely* as a product of primes:

$$12 = 2^2 \cdot 3$$

$$51840 = 2^7 \cdot 3^4 \cdot 5.$$

So just as all molecules are built of atoms, all integers are built out of prime numbers. They are the *building blocks* of numbers.

Lemma 5.5. Suppose p is a prime number and $n \in \mathbb{N}$. Then we cannot have both $p \mid n$ and $p \mid n+1$.

Theorem 5.6 (Euclid, 300BC)**.** There are infinitely many prime numbers.

Proof. Suppose towards a contradiction there were only finitely many. Write them out as $p_1, \ldots, p_k$. Let

$$N = p_1 \cdot p_2 \cdots p_k + 1.$$

Then by the lemma, N is not divisible by any of the primes $p_1, \ldots, p_k$. However every number decomposes uniquely into primes — this means that N is divisible by some prime *not on our list.* This contradicts that $p_1, \ldots, p_k$ were the only prime numbers. □

Definition 5.7. A number is *rational* if it is of the form $\frac{a}{b}$ for $a, b \in \mathbb{Z}$ and $b \neq 0$. A number is *irrational* if it is not of this form.

Observe that we can always write a rational number in *reduced* form, which means a and b have no common multiples. For instance $\frac{16}{12}$ isn't reduced, since the top and bottom share a factor of 4, but we can write it in a reduced form as $\frac{4}{3}$.

Lemma 5.8. (Exercise on the homework) Let $a \in \mathbb{Z}$. If a^2 is even, then a is even.

Theorem 5.9. The quantity $\sqrt{2}$ is not rational.

Proof. Suppose towards a contradiction that $\sqrt{2}$ was rational. Then we can write $\sqrt{2} = \frac{a}{b}$, and we can assume this fraction is reduced (so that a and b have no common factors). Then

$$2 = \frac{a^2}{b^2},$$

so we have that

$$2b^2 = a^2.$$

This means that a^2 is even. By the lemma this implies a is even, so we can write $a = 2k$ for some $k \in \mathbb{Z}$. Let's expand the equation above:

$$2b^2 = a^2 = (2k)^2 = 4k^2.$$

We can divide by a 2 on both sides, and we get

$$b^2 = 2k^2.$$

This implies b^2 is even, which by the lemma implies b is even. So a and b are both even, meaning they are both divisible by 2. This contradicts our assumption that a/b was a reduced fraction. □

6. Modular arithmetic

When we write $22/7 = 3\frac{1}{7}$, we are really using the fact that

$$22 = 3 \cdot 7 + 1.$$

Here this 1 is the *remainder* when we divide 22 by 7. We're going to state a theorem that shows such an expression is always unique.

Theorem 6.1 (Euclidean Division Algorithm)**.** Let n and $d \geq 1$ be integers. Then there exist uniquely determined integers q and r so that

$$n = qd + r$$

for $0 \leq r < d$.

Proof. We are given n and d and want to find q and r. In particular we want r to be as small as possible, so we want to minimize the value $n - qd$, while still keeping it non-negative. To that end, let's take the set

$$X = \{n - td \colon t \in \mathbb{Z},\ n - td \geq 0\}.$$

We first claim X is nonempty. This is true because if $n \geq 0$, then $n = n - 0d$ is in X, and if $n < 0$ then $n - nd = n(1-d)$ is in X.

Since X is nonempty we can let r be the smallest member of X. Then $r = n - qd$ for some $q \in \mathbb{Z}$. We still have to show that

(1) $0 \leq r < d$, and
(2) r and q are *uniquely determined*.

For the first step, suppose towards a contradiction that $r \geq d$. Then we can write

$$0 \leq r - d = n - (q+1)d.$$

Hence $r - d$ is in X, but $r - d < r$, contradicting the minimality of r. Hence we conclude that $0 \leq r < d$.

To show uniqueness, suppose that $n = q'd + r'$ with $0 \leq r' < d$. Let's first assume that $r \leq r'$. Then we have that

$$(q - q')d = r' - r \leq r' < d$$

so $r' - r$ is a nonnegative multiple of d which is less than d, which can only happen if $r' - r = 0$. Hence $r' = r$ and $q = q'$. The case where $r \geq r'$ can be done similarly. □

Definition 6.2. We say two integers a and b are *congruent modulo* n if $n \mid (a - b)$. In this case we write

$$a \equiv b \pmod{n}.$$

Corollary 6.3. Every integer a admits a unique *remainder* modulo n, which we denote by $\bar{a}$, for which $0 \leq \bar{a} < n$.

Proof. By the division algorithm, we have a unique expression

$$a = qn + \bar{a}$$

for some $0 \leq \bar{a} < n$. □

Let's define a new set

$$\mathbb{Z}/n\mathbb{Z} := \left\{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}\right\}.$$

We can add and multiply things in this set, and this is called *modular arithmetic*.[12]

Example 6.4. Working with $\mathbb{Z}/12$ is essentially what we do when we tell time – $\bar{9} + \bar{4} = \overline{13} = \bar{1}$ in $\mathbb{Z}/12$, since four hours past nine, it will be one o'clock.

We'll get into modular arithmetic a lot more on the worksheet and after the midterm when we study group theory.

[12] Really what we're doing is taking the set of all $\bar{k}$ for $k \in \mathbb{Z}$, and then saying $\bar{k} = \overline{k+n} = \overline{k+2n} = \cdots$. This comes from something called an *equivalence relation* that we won't get to in this class.

6.1. **Mathematical induction.** Suppose we're bored in class and we start adding odd numbers starting at one:

$$\begin{gathered}1 = 1\\ 1+3 = 4 = 2^2\\ 1+3+5 = 9 = 3^2\\ 1+3+5+7 = 16 = 4^2.\end{gathered}$$

We start to see a pattern, and we might want to guess that it is always true:

Conjecture 6.5. The sum of the first n odd natural numbers equals n^2. In other words,

$$1 + 3 + \ldots + (2n-1) = n^2.$$

It's not obvious how to prove something like this – we will discuss a *new proof technique.*

Induction: Let $P(n)$ be a mathematical statement about $n \in \mathbb{N}$ that we want to prove for all $n \in \mathbb{N}$. In the example above, $P(n)$ might be the statement "the sum of the first n odd natural numbers equals n^2."

Suppose we can prove that

(1) $P(1)$ is true (the *base case*)
(2) $P(n)$ being true implies that $P(n+1)$ is true (the *inductive step*)

Then we claim we know that $P(n)$ is true for all $n \in \mathbb{N}$! Why is this true? Because we know $P(1)$ is true, and since $P(1)$ implies $P(2)$ we know that $P(2)$ is true as well, and since $P(2)$ implies $P(3)$, we know that $P(3)$ is true as well, and so on...

We think about this like dominoes falling.

So let's prove the conjecture!

Proof. **Base case**: If $n = 1$, then it is clear that $1 = 1^2$ so we are done.[13]

Inductive step: Suppose we know that

$$1 + 3 + \cdots + (2n-1) = n^2.$$

We want to argue that this implies that $1 + 3 + \ldots + (2n-1) + (2n+1)$ is equal to $(n+1)^2$. Let's expand:

$$(n+1)^2 = n^2 + 2n + 1.$$

We can *apply the inductive hypothesis* to plug in the value for n^2, and we get

$$\begin{aligned}(n+1)^2 &= n^2 + 2n + 1\\ &= (1 + 3 + \ldots + (2n-1)) + (2n+1).\end{aligned}$$

Hence $(n+1)^2$ is equal to the sum of the first $n+1$ odd numbers, and we are done! □

Proposition 6.6. For every $n \geq 4$, we have that $2^n < n!$.

Proof. **Base case**: If $n = 4$, then $2^4 = 16 < 24 = 4!$.

Inductive step: Suppose $2^n < n!$. Then

$$2^{n+1} = 22^n < 2n! < (n+1)n! = (n+1)!,$$

since $2 < n+1$ for $n \geq 2$ (in particular for $n \geq 4$). □

[13] Base cases can often be the easiest part of a proof by induction, but they are necessary to include!

Strong induction: In trying to prove $P(n)$ implies $P(n+1)$ we might sometimes want to assume not just that $P(n)$ holds but that $P(k)$ holds for all $k \leq n$. This is also a valid proof technique called *strong induction*

(1) $P(1)$ is true (base case)
(2) $P(1) \wedge P(2) \wedge \cdots \wedge P(n)$ implies $P(n+1)$ (strong inductive step).

Proposition 6.7. Every integer $n \geq 2$ is a product of primes.

Proof. **Base case**: $n = 2$ is itself a prime.

Inductive step. Suppose we know every integer $k \leq n$ is a product of primes. We want to show $n+1$ is a product of primes. If $n+1$ is itself a prime, then we are done. If not, then it factors as

$$n+1 = ab,$$

where $2 \leq a, b < n$. Then each of a and b factor as a product of primes by the inductive hypothesis, hence $n+1$ is a product of all the primes in a and in b. □

Proposition 6.8. We have that $5^{2n} \equiv 1 \pmod{24}$ for each $n \geq 0$.

Proof. We prove by strong induction. Suppose we know that $5^{2k} \equiv 1 \pmod{24}$ for each $k \leq n$. We want to prove it for $n+1$. We can write

$$5^{2(n+1)} - 1 = 5^{2n+2} - 1 = (5^{n+1} - 1)(5^{n+1} + 1).$$

By strong induction, $5^{n+1} \equiv 1 \pmod{24}$, hence $24 \mid (5^{n+1} - 1)$, and hence 24 divides the entire product above. □

7. Functions

Given two sets X and Y, a *function* from X to Y is a way to take as input elements in X and output elements in Y. We write $f\colon X \to Y$ to say "f is a function from X to Y."

Notation 7.1. We write $f(x) = y$ to mean that, for the function $f\colon X \to Y$, x is sent to y. We might also write $x \mapsto y$. This arrow $\mapsto$ is
`mapsto` in LaTeX, and means x gets "mapped to" some value by a function.

[draw some functions, see which are valid]

Q: What are some properties that functions can have?

Definition 7.2. A function $f\colon X \to Y$ is *surjective* or *onto* if every element $y \in Y$ is mapped to by some $x \in X$. More explicitly, for every $y \in Y$ there exists some $x \in X$ (maybe more than one) so that $f(x) = y$.

Question 7.3. If X and Y are finite, and $f\colon X \to Y$ is some surjective function, what can we say about the *sizes* of X and Y?

Definition 7.4. A function $f\colon X \to Y$ is *injective* or *one-to-one* if no two elements in X map to the same element in Y. In other words if $a, b \in X$ and $a \neq b$ then $f(a) \neq f(b)$.

Giving a direct proof of injectivity using this definition is a little hard when X is a very big set. Easier is to prove the contrapositive! That is, you suppose $f(a) = f(b)$ for some $a, b \in X$ and then argue that $a = b$.

Question 7.5. If X and Y are finite, and $f\colon X \to Y$ is some *injective* function, what can we say about the *sizes* of X and Y?

Definition 7.6. We say a function $f\colon X \to Y$ is *bijective* if it is both injective and surjective.

Question 7.7. If X and Y are finite, and $f\colon X \to Y$ is some *bijective* function, what can we say about the *sizes* of X and Y?

The following example is goofy?

Proposition 7.8. (From Hammack, p.234) There are two people in the state of Texas with the same number of hairs on their head.

Proof. There exists some function

$$\{\text{people in Texas}\} \to \mathbb{N},$$

counting the number of hairs on every Texan's head. Biology tells us every human has less than 1 million hairs on their head, so let's actually revise this function to write

$$\{\text{people in Texas}\} \to \{1, 2, 3, 4, \ldots, 1 \text{ million}\}.$$

And the population of Texas is around 31.29 million (in 2024), so the size of the set on the left is much bigger than the size of the set on the right! If this function were injective, then because both sets are finite, the size of the set on the left would have to be smaller than the size of the set on the right. Hence this function is not injective! Meaning there are two people in Texas with the same number of hairs on their head. □

This is what's known as the *pigeonhole principle*

Pigeonhole principle: If X and Y are finite sets and $f\colon X \to Y$ is any function, then

(1) If $|X| > |Y|$ then f is not injective
(2) If $|X| < |Y|$ then f is not surjective.

Example 7.9. Most things you've seen in calculus are *functions* from $\mathbb{R} \to \mathbb{R}$. We write $f(x) = \cos(x)$ for instance to refer to the function $\mathbb{R} \to \mathbb{R}$ sending $x \mapsto \cos(x)$.

Example 7.10. The function

$$\begin{aligned} f\colon \mathbb{R} \smallsetminus \{0\} &\to \mathbb{R} \\ x &\mapsto \frac{1}{x} + 1 \end{aligned}$$

is injective but not surjective.

Proof. Suppose we have $a, b \in \mathbb{R} \smallsetminus \{0\}$ so that $f(a) = f(b)$. That is,

$$\begin{aligned} \frac{1}{a} + 1 &= \frac{1}{b} + 1 \\ \frac{1}{a} &= \frac{1}{b}. \end{aligned}$$

Inverting we get $a = b$. It is not surjective because 1 is not hit. □

7.1. **Bijectivity and size.** We have seen at least for finite sets that the existence of a bijection $X \to Y$ means that X and Y have the same size. To that end, bijections are a great way to talk about the *size* of two sets. If you can prove that no bijection exists between X and Y then they must have different sizes.

What about infinite sets?

Definition 7.11. We say a set X is *countable* or *countably infinite* if there exists a bijection $f\colon \mathbb{N} \to X$.

We say this because this means all the elements of X can be labeled as $f(0), f(1), f(2), f(3)$, etc. In other words they can be *counted.*

Question 7.12. Does there exist a bijection $\mathbb{N} \to [0,1]$? Certainly there exist many injections, one of which is given by sending $n \mapsto 1/n$ and $0 \mapsto 0$. This means that $|\mathbb{N}| \leq |[0,1]|$.

Theorem 7.13 (Cantor diagonalization)**.** *There exists no bijection*

$$\mathbb{N} \to [0,1].$$

Proof. Suppose towards a contradiction that there did exist some bijection, call it f. Then we will derive a contradiction by explicitly constructing an element $w \in [0,1]$ which is *not of the form* $f(n)$ for any $n \geq 0$. Here's how we do this – we write out the decimal expansion of each value of f.

$$\begin{aligned}
f(0) &= 0.373294170300213890021...\\
f(1) &= 0.721092380414890237298...\\
f(2) &= 0.139174912828949921322...\\
f(3) &= 0.775999030321390321321...\\
f(4) &= 0.389020391888329923021...\\
&\vdots
\end{aligned}$$

We then go through and build a new number w whose decimal expansion is different than $f(n)$ at level n.

$$\begin{aligned}
f(0) &= 0.\mathbf{3}73294170300213890021...\\
f(1) &= 0.7\mathbf{2}1092380414890237298...\\
f(2) &= 0.13\mathbf{9}174912828949921322...\\
f(3) &= 0.772\mathbf{5}99903032139032132...\\
f(4) &= 0.3891\mathbf{0}20391888329923021...\\
&\dots\\
w &= 0.\mathbf{43061}\dots
\end{aligned}$$

□

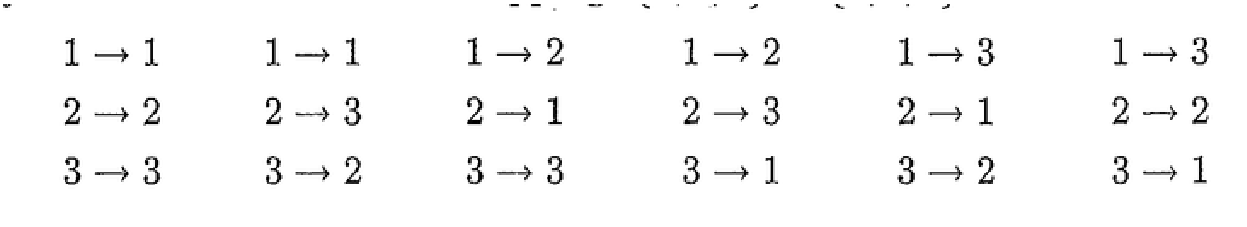
What does this tell us? $\mathbb{N}$ and $[0,1]$ are clearly both infinite sets, but one of them is *strictly smaller* than the other. We might say $[0,1]$ is *uncountably infinite.* We have proven that there exist different sizes of infinity!

8. Permutations

A *permutation* is a way to rearrange sets.

Definition 8.1. A *permutation* of a set X is any bijection from a set X to itself.

Example 8.2. Let $n = 3$, then we can study all the permutations from $\{1,2,3\}$ to itself:



$1 \to 1$	$1 \to 1$	$1 \to 2$	$1 \to 2$	$1 \to 3$	$1 \to 3$
$2 \to 2$	$2 \to 3$	$2 \to 1$	$2 \to 3$	$2 \to 1$	$2 \to 2$
$3 \to 3$	$3 \to 2$	$3 \to 3$	$3 \to 1$	$3 \to 2$	$3 \to 1$

Notation 8.3. We denote by Σ_n the set of permutations of $\{1,2,3,\dots,n\}$. In other words

$$\Sigma_n := \{f\colon \{1,2,\dots,n\} \to \{1,2,\dots,n\} : f \text{ is bijective}\}.$$

What's a concise way to write a permutation? We can write the numbers $\{1,\dots,n\}$ across the top row, and where each of them is mapped to along the bottom row. Then our 6 permutations above can all be written as

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}.$$

Definition 8.4. If $\sigma,\tau\in\Sigma_n$, we define their *multiplication* $\sigma\cdot\tau$ to be the permutation sending $k\in\{1,\dots,n\}$ to $\sigma(\tau(k))$.

Example 8.5. If $\sigma=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$ and $\tau=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$, then

$$\begin{aligned}(\sigma\cdot\tau)(1)&=\sigma(\tau(1))=\sigma(2)=1\\(\sigma\cdot\tau)(2)&=\sigma(\tau(2))=\sigma(1)=3\\(\sigma\cdot\tau)(3)&=\sigma(\tau(3))=\sigma(3)=2.\end{aligned}$$

So we have that

$$\sigma\cdot\tau=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$$

We could also see this by stacking τ on top of σ!

Let $e=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$. This is called the *identity permutation* because it doesn't do anything.

Note that $e\cdot\sigma=e$ and $\sigma\cdot e=\sigma$. This should remind us of multiplying by 1.

We say another element τ in Σ_n is the *inverse* of σ if $\tau\sigma=e$ and $\sigma\tau=e$.

8.1. **Cycles.**

Consider

$$\sigma=\begin{pmatrix}1&2&3&4&5&6\\1&5&6&4&3&2\end{pmatrix}.$$

What does σ do if we repeat it over and over? It cycles $2\mapsto5\mapsto3\mapsto6\mapsto2\mapsto\cdots$, and 1 and 4 stay fixed.

For these kinds of permutations, we have some more concise notation.

Definition 8.6. If $\{k_1,\dots,k_r\}\subseteq\{1,\dots,n\}$, we write

$$\sigma=(k_1\ \dots\ k_r)$$

for the permutation in Σ_n satisfying

▷ $\sigma(k_i)=k_{i+1}$ if $1\le i<r$
▷ $\sigma(k_r)=k_1$
▷ $\sigma(k)=k$ if $k\notin\{k_1,\dots,k_r\}$

So in *cycle notation*, we have that

$$\begin{pmatrix}1&2&3&4&5&6\\1&5&6&4&3&2\end{pmatrix}=(2\ 5\ 3\ 6).$$

It doesn't matter where we start the notation, we have that

$$(2\ 5\ 3\ 6)=(5\ 3\ 6\ 2)=(3\ 6\ 5\ 2)=(6\ 5\ 2\ 3)\in\Sigma_6.$$

We would say σ is a 4-*cycle*.

Remark 8.7. Not every permutation is a cycle!

Theorem 8.8. *If $\sigma\in\Sigma_n$ is an r-cycle, then σ^{-1} is an r-cycle.*

The proof follows by arguing that

$$(k_1\ k_2\ \cdots\ k_r)^{-1} = (k_1\ k_r\ k_{r-1}\ \cdots\ k_2)$$

What about the following permutation?

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 1 & 6 & 7 & 5 \end{pmatrix}.$$

It's not a single cycle, but the *product of two cycles*. In other words, it can be thought of as the *product* of the two cycle permutations

$$(1\ 2\ 3\ 4) \cdot (5\ 6\ 7).$$

8.2. **Cycle composition.** What happens when we take a product of cycles? We now bump into some annoying notation, for instance consider:

$$(1\ 3\ 2\ 5) \cdot (5\ 7\ 4) \in \Sigma_7.$$

We have $\sigma = (1\ 3\ 2\ 5)$ and $\tau = (5\ 7\ 4)$, so the induced permutation will be

$$\sigma \cdot \tau = (1\ 3\ 2\ 5\ 7\ 4).$$

Definition 8.9. We say two cycles $(k_1\ \cdots k_r)$ and $(m_1\ \cdots\ m_s)$ in Σ_n are *disjoint* if $k_i \neq m_j$ for every $1 \leq i, j \leq n$.

Exercise 8.10. If σ and τ are disjoint cycles then $\sigma\tau = \tau\sigma$.

Definition 8.11. A *cycle decomposition* of a permutation $\sigma \in X_n$ is an expression for it as a product of disjoint cycles.

We should think of *cycles* as analogous to primes in some sense – just as every number factors uniquely into a product of primes, we want to see that every permutation factors uniquely as a product of disjoint cycles.

Theorem 8.12 (Cycle decomposition theorem)**.** If $\sigma \in \Sigma_n$ is a non-identity permutation, then σ is a product of (one or more) disjoint cycles of length at least two. This factorization is unique up to the order of the factors.

Proof. If $n = 2$, we're done, because there is only one non-identity permutation, which is $(1\ 2)$, clearly a cycle.

We proceed by induction. Suppose we've proven it for S_{n-1}, and take $\sigma \in S_n$. If $\sigma(n) = n$, then σ lies in S_{n-1}, and we are done! Since σ is a bijection, it sends *something* to n, and we've assumed this something isn't n itself, so there is some $1 \leq m < n$ so that $\sigma(m) = n$. Let's write $\gamma = (m\ n)$ and define $\tau = \sigma\gamma$. Observe that

$$\tau(n) = \sigma(\gamma(n)) = \sigma(m) = n.$$

So $\tau \in S_{n-1}$, so τ is a product of disjoint transpositions. Note also that

$$\tau\gamma = \sigma\gamma^2 = \sigma \cdot e = \sigma,$$

since $\gamma^2 = e$. We now consider two cases:

Case 1: $\tau(m) = m$. Then $\gamma = (m\ n)$ and τ are disjoint, and $\sigma = \tau\gamma$ is a product of disjoint cycles!

Case 2: $\tau(m) \neq m$. Then m is moved by exactly one cycle factor of τ, so we can write

$$\tau = \mu \cdot (m\ k_1\ \cdots\ k_r)$$

for some k_i's, where μ is a product of disjoint cycles fixing $k_1, \ldots, k_r$ and also fixing n. Then

$$\sigma = \tau\gamma = \mu(m\ k_1\ \cdots\ k_r)(m\ n) = \mu(m\ n\ k_1\ \cdots\ k_r).$$

We now have to show uniqueness, which is on the homework! □

9. Monoids and groups

We've seen that the set Σ_n of bijections $\{1, \dots, n\}$ can be given a *multiplication*, defined by composing permutations. This has an identity element, has inverses, etc.

We'd like to abstract this a bit.

Definition 9.1. A *binary operation* $\star$ on a set X is a function

$$X \times X \to X.$$

We denote by $\star(x_1, x_2)$ as $x_1 \star x_2$.

Example 9.2. Addition and multiplication are binary operations on $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, etc.

Definition 9.3. We say a binary operation $\star$ on a set X is *associative* if

$$(x_1 \star x_2) \star x_3 = x_1 \star (x_2 \star x_3)$$

for every $x_1, x_2, x_3 \in X$.

Every binary operation we will deal with in this class will be associative, but there exist interesting operations which aren't! (look up the *octonions* if you want a particular example).

Definition 9.4. If $(X, \star)$ is a set with a binary operation, we say it has a *unit*, or it is *unital* if there exists an element $e \in X$ for which

$$e \star x = x = x \star e$$

for every $x \in X$.

Theorem 9.5. If $(X, \star)$ is unital, then that unit is unique.

Proof. Suppose $e_1, e_2 \in X$ are both units for $\star$. Then

$$e_1 = e_1 \star e_2 = e_2,$$

by unitality. □

Definition 9.6. We say a set with a binary operation $(X, \star)$ is a *monoid* if it is associative and unital.

Definition 9.7. We say a monoid $(X, \star)$ is *commutative* if $x \star y = y \star x$ for every $x, y \in X$.

Example 9.8. The set of natural numbers with $(\mathbb{N}, +)$ is a commutative monoid[14] with binary operation given by addition

$$\begin{aligned} \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ (a, b) &\mapsto a + b. \end{aligned}$$

Proof. We first verify addition is well-defined, this is true since the addition of any two natural numbers is also a natural number. We also check that addition is associative, which is true. Finally, we see $0 + n = n = n + 0$ for any $n \in \mathbb{N}$, and that $x + y = y + x$, so we have a commutative monoid. □

Example 9.9. The set Σ_n of permutations is a monoid (not necessarily commutative)!

Proof. We verify the composition of two permutations is indeed a permutation, so the binary operation of permutation composition is well-defined. Permutation composition is associative, as we can check, and the identity permutation is an identity element in the monoid sense. □

Example 9.10. If X is a set, then $\mathcal{P}(X)$ is a monoid under union.

[14] This is kind of *why* we want to include 0 as a natural number.

Example 9.11. If X is a set, then $\mathcal{P}(X)$ is a monoid under intersection.

Example 9.12. We have that $(\mathbb{Z}, \cdot)$ is a monoid under multiplication. The unit is 1.

Example 9.13. For the CS people – come up with monoid operations on certain data structures!

If we have a binary operation on a set, we can *check* whether it is a monoid or not. In this sense the binary operation is *structure*, whereas being a monoid is a *property* that is either satisfied or failed by that structure.

Question: How many binary operations exist on a finite set X?

If X has n elements, then we are asking how many functions

$$X \times X \to X$$

exist. Note that $X \times X$ has n^2 elements, and for each of these we have n choices of where it can be sent in X. So altogether there are exactly n^{n^2} binary operations on a set X. Let's look at how fast these grow

n	1	2	3	4
n^{n^2}	1	16	$19,683$	over 4 billion

Question: How many of these are monoids?

Answer: Wide open.

This question is a little too hard. It turns out if we impose one more condition, this type of problem becomes more approachable.

Definition 9.14. Let $(X, \star, e)$ be a monoid, and let $x \in X$ be an arbitrary element. We say $y \in X$ is an *inverse* to x if

$$x \star y = e, \text{ and } y \star x = e.$$

Theorem 9.15. If $x \in X$ is an element in a monoid which has an inverse, then that inverse is unique.

Proof. Suppose both y and z were inverses to x. Then we can write

$$y = y \star e = y \star (x \star z) = (y \star x) \star z = e \star z = z.$$

So $y = z$. □

Definition 9.16. We say a monoid $(X, \star, e)$ is a *group* if each element admits an inverse.

Example 9.17. $(\mathbb{Z}, +)$ is a group with identity 0, and the inverse to x given by $-x$.

Example 9.18. $(\mathbb{N}, +)$ is not a group, since elements don't have inverses under addition.

Example 9.19. The nonzero rational numbers $(\mathbb{Q} \setminus \{0\}, \cdot)$ form a group under multiplication.

Example 9.20. The set Σ_n of permutations of the set $\{1, 2, \ldots, n\}$ form a group. This is a really important example – we call it the *symmetric group.*

Definition 9.21. If X is a set with n elements, then a *Latin square* for X is an $n \times n$ grid where each element $x \in X$ appears exactly once on each row and column.

Example 9.22. For $X = \{a, b\}$, there are two Latin squares for X, which are

a	b
b	a

b	a
a	b

Given a group, we can write out its multiplication table, sometimes called a *Cayley table.* These differ from monoids in the following super key way!

Proposition 9.23. If $(X, \star)$ is a group, then its Cayley table is a Latin grid.

Proof. We prove it for cols and see that . is similar. Recall each entry is row$\star$column.

In the column for y, we claim that x appears once. This is because we can look at the row corresponding to $z = x \star y^{-1}$. Then

$$z \star y = (x \star y^{-1}) \star y = x \star (y^{-1} \star y) = x \star e = x.$$

We claim x doesn't appear two (or more) times in the column. Indeed suppose x appeared in the z_1 and z_2 rows in the column corresponding to y. Then we have that

$$x = z_1 y = z_2 y.$$

Multiplying on the right by y^{-1}, we get

$$z_1 = z_1 y y^{-1} = z_2 y y^{-1} = z_2.$$

So $z_1 = z_2$, and x is only in that row! □

Suppose we want to classify group structures on a set with n elements, this lets us dramatically reduce from the number of binary operations n^{n^2} down to the number of Latin grids!

n	1	2	3	4	$\cdots$
number of binary operations, i.e. n^{n^2}	1	16	$19,683$	over 4 billion	$\cdots$
number of $n \times n$ Latin squares	1	2	12	576	$\cdots$

There is no known formula for the number of Latin squares in terms of n. We know it is at least $\geq \frac{(n!)^{2n}}{n^{n^2}}$.

10. Homomorphisms and isomorphisms

$\circ$	e	a	b	c	d
e	e	a	b	c	d
a	a	e	d	b	c
b	b	c	e	d	a
c	c	d	a	e	b
d	d	b	c	a	e

then $(ab)c = dc = a$ and $a(bc) = ad = c$. not a group!

As we have seen last time, the *Cayley table* of any group is a Latin square.

Let's write down some 4×4 Latin squares:

a	b	c	d
b	c	d	a
c	d	a	b
d	a	b	c

b	c	d	a
c	d	a	b
d	a	b	c
a	b	c	d

a	b	c	d
b	a	d	c
c	d	a	b
d	c	b	a

how are these different? How are they similar?

Definition 10.1. Let G and H be groups. A function

$$f \colon G \to H$$

is called a *homomorphism* if

$$f(g_1 \cdot g_2) = f(g_1) \cdot f(g_2)$$

for every $g_1, g_2 \in G$.

Theorem 10.2. Let

$$f\colon G \to H$$

be a group homomorphism. Then

(1) $f(1_G) = 1_H$, in other words f preserves the unit
(2) $f(g^{-1}) = f(g)^{-1}$ for all $g \in G$
(3) $f(g^k) = f(g)^k$ for any $k \geq 1$.

Proof. For the first one

$$f(1_G) \cdot f(1_G) = f(1_G \cdot 1_G) = f(1_G) = f(1_G) \cdot 1_H$$

Cancelling in H, we get $f(1_G) = 1_H$.

For the second, we see that

$$f(g^{-1})f(g) = f(g^{-1}g) = f(1_G) = 1_H.$$

Therefore $f(g^{-1})$ is the inverse to $f(g)$ in H, so $f(g^{-1}) = f(g)^{-1}$, since we proved last week that inverses are unique.

For the last point, we induct! Clearly the $k = 1$ case is true. Suppose we have proven it for k and want to see it for $k + 1$, then

$$f(g^{k+1}) = f(g \cdot g^k) = f(g) \cdot f(g^k) = f(g) \cdot f(g)^k = f(g)^{k+1}.$$

So we're done. □

Definition 10.3. A homomorphism which is bijective is called an *isomorphism.*

Let $G = \{1, g, h\}$ be a set with three elements, and consider the binary operation given by the Cayley table

	1	g	h
1	1	g	h
g	g	h	1
h	h	1	g

We claim this is a group. Clearly 1 is the identity, g and h are inverses, and we can verify that multiplication is associative.

We can also define a group

$$H = \{a, b, c\}$$

with the Cayley table

	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

Is this a group? Why or why not?

Proposition 10.4. The bijection

$$\begin{aligned} f\colon G &\to H \\ 1 &\mapsto a \\ g &\mapsto b \\ h &\mapsto c \end{aligned}$$

is an *isomorphism.*

Proof. todo □

Given two groups G and H like those above, we write $G \cong H$ to mean they are *isomorphic.*

Studying groups *up to isomorphism* means we only consider two groups to be different if they are *not isomorphic.*

Theorem 10.5. Up to isomorphism, there is only one group structure on a set with two elements.

Proof. We define $G = \{e, g\}$ to be a set with two elements, where e is the identity, and $g^2 = e$. It is clear this is a group.

Let $X = \{x_1, x_2\}$ be a set with two elements, and suppose it has a group structure. Since its Cayley table must be a Latin grid, there are only two possibilities for its Cayley table, namely:

	x_1	x_2
x_1	x_1	x_2
x_2	x_2	x_1

or

	x_1	x_2
x_1	x_2	x_1
x_2	x_1	x_2

In the first case, we observe that x_1 is the identity, so we claim that

$$\begin{aligned} f \colon X &\to G \\ x_1 &\mapsto e \\ x_2 &\mapsto g \end{aligned}$$

is a bijection. Indeed we check

$$\begin{aligned} f(x_1 \cdot x_1) &= f(x_1) = e = e \cdot e = f(x_1) \cdot f(x_1) \\ f(x_1 \cdot x_2) &= .... \end{aligned}$$

□

The existence of group isomorphisms means *it doesn't matter what we call the elements of our set*! The only thing that matters is how they multiply, i.e. how their Cayley tables look.

Example 10.6. Recall there are 16 logic gates, corresponding to binary operations on the set $B = \{T, F\}$. Since there are only two 2×2 Latin squares, and both give group structures, we have that exactly two of them give group structures. These correspond to the following Cayley tables, which we will also write as truth tables:

	T	F
T	F	T
F	T	F

which is

P	Q	$P \oplus Q$
T	T	F
T	F	T
F	T	T
F	F	F

and

	T	F
T	T	F
F	F	T

which is

P	Q	$P \odot Q$
T	T	T
T	F	F
F	T	F
F	F	T

These notations $\oplus$ and $\odot$ are also called XOR and XNOR. In terms of basic operations, we see that $P \oplus Q$ means "P or Q but not both." In symbolic logic, this is:

$$P \oplus Q = (P \vee Q) \wedge \neg(P \wedge Q).$$

The exclusive nor operation XNOR is the negation of XOR. We can also think of it as "P and Q or not P and not Q." That is,

$$P \odot Q = (P \wedge Q) \vee (\neg P \wedge \neg Q).$$

These two groups are isomorphic, and since the unit must be sent to the unit under a group isomorphism, the function exhibiting the isomorphism is negation $\neg\colon B \to B$. The fact that negation is a group isomorphism is precisely a case of deMorgan's laws for exclusive or and exclusive nor:

$$\neg(P \oplus Q) = \neg P \odot \neg Q$$
$$\neg(P \odot Q) = \neg P \oplus \neg Q.$$

11. Subgroups

We saw on the homeworks some examples of groups living "inside" of other groups. Let's make this precise.

Definition 11.1. Let G be a group. A subset $H \subseteq G$ is called a *subgroup* if H is a group itself, with group operation defined by the group structure on G.

Example 11.2. If G is a group, then $G \subseteq G$ is a trivial subgroup of itself. Also the subset containing only the identity is a subgroup $\{1\} \subseteq G$.

Example 11.3. We have that $(\mathbb{Z}, +) \subseteq (\mathbb{Q}, +)$ is a subgroup.

Example 11.4. The even numbers form a subgroup of $(\mathbb{Z}, +)$. The odd numbers don't (because they're not closed under addition, *and* they don't have zero).

Example 11.5. We have that $(\{+1, -1\}, \cdot) \subseteq (\mathbb{Q} \smallsetminus \{0\}, \cdot)$ is a subgroup.

Example 11.6 (from HW5)**.** We have that

$$C_4 = \{\epsilon,\ (1\ 2\ 3\ 4),\ (1\ 3)(2\ 4),\ (1\ 4\ 3\ 2)\}, \text{ and}$$
$$K_4 = \{\epsilon,\ (1\ 2),\ (3\ 4),\ (1\ 2)(3\ 4)\}$$

are subgroups of Σ_4.

Theorem 11.7 (Subgroup test)**.** If G is a group, then a nonempty subset $H \subseteq G$ is a subgroup if and only if three things hold:

(1) H contains the unit, meaning $1_G \in H$
(2) H is closed under multiplication, meaning if $h_1, h_2 \in H \subseteq G$, then $h_1 \cdot h_2 \in H$.
(3) H contains inverses: if $h \in H$ then $h^{-1} \in H$, where h^{-1} means the inverse of h in G.

Proof. If H satisfies all the conditions, then it is closed under multiplication by (2), it has a unit by (1), has inverses by (3) and the group operation is associative since the group structure in G was associative.

Conversely, suppose H is a subgroup, we claim it satisfies all the conditions. For (1), we have that

$$1_H^2 = 1_H = 1_H \cdot 1_G,$$

so $1_H = 1_G$. Property (2) is part of the data of being a subgroup. For property (3) if h' is the inverse of h in H and h^{-1} is its inverse in G, then

$$h'h = 1_H = 1_G = h^{-1}h,$$

so by cancellation in G, we have that $h' = h^{-1}$. □

When H is finite, we have a much simpler test to be a subgroup!

Theorem 11.8 (Finite subgroup test)**.** Let G be a group, and $H \subseteq G$ a nonempty finite subset. Then H is a subgroup if and only if H is closed under the group operation.

Proof. The forwards direction is immediate, so we want to show that if H is closed under operation, then it is a subgroup. We will do this via the subgroup test — we will show that H contains the unit and contains inverses.

Let $h \in H$ be any element, and start taking powers of it:

$$\{h, h^2, h^3, h^4, \ldots\} \subseteq H.$$

Since H is finite, the set with the powers of H must be finite as well. This implies that $h^n = h^{n+m}$ for some $m, n \geq 1$. By cancellation in G, this means that $1 = h^m$. Since $h^m \in H$, this implies $1 \in H$.

Since $1 = h^m$, we have that $h(h^{m-1}) = 1$. In other words, $h^{m-1} = h^{-1}$ in G. Thus since $h^{m-1} \in H$, we have that $h^{-1} \in H$, so H is closed under taking inverses. □

Exercise 11.9. Determine all the subgroups of

$$(\mathbb{Z}/6, +) = \{0, 1, 2, 3, 4, 5\}.$$

We ask which subsets are closed under addition mod 6. There are two *trivial* cases, namely $\{0\}$ and the whole set. Are there any others?

Some other $H \subseteq \mathbb{Z}/6$ has to contain 0, and we notice that if it contains 1 it is closed under addition, so it contains $1 + 1 = 2$, and 3, and so on and we get the whole group. So let's assume it doesn't contain 1.

It could contain 2, and then it must also contain $2 + 2 = 4$. We claim that

$$\{0, 2, 4\}$$

is a subgroup!

Similarly, $\{0, 3\}$ is a subgroup. These are all the subgroups.

12. Groups via generators and relations

One elegant way to present groups is to say they are *generated* by elements and relations.

Definition 12.1. A group G is *generated* by some elements $a_1, a_2, \ldots, a_n \in G$ if every element in G can be expressed as a product of the a_i's and a_i^{-1}'s.

That is, we write

$$G = \langle \text{generators} | \text{relations} \rangle.$$

If you go to the group theory wiki to learn anything, this is how you'll see it presented!

Example 12.2. We can let

$$G = \left\langle a, b \colon a^3 = b^2 = 1,\ ab = ba^2 \right\rangle.$$

This means take the group given by all the powers of a, all the powers of b, all their inverses, and all products of stuff in them. The relations let us simplify long expressions in a's and b's. For instance

$$a^5 b^{-7} a^3 bab^{-1}$$

can be rewritten using the relations as

$$\begin{aligned} a^3a^2b^{-6}b^{-1}a^3bab^{-1} &= 1a^2(b^2)^{-3}b^{-1}1bab^{-1} \\ &= a^2b^{-1}bab^{-1} \\ &= a^2ab^{-1} \\ &= a^3b^{-1} \\ &= b^{-1} \\ &= b. \end{aligned}$$

We can prove that G is, as a set, equal to

$$\{1, a, b, a^2, ba, ba^2\}.$$

Definition 12.3. If a group has a single generator, it is called a *cyclic group.*

Example 12.4. We can consider the group

$$C_6 = \langle g \colon g^6 = 1 \rangle .$$

This has as elements $\{1, g, g^2, g^3, g^4, g^5\}$.

We can check that

$$\begin{aligned} \mathbb{Z}/6 &\to G \\ n &\mapsto g^n \end{aligned}$$

is a group isomorphism!

Notation 12.5. We will write C_n for the group

$$C_n = \langle g \colon g^n = 1 \rangle .$$

This is called a *cyclic group of order n.*

13. Orders of elements and groups

Definition 13.1. If G is a finite group, we say its *order* is its size. If G is an infinite group, we say it has *infinite order.*

Unfortunately the word "order" is overloaded – we use order to talk both about groups as well as their elements.

Definition 13.2. If $g \in G$, its *order*, denoted $o(g)$, is the smallest integer $k \geq 1$ so that $g^k = 1$, unless no such integer exists, in which case we say $o(g) = \infty$.

Example 13.3. The order of the unit element is always 1.

Example 13.4. If $o(g) = 2$ then $g^2 = 1$, meaning that $g = g^{-1}$. An element has order two if and only if it is its own inverse.

Example 13.5. Consider the group

$$\{1, -1, i, -1\},$$

under multiplication of complex numbers. Then $o(1) = 1$, and $o(-1) = 2$, and $o(i) = o(-i) = 4$.

Example 13.6. In $(\mathbb{Z}, +)$, the order of 1 is infinite, since $1 + \ldots + 1$ will never equal zero.

Example 13.7. If $\sigma \in \Sigma_n$ is a cycle of the form

$$\sigma = (a_1 \; a_2 \; \cdots \; a_r),$$

then $o(\sigma) = r$.

Theorem 13.8. If $g \in G$ has order n, then

(1) $g^k = 1$ if and only if $n \mid k$
(2) $g^a = g^b$ if and only if $a \equiv b \pmod{n}$
(3) All the elements $\{1, g, g^2, \dots, g^{n-1}\}$ are distinct, and they form a cyclic subgroup of G.

Proof. For the first point, if $n \mid k$ then $k = mn$ for some m, so $g^k = g^{mn} = (g^n)^m = 1^m = 1$. Conversely if $g^k = 1$, we use the division algorithm to write $k = qn + r$ for some $0 \leq r < n$. Then

$$1 = g^k = g^{qn+r} = (g^n)^q g^r.$$

So $g^r = 1$. If $r > 0$ then this contradicts n being the order of g. So we must have $r = 0$ and therefore $k \mid n$.

For the second part, if $g^a = g^b$, then $g^{a-b} = 1$, hence $n \mid (a - b)$ meaning that $a \equiv b \pmod{n}$, and vice versa.

Finally the division algorithm lets us show the final part. We can write any g^k as $g^{qn+r} = g^r$ where $0 \leq r \leq n - 1$. □

Corollary 13.9. Let $g \in G$, and let $\langle g \rangle \subseteq G$ be the cyclic subgroup generated by g. Then $|\langle g \rangle| = o(g)$. That is, this group is cyclic of order equal to the order of g.

This corollary is (probably) the reason for the conflation between the order of a group and the order of an element.

Question 13.10. If $g, h \in G$ have orders $o(g) = a$ and $o(h) = b$, what is the order of gh?

In general this is quite hard to answer, and there's no general formula, it really depends on the context!

Definition 13.11. We say g and h *commute* if $gh = hg$.

Example 13.12. Suppose g and h commute, and $o(g) = 2$ and $o(h) = 3$. Then

$$\begin{aligned}(gh)^2 &= g^2h^2 = h^2\\ (gh)^3 &= g^3h^3 = g\\ (gh)^4 &= g^4h^4 = h\\ (gh)^5 &= g^5h^5 = gh^2\\ (gh)^6 &= g^6h^6 = 1.\end{aligned}$$

So $o(gh) = 6$.

Example 13.13. In Σ_3, we have that $o((1\ 2)) = 2$ and $o((1\ 2\ 3)) = 3$. However

$$(1\ 2)(1\ 2\ 3) = (2\ 3),$$

which has order two. This is because these two elements *don't commute*:

$$\begin{aligned}(1\ 2)(1\ 2\ 3) &= (2\ 3)\\ (1\ 2\ 3)(1\ 2) &= (1\ 3).\end{aligned}$$

If cycles are *disjoint*, we have a nicer answer.

Definition 13.14 (lcm)**.**

(1) Given two positive natural numbers a and b, their *least common multiple*, denoted $\operatorname{lcm}(a, b)$, is the smallest positive integer k which is a multiple of a and b, meaning the smallest k for which $a \mid k$ and $b \mid k$.

(2) Given n positive natural numbers $d_1, \dots, d_n$, their least common multiple $\operatorname{lcm}(d_1, \dots, d_n)$ is the smallest positive integer k for which $d_i \mid k$ for each $1 \leq i \leq n$.

Proposition 13.15. If $d_1, \dots, d_n > 0$, and $k \in \mathbb{N}$ satisfies $d_i \mid k$ for each i, then $\operatorname{lcm}(d_1, \dots, d_n) \mid k$.

Proof. If $\operatorname{lcm}(d_1, \dots, d_n) = k$ then we're done. If not, then $\operatorname{lcm}(d_1, \dots, d_n) < k$, since lcm was the *least* common multiple. By the division algorithm we get

$$k = q \cdot \operatorname{lcm}(d_1, \dots, d_n) + r,$$

where $0 \leq r < \operatorname{lcm}(d_1, \dots, d_n)$. We can rewrite

$$r = k - q \cdot \operatorname{lcm}(d_1, \dots, d_n).$$

Since $d_i \mid k$ and $d_i \mid \operatorname{lcm}(d_1, \dots, d_n)$ for each i, this implies $d_i \mid r$. So r is a common multiple of the d_i's, and $r < \operatorname{lcm}(d_1, \dots, d_n)$. This implies that $r = 0$. □

Theorem 13.16. If $\sigma \in \Sigma_n$ has a disjoint cycle decomposition as

$$\sigma = \alpha_1 \alpha_2 \cdots \alpha_r,$$

with α_i a d_i-cycle, then $o(\sigma) = \operatorname{lcm}(d_1, \dots, d_r)$.

Proof. Write $\ell = \operatorname{lcm}(d_1, \dots, d_n)$ for simplicity. We claim that $o(\sigma) \mid \ell$ and $\ell \mid o(\sigma)$, and this will prove they are equal (since they are both positive integers).

To see that $o(\sigma) \mid \ell$, we will argue that $\sigma^\ell = 1$, after which we can use the previous theorem. Since each $d_i \mid \ell$, we can write $\ell = q_i d_i$ for some $q_i \in \mathbb{Z}$. Then

$$\sigma^\ell = (\alpha_1 \cdots \alpha_r)^\ell.$$

Since the α_i's commute, we can shuffle everything around and write the above as

$$\begin{aligned} &= \alpha_1^\ell \cdots \alpha_r^\ell \\ &= \alpha_1^{q_1 d_1} \alpha_2^{q_2 d_2} \cdots \alpha_r^{q_r d_r} \\ &= (\alpha_1^{d_1})^{q_1} (\alpha_2^{d_2})^{q_2} \cdots (\alpha_r^{d_r})^{q_r} \\ &= 1^{q_1} \cdots 1^{q_r} \\ &= 1. \end{aligned}$$

This proves $o(\sigma) \mid \ell$.

To show that $\ell \mid o(\sigma)$, it suffices to see that $o(\sigma)$ is a multiple of each d_i, since the lcm of the d_i's will divide any multiple of all the d_i's by that previous proposition. Again by the theorem on orders, we want to show that $\alpha_i^{o(\sigma)} = 1$ for any i and we will be done. We will argue for $i = 1$ and see it is true for the remaining ones by a similar argument.

If k is fixed by α_1, then it is also fixed by $\alpha_1^{o(\sigma)}$. If k is not fixed by α_1, then it *is* fixed by all the other α_i's. So we see that

$$\begin{aligned} k = \sigma^{o(\sigma)}(k) &= \alpha_1^{o(\sigma)} \cdots \alpha_r^{o(\sigma)}(k) \\ &= \alpha_1^{o(\sigma)}(k). \end{aligned}$$

Hence $\alpha_1^{o(\sigma)}$ fixes k no matter what, and therefore $\alpha_1^{o(\sigma)}$ is the identity. □

Remark 13.17. We really needed that we were talking about *permutations* here. This type of argument doesn't hold in general for groups (i.e. if $g_1, \dots, g_n$ pairwise commute in some group G, it is *not* true that the order of their product is the lcm of the orders of the g_i's. This is kind of a subtle question in group theory and requires a bit more machinery than we currently have to solve it even in the context of abelian groups).