Math 3310: Intro to Proofs

Stephen Peña

Contents

1 Introduction: Why proofs?

Many of us are familiar with a myriad of mathematical statements, and we may even be convinced of their veracity. Beliefs aside, how can we convince ourselves a statement is genuinely true? Most importantly, how can you convince your fellow mathematicians that something is true? The answer is by presenting a coherent mathematical argument involving already established facts. Such a coherent mathematical argument is called a *proof*.

Example 1.1. Suppose we know for a fact that 1+1=2. We are tasked with proving the following: $2 + 1 = 3$. One way we could proceed is as follows:

Proposition 1.1. *2+1=3.*

Proof. We have that 1+1=2. Note that $3 = 1 + 1 + 1$. Substituting in $1 + 1 = 2$, we have $3 = 2 + 1 = 1 + 2$. □

There are a few things to note about the above proof. First, note the use of complete sentences. Second, there aren't any meaningless sentences. There is no need for flowery language or special prose. You simply want to present a linear argument of mathematical statements that lead to the result you desire. Here is a (slightly) less trivial example.

Example 1.2. Prove that

$$\frac{d}{dx}\int_0^x \cos(t)dt = \cos(x).$$

Proof. First, note that $\cos(t)$ is continuous on the interval $[0, x]$. We thus have by the fundamental theorem of calculus that

$$\begin{aligned}\frac{d}{dx}\int_0^x \cos(t)dt &= \frac{d}{dx}(\sin(x) - \sin(0)) \\ &= \frac{d}{dx}\sin(x) - \frac{d}{dx}\sin(0) \\ &= \frac{d}{dx}\sin(x) - 0 \\ &= \cos(x).\end{aligned}$$

□

In this class we will learn many different techniques which are considered "standard" I intend to teach you these techniques in the context of genuine mathematics, rather than as isolated ideas. My hope is that you will come to understand (broadly) what mathematics is, and how it is done.

2 A Primer: Propositional Logic

2.1 The Language

Oversimplifying, logic is what underlies most of the mathematics in existence. In fact, a proof is actually more of a "representative" of the *actual* proof written completely in logic.

Definition 2.1. A **statement** in logic is a declarative sentence, expressing a single idea, that can be either true or false.

Example 2.1.

1. His dog is brown.
2. The sky is blue.
3. (Non-example) What time is it?
4. (Non-example) Sit down.

Notice that the definition of a statement includes *any* declarative sentence that can only be true or false. Importantly, it doesn't matter whether the sentence can be true in the way that we typically perceive something being true.

Example 2.2.

1. I have the recipe for concentrated dark matter.
2. The color blue sleeps soundly.

Both of the above sentences are perfectly acceptable statements in logic. We will not be concerned with facts as much as *truth.* I hope to make this clearer as we proceed. We will need a way to relate different statements.

Definition 2.2. A **compound** statement is a statement involving two or more ideas.

A compound statement is formed by connecting two (simple) statements with what are called logical connectives. We use connectives frequently in our everyday lives. The simplest examples are *and* and *or*.

Example 2.3.

1. My dog is brown and Georgia is above sea level.
2. Jake's sandwich is moldy or I am an otter.

We also want to be able to express an idea *not* being true.

Definition 2.3. A **negation** is a logical statement involving the word *not.* We will denote logical negation by $\neg$.

Example 2.4.

1. John is not my friend.
2. You are not a dinosaur.

We can negate a statement by transforming the original statement into a negation, typically in the most obvious way.

Example 2.5. ¬(My dog has fleas)=My dog does not have fleas.

The last topic we need to cover before we begin computing truth values is quantifiers.

2.2 Quantifiers

There are two kinds of quantifiers we will discuss, namely *existential* quantifiers and *universal* quantifiers.

Definition 2.4. An **existential** quantifier (denoted ∃) expresses the idea of existence of one or more things, satisfying one or more properties.

Example 2.6. Once again, we are familiar with many existential quantifiers in everyday life. Saying something like"there is a bug on me" is asserting the existence of a bug satisfying a certain property. Another example: "Some people eat carrots." This asserts the idea that there exists *at least* one person that eats carrots. The "at least one" concept will become important later on, so make sure you understand this.

Definition 2.5. A **universal** quantifier (denoted ∀) expresses the idea that *all* of a certain object satisfy one or more properties.

Example 2.7. The sentence "all cows eat grass" expresses the idea that *every* cow eats grass. Similarly, shortening to "men are pigs" expresses the idea that *every* man is a pig.

Universal and existential quantifiers have a unique relationship, and in fact this relationship is our first example of what is called *duality* in mathematics. Let's see what happens when we try to negate an existential quantifier. Consider the statement "There is a bear that plays piano." This is more subtle than negating simple statements in the way that we did earlier. You can think about this as follows: suppose a being tells you that there is a bear that plays piano. This being offers you infinite wishes if you can show them that they are wrong. How could you go about this? A reasonable first step is to find a bear that doesn't play piano. You find such a bear and present it to the being. The being says "you have showed me that there is a bear that doesn't play piano. I never said *all* bears played piano. I only said there was a bear that did." What is our next step?

Obviously showing the being 2 bears that do not play piano suffers from the same flaw. The same goes for 3, 4, etc. We thus can only obtain our wishes if we show that there

are *no* bears that play piano. We have to show the being that every bear in existence does not play piano. You also realize the being is kind of jerk for this.

Do you see what happened? We were tasked with negating an existential quantifier, and we ended up with a statement concerning *all* of a certain thing. In other words, the existential quantifier, when negated, became a *universal* quantifier. In other words, to negate "There is a bear that plays piano", we were led to the sentence "there are no bears that play piano," or "for every bear, the bear does not play piano." This in fact is a general phenomenon which will become our first theorem:

Theorem 2.1. $\neg\exists = \forall$. *In other words, the negation of an existential quantifier is a universal quantifier.*

Let x be some object, and let $p(x)$ mean "x satisfies property p." Then we can write the above theorem as

$$\neg[\exists x, p(x)] = \forall x, \neg p(x).$$

Our bear example fits in nicely here, where x is a bear and $p(x)$ means "the bear plays piano."

Now, let's see what happens in the opposite situation. Suppose we want to negate the following statement: "All cows eat grass." If we assume the situation above, how can we earn our wishes this time? We have to produce a counterexample. If we show the being that there is at least one cow that does not eat grass, we will earn our wishes. In other words, we need to show that there exists a cow that doesn't eat grass. So, we started with a statement about "all" of something, and ended up with a statement about existence. This assembles into a similar theorem as above:

Theorem 2.2. $\neg\forall = \exists$. *In other words, the negation of a universal quantifier is an existential quantifier.*

We can also represent this theorem as:

$$\neg[\forall x, p(x)] = \exists x, \neg p(x).$$

Thus, in a precise sense, universal and existential quantifiers are dual to one another.

2.3 Truth Tables

We may now begin our study of logical equivalence, which will be precisely defined below.

We have several logic operations on statements we can perform. Let p, q be statements. The expression $p \vee q$, read "p or q", expresses the idea "or". We want to examine when $p \vee q$ corresponds to "true." Consider the following statement "I am a gorilla or an artist." We want to determine when this statement is a lie. Suppose I am a gorilla, and I made the statement. Would I be lying? No, because I said I was a gorilla or an artist. Same for if I were an artist. What if I were *both* a gorilla and an artist? Would the statement then be a lie? The answer is no. I didn't say I was *either* a gorilla or an artist. I only

said I was "at least" one. In fact, the only way this statement would be a lie would be if I was neither a gorilla nor an artist. What we have just done is computed the following truth table:

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Let's examine this table more closely. The first two columns correspond to the total number of independent statements (I will often call these independent variables). The third is a combination of two, so would not be considered an independent statement. This terminology comes from expressions like $f(x) = x^2$. We see that the numerical value of $f(x)$ is determined by what value x takes, while x has no restriction. So $f(x)$ is dependent on x, while x is independent. The rows corresponding to the first two columns have values that range over every possible scenario we could have. For instance, if $p :=$ "I am a gorilla" and $q :=$ "I am an artist", then the scenario where I am a gorilla but not an artist corresponds to a T in the first column and an F in the second column, etc. Then, the third column contains the computations we made. We compute these values algebraically, according to the following rules:

$$\begin{aligned} T \vee T &= T, \\ T \vee F &= T, \\ F \vee T &= T, \\ F \vee F &= F. \end{aligned}$$

We see that these are precisely the truth values we "computed" in our discussion above. Truth table construction can be expressed via the following algorithm:

Input data: A logical expression involving at least two independent variables.

1. Identify all independent variables involved, $p_1, p_2, \ldots, p_k$.
2. List all independent variables from left to right, in no particular order, on the top of the table, reserving one column for each variable.
3. Break apart the given expression into the simplest expressions you can, and then list them on top of the table after the independent variables, complexity increasing from left to right. The last column should always be the entire expression.
4. List every possible combination of $T's$ and $F's$ that the *indepenent* variables can take, adding one row for each possible combination. In general you will have 2^k rows, where k is the total number of independent variables.

5. Compute the rest of the rows, column by column, and then descend down to the next row.

There is another central operation for statements, an operation expressing the idea of "and."The notation is $p \wedge q$, and is read "p and q." The truth values of this logical operation are also computed algebraically, according to the following rules:

$$T \wedge T = T,$$
$$T \wedge F = F,$$
$$F \wedge T = F,$$
$$F \wedge F = F.$$

These rules yield the following truth table:

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

Exercise 2.1. Compute the truth table for the expression $(p \vee q) \wedge r$.

Definition 2.6. Two statements are **logically equivalent** if they contain the same independent variables and have the same final columns in their truth tables (up to reordering of rows). I will denote this via $\overset{LE}{=}$.

We are now ready for our first proof.

Theorem 2.3 (De Morgan's Laws)**.** *Let p and q be logical statements. Then*

$$\neg(p \vee q) \overset{LE}{=} \neg p \wedge \neg q, \quad \textit{and } \neg(p \wedge q) \overset{LE}{=} \neg p \vee \neg q.$$

Proof. We prove the equivalence on the left. The first truth table is:

p	q	$p \vee q$	$\neg(p \vee q)$
T	T	T	F
T	F	T	F
F	T	T	F
F	F	F	T

.

The second is:

p	q	$\neg p$	$\neg q$	$\neg p \wedge \neg q$
T	T	F	F	F
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

.

We thus have $\neg(p \vee q) \overset{LE}{=} \neg p \wedge \neg q$. □

Exercise 2.2. Prove that $\neg(p \wedge q) \overset{LE}{=} \neg p \vee \neg q$.

2.4 Conditional and Biconditional

We are now prepared to study the conditional and biconditional, which will give rise to many of the proof techniques we will see in this class.

Definition 2.7 (Conditional)**.** Let p and q be statements. Then, the "conditional", written $p \implies q$, is the operation corresponding to the idea "if p then q" or "p implies q." $p \implies q$ can be read "if p then q," "p implies q," or "p is sufficient for q" (this will be explained later). $p \implies q$ is defined via the following truth table:

p	q	$p \implies q$
T	T	T
T	F	F
F	T	T
F	F	T

Remark. It is often useful to keep a specific example in mind when thinking of the conditional. The one I have taught in the past with success is: "If it is raining, I will take my umbrella." Determining when this statement is a *lie* will give you precisely the above truth table. For instance, if it was raining, and I did *not* take my umbrella, this statement would be a lie, so that on the truth table we would write an F. The last two rows are slightly more subtle. If it wasn't raining, but I took my umbrella anyways, would I have been lying when I said "If it is raining, I will take my umbrella"? It turns out this is not a lie. If someone says "If x then y" and x never happens, then one cannot determine whether they were lying. Thus, the value defaults to true.

Exercise 2.3. Prove that $\neg(p \implies q) \overset{LE}{=} p \wedge \neg q$.

Definition 2.8 (Bi-Conditional)**.** Let p and q be statements. Then, the "biconditional", written $p \iff q$, is the operation corresponding to the idea "p if and only if q," i.e., that you cannot have p without q, and vice-versa. They always come in a pair. $p \iff q$ can be read "p if and only if q," or "p is necessary and sufficient for q". $p \iff q$ is defined via the following truth table:

p	q	$p \iff q$
T	T	T
T	F	F
F	T	F
F	F	T

Not that the biconditional is true as long as the values of p and q are the same.

Remark. The running example I have used in the past for the biconditional is "The sun is alive if and only if the sun is shining" It is an extremely rigid relationship.

Exercise 2.4. Prove that $p \iff q \overset{LE}{=} (p \implies q) \wedge (q \implies p)$.

Definition 2.9 (Contrapositive)**.** Let p and q be statements. The **contrapostiive** is defined as $\neg q \implies \neg p$

Theorem 2.4. *Let p and q be statements. Then,*

$$(\neg q \implies \neg p) \overset{LE}{=} (p \implies q).$$

Proof. standard truth tables argument. □

Remark. Now, to explain calling things "necessary" or "sufficient," consider $p \implies q$, and assume its value is "true". Then, if p is true, then q *must* be true, or the entire expression would be false. Thus, we say q is **necessary** for p, in that we cannot have p without q. Similarly, we say p is **sufficient** for q, since if we want to know if q has happened, it suffices to show that p happened.

You will not see another truth table in this class. The entire point of this section was to familiarize you with what something being "true" means in mathematics. It means one has constructed a proof, using the underlying logic we have covered above. We will frequently encounter conditional proofs, biconditional proofs, and proofs by contrapositive in what follows, so ensure you understand "big picture" what these concepts really mean.

3 Sets and Functions

3.1 The Basics of Sets

The "standard" first definition one sees of a set is as follows:

Definition 3.1 (Set)**.** A **set** is a a collection of objects. Objects in a set S are called **elements**, with the convention that each element is listed only once.

Now that you are beginning your mathematical life, I can tell you that there is a more rigorous definition of a set. The above definition is essentially what is used in what is called *naïve set theory*. We do not have the tools to understand the rigorous definition (let alone the different models of set theory that exists), but we can understand why there is a need for a nontrivial formulation. What follows is what is known as Russel's paradox:

Let S be the set of all sets that are not elements of themselves. Then, if S is an element of S, then S cannot be in S. Similarly, if S is not in S, then S is in S.

This paradox led mathematicians to restrict what kind of collections constitute a set, ruling out situations such as the above. For our purposes, the naïve definition will work fine. For more details, the book *Set Theory* by Jech is a classic.

Some common examples of sets we are familiar with: $\mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{Q}, \mathbb{N}$, the set of all baseball players, the set of stars in the universe, etc. For the set of numbers 1-5, we can write $\{1, 2, 3, 4, 5\}$, where $\{\}$ indicates that we are looking at a set. If x is an element of a set A, we can write this in shorthand as $x \in A$. We can build new sets from already established sets via set-builder notation a follows:

$$S = \{x \in A : x \text{ satisfies some condition } \}.$$

For instance, we can represent the even natural numbers via

$$\{x \in \mathbb{N} : x = 2k, k \in \mathbb{N}\}.$$

We can also write some of the common sets above in this notation, e.g.

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}, \qquad \mathbb{Z} = \{x : (x \in \mathbb{N}) \vee (-x \in \mathbb{N})\}.$$

In the construction of $\mathbb{Z}$ as above, we assumed $0 \in \mathbb{N}$. However, whether or not 0 is a natural number is a hotly debated topic.

Suppose you have a bag of skittles. The skittles in the bag form a set. Now what if you only want to consider the red skittles? What kind of relation at the level of set theory does this correspond to?

Definition 3.2 (Subset)**.** Let A and B be sets. We say that A is a **subset** of B, written $A \subset B$ if, for all $x \in A$, $x \in B$.

Example 3.1. $\mathbb{N} \subset \mathbb{Z}, \mathbb{R} \subset \mathbb{C}$. Also note that every set is a subset of itself.

Definition 3.3. Let A be a finite set. Then $\#A$ is defined as the number of elements, or *cardinality*, of A.

Definition 3.4. Let A and B be sets. Then, we say $A = B$ if A and B have the same elements.

Theorem 3.1. *Let A and B be sets. Then, $A = B$ if and only if $A \subset B$ and $B \subset A$.*

Before we prove this, we need to make an observation. This statement contains an "if and only if." Remember that $p \iff q$ is logically equivalent to $(p \implies q) \wedge (q \implies p)$. So, to prove the above theorem, we must prove that $A = B$ implies $A \subset B$ and $B \subset A$, and we must also prove that if $A \subset B$ and and $B \subset A$, then $A = B$. Now you see why this logical concept is called bicondtional. We have literally broken down a biconditional into two conditionals. We now proceed with the proof:

Proof. First, suppose $A = B$. Recall that every set is a subset of itself. We thus have $A \subset A = B$ so that $A \subset B$, and similarly we have $B \subset B = A$ so that $B \subset A$. This completes the first "direction."

For the converse, assume that $A \subset B$ and $B \subset A$. A nice way to proceed is: suppose that $A \neq B$. What does this negation mean? For the sets to not be equal, the have

to differ by *at least one* element. So, let's pick one such element. Let $x \in A$ such that $x \notin B$. But $A \subset B$ so that $x \in B$! What just happened? We showed that supposing $A \neq B$ implies that a statement was both true and false, which cannot happen. So A must be equal to B. This is called a *proof by contradiction.* $\square$

Proof by contradiction is an important technique in mathematics. At the level of logic, we assumed that $R =$"$A \subset B$ and $B \subset A$" was true, and that $P =$ "$A = B''$ was false, i.e. that $\neg P = T$, and then we showed that $\neg P \implies \neg R$ so that R was both true and false. This is impossible, so our supposition that $A \neq B$ had to have been wrong. This means that we must have $A = B$. For the rest of the course, we won't bother to break down what is happening at the level of logic for our proofs by contradiction. You only need to understand that the big idea, namely, we make a supposition, and show it directly implies something being true and false. This means the original supposition must have been wrong. It may seem kind of like magic, which is why, this first time, I broke it down to the level of logic. The proofs we have done before this moment have all been what are colloquially called "direct" proofs. This loosely may be interpreted as "without using contradiction or contrapositive."

Continuing our study of the basics of set theory, we now discuss some operations one can do on sets.

Definition 3.5 (Unions and Intersections)**.** Let A and B be sets. We then define $A \cup B$, read "A union B", as

$$A \cup B = \{x : (x \in A) \vee (x \in B)\}.$$

We define similarly $A \cap B$, read "A intersect B", as

$$A \cap B = \{x : (x \in A) \wedge (x \in B)\}.$$

Example 3.2. Let $A = \{a, b, c, d, e, f\}$ and $B = \{e, f, g, h, i, j, k\}$. Then, we have

$$A \cup B = \{a, b, c, d, e, f, g, h, i, j, k\}, \quad A \cap B = \{e, f\}.$$

Example 3.3. Let $A = \{1, 3, 5, 7, \dots\}$ and $B = \{0, 2, 4, 6, \dots\}$. Then $A \cup B = \mathbb{N}$.

In the above example, what happens if we take $A \cap B$? For this to be a well-defined operation on sets, this needs to have a value. What could it be? Our answer comes in the form of the *empty set.*

Definition 3.6. The set containing no elements, written $\emptyset$, is defined as $\emptyset = \{\}$. This set is named the **empty set**.

This set may seem silly at first, but its importance cannot be overstated. You will use it for the rest of your mathematical career.

Remark. Note that, by definition, the empty set is a subset of every set. Note also that $\emptyset \neq \{\emptyset\}$.

Exercise 3.1. Prove the following relations:

1. $A \cap B \subset A$
2. $A \cap B \subset B$
3. $A \cap B = B \cap A$.

There is one more operation that will be essential for us:

Definition 3.7. Let $A \subset U$. Then the complement of A in U, written $U \backslash A$, is defined as

$$U \backslash A = \{x : (x \in U) \wedge (x \notin A)\}.$$

When the ambient set is understood, we sometimes write the complement of A as A^c.

Example 3.4. Let $X = \mathbb{N}$. Then $X \backslash \{0, 2, 4, 6, \dots\} = \{1, 3, 5, 7 \dots\}$.

Lemma. *Let $A \subset X$. Then $(A^c)^c = A$.*

Proof. We use now, and in almost every proof of set equality we will do, Theorem 3.1. To that end, let $x \in (A^c)^c$. Then $x \notin A^c$, so that $x \in A$. Hence $(A^c)^c \subset A$.

For the reverse inclusion, let $x \in A$. Then $x \notin A^c$, so that $x \in (A^c)^c$. Hence $A \subset (A^c)^c$. Thus, by Theorem 3.1, $(A^c)^c = A$. □

Theorem 3.2 (De Morgan's Laws for Sets)**.** *Let $A, B \subset X$ be subsets of some set X. Then, we have:*

$$(A \cup B)^c = A^c \cap B^c, \quad \textit{and } (A \cap B)^c = A^c \cup B^c.$$

Proof. We prove the left equality first. Let $x \in (A \cup B)^c$. We want to prove that $x \in A^c \cap B^c$. Since $x \in (A \cup B)^c, x \notin A \cup B$ by the above lemma. Thus $x \notin A$ and $x \notin B$ by the definition of a union, so that $x \in A^c$ and $x \in B^c$. Hence $x \in A^c \cap B^c$, as desired.

For the reverse inclusion, let $x \in A^c \cap B^c$. Then x is neither in A nor B, so that $x \notin A \cup B$. Hence $x \in (A \cup B)^c$. Therefore, by Theorem 3.1, $(A \cup B)^c = A^c \cap B^c$. □

Exercise 3.2. Prove that $(A \cap B)^c = A^c \cup B^c$ (Hint: You proof will look similar in spirit to the above proof).

3.2 Functions

Definition 3.8 (Cartesian Product)**.** Let A and B be sets. The **Cartesian product**, denoted $A \times B$, is defined as:

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

Example 3.5. The most familiar examples from calculus: $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n}$,

$\mathbb{C}^n = \underbrace{\mathbb{C} \times \cdots \times \mathbb{C}}_{n}$.

Definition 3.9 (Relation)**.** Let A and B be sets. A **relation** (or binary relation) on A and B is a triple (R, A, B) where $R \subset A \times B$. A is called the **domain**, while B is called the **codomain**. If $(a, b) \in R$, we sometimes will write aRb for (a, b)

Example 3.6. Let $A = \mathbb{R}$ and $B = \mathbb{N}$. An example of a relation (albeit not a very interesting one) is the set $\{(\pi, 4), (e, 7), (\frac{1}{2}, 1)\}$. Note there are no restrictions on the subset.

Example 3.7. A less trivial example. Let $A = B = \mathbb{Z}$. Define the relation $R = \leq$ as aRb if and only if $a \leq b$.

Definition 3.10 (Function)**.** A **function** is a triple (R, A, B), where A and B are sets and R is a relation on A and B such that:

i. $(x_1, y_1) = (x_1, y_2)$ if and only if $y_1 = y_2$. I.e., each $x \in A$ is associated to a single $y \in B$.

ii. For every $x \in A$, there is some $y \in B$ such that $(x, y) \in R$.

A is called the **domain**, while B is called the **codomain**.

Example 3.8. Let $A = \{1, 2, 3, 4, 5\}$ and $B = \{a, b\}$. Then, an example of a function would be $\{(1, a)(2, b)(3, b)(4, a)(5, a)\}$. A non-example would be $\{(1, a), (1, b), (3, b)\}$.

Functions are omnipresent in mathematics, and we will be concerned with them in one way or another for the rest of the semester/your mathematical career. This definition may seem a little obscure at first, but after spending some time with it, you will see it coincides with the way we usually think of functions, as maps between sets:

Definition 3.11 (Functions 2)**.** Let A and B be sets. A **function** is a triple (f, A, B) where f assigns to each element of A a unique element of B. Note that multiple elements of A can map to the same element of B. We only require that each $x \in A$ produces only one $y \in B$. Standard notation is that $f(x) \in B$ is the element to which f has assigned x. A is called the **domain**, while B is called the **codomain**.

Remark. Note that the definition of a function requires you to specify your domain and codomain. When you understand what a function truly is at the level of set theory (our first definition of a function), it makes perfect sense why this must be the case.

We will almost always (there may be one or two exceptions at the very end of the course) think of functions as in our second definition.

Definition 3.12. Let $f, g : A \to B$ be functions. Then $f = g$ if and only if for all $x \in A$, $f(x) = g(x)$.

Example 3.9. Some examples of functions you have undoubtedly encountered before:

- $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = x^2$.

- $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = \sin(x)$.
- $f : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ defined as $f(x, y) = x + iy$.

Definition 3.13 (Function Composition)**.** If $f : A \to B$ and $g : B \to C$ are functions, we can form their composite, a function $g \circ f : A \to C$, defined as $x \mapsto f(x) \mapsto g(f(x))$.

Theorem 3.3. *Function composition is an associative operation. That is, If $f : A \to B$, $g : B \to C$, and $h : C \to D$ are functions, then $h \circ (g \circ f) = (h \circ g) \circ f$.*

Definition 3.14. Let $f : X \to Y$ be a function. Let $B \subset Y$. We define the **preimage** of B, written $f^{-1}(B)$ (not to be confused with the inverse of f), as

$$f^{-1}(B) = \{x \in X : f(x) \in B\}.$$

Theorem 3.4. *Let A and B be subsets of X, and let let $f : X \to Y$ be a function. Then, $f(A \cup B) = f(A) \cup f(B)$.*

Proof. We prove first that $f(A \cup B) \subset f(A) \cup f(B)$. Let $x \in f(A \cup B)$. Then, there exists an $\tilde{x} \in A \cup B$ such that $f(\tilde{x}) = x$. Since $\tilde{x} \in A \cup B$, we have that $\tilde{x} \in A$ or $\tilde{x} \in B$, possibly both. We can, without loss of generality, assume $\tilde{x} \in A$. Then $x \in f(A)$ so that $x \in f(A) \cup f(B)$.

For the reverse inclusion, let $x \in f(A) \cup f(B)$. Then, there is an $\tilde{x} \in X$ such that $f(\tilde{x}) = x$. Since $x \in f(A) \cup f(B)$, $\tilde{x}$ is in either A or B, possibly both. Hence $\tilde{x} \in A \cup B$, so that $f(\tilde{x}) = x \in f(A \cup B)$. Therefore, by Theorem 3.1, $f(A \cup B) = f(A) \cup f(B)$. □

A natural question to ask now is whether or not intersections are also preserved under functions. I.e., does $f(A \cap B) = f(A) \cap f(B)$? It turns out that this is not the case. In face, this is only true when f satisfies a certain property, which we will study very soon. Inverse images, on the other hand, are well-behaved with respect to set theoretic operations:

Exercise 3.3. Let $A, B \subset Y$, with $f : X \to Y$ a function.

- Prove that $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.
- Prove that $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.

3.3 Injective and Surjective Functions

Up to now, we have been considering all functions. We now consider certain "types" of functions that are useful in mathematics.

Definition 3.15 (Injective)**.** Let $f : A \to B$ be a function. We say f is **injective** if, for $x \neq y \in A, f(x) \neq f(y)$. It is often useful to use the contrapositive, which is: f is injective if $f(x) = f(y) \implies x = y$. f is called an injection, and we say f maps A *injectively* into B. Injectivity is also called being $1 - 1$.

Example 3.10. First, a non-example. Consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. We see that this function is not injective, because both x and $-x$ map to x^2.

Now, consider $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = x + 3$. Then, we have:

$$\begin{aligned} g(x) = g(y) &\implies x + 3 = y + 3 \\ &\implies x = y. \end{aligned}$$

So g is in fact injective.

Lemma. *Let $A \subset X$, and let $f : X \to B$ be a function. Then, $A \subset f^{-1}(f(A))$.*

Proof. Let $x \in A$. Then $f(x) \in f(A)$, so that $x \in f^{-1}(f(A))$. □

This isn't very exciting. However, using injectivity, we can do something pretty neat:

Theorem 3.5. *Let $f : X \to Y$ be a function. Then, $A = f^{-1}(f(A))$ for all subsets $A \subset X$ if and only if f is injective.*

Proof. This is an"if and only i'f' proof, so we must prove both "directions": For the first direction, assume f is injective. By the preceding lemma, we have that $A \subset f^{-1}(f(A))$. We thus have only to prove that $f^{-1}(f(A)) \subset A$. To that end, let $x \in f^{-1}(f(A))$. Then $f(x) \in f(A)$. So, there is some $y \in A$ such that $f(x) = f(y)$. By injectivity, we have $x = y$, so that $x \in A$. Hence $A = f^{-1}(f(A))$.

Conversely, assume $A = f^{-1}(f(A))$ for all $A \subset X$. Suppose $f(x) = f(y)$. Then, we have:

$$\{x\} = f^{-1}(f(\{x\})) = f^{-1}(f(\{y\})) = \{y\},$$

so that $x = y$. □

Exercise 3.4. Let $f : A \to B$ and $g : B \to C$ be functions. Show that if $g \circ f$ is injective, then f is injective.

Definition 3.16. Let A be a set. Then, the function from A to itself that maps each $x \in A$ to itself is called the **identity** function, written id_A.

Theorem 3.6. *Let $f : A \to B$ be a function, where neither A nor B are empty. Then f is injective if and only if f has a left inverse. That is, if there exists some $g : B \to A$ such that $g \circ f = id_A$.*

Proof. Assume f has a left inverse. Let $x, y \in A$ such that $f(x) = f(y)$. Then, we have:

$$x = (g \circ f)(x) = g(f(x)) = g(f(y)) = y,$$

so that $x = y$ and f is injective.

Conversely, assume f is injective. We must construct a left inverse g. By definition, for every $b \in f(A)$, there is a unique $x_b \in A$ such that $f(x_b) = b$. Fix some element $b_0 \in B$. Define $g : B \to A$ as follows:

$$g(b) = \begin{cases} x_b & b \in f(A) \\ b_0 & otherwise \end{cases}.$$

Then, for all $x \in A$, we have $(g \circ f)(x) = g(f(x)) = x$, so that $(g \circ f)(x) = \mathrm{id}_A$ as desired. □

Definition 3.17 (Surjective)**.** Let $f : A \to B$ be a function. We say f is **surjective** if, for all $b \in B, \exists x \in A$ such that $f(x) = b$. f is called a surjection, and we say f maps A *surjectively* onto B. Surjectivity is also called being *onto*.

Example 3.11. First, a non-example. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 0$ for all x. Then this map is obviously not surjective.

Now, let $g : \mathbb{R} \to \mathbb{R}$ be given by by $g(x) = 7x + 3$. Say we have a specific y we want to map to. Then, setting $y = 7x + 3$ and solving for x, we see that $x = \frac{y-3}{7} \in \mathbb{R}$ maps to our desired y. Hence g is surjective.

Exercise 3.5. Let $f : X \to Y$ be a function. Prove that $f(f^{-1}(B)) = B$ for any subset $B \subset Y$ if and only if f is surjective.

Theorem 3.7. *Let $f : A \to B$ be a function, where A, B are nonempty. Then f is surjective if and only if it has a right inverse.*

Proof. Assume f has a right inverse g. Let $b \in B$. Then $b = (f \circ g)(b) = f(g(b))$, so that f maps $g(b)$ to b.

Conversely, assume f is surjective. Then $f^{-1}(b) \neq \emptyset$ for every $b \in B$. Choose one element x_b in the preimage of each b. We then define a right inverse g by $g(b) = x_b$. Then $(f \circ g)(b) = f(g(b)) = f(x_b) = b$, as desired. □

There is a special name for functions that are both injective and surjective:

Definition 3.18 (Bijection)**.** Let $f : A \to B$ be a function. If f is both injective and surjective, then f is called a **bijection**.

Corollary 3.1. *A function is bijective if and only if it has a left and right inverse.*

Example 3.12. There is always an obvious bijection from a set A to itself, namely id_A.

As a less trivial example, consider $f : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ defined by $f(x, y) = x + iy$. This is clearly a bijection.

From the above corollary, we see that the existence of inverses is significant. We can go even further:

Theorem 3.8. *Let $f : A \to B$ be a function. If f has a left inverse g and a right inverse h, then $g = h$.*

Proof. We have, for all $x \in B$, that

$$\begin{aligned} h(x) &= (g \circ f)(h(x)) \quad \text{(since } g \text{ is a left inverse)} \\ &= [(g \circ f) \circ h](x) \\ &= [g \circ (f \circ h)](x) \quad \text{(by associativity)} \\ &= g(x) \quad \text{(since } h \text{ is a right inverse).} \end{aligned}$$

Hence $g = h$. □

Thus, for a bijection f, we can talk about *the* inverse of f, denoted f^{-1} (not to be confused with inverse image).

We can now rephrase the above corollary as something stronger:

Proposition 3.1. *A function is bijective if and only if it has a two-sided inverse.*

Being bijective is a very strong property. This topic gives us an opportunity to acknowledge a major theme in mathematics: if you want to learn about some object, study maps going into and out of the object. For example:

Theorem 3.9. *If $f : A \to B$ is a bijection, then $\#A = \#B$.*

Proof. We proceed by contrapositive. Suppose that $\#A \neq \#B$. Then one of the cardinalities must be greater. We may assume without loss of generality that $\#A < \#B$. Then, any function from A to B must miss at least one element, or we would have one element of A mapping to multiple elements of B, which violates the definition of a function. Hence there are no surjective functions from A to B, so that in particular, there are no bijections. □

In fact, if you take a dedicated set theory class, you will see that the *definition* of two sets having the same cardinality is the existence of a bijection between them.

Exercise 3.6.

1. Prove that the composition of injective functions is injective.
2. Prove that the composition of surjective functions is surjective.
3. Deduce that the composition of bijections is bijective.

Exercise 3.7. Let $f : X \to Y$ be a function. Prove that $f(A \cap B) = f(A) \cap f(B)$ for all $A, B \subset X$ if and only if f is injective. [Hint: $f(A \cap B) \subset f(A) \cap f(B)$ for ANY function f, not just injective functions. So you need only show the reverse containment for this proof.]

4 Analysis

4.1 Continuity

Up to now, our sets have had no extra structure on them. We have been dealing with them as *just* sets, and the maps between as just a relation. It is usually useful to consider sets with extra structure. We begin with a concept you have encountered before, although maybe not in a mathematically rigorous sense. You may have seen a definition such as the following: "a function is continuous if you can draw its graph without picking up your pencil." For many reasons, this definition is inadequate. The most obvious reason is

probably that it is hard to work in the concept of what drawing on paper means in terms of logic. In this section we will see the "actual" definition. First some preliminaries.

For many sets you have encountered, such as $\mathbb{R}$ or $\mathbb{C}$, it makes sense to think about the "distance" between certain points. In order to do that, we need a way to assign a numerical distance to a set.

Definition 4.1 (Metric)**.** Let X be a set. A metric on x is a function $d : X \times X \to [0, \infty)$ such that, for all $x, y, z \in X$, the following are satisfied:

i. $d(x, y) = 0 \iff x = y$

ii. $d(x, y) = d(y, x)$

iii. $d(x, z) \leq d(x, y) + d(y, z)$

Given the above three items are satisfied, we can deduce that $d(x, y) \geq 0$ for all $x, y \in X$.

Definition 4.2 (Metric Space)**.** A **metric space** is a pair (X, d), where X is a set and d is a metric on X.

Example 4.1. The most familiar example is most likely $(\mathbb{R}, |\cdot|)$, where $|\cdot|$ is the absolute value.

Example 4.2. Consider $(\mathbb{C}, |\cdot|)$, where $|\cdot|$ is the modulus defined as $|a + bi| = \sqrt{a^2 + b^2}$. This is also a metric.

We are now ready for the definition of continuity. It may seem a little technical at first, but we will go over the definition in detail.

Definition 4.3 (Continuity)**.** Let (X, d) and (Y, ρ) be metric spaces, with $f : X \to Y$ a function. Then, we say f is **continuous at a point** $x_0 \in X$ if for every $\epsilon > 0$, there is a $\delta > 0$ such that for all $y \in X$ such that $d(x_0, y) < \delta$, we have $\rho(f(x_0), f(y)) < \epsilon$. We say f is **continuous** if it is continuous at every point of X.

We examine this definition in a familiar case. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = 3x + 4$. Let $\epsilon = \frac{1}{2}$. Fix $x_0 \in \mathbb{R}$. We want to find a δ such that for $|x_0 - x| < \delta$, $|f(x_0) - f(x)| < \frac{1}{2}$. We make the following computations:

$$\begin{aligned} |f(x_0) - f(x)| &= |3x_0 + 4 - (3x + 4)| \\ &= |3x_0 + 4 - 3x - 4| \\ &= |3x_0 - 3x| \\ &= |3(x_0 - x)| \\ &= 3|x_0 - x|. \end{aligned}$$

So, if we are to have $|f(x_0) - f(x)| < \frac{1}{2}$, we must then have $3|x_0 - x| < \frac{1}{2}$. Dividing both sides by 3, we have our desired requirement: $|x_0 - x| < \frac{1}{6}$. Let's test this out: set

$x_0 = 1$. We need a number x such that $|1 - x| < \frac{1}{6}$. Let try $\frac{7}{8}$. Then $\left|1 - \frac{7}{8}\right| = \frac{1}{8} < \frac{1}{6}$. Hence,

$$\left|f(1) - f\left(\frac{7}{8}\right)\right| = 3 + 4 - (3\left(\frac{7}{8}\right) + 4) = 3 - 3\frac{7}{8} = 3 - \frac{21}{8} = \frac{24}{8} - \frac{21}{8} = \frac{3}{8} = .375.$$

Since this value is indeed less than $\frac{1}{2}$, our process seems to have worked. Wonderful!

We can generalize this method with any positive real number. Let $\epsilon > 0$ (we pretend it is small, because we want to prove continuity for all numbers, no matter how small). Following the exact same argument as before, we have that we require $\delta < \frac{\epsilon}{3}$. Indeed, for $x, y \in \mathbb{R}$ such that $|x - y| < \delta$, we have $|f(x) - f(y)| \leq 3|x - y| < \epsilon$.

Exercise 4.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a function of the form $f(x) = ax + b$. Prove that f is continuous on all of $\mathbb{R}$. (Hint: Two cases are necessary, $a = 0$ and $a \neq 0$.)

Exercise 4.2. Let (X, d) and (Y, ρ) be metric spaces, and let $f : (X, d) \to (Y, \rho)$ be defined by $f(x) = a$ for all $x \in X$, where $a \in Y$ is a fixed constant. Prove that f is continuous.

4.2 Continuity of Polynomials

Theorem 4.1. *Let $f, g : (X, d) \to \mathbb{R}$ be continuous functions. Then $(f+g) : (X, d) \to \mathbb{R}$, defined by $(f + g)(x) = f(x) + g(x)$, is continuous.*

Proof. Let $\epsilon > 0$. Fix $x \in X$. Since both f and g are continuous, there exist δ_1, δ_2 such that for all $y \in X$, we have $d(x, y) < \delta_1 \implies \rho(f(x), f(y)) < \epsilon$, and $d(x, y) < \delta_2 \implies \rho(g(x), g(y)) < \epsilon$. Let $\delta = \min(\delta_1, \delta_2)$. Then, we have:

$$\begin{aligned} |(f + g)(x) - (f + g)(y)| &= |f(x) - f(y) + g(x) - g(y)| \\ &\leq |f(x) - f(y)| + |g(x) - g(y)| \\ &< 2\epsilon \end{aligned}$$

for all y such that $d(x, y) < \delta$. □

Theorem 4.2. *Let $f : (X, d) \to \mathbb{R}$ be a continuous function. Then $f^2 : (X, d) \to \mathbb{R}$, defined by $f^2(x) = f(x)^2$, is continuous.*

Proof. Fix $x_0 \in X$. Let $\epsilon > 0$. By assumption, there is some $\delta > 0$ such that for any $y \in X$ with $d(x_0, y) < \delta$, we have $|f(x_0) - f(y)| < \epsilon$. For such a y, we have

$$\begin{aligned} |f(y) + f(x_0)| &= |f(z) - f(x_0) + 2f(x_0)| \\ &\leq |f(z) - f(x_0)| + |2f(x_0)| \\ &\leq \epsilon + 2|f(x_0)|. \end{aligned}$$

Define C as the constant $\epsilon + 2|f(x_0)|$. It is important to note here that C depends only on x_0. Then, we have:

$$\begin{aligned} |f^2(x_0) - f^2(y)| &= |f(x_0) + f(y)||f(x_0) - f(y)| \\ &\leq \epsilon \cdot C. \end{aligned}$$

Hence, requiring that $|f(x_0) - f(y)| < \frac{\epsilon}{C}$ yields the desired result. □

Corollary 4.1. *Let $f, g : (X, d) \to \mathbb{R}$ be continuous functions. Then $(fg) : (X, d) \to \mathbb{R}$, defined by $(fg)(x) = f(x)g(x)$, is continuous.*

Proof. Note that $fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$. Hence, the above two theorems imply fg is continuous. □

We are now ready for the main result of this section. First, a definition you are probably familiar with:

Definition 4.4. A **polynomial** with real coefficients in a single indeterminate x is defined as

$$\sum_{k=0}^{n} a_k x^k = a_n x^n + \cdots + a_1 x + a_0,$$

where $a_k \in \mathbb{R}$ for all k. A real valued **polynomial function** with real coefficients is a function $f : \mathbb{R} \to \mathbb{R}$ such that the value at a point $t \in \mathbb{R}$ is given by

$$f(t) = a_n t^n + \cdots + a_1 t + a_0.$$

Theorem 4.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a polynomial function. Then, f is continuous.*

Proof. Since, by definition, any polynomial is a finite combination of addition and multiplication of continuous functions, any polynomial is continuous. □

Exercise 4.3. Let (X, d) be a metric space, and fix $a \in X$. Define $f : (X, d) \to (\mathbb{R}, |\cdot|)$ as $f(x) = d(a, x)$. Prove that f is continuous.

4.3 Further Reading

1. For a nice book on "thinking analytically", there is *The Way of Analysis* by Robert Strichartz.
2. The book used here for "Advanced Calculus" is *Introduction to Real Analysis* by Robert Bartle and Donald Sherbert.
3. A beginning graduate/advanced undergraduate analysis book is the beautiful *Principles of Mathematical Analysis* by Walter Rudin, considered by many to be the gold standard for introductory analysis books. While it can be terse at times, it is certainly worth reading.

5 Algebra

5.1 Groups

In the last section we investigated putting one type of structure on a set, i.e., metrics. In algebra, we will be concerned with considering structures on sets corresponding to algebraic manipulation. We begin with the following definition:

Definition 5.1. Let A be a set. A **binary operation** is a map[1] $m : A \times A \to A$. We typically right $m(x, y) \in A$ as $xy, x \cdot y$ or $x * y$. A binary operation is *associative* if $x(yz) = (xy)z$ for all $x, y, z \in A$. A binary operation is *commutative* if $xy = yx$ for all $x, y \in A$. A pair consisting of a set S and a binary operation $*$ on S is called a **magma**. A magma with an associative binary operation is called a **semigroup**.

Definition 5.2. Let $(A, \times)$ be a magma. An **identity element** is a element e of A such that $ex = xe = x$ for all $x \in A$.

Theorem 5.1. *Let $(A, \times)$ be a magma. If an identity for the binary operation exists, it is unique.*

Proof. Let e_1, e_2 be two identity elements. Then, we have:

$$e_1 = e_1 e_2 = e_2.$$

□

Definition 5.3. Let $(A, \times)$ be a magma with identity. Let $x \in A$. Then an **inverse** of x is an element $y \in A$ such that $xy = e = yx$.

Theorem 5.2. *Let $(A, \times)$ be a semigroup with identity. If $x \in A$ has an inverse, then the inverse is unique.*

Proof. Let $y, y' \in A$ be inverses of x. Then, we have

$$y = ey = (y'x)y = y'(xy) = y'e = y'.$$

□

Definition 5.4. A **monoid** is a semigroup with identity.

We write the unique inverse of an element a as a^{-1}.

Definition 5.5 (Group)**.** A **group** is a monoid in which every element has an inverse.

Example 5.1. The following describe group structures:

- Underlying set: $\mathbb{Z}$ Binary operation: $+$ Identity: 0
- Underlying set: $\mathbb{R}$ Binary operation: $+$ Identity: 0

[1] the quality of having the output of an operation remain in the set is called *closure* under the operation.

- Underlying set: $\mathbb{R}\backslash\{0\}$ Binary operation: $\times$ Identity: 1
- Underlying set: $\mathrm{Mat}(n, \mathbb{R})$ Binary operation: $+$ Identity: 0 matrix
- Underlying set: $\mathrm{GL}(n, \mathbb{R})$ Binary operation: $\times$ Identity: Identity Matrix

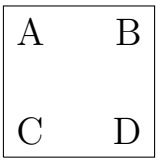
The following do **not** describe group structures:

- Underlying set: $\mathbb{Z}$ Binary operation: $\times$
- Underlying set: $\mathbb{N}$ Binary operation: $+$
- Underlying set: $\mathrm{Mat}(n, \mathbb{R})$ Binary operation: $\times$

Exercise 5.1. Let $(G, \times)$ be a group. Let $a, b \in G$.

1. Prove that $(ab)^{-1} = b^{-1}a^{-1}$.
2. Prove that $(a^{-1})^{-1} = a$.

Now, we construct a group you may not have come across before. Consider the following picture:



Consider 90 degree clockwise rotations of the square taking one corner to another. I.e., one rotation takes A to B, B to D, etc. Let r denote such a rotation. First, notice that there is an obvious inverse to r. The counterclockwise rotation by 90 degrees taking A back to the top left corner, etc. Is there an identity rotation? Indeed, the "trivial" rotation that leaves all points fixed. Is this the only type of symmetry the square has?

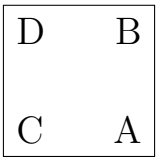
Imagine "reflecting" across the vertical line splitting the square into a left and right side. Is there an inverse to this operation? Indeed, performing this same reflection again will return the corners to where they were originally. Reflecting across the horizontal line splitting the square, as well as the two diagonals are both involutive operations. Assembling what we have so far:

Definition 5.6 (D_8)**.** The **Dihedral group** on 4 vertices, otherwise known as D_8, is the group with the following elements:

$$r, r^2, r^3, e, v, h, d, d'.$$

Proposition 5.1. *The previous definition defines a group structure.*

Proof. Each r and each reflection has an obvious inverse. We thus need to investigate as to what happens when we multiply rotations and reflections. First consider rv, where this is defined by first performing v and then r. We arrive at



which is none other than d. Further, consider v. We have $v =$

B	A
D	C

which is dr! Investigating this pattern further, we have that $h =$

C	D
A	B

which is $dr^{-1} = dr^3$. Finally, we have $d' =$

A	C
B	D

which is none other than dr^2.

So we have that we can rewrite the group elements as

$$D_8 = \{e, r, r^2, r^3, d, dr, dr^2, dr^3\}.$$

□

We can now check closure and the existence of inverses in a systematic fashion. Clearly we have r and its own multiples remain in the group and have inverses. We proceed with the nontrivial multiplications:

We have $rdr = d \in D_8$, so that not only is rdr in D_8, but rdr is its own inverse as well. Indeed, $(rdr)^2 = rdrrdr = e$.

Exercise 5.2. Finish proving that D_8 is a group. I.e., show that all group multiplications remain in D_8, and exhibit the inverses of each element. Explicitly, fill in the following multiplication table (note that a lot of them have been done for you already, check your notes from class):

$\cdot$	e	r	r^2	r^3	d	dr	dr^2	dr^3
e								
r								
r^2								
r^3								
d								
dr								
dr^2								
dr^3								

Make sure you understand this group, as we will use it for many examples.

Example 5.2. Here, we consider the *Klein 4-group*, denoted V_4, the group on four elements e, a, b, c where each non-identity element is involutive, with multiplication table:

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Note that there is something fundamentally different about the operation in V_4 as opposed to D_8. For any elements $x, y \in D_8$, we cannot assume $xy = yx$ (for instance, $rd \neq dr$), whereas for all $x, y \in V_4$, we have $xy = yx$.

Definition 5.7. Let $(G, \times)$ be a group. If, for all $a, b \in G$, we have $ab = ba$, we say that G is **abelian**. Typically, if G is an abelian group, we write the operation as addition.

Example 5.3. The following are abelian groups:

1. $(\mathbb{Z}, +)$
2. $(\mathbb{R}, +)$
3. $(\mathbb{R}\backslash\{0\}, \times)$
4. $(\mathrm{Mat}(m \times n, \mathbb{R}), +)$

The following are not abelian groups:

1. $(D_8, \times)$ (in fact, $r^i s = sr^{-i}$ for $0 \leq i \leq 4$)
2. $(\mathrm{GL}(n, \mathbb{R}), \times)$

Now that we are familiar with the structure of groups, we can talk about the maps between them that respect the group structure.

5.2 Homomorphisms

Definition 5.8. Let $(G, *_1)$ and $(H, *_2)$ be groups. A function $\phi : G \to H$ is a **homomorphism** if, for all $x, y \in G$, we have

$$\phi(x *_1 y) = \phi(x) *_2 \phi(y).$$

So, we have that a homomorphism between two groups is a map of sets that respects the group structures of the domain and codomain.

Proposition 5.2. *Let $(G, *_1)$ and $(H, *_2)$ be groups with identity elements e_1 and e_2, respectively. If $\phi : (G, *_1) \to (H, *_2)$ is a homomorphism, then $\phi(e_1) = e_2$.*

Proof. We have that, for $x \in G$,

$$\phi(x) *_2 e_2 = \phi(x) = \phi(x *_1 e_1) = \phi(x) *_2 \phi(e_1),$$

so that $\phi(x)*_2 e_2 = \phi(x)*_2\phi(e_1)$. Multiplying by $\phi(x)^{-1}$ on both sides, we have $e_2 = \phi(e_1)$, as desired. □

Example 5.4. Consider $\phi : (\mathbb{Z}, +) \to (\mathbb{R}, +)$ given by $\phi(x) = 4x$. Then $\phi(0) = 4(0) = 0$, which is a good start. Then, for any arbitrary $x, y \in \mathbb{Z}$, we have

$$\phi(x + y) = 4(x + y) = 4x + 4y = \phi(x) + \phi(y).$$

Hence ϕ is a homomorphism.

Exercise 5.3. Let $(G, *_1)$ and $(H, *_2)$ be groups, and let $\phi : G \to H$ be a homomorphism. Prove that $\phi(x^{-1}) = \phi(x)^{-1}$ for all $x \in G$.

We now are ready to state what is means for two groups to be the "same." It is generally too restrictive to consider mathematical objects the same only when they are literally the same at the level of sets and functions. Each topic we have covered so far has its own version of sameness. We consider sets the same if there is a bijection between them. Metric spaces are the same when there is a continuous bijection between them with continuous inverse. Such a map is called a *homeomorphism*. For groups, we have the following:

Definition 5.9. Let $\phi : (G, *_1) \to (H, *_2)$ be a homomorphism of groups. If this homomorphism is bijective, we say ϕ is an **isomorphism** between G and H, and that G and H are **isomorphic** as groups.

Example 5.5. For any group G, id_G is an isomorphism from G to itself. Less trivially, consider $\exp : (\mathbb{R}, +) \to (\mathbb{R}^+, \times)$. We saw last time that this was a homomorphism. When we restrict to the positive real numbers, exp has an inverse, namely $\ln$.

Exercise 5.4.

1. Prove that the composition of homomorphisms is again a homomorphism.
2. Let $\hom_{iso}(A, A)$ denote the set of all group isomorphisms from a group A to itself. Prove that $\hom_{iso}(A, A)$ is itself a group.

Before we move on to subgroups, we will make a slight detour.

5.3 Equivalence Relations

Here we recall the definition of a relation:

Definition 5.10 (Relation)**.** Let A and B be sets. A **relation** on A and B is a triple (R, A, B) where $R \subset A \times B$. A is called the **domain**, while B is called the **codomain**. If $(a, b) \in R$, we sometimes will write aRb for (a, b).

Definition 5.11. Let (A, A, R) be a relation. We say (A, A, R) is an **equivalence relation** if:

1. For all $a \in A$, aRa (reflexivity);
2. If aRb, then bRa (symmetry);
3. If aRb and bRc, then aRc (transitivity).

If R defines an equivalence relation, we sometimes write $a \sim b$ if aRb.

Example 5.6. The standard example of an equivalence relation is equality.

Example 5.7. Let $f : A \to B$ be a map of sets. Let $a, a' \in A$. Define the following relation: $a \sim a'$ iff $f(a) = f(a')$.

Exercise 5.5. Let $\mathcal{C}$ be a collection of sets. Prove that the relation "$A \sim B$ iff there is a bijection from A to B" is an equivalence relation on $\mathcal{C}$, where $A, B \in \mathcal{C}$.

Definition 5.12. Let $\sim$ be an equivalence relation on A, and let $a \in A$. The **equivalence class** of a, denoted $[a]_\sim$ or $\bar{a}$, is defined as

$$[a]_\sim := \{b \in A : a \sim b\}.$$

First, an observation. Note that $[a]_\sim \notin A$, but rather $[a]_\sim \in \mathcal{P}(A)$, the power set of A. We denote the set of all equivalence classes as $A/\sim$, which is a subset of $\mathcal{P}(A)$.

Definition 5.13. Let S be a set. A **partition** of S is a collection of nonempty subsets of S, $\{A_i\} \subset \mathcal{P}(S)$, such that:

1. $S = \cup_i A_i$;
2. $A_i \cap A_j = \emptyset$ if $i \neq j$.

Theorem 5.3. *Let $\sim$ be an equivalence relation on a set A. Then $A/\sim$ is a partition of A.*

Proof. First, observe that since equivalence relations are reflexive, $\cup_a [a]_\sim = A$. We thus must only prove that equivalence classes are pairwise disjoint. Indeed, let $c \in [a]_\sim \cap [b]_\sim$. Then $a \sim c$ and $c \sim b$, so that by transitivity, $a \sim b$ and $[a]_\sim = [b]_\sim$. Thus, if $a \not\sim b$, then $[a]_\sim \cap [b]_\sim = \emptyset$. □

Exercise 5.6. Let $\{A_i\}$ be a partition of a set S. Construct an equivalence relation on S such that $S/\sim = \{A_i\}$.

Definition 5.14. Let $a, b \in \mathbb{Z}$, $n \in \mathbb{Z}^+$. We say a is congruent to b mod n if there is some $k \in \mathbb{Z}$ such that $a - b = nk$.

Proposition 5.3. *Consider the relation on $\mathbb{Z}$ defined as $x \sim y$ if and only if $x \equiv y$ mod n. This defines an equivalence relation on $\mathbb{Z}$.*

Proof. First, we show reflexivity. Note that $a - a = 0 = 0(n)$, so that a is congruent to itself mod n.

Next, suppose $x \sim y$. Then there is some $k \in \mathbb{Z}$ such that $x - y = kn$. Then, $-(y - x) = kn$, so that $y - x = -kn$. Hence $y \sim x$.

Lastly, we prove transitivity. Suppose $a \sim b$ and $b \sim c$. Then, we have $a - b = kn$ and $b - c = k'n$. Then, $b = c + k'n$. Plugging in this expression for b, we have $a - (c + k'n) = kn$, so that $a - c = kn + k'n = (k + k')n$. Hence, $a \sim c$, as desired. □

Let's examine the equivalence classes of a small n. Let $n = 3$. Then, we have $4 - 1 = 1(3)$, so $4 \in [1]_\sim$. $5 - 1 = 4, 6 - 1 = 5, 7 - 1 = 6 = 2(3)$. So far we have $[1]_\sim = \{1, 4, 7\}$. Notice a pattern? Indeed, we have

$$\{1, 4, 7, \dots\} \subset [1]_\sim.$$

Now, we are partitioning all of the integers, so what can we do with negative numbers? To solve this consider the expression $a - b = kn$. Adding b to both sides, we have $a = kn + b$. Does this trivial manipulation help us? Indeed, this relates modular arithmetic to the *Euclidean algorithm*:

Theorem 5.4. *Let $a, b \in \mathbb{Z}, b \neq 0$. Then, there exist unique integers q and r such that $a = bq + r$, where $0 \leq r < |b|$. q is called the quotient and r is the remainder.*

We do not prove this now, but instead investigate its relation to modular arithmetic. We saw above that $1 \equiv 4 \bmod 3$. Interestingly enough, this is equivalent to saying $4 = 3(1) + 1$. We also saw $1 \equiv 7$. Once again, $7 = (2)3 + 1$. So, the equivalence class of a number is in fact the remainder in the Euclidean algorithm. Why does this help us? Because now we can easily compute the equivalence class of negative numbers. Indeed, consider -2. We can write $-2 = -1(3) + 1$, so that $-2 \equiv 1 \bmod 3$. You may have noticed that adding 3 or subtracting 3 takes you to something in the same equivalence class. We now know this is easily explained by Euclid's algorithm. We can thus simplify our lives with the following proposition:

Proposition 5.4. *Let $n \in \mathbb{Z}^+$. Then, the equivalence class of a number $a \in \mathbb{Z}$ mod n can be computed as*

$$[a]_\sim = \{a + kn : k \in \mathbb{Z}\}.$$

Proof. Let $x \in [a]_\sim$. Then $a - x = kn$ for some k, so that $a - kn = x$. Then, defining $k' = -k$, we have $x = a + k'n$, so that $x \in \{a + kn : k \in \mathbb{Z}\}$ and $[a]_\sim \subset \{a + kn : k \in \mathbb{Z}\}$.

Let $x \in \{a + kn : k \in \mathbb{Z}\}$. Then $x = a + kn$ for some $k \in \mathbb{Z}$. Hence $x - a = kn$, and $x \in [a]_\sim$. Thus $\{a + kn : k \in \mathbb{Z}\} \subset [a]_\sim$.

Therefore $[a]_\sim = \{a + kn : k \in \mathbb{Z}\}$. □

We now reach the climax of our discussion on equivalence relations:

Theorem 5.5. *Let $\mathbb{Z}/n\mathbb{Z}$ denote the set of equivalence classes of $\mathbb{Z}$ mod n. Then, under the binary operation $[x]_\sim + [y]_\sim = [x + y]_\sim$, $\mathbb{Z}/n\mathbb{Z}$ is a group.*

Proof. Associativity is obvious, as well as the fact that $[0]_\sim$ acts as the identity(if you doubt this, see our discussion on the fact that adding or subtracting n leaves the remainder in the division algorithm unchanged). We check the existence of inverses. Let $[-x]_\sim$ denote the equivalence class of the inverse of $x \in \mathbb{Z}$. Then $[x]_\sim + [-x]_\sim = [x-x]_\sim = [0]_\sim$. Is this well defined on a different representative of the equivalence class of $[-x]_\sim$? Let $z \in [-x]_\sim$. Then $-x - z = kn$ for some k. Then $[z]_\sim + [x]_\sim = [z+x]_\sim = [-kn-x+x]_\sim = [-kn]_\sim = [0]_\sim$. So with another representative, we still get the equivalence class of 0, as desired. □

Example 5.8. Let $n = 3$ once again. $\mathbb{Z}/3\mathbb{Z} = \{[0], [1], [2]\}$. Indeed, for any $n \in \mathbb{Z}^+$, $\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \ldots, [n-1]\}$. These are the possible remainders when dividing by n. Once again, the division algorithm shows up.

Exercise 5.7. Let $(G, \star)$ and $(H, *)$ be groups. Prove that $G \times H$, the cartesian product of G and H, is a group. Hint: The binary operation is defined as $(g_1, h_1) \cdot (g_2, h_2) = (g_1 \star g_2, h_1 * h_2)$.

5.4 Subgroups and Normal Subgroups

Definition 5.15. Let $(G, \times)$ be a group, and let $H \subset G$. If $(H, \times)$ is itself a group, we say H is a subgroup of G.

Example 5.9. $SL(n, \mathbb{R})$ is a subgroup of $GL(n, \mathbb{R})$.

Example 5.10. Define $n\mathbb{Z}$ as $\{nk : k \in \mathbb{Z}\}$. This is a subgroup of $\mathbb{Z}$. Indeed, if $x, y \in n\mathbb{Z}$, then $x + y = an + bn = (a+b)n \in n\mathbb{Z}$. If $x = an \in n\mathbb{Z}$, then $an + (-a)n = (a-a)n = 0$, so that we have inverses.

Exercise 5.8. Let $H \subset G$ be a subgroup of G. Prove that the identity element $e \in G$ is also the identity element of H.

Definition 5.16. Let G be a group and let $H \subset G$ be a subgroup of G. The **left coset** with respect to $g \in G$ is defined as

$$gH = \{gh : h \in H\}.$$

Proposition 5.5. *Let G be a group and let $H \subset G$ be a subgroup of G. The collection of left cosets of H in G, denoted G/H, is a partition of G.*

Proof. First, note that since $e \in H$, $g \in gH$ for all $g \in G$. So every $g \in$ is contained in some coset. Now, let $x \in gH \cap g'H$, where $g, g' \in G$. Then, there are $h, h' \in H$ such that $x = gh = g'h'$. We then have $g = g'h'h^{-1}$. Let $h'h^{-1} = s \in H$. Then, for any $gt \in gH$, $t \in H$, we have $gt = g'st \in g'H$. Hence $gH \subset g'H$. Interchanging g and g', the same argument yields $g'H \subset gH$, so that $gH = g'H$. □

We can similarly define a right coset in the obvious way.

Definition 5.17. Let G be a group and let $H \subset G$ be a subgroup of G. Then H is normal in G if for all $g \in G$, we have

$$gH = Hg,$$

or equivalently $gHg^{-1} = H$.

We are interested in the existence of normal subgroups due to the following theorem:

Theorem 5.6. *Let G be a group and let $H \subset G$ be a subgroup of G. Then G/H is a group, under the operation $gHg'H = (gg')H$, if and only if H is normal in G.*

Proof. First, we show that the binary operation being well defined implies normality of H. Indeed, let $s, t \in G$, and $s' \in sH$, $t' \in tH$. Since we assume we have we well-definedness, $stH = s't'H$ for all s, t, s', t' defined as above. We show $gHg^{-1} = H$ for all $g \in G$. To that end, let $g \in G$. Let $s = e, s' = h \in H$, $t = t' = g^{-1}$. Then, by assumption, $g^{-1}H = hg^{-1}H$. We have that $hg^{-1} \in hg^{-1}H$, so that by the prior equality, there is some $r \in H$ such that $g^{-1}r = hg^{-1}$. We then have $r = ghg^{-1} \in H$, so that h is indeed normal.

We now show that, for N normal, the operation defined above is a well-defined group operation. Associativity, inverses, and identity are clear from the fact that G is a group. We need only check well-definedness. To that end, let $u, v \in G$, with $u_1 \in uH, v_1 \in vH$. We then have $uvH = u(vH) = u(Hv) = u(Hv_1) = (uH)v_1 = (u_1H)v_1 = u_1v_1H$. Hence, the operation is indeed well defined. □

Definition 5.18. Let $\phi : A \to B$ be a group homomorphism, with e_B the identity element of B. The **kernel** of a homomorphism is the following set:

$$\ker \phi = \{a \in A : \phi(a) = e_B\}.$$

Proposition 5.6. *Let $\phi : A \to B$ be a group homomorphism. Then $\ker \phi$ is a subgroup of G.*

Proof. First, recall that $\phi(e_A) = e_B$, so that $e_A \in \ker \phi$. Next, let $x, y \in \ker \phi$. Then, $\phi(xy) = \phi(x)\phi(y) = e_B$, so that $xy \in \ker \phi$. If $x \in \ker \phi$, we have $\phi(x) = e_B = \phi(x)\phi(x)^{-1}$. In particular, $\phi(x) = \phi(x)\phi(x)^{-1}$, so that, multiplying both sides by $\phi(x)^{-1}$ on the left, we have $e_B = \phi(x)^{-1} = \phi(x^{-1})$. So $x^{-1} \in \ker \phi$ and $\ker \phi$ is a group. □

Exercise 5.9. Let $\phi : A \to B$ be a group homomorphism. Prove that $\ker \phi$ is a normal subgroup of A.

We now state the main result of this section, which justifies a shift in thinking:

Theorem 5.7. *Let $H \subset G$ be a subgroup of G. Then H is normal if and only f H is the kernel of some homomorphism.*

Proof. Assume H is the kernel of some homomorphism. Then by the previous exercise, H is normal.

Conversely, let H be normal. Consider the map $\pi : G \to G/H, x \mapsto xH$. By the way we defined multiplication on G/H, this is a homomorphism. We must prove $\ker \pi = H$. To that end, let $x \in H$. Then $\pi(x) = xH = H = eH$, so that indeed $H \subset \ker \pi$. For the reverse inclusion, assume $\pi(x) = eH$. Then $H = xH$. Since $e \in H$, $x = xe \in H$, so that $\ker \pi \subset H$. Therefor $\ker \pi = H$, as desired. □

We have thus proven that, in order to study the structure of a group, we can study homomorphisms out of the group. This is a common theme in mathematics: to learn about an object, study maps to and from the object. It is not unlike how scattering experiments work in particle physics. Though we may not be able to observe a particle directly, we can deduce its presence via manufactured collisions.

6 Topology

I mentioned in class that we would revisit continuity and generalize the ϵ-δ definition. We begin this process now:

Definition 6.1. Let (X, d) be a metric space, $x \in X$, and $\epsilon > 0$. Then, the **open ball** of radius ϵ centered at x, denoted $B(x, \epsilon)$, is defined as

$$B(x, \epsilon) = \{y \in X : d(x, y) < \epsilon\}.$$

Note that $B_\epsilon(x)$ is also common notation.

Definition 6.2. Let (X, d) be a metric space. A subset $U \subset X$ is said to be **open** if, for any $x \in U$, there is some $\epsilon > 0$ such that $B(x, \epsilon) \subset U$.

Example 6.1. Consider the metric space $(\mathbb{R}, |\cdot|)$. Then open balls in $\mathbb{R}$ are precisely the open intervals you have been dealing with since college algebra. Indeed, let (a, b) be such an interval. Then $(a, b) = B(\frac{a+b}{2}, |\frac{b-a}{2}|)$. If we consider $\mathbb{R}^2$ under the metric $d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$, we see $B(x, \epsilon)$ is a circle. For $\mathbb{R}^3$, we literally have a ball.

We will see right now that a metric space and the open sets defined above have the structure of a *topological space*.

6.1 Topological Spaces

Definition 6.3. A **topological space** is is a pair (X, τ), where X is a set and $\tau \subset \mathcal{P}(X)$ satisfying the following axioms:

1. X,$\emptyset \in \tau$;
2. any arbitrary union (finite or otherwise) of elements of τ is again in τ;
3. any finite intersection of elements of τ is again in τ.

The elements of τ are called **open** sets and τ is called a topology on X.

Example 6.2. Let $X = \{1, 2, 3, 4\}$. Let $\tau_1 = \{\emptyset, X, \{2\}, \{1, 2\}, \{2, 3\}, \{1, 2, 3\}\}$ and $\tau_2 = \{\emptyset, X\}$. Both of these are topologies on X, the latter called the trivial or discrete topology.

Exercise 6.1. Let X be any set. Let $\tau = \mathcal{P}(X)$, the power set of X. Prove that (X, τ) is a topological space.

Before we begin studying the topology on generic metric spaces, we study a definition that will make our life easier.

Definition 6.4. Let X be a set. A collection $\mathcal{B}$ is a **base** for some topology on X if the following hold:

1. For all $x \in X$, there is some $B \in \mathcal{B}$ such that $x \in B$;
2. for all $x \in B_1 \cap B_2, B_1, B_2 \in \mathcal{B}$, there is some $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.

Proposition 6.1. *Let (X, d) be a metric space. Then $\mathcal{B} = \{B(x, \epsilon) : x \in X, \epsilon \in \mathbb{R}_+^\times\}$ is a basis for a topology on X.*

Proof. We clearly have the first basis condition. We prove the second. Let $\epsilon, \delta > 0$. Let $x, y \in X$ and $a \in B(x, \epsilon) \cap B(y, \delta)$. Then, choosing $\rho = \min(\epsilon - d(a, x), \delta - d(a, y))$, we have $B(a, \rho) \subset B(x, \epsilon) \cap B(y, \delta)$. □

We can now succinctly write the topology generated by a basis $\mathcal{B}$ as

$$\tau = \{\cup_\alpha B_\alpha : B_\alpha \in \mathcal{B}, \alpha \in A\},$$

where A is an appropriate indexing set. This should be understood as taking all possible unions of elements of $\mathcal{B}$.

Example 6.3. Let (X, d) be a metric space. Then the topology generated by $\mathcal{B} = \{B(x, \epsilon) : x \in X, \epsilon \in \mathbb{R}_+^\times\}$ is called the *usual topology.* I will denote it $\mathcal{U}$.

Example 6.4. Consider $\mathbb{R}$ with the absolute value. Then the usual topology is precisely the topology generated by the open intervals. Similarly with open discs in $\mathbb{R}^2$, etc.

6.2 Continuous Maps of Topological Spaces

Definition 6.5. Let (X, τ) and (Y, σ) be topological spaces. A function $f : X \to Y$ is said to be **continuous** if, for every $A \in \sigma$, we have $f^{-1}(A) \in \tau$. In words, the preimage of open sets are open.

Example 6.5. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Then, we have $f^{-1}(\mathbb{R}) = \{x \in \mathbb{R} : f(x) \in \mathbb{R}\} = \mathbb{R} \in \mathcal{U}$. We have from above that the usual topology on $\mathbb{R}$ is generated by open intervals. So, we check that the preimage of open intervals is open. Indeed, we have $f^{-1}((a, b)) = \{x \in \mathbb{R} : f(x) \in (a, b)\} = \{x \in \mathbb{R} : x^2 \in (a, b)\} = (-\sqrt{b}, -\sqrt{a}) \cup (\sqrt{a}, \sqrt{b})$.

Theorem 6.1. *Let (M_1, d_1) and (M_2, d_2) be metric spaces. A function $f : M_1 \to M_2$ is continuous (in the ϵ-δ) sense if and only if $f : (M_1, \mathcal{U}) \to (M_2, \mathcal{U})$ is continuous in the topological sense.*

Proof. Assume f is continuous in the δ-ϵ sense. Let $x \in M_2$ and let $\epsilon > 0$. Let $b \in f^{-1}(B(x, \epsilon))$. Then, $f(b) \in B(x, \epsilon)$. Since $B(x, \epsilon)$ is open, there is a $\gamma > 0$ such that $B(f(b), \gamma) \subset B(x, \epsilon)$. By definition, there exists a $\delta > 0$ such that for $y \in M_1$ with $d_1(b, y) < \delta$, we have $d_2(f(b), f(y)) < \gamma$. Then, $f(B(b, \delta)) \subset B(f(b), \gamma)$, so that $B(b, \delta) \subset f^{-1}(B(f(b), \gamma)) \subset B(x, \epsilon)$. Hence $f^{-1}(B(x, \epsilon))$ is open, as desired.

For the converse, assume f is continuous in the topological sense. Fix $x_0 \in M_1$, and let $\epsilon > 0$. Consider $B(f(x_0), \epsilon)$. We then have $f^{-1}(B(f(x_0), \epsilon))$ is open by definition of topological continuity. Since $f^{-1}(B(f(x_0), \epsilon))$ is open, there is an open ball $B(x_0, \gamma) \subset f^{-1}(f(x_0), \epsilon))$. Choosing $\delta = \gamma$, we have that for all $y \in M_1$ such that $d_1(x_0, y) < \delta$, we have $d_2(f(x_0), f(y)) < \epsilon$. □

We thus have that the ϵ-δ definition is contained within the topological definition. So we have indeed abstracted the ϵ-δ definition of continuity.

A Countable vs Uncountable

Definition A.1. Let X be a set. We say X is **countable** if at least one of the following two hold:

1. there is an injection $f : X \to \mathbb{N}$;
2. there is a surjection $g : \mathbb{N} \to X$.

If there is a bijection $\phi : X \to \mathbb{N}$, X is called **countably infinite**.

Definition A.2. A set X is **uncountable** if it is not countable.

Here are some nice results:

1. any subset of a countable set is countable;
2. if S is countable, then $S \cup \{x\}$;
3. if A and B are countable, then $A \cup B$ is countable;
4. the cartesian product of two countable sets is countable;
5. the integers and rational numbers are countable.

B Proof by Induction

The principle of mathematical induction is a proof technique that exploits the structure of the natural numbers. Specifically, we exploit the fact that $<$ is irreflexive and contains

no infinite descending chains. To quote Donald Knuth: "mathematical induction proves that we can climb as high as we like on a ladder, by proving that we can climb onto the bottom rung (the basis) and that from each rung we can climb up to the next one (the step)."

We use induction when we are proving a relation involving a natural number holds for all natural numbers. The standard structure of such an argument is as follows:

1. the initial, or base case, where we prove the relation holds for $n = 0$ or $n = 1$;
2. the induction step, where we show that if the relation holds for some $k \in \mathbb{N}$, then it must hold for $k + 1$.

In the second step, assuming that the relation holds for some arbitrary k is called the *induction hypothesis*. So, we see altogether that these two steps show a relation is true for all natural numbers.

Here is a nice example:

Proposition B.1. *For any $n \in \mathbb{N}$,*

$$0 + 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Proof. We proceed by induction. First, for $n = 0$, we have $\frac{0(1)}{2} = 0$. This completes the first step. Now assume the above relation holds for an arbitrary $k \in \mathbb{N}$ (induction hypothesis). We show that this implies the relation for $k + 1$. Indeed, assuming

$$0 + 1 + \cdots + k = \frac{k(k+1)}{2},$$

we see that:

$$\begin{aligned} 0 + 1 + \cdots + k + (k+1) &= \frac{k(k+1)}{2} + k + 1 \\ &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \\ &= \frac{(k+1)((k+1)+1}{2}. \end{aligned}$$

Hence, by the principle of mathematical induction, $0 + 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ for any $n \in \mathbb{N}$. □

Why do we care? Presumably, mathematicians do not learn induction for the purpose of torturing college algebra students or computer scientists. Despite its apparent simplicity, this is quite a formidable tool. When one studies group theory, you discover that integers and natural numbers have a lot to do with interesting structures, especially the

theory of solvable groups, Nilpotent groups, and proving results about finite groups in general.

For those interested in computer science, you prove the correctness of an algorithm using what is essentially proof by induction. You construct what is called a "loop invariant", which is essentially the heart of the information your algorithm encodes. You then perform induction on the invariant. Specifically, you prove the invariant is true before any iterations of the algorithm, and then you show that passing from the kth iteration to the $k+1$st iteration leaves your invariant true. Finally, you prove that the loop terminates and the invariant is true on the entire input. The steps are called initialization, maintenance, and termination, respectively.

As an example, the standard bubblesort algorithm receives an array of integers A of size n and outputs A in sorted order. The loop invariant for the standard algorithm is "At iteration k, the subarray $A[1,\ldots,k]$ is sorted and any element of $A[k+1,\ldots,n]$ is greater than or equal to any element in $A[1,\ldots,k]$." This indeed is true at all 3 steps, and so bubblesort is a correct algorithm.