Pre-Spivak Notes for Honors Calculus I

I decided to write up some notes for the portion of this class that comes before we start following Spivak. Spivak starts with the real numbers, and our goal is to build up some basic language so we can appreciate what he does.

1 Day 2

The basic idea is that we want to start with something very primitive and build our way up to the real numbers. The progression is as follows:

1. The real numbers $\mathbb{R}$ (e.g. numbers like -4 and π and $4/5$) are built from the rational numbers $\mathbb{Q}$ (i.e. fractions like $4/5$; whole numbers like -4 also count as fractions since we can write $-4/ = -4/1$).
2. The rational numbers $\mathbb{Q}$ are built from the integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
3. The integers $\mathbb{Z}$ are built from the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$.
4. The natural numbers $\mathbb{N}$ are built from the theory of *sets*.

The theory of sets is the primitive language underlying all modern mathematics. We will not give a formal treatment of it since that would be a course unto itself, but we will try to give you a taste of what is going on.

A set is simply a collection of things. Here are some examples:

- The set of all Honors Calculus students. Here we specify the set by simply giving a verbal description of it.
- The set {apple, orange, banana} is a finite set whose elements we list.
- The sets $\{1, 3, 5, 7, \ldots\}$ and $\{2, 4, 6, 8, \ldots\}$ of odd and even natural numbers. These sets are infinite, so we can't list them all, but the ... indicate that the pattern continues.
- The *empty set* $\emptyset$, which has no elements in it.

One important piece of notation we will use is *set-builder notation.* I will illustrate this using the even and odd numbers:

$$\{2, 4, 6, 8, \ldots\} = \{n \mid \text{there exists } k \geq 1 \text{ such that } n = 2k\}$$

and

$$\{1, 3, 5, 7, \ldots\} = \{n \mid \text{there exists } k \geq 1 \text{ such that } n = 2k - 1\}.$$

Here you should read the "|" symbols as "such that".

We will write $\in$ to say that an element is a member of a set, and $\notin$ to say that it is not. For instance,

$$2 \in \{2, 4, 6, 8, \ldots\} \quad \text{but} \quad 3 \notin \{2, 4, 6, 8, \ldots\}.$$

One important feature of sets is that there elements can be very general; in fact, it is important to allow sets to contain sets as elements! For instance, we have a set

$$\{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$$

containing all sets of size 2 consisting of numbers from $1, 2, 3$. We can even have the empty set as an element of a set. For instance, we might have

$$\{\emptyset, \{1, 2\}, \{1, 3\}, \{2, 3\}\}$$

which is different from the above set since it contains a new element:

$$\emptyset \in \{\emptyset, \{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

We will encounter lots of different kinds of sets as the course progresses, and you will gain experience with them. They seem pretty abstract at first!

Here is an important cautionary example.

Example 1.1. This example is called Russell's Paradox. Let X the set of all sets S such that $S \notin S$. In other words,

$$X = \{S \mid S \text{ a set such that } S \notin S\}.$$

Question: does X contain itself? If X does not contain itself, then by definition $X \in X$, which means that it does contain itself, which means that $X \notin X$, etc. We keep going round and round in logical circles! □

The point of this example is that we have to be somewhat careful when specifying sets if we want to avoid introducing contradictions to the foundations of mathematics. The set of rules that most mathematicians follow is called the "Zermelo-Frankel Axioms", and they are set up so that you are not allowed to talk about things like "the set of all sets", or introduce self-reference. We will not state these axioms, but we promise that all the constructions we do this semester will be consistent with them!

We now introduce some basic language for manipulating sets. We start with the following definition:

Definition 1.2. For sets A and B, the *union* of A and B, denoted $A \cup B$, is the set of all things that lie in either A or B. In set-builder notation, this is

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

In this definition, I want to emphasize that we are using the *inclusive or*: when we say that $x \in A$ or $x \in B$, we allow x to be contained in both A and B. We will use this sense of "or" throughout this class. It is helpful to think about it using a *true table*. For statements P and Q (e.g. we might have P equal to $x \in A$ and Q equal to $x \in B$), we have that "P or Q" is true if P is true and Q is false, if P is true and Q is true, and if P is false but Q is true; however, we have that "P or Q" is false if P is false and Q is false. It is traditional to use $\vee$ as a shorthand for "or", so the above rules are summarized in the following table:

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

Here are some examples of unions:

- If $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$, then $A \cup B = \{1, 2, 3, 4\}$.
- If E is the set of even natural numbers and O is the set of odd natural numbers, then $E \cup O = \mathbb{N}$.

The next operation is as follows:

Definition 1.3. For sets A and B, the *intersection* of A and B, denoted $A \cap B$, is the set of all things that lie in either A or B. In set-builder notation, this is

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}\,. \qquad \square$$

The logical connective "and" here functions just like you think it does, but just in case here is a truth table for it. In this table, we use $\wedge$ to denote "and":

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

Here are some examples of intersections:

- If $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$, then $A \cap B = \{2, 3\}$.
- If E is the set of even natural numbers and O is the set of odd natural numbers, then $E \cap O = \emptyset$. By the way, in this case we say that E and O are *disjoint*.

We close with the first small lemma of this course. Its proof is a good illustration of how you prove simple things using the definitions.

Lemma 1.4. *If A and B and C are sets, then $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.*

Proof. We have to show two things: every element of $A \cap (B \cup C)$ lies in $(A \cap B) \cup (A \cap C)$, and every element of $(A \cap B) \cup (A \cap C)$ lies in $A \cap (B \cup C)$.

We start by showing that every element of $A \cap (B \cup C)$ lies in $(A \cap B) \cup (A \cap C)$. Consider some $x \in A \cap (B \cup C)$. We want to show that $x \in (A \cap B) \cup (A \cap C)$. Since $x \in A \cap (B \cup C)$, we have $x \in A$ and $x \in B \cup C$. Since $x \in B \cup C$, we have that $x \in B$ or $x \in C$. We consider these two possibilities separately:

- If $x \in B$, then since we know $x \in A$ we can deduce that $x \in A \cap B$, and thus that $x \in (A \cap B) \cup (A \cap C)$.
- If $x \in C$, then since we know $x \in A$ we can deduce that $x \in A \cap C$, and thus that $x \in (A \cap B) \cup (A \cap C)$.

In both cases, we have achieved our goal of showing that $x \in (A \cap B) \cup (A \cap C)$.

We next show that every element of $(A \cap B) \cup (A \cap C)$ lies in $A \cap (B \cup C)$. Consider some $x \in (A \cap B) \cup (A \cap C)$. We want to show that $x \in A \cap (B \cup C)$. Since $x \in (A \cap B) \cup (A \cap C)$, we either have $x \in A \cap B$ or $x \in A \cap C$. We consider these two possibilities separately:

- If $x \in A \cap B$, we have $x \in A$ and $x \in B$. Since $x \in B$, we have $x \in B \cup C$. Since $x \in A$ and $x \in B \cup C$, we can conclude that $x \in A \cap (B \cup C)$.
- If $x \in A \cap C$, we have $x \in A$ and $x \in C$. Since $x \in C$, we have $x \in B \cup C$. Since $x \in A$ and $x \in B \cup C$, we can conclude that $x \in A \cap (B \cup C)$.

In both cases, we have achieved our goal of showing that $x \in A \cap (B \cup C)$. $\square$

This proof is perhaps not the most exciting proof, but it illustrates some important things:

1. The first step to proving something is usually to try to carefully state your goal.
2. You then carefully state what your assumptions are.
3. You then use logic to carefully transform your assumptions into your goal.

Often quite a bit of creativity is needed along the way, but it is important to get used to boring examples like the one above first! Over the next week or so, we will give a number of other simple examples of set-theoretic proofs (plus a couple of interesting and surprising ones to entertain you and make this worth your intellectual time!). You will also have to produce some in the homework.

2 Day 3

Last time we introduced sets and discussed intersections (denoted with $\cap$) and unions (denoted with $\cup$). We also discussed some logical symbols ($\wedge$ for "and", $\vee$ for "or"). In today's lecture, we will spend some more time discussing basic logic.

Our first order of business will be to talk about *negation.* If P is a statement (i.e. a sentence with a definite truth value), then $\neg P$ will denote the statement with the opposite truth value. You should read "$\neg P$" as "not P". For instance, if P is the true statement "Humans are animals", then $\neg P$ is the false statement "Humans are not animals".

To further illustrate this, consider the following examples. Let P be the statement "You are rich" and Q be the statement "You are happy".

- The statement $P \wedge Q$ is "You are rich and you are happy". Its negation $\neg(P \wedge Q)$ is the statement “You are *not* both rich and happy”, or equivalently “You are not rich or you are not happy”. We have discovered that $\neg(P \wedge Q)$ is completely equivalent to $(\neg P) \vee (\neg Q)$. This could also be illustrated by examining truth tables as follows:

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$	$\neg P$	$\neg Q$	$(\neg P) \vee (\neg Q)$
T	T	T	F	F	F	F
T	F	F	T	F	T	T
F	T	F	T	T	F	T
F	F	F	T	T	T	T

You'll observe that the columns corresponding to $\neg(P \wedge Q)$ and $(\neg P) \vee (\neg Q)$ are identical.

- The statement $P \vee Q$ is “You are rich or you are happy”. Its negation $\neg(P \vee Q)$ is the statement “You are not rich or happy”, or equivalently “You are not rich and you are not happy”. We have discovered that $\neg(P \vee Q)$ is completely equivalent to $(\neg P) \wedge (\neg Q)$. This could also be illustrated by examining truth tables as follows:

P	Q	$P \vee Q$	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$(\neg P) \wedge (\neg Q)$
T	T	T	F	F	F	F
T	F	T	F	F	T	F
F	T	T	F	T	F	F
F	F	F	T	T	T	T

You'll observe that the columns corresponding to $\neg(P \vee Q)$ and $(\neg P) \wedge (\neg Q)$ are identical.

The next logical statement we will introduce is $\Rightarrow$, which means “implies”. The statement $P \Rightarrow Q$ means that P being true implies that Q is true, or equivalently “If P, then Q”. Here are some examples:

- Returning to the case where P is “You are rich” and Q is “You are happy”, the statement $P \Rightarrow Q$ means “You are rich implies you are happy”, or equivalently “If you are rich, then you are happy”. This is (according to most philosophical systems) a false statement!
- Here is a more mathematical example. Let P be the statement “n is an even integer” and let Q be the statement “n is an integer”. The statement $P \Rightarrow Q$ means “If n is an even integer, then n is an integer”. This is true, albeit not very interesting!
- This last example illustrates an important and at-first subtle point. Let P be the statement “n is an even integer and an elephant” and Q be the statement “n is rhino”. The statement $P \Rightarrow Q$ means “If n is an even integer and an elephant, then n is a rhino”. Is this true or false? It is actually **true**, but for silly reasons: the statement “n is an even integer and an elephant” is false, so it is true that $P \Rightarrow Q$ is true, but just because it is completely empty! This is the rare fact in mathematics that is so trivial that it is confusing...

These examples illustrate what it means for $P \Rightarrow Q$ to be true: if P is true, then Q must be true, but if P is false then Q can be either true or false. So it has the following truth table:

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

Just to emphasize again: if P is false, then $P \Rightarrow Q$ is true no matter what truth-value Q has!!!

If you reflect on the meaning of $P \Rightarrow Q$, you'll see that it is true precisely when either P is false or when Q is true. So $P \Rightarrow Q$ is equivalent to the statement $(\neg P) \vee Q$, as is illustrated in the following truth table:

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

Let's now think about what happens when we negate $P \Rightarrow Q$. Since $P \Rightarrow Q$ is true when either P is false or Q is true, $\neg(P \Rightarrow Q)$ is true when P is true and Q is false. In other words, $\neg(P \Rightarrow Q)$ is equivalent to the logical statement $P \wedge (\neg Q)$, as is illustrated in the following truth table:

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

The *converse* of an implication $P \Rightarrow Q$ is the implication $Q \Rightarrow P$. These are very different statements, as is illustrated in the following examples:

- Let P be the statement "n is an even integer" and Q be the statement "n is an integer". The implication $P \Rightarrow Q$ is "If n is an even integer, then n is an integer". This is true. However, its converse $Q \Rightarrow P$ is "If n is an integer, then n is an even integer". This is false!!!
- Here is a silly example. If P is the statement "A person is a Texan" and Q is the statement "A person owns a gun", then $P \Rightarrow Q$ is the statement "If a person is a Texan, then they own a gun". This very well might be true (I lived in Texas for 10 years, and I think every native Texan I met owned a gun...). However, its converse $Q \Rightarrow P$ is the statement "If a person owns a gun, then they are a Texan". This is definitely false.

We can illustrate the difference between $P \Rightarrow Q$ and $Q \Rightarrow P$ by examining their quite different truth tables:

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

The final logical notion we will introduce is the symbol $\iff$. The statement $P \iff Q$ should be read "P is true if and only if Q is true". It means that P is true/false precisely when Q is true/false. Another way of saying it is that $P \iff Q$ is true if *both* $P \Rightarrow Q$ and its converse $Q \Rightarrow P$ are true. Here are some examples:

- Let P be the statement "n is an even integer" and Q be the statement "n is an integer". The statement $P \iff Q$ means that "n is an even integer if and only if it is an integer", which is utterly false!
- Here is a true $\iff$ statement from the theory of sets. You should reflect on why it is true. Let P be the statement "The sets A and B are equal" and let Q be the statement "$A \cap B = A$ and $A \cap B = B$". Then $P \iff Q$ means that "Sets A and B are equal if and only if $A \cap B = A$ and $A \cap B = B$". You might want to try to write a formal proof of this, but at the very least you should convince yourself that it is true.

To close, here is the truth table for $P \iff Q$ and $(P \wedge Q) \vee ((\neg P) \wedge (\neg Q))$ to illustrate the fact that they are equivalent statements:

P	Q	$P \iff Q$	$P \wedge Q$	$(\neg P) \wedge (\neg Q)$	$(P \wedge Q) \vee ((\neg P) \wedge (\neg Q))$
T	T	T	T	F	T
T	F	F	F	F	F
F	T	F	F	F	F
F	F	T	F	T	T

3 Day 4

Our last lecture was devoted to logic. The main topic was implications like $P \Rightarrow Q$. Today we will start with a little bit of set theory to illustrate this. We start with the following definition.

Definition 3.1. A set A is a *subset* of a set B if every element of A is also an element of B. We will write this by $A \subset B$. We allow the possibility that $A = B$. If $A \subset B$ but $A \neq B$, then we will say that A is a *proper subset* of B and write $A \subsetneq B$. □

Example 3.2. If B is the set of Honors Calculus students and A is the set of Honors Calculus students who live in Alumni Hall, then $B \subsetneq A$. □

Example 3.3. For a more mathematical example, we have the following inclusions between the sets of numbers we have briefly talked about:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

It is clear that $\mathbb{N} \subsetneq \mathbb{Z}$ since $-1 \in \mathbb{Z}$ but $-1 \neq \mathbb{N}$. Similarly, it is clear that $\mathbb{Z} \subsetneq \mathbb{Q}$. It is also true that $\mathbb{Q} \subsetneq \mathbb{R}$. You've probably been told that numbers like $\sqrt{2}$ and π are not rational numbers, but this is not obvious! Rather than asking you to trust an authority figure like me, I promise to prove these facts to you during the course! $\square$

This is a good opportunity to introduce another logical symbol. We will use the symbol $\forall$ to denote "for all". Its use is illustrated in the following symbolic statement that is true precisely when $A \subset B$:

$$\forall x(x \in A \Rightarrow x \in B).$$

You should reflect on why this is *exactly* the same as the definition of $A \subset B$.

If A is *not* a subset of B, then we will write $A \not\subset B$. How can we express this using our logical notation? One possibility would be to simply write

$$\neg\forall x(x \in A \Rightarrow x \in B).$$

This says that it is false that for all $x \in A$, we have $x \in B$. An equivalent statement would be that there exists some $x \in B$ such that $x \notin A$. We will use the mathematical symbol $\exists$ to express the statement "there exists", so $A \not\subset B$ is equivalent to the assertion that

$$\exists x(x \in B \wedge x \notin A).$$

Here is another true logical statement involving $\exists$:

$$\exists x(x \in A \cap B) \iff A \cap B \neq \emptyset.$$

You should reflect on the meaning here to see why it is true!

We now turn to the final abstract topic we will discuss before we launch into our discussion of numbers. Much of your previous education was devoted to functions like $f(x) = \cos(x)$ or $f(x) = \sqrt{x}$. You might have been led to think that a function is the same thing as a formula. This is not true at all, and to think clearly about the foundations of calculus it will be important to adopt a more abstract definition of a function.

So what is a function? Informally, it is a rule that associates to every element x of a set X an element $f(x)$ of a set Y. We will write such a function as $f\colon X \to Y$ and call X the *domain* of f and Y the *codomain*. Here are a bunch of examples:

Example 3.4. If X is the set of Honors Calculus students and Y is the set of countries in the world, then there is a function $f\colon X \to Y$ with $f(x) \in Y$ the birthplace of student $x \in X$. $\square$

Example 3.5. That might not sound very mathematical. We can definitely define functions like $f\colon \mathbb{R} \to \mathbb{R}$ with $f(x) = \cos(x)$. We could also define $g\colon \mathbb{R} \to [-1, 1]$ with $g(x) = \cos(x)$. This is a *different* function than f since its codomain is different. $\square$

Example 3.6. There are also numerical functions that are not given by formulas. For instance, we could define $\phi\colon \mathbb{N} \to \mathbb{N}$ by letting $\phi(n) \in N$ be the n^{th} decimal digit of π. While we can easily calculate this, there is no known formula for it! $\square$

Example 3.7. Here is another interesting numerical function. Define $\psi\colon \mathbb{R} \to \mathbb{R}$ via the rule

$$\psi(x) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

You might try to imagine what the graph of this looks like! $\square$

Example 3.8. There are also functions that are clearly important in math, but whose domains/codomains are not sets of numbers. Let $\mathbb{R}[t]$ be the set of real polynomials in one variable t, so for instance

$$3t^2 - 7t + \pi \in \mathbb{R}[t].$$

There is then a function $D\colon \mathbb{R}[t] \to \mathbb{R}[t]$ defined by the formula $D(f(t)) = f'(t)$, where f' is the derivative of f. Thus

$$D(3t^2 - 7t + \pi) = 6t - 7.$$ $\square$

Example 3.9. An important function is the *identity function.* For a set Z, the identity function of Z is the function $\mathbb{1}_Z\colon Z \to Z$ defined via the formula $\mathbb{1}_Z(x) = x$ for all $x \in Z$. $\square$

Remark 3.10. See Prof. Nicolaescu's notes for how to interpret functions as special kinds of sets. In math, it is sets all the way to the bottom!

Now consider a function $f\colon X \to Y$. For $A \subset X$, we write

$$f(A) = \{y \in Y \mid \exists x(x \in X \wedge f(x) = y\}.$$

In other words, $f(A) \subset Y$ is the set of all elements of Y that are "hit" by f. Here we can take $A = X$, in which case we call $f(X)$ the *range* of f. Note that this is different from the codomain! Here is an example:

Example 3.11. Let's return to the example $\psi\colon \mathbb{R} \to \mathbb{R}$ defined by

$$\psi(x) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

We then have $\psi(\mathbb{Q}) = \{1\}$, and the range of ψ is $\psi(\mathbb{R}) = \{0, 1\}$. Notice that this range is *different* from the codomain. $\square$

4 Day 5

Today we continue our discussion of functions. Consider a function $f\colon X \to Y$. Last time we discussed the image $f(A)$ of a subset $A \subset X$. We now reverse this. For $y \in Y$, define

$$f^{-1}(y) = \{x \mid x \in X \text{ and } f(x) = y\}.$$

In other words, $f^{-1}(y)$ is all elements of X that map to y. Here is an example.

Example 4.1. Let $\psi\colon \mathbb{R} \to \mathbb{R}$ be

$$\psi(x) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

Then $\psi^{-1}(1) = \mathbb{Q}$ and $\psi^{-1}(0) = \mathbb{R} \setminus \mathbb{Q}$ and $\psi^{-1}(1/2) = \emptyset$. $\square$

More generally, consider a subset $B \subset Y$. Define

$$f^{-1}(B) = \{x \mid x \in X \text{ and } f(x) \in B\}$$
$$= \bigcup_{b \in B} f^{-1}(b).$$

Example 4.2. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined via $f(x) = \sin(x)$. Then

$$f^{-1}([0,1]) = \bigcup_{k=-\infty}^{\infty} [2k\pi, (2k+1)\pi].$$

The similar-looking function $g\colon [-\pi, \pi] \to \mathbb{R}$ defined via $g(x) = \sin(x)$ would have

$$g^{-1}([0,1]) = [0, \pi].$$

So the domain is important! □

We now define the composition of two functions.

Definition 4.3. The *composition* of functions $f\colon X \to Y$ and $g\colon Y \to Z$ is the function $g \circ f\colon X \to Z$ defined by $g \circ f(x) = g(f(x))$ for $x \in X$. □

Observe that for this to make sense, it is important for the codomain of f to equal the domain of g. Here is an example:

Example 4.4. Let $f\colon \mathbb{Q} \to \mathbb{N}$ be the function $f(a/b) = a^2 + 1$, where $a/b \in \mathbb{Q}$ is in lowest terms. Let $g\colon \mathbb{N} \to \mathbb{R}[t]$ be the function $g(n) = (t-1)^n$. Then $g \circ f\colon \mathbb{Q} \to \mathbb{R}[t]$ is the function

$$g \circ f(a/b) = (t-1)^{a^2+1},$$

where $a/b \in \mathbb{Q}$ is in lowest terms. □

We now discuss three important properties that function might have. The first is as follows:

Definition 4.5. A function $f\colon X \to Y$ is *injective* if for all $x, x' \in X$,

$$f(x) = f(x') \Rightarrow x = x'.$$ □

The equivalent contrapositive of this might be more intuitive:

$$x \neq x' \Rightarrow f(x) \neq f(x').$$

In other words, different elements of X must map to different elements of Y. Here is a sample proof involving injectivity:

Lemma 4.6. *If $f\colon X \to Y$ and $g\colon Y \to Z$ are functions such that $g \circ f\colon X \to Z$ is injective, then f is injective.*

Proof. Consider $x, x' \in X$, and assume that $f(x) = f(x')$. We must prove that $x = x'$. Since $f(x) = f(x')$, we have $g \circ f(x) = g \circ f(x')$. Since $g \circ f$ is injective, this implies that $x = x'$, as desired. □

The second is as follows:

Definition 4.7. A function $f\colon X \to Y$ is *surjective* if for all $y \in Y$, there exists some $x \in X$ such that $f(x) = y$. $\square$

A sample proof involving surjectivity is as follows:

Lemma 4.8. *If $f\colon X \to Y$ and $g\colon Y \to Z$ are functions such that $g \circ f\colon X \to Z$ is surjective, then g is surjective.*

Proof. Consider some $z \in Z$. We must find some $y \in Y$ such that $g(y) = z$. Since $g \circ f\colon X \to Z$ is surjective, there exists some $x \in X$ such that $g \circ f(x) = z$. Setting $y = f(x)$, we have $g(y) = g \circ f(x) = z$, as desired. $\square$

The third condition is as follows:

Definition 4.9. A function $f\colon X \to Y$ is *bijective* if there exists a *two-sided inverse*, i.e. a function $F\colon Y \to X$ such that $F \circ f = \mathbb{1}_X$ and $f \circ F = \mathbb{1}_Y$. $\square$

Remark 4.10. It is traditional to write the two-sided inverse of a bijective function $f\colon X \to Y$ as $f^{-1}\colon Y \to X$. We will eventually start doing this, but we do not do so right now since we do not want to confuse this with the preimage of f.

What does this have to do with being injective and surjective? In the homework, you will prove the following:

Lemma 4.11. *Let $f\colon X \to Y$ be a function. The following then hold:*

- *The function f is injective if and only if there exists some $G\colon Y \to X$ such that $G \circ f = \mathbb{1}_X$.*
- *The function f is surjective if and only if there exists some $H\colon Y \to X$ such that $f \circ H = \mathbb{1}_Y$.*

These functions G and H are known as *one-sided inverses*. We then have the following:

Lemma 4.12. *A function $f\colon X \to Y$ is bijective if and only if f is both injective and surjective.*

Proof. We first prove that if f is bijective, then f is injective and surjective. Assume that f is bijective, and let F be its two-sided inverse. We thus have $F \circ f = \mathbb{1}_X$, so by Lemma 4.6 the function f is injective. We also have $f \circ F = \mathbb{1}_Y$, so by Lemma 4.8 the function f is surjective. We have proven that f is both injective and surjective, as desired.

We now prove that if f is both injective and surjective, then f is bijective. Assume that f is both injective and surjective. Since f is injective, Lemma 4.11 shows that there exists a one-sided inverse $G\colon Y \to X$ such that $G \circ f = \mathbb{1}_X$. since f is surjective, Lemma 4.11 shows that there exists a one-sided inverse $H: Y \to X$ such that $f \circ H = \mathbb{1}_Y$. We want to prove that f is bijective, i.e. that f has a two-sided inverse F. To do this, it is enough to prove that $G = H$ since then $F = G = H$ is a two-sided inverse. We check this as follows:

$$G = G \circ \mathbb{1}_Y = G \circ (f \circ H) = (G \circ f) \circ H = \mathbb{1}_X \circ H = H.$$

This is what we were trying to show, so the lemma follows. □

5 Day 6

Our goal today is to answer the following basic question:

Question 5.1. *How can we measure the number of elements in a set?*

It is clear what this means for sets containing finitely many elements. For instance, the set {apple, orange, banana} contains 3 elements. But how about infinite sets like $\mathbb{N}$? Is there an interesting way to say that one infinite set is "larger" than another?

The first basic thing we need to agree on is what it means for two infinite sets to have the same size. We will use the following definition:

Definition 5.2. We say that a set X has the *same cardinality* as a set Y if there exists a bijection $f: X \to Y$. □

Here are a few trivial observations about this:

- X has the same cardinality as itself: use the bijection $\mathbb{1}_X: X \to X$.
- If X has the same cardinality as Y, then Y has the same cardinality as X. Indeed, if $f: X \to Y$ is a bijection, then its inverse $f^{-1}: Y \to X$ is a bijection.
- If X has the same cardinality as Y and Y has the same cardinality as Z, then X has the same cardinality as Z. Indeed, if $f: X \to Y$ and $g: Y \to Z$ are bijections, then $g \circ f: X \to Z$ is a bijection.

Later on in your mathematical education you will learn that these three properties mean that the relationship "has the same cardinality" is an *equivalence relation* on sets. We will simply observe that if we use the notation $|X| = |Y|$ to mean that X has the same cardinality as Y, then the above three properties imply that this use of "=" has all the properties of the more familiar uses of the equals sign.

Here is a first example.

Example 5.3. Letting $[n] = \{1, \ldots, n\}$, if a set X satisfies $|X| = |[n]|$, then X has n elements in the usual sense. For instance, here is a bijection $f: \{\text{apple, orange, banana}\} \to [3]$

$$f(\text{apple}) = 1$$
$$f(\text{orange}) = 2$$
$$f(\text{banana}) = 3$$

□

So how about infinite sets? It turns out that most people's intuition about the cardinalities of infinite sets are **completely wrong**. The rest of this lecture is really a series of examples illustrating the subtle issues here.

Let's first consider the set

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

Your intuition probably says that $\mathbb{N}$ is larger than its subset

$$Y = \{3, 4, 5, \ldots\}.$$

However, this is false, and $|\mathbb{N}| = |Y|$. Here's a bijection $f\colon \mathbb{N} \to Y$:

$$\begin{aligned} f(1) &= 3 \\ f(2) &= 4 \\ f(3) &= 5 \\ &\vdots \end{aligned}$$

The point here is that the "names" of the elements of a set are irrelevant to its cardinality: the bijection in some sense allows us to change them in arbitrary ways!

Here's a more drastic example. Let

$$E = \{2, 4, 6, \ldots\}$$

be the set of even integers. Is $\mathbb{N}$ bigger than E? Again, they have the same cardinality! Indeed, the function $f\colon \mathbb{N} \to E$ with $f(n) = 2n$ is a bijection.

We can play this kind of game forever, and in the homework you will prove the ultimate result in this direction:

Lemma 5.4. *Let $Y \subset \mathbb{N}$ be any infinite subset. Then $|\mathbb{N}| = |Y|$.*

Sets with the same cardinality as $\mathbb{N}$ are important enough to have a special name:

Definition 5.5. A set X is *countably infinite* if $|\mathbb{N}| = |X|$. A set X is *countable* if it is either finite or countably infinite. □

This terminology arises because a set is countable if we can enumerate its elements in sequential order (i.e. "count" its elements). For instance, if O is the set of odd numbers, then O is countably infinite since we can enumerate O as

$$1, 3, 5, 7, 9, \ldots$$

This establishes a bijection between $\mathbb{N}$ and O that takes $n \in \mathbb{N}$ to the n^{th} term in the list.

Many sets that appear to be "larger" than $\mathbb{N}$ are themselves countable. For instance, we have the following:

Lemma 5.6. *The set $\mathbb{Z}$ of integers is countable.*

Proof. We can list the elements of $\mathbb{Z}$ as

$$0, 1, -1, 2, -2, 3, -3, \ldots$$

This corresponds to the bijection $f\colon \mathbb{N} \to \mathbb{Z}$ given by

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n-1}{2} & \text{if } n \text{ is odd.} \end{cases}$$ □

A more interesting example is the set of rational numbers $\mathbb{Q}$. Surely this is not countable! But it is:

Lemma 5.7. *The set $\mathbb{Q}$ is countable.*

Proof. We enumerate elements of $\mathbb{Q}$ by first writing down the fractions $\pm a/b$ with $a + b = 1$, then those with $a + b = 2$, then those with $a + b = 3$, etc. Here that is done explicitly (to simplify our notation, we will write $\pm\frac{a}{b}$ to mean $\frac{a}{b}, -\frac{a}{b}$):

$$\begin{array}{ll} \frac{0}{1} & \\ \pm\frac{1}{1} & \\ \pm\frac{2}{1}, \pm\frac{1}{2} & \\ \pm\frac{3}{1}, \pm\frac{1}{3} & \qquad (\pm\frac{2}{2} \text{ not needed since not in lowest terms!}) \\ \pm\frac{4}{1}, \pm\frac{3}{2}, \pm\frac{2}{3}, \pm\frac{1}{4} & \\ \vdots & \end{array}$$

Here each rational number appears in some row, the the bijection $\mathbb{N} \to \mathbb{Q}$ first goes along the first row, then the second, then the third, etc. □

At this point, you might be tempted to conjecture that *every* set is countable. But this is wrong:

Theorem 5.8. *The set $\mathbb{R}$ is not countable.*

This theorem was first proved by Cantor via what is called his *diagonal argument*. I think this is one of the most beautiful proofs in undergraduate mathematics, and its discovery was one of the events that precipitated the modern approach to set theory (and the foundations of math more generally). I will present it on Monday.

6 Day 7

Recall from last time that if there exists a bijection $f\colon X \to Y$ between two sets, then we say that X and Y have the *same cardinality*, in which case we write $|X| = |Y|$. If $|X| = |\mathbb{N}|$, then we say that X is *countably infinite*. We proved that many sets are countably infinite; for instance, the set of even natural numbers, the set of integers, and the set of rational numbers. In the homework you will prove that many other natural sets are countably infinite.

Our first goal today is to prove the following beautiful theorem of Cantor, which provides our first example of an infinite set that is so large that it is *not* countably infinite.

Theorem 6.1. *The set $\mathbb{R}$ is not countable.*

Proof. In the homework, you will show that $|\mathbb{R}| = |(0,1)|$, so it is enough to prove that the open interval $(0,1)$ is not countably infinite (this is just for convenience – the same argument works for $\mathbb{R}$, but requires slightly more complicated notation). To prove that $(0,1)$ is not countably infinite, it is enough to prove that all functions $f\colon \mathbb{N} \to (0,1)$ are non-surjective.

So consider some function $f\colon \mathbb{N} \to (0,1)$. We want to produce some $x \in (0,1)$ such that x is not in the image of f. Consider the decimal expansions of the $f(n)$:

$$\begin{aligned} f(1) &= 0.d_{11}d_{12}d_{13}d_{14}\cdots \\ f(2) &= 0.d_{21}d_{22}d_{23}d_{24}\cdots \\ f(3) &= 0.d_{31}d_{32}d_{33}d_{34}\cdots \\ f(4) &= 0.d_{41}d_{42}d_{43}d_{44}\cdots \\ &\ \ \vdots \end{aligned}$$

Thus $d_{ij} \in \{0, 1, \ldots, 9\}$ is the j^{th} decimal digit of $f(i)$. Here our convention is that a decimal expansion can *not* end in infinitely many 9's (this is needed to make sure that decimal expansions are unique!). For each n, let e_n be a decimal digit that is *different* from d_{nn} and does not equal 9. Setting

$$x = 0.e_1e_2e_3e_4\cdots,$$

we cannot have $f(n) = x$ for any n since the digits d_{nn} and e_n are different. It follows that x is not in the image of f and thus that f is not surjective, as desired. □

In the homework, you will prove that for all infinite sets $X \subset \mathbb{N}$, the set X is countably infinite. The above theorem says in some sense that $\mathbb{R}$ is "bigger" than $\mathbb{N}$. Informally, the following question asks whether there are infinities that lie strictly between $\mathbb{N}$ and $\mathbb{R}$:

Question 6.2 (Continuum hypothesis, Cantor 1878)**.** *Let $X \subset \mathbb{R}$ be an infinite set. Must it be the case that either $|X| = |\mathbb{N}|$ or $|X| = |\mathbb{R}|$?*

This question spurred a huge amount of work in the foundations of mathematics, and appeared in Hilbert's famous list of the most important open problems in math at the beginning of the 20th century. It was resolved in amazing work of Gödel and Cohen:

Theorem 6.3 (Gödel, Cohen)**.** *The continuum hypothesis cannot be proven or disproven from the ZFC axioms of set theory.*

It is thus not clear at all whether or not the continuum hypothesis is true. It might be that we are missing some basic axioms, but no one has managed to find any that are universally accepted. Figuring out what kind of truth value to assign to the continuum hypothesis remains a fundamental question in the philosophy of mathematics.

We now return to earth. We have been talking about what it means for two sets to have the same cardinality, i.e. for $|X| = |Y|$. How can we make precise what it would mean for Y to have a *larger* cardinality than X? We make the following definition.

Definition 6.4. For sets X and Y, we say that the cardinality of Y is *at least* the cardinality of X if there exists an injection $\iota\colon X \to Y$. We write this as $|X| \leq |Y|$. □

Here is some examples:

Example 6.5. We have $|\mathbb{N}| \leq |\mathbb{Z}|$. Indeed, the inclusion $\iota\colon \mathbb{N} \to \mathbb{Z}$ map is an injection. In this case, we actually have $|\mathbb{N}| = |\mathbb{Z}|$ since in the last class we wrote down a bijection between $\mathbb{N}$ and $\mathbb{Z}$. □

Example 6.6. We have $|\mathbb{N}| \leq |\mathbb{R}|$. Indeed, the inclusion $\iota\colon \mathbb{N} \to \mathbb{R}$ is an injection. Cantor's theorem above says that $|\mathbb{N}| \neq |\mathbb{R}|$. □

The following theorem says that this notion of inequality behaves like our ordinary notion.

Theorem 6.7 (Cantor–Schröder–Bernstein)**.** *For sets X and Y, if there exist injections $f\colon X \to Y$ and $g\colon Y \to X$, then there exists a bijection $h\colon X \to Y$.*

In other words, if $|X| \leq |Y|$ and $|Y| \leq |X|$, then we have $|X| = |Y|$. We will not use this theorem, so we will not prove it (though the proof does not use any technology we do not know). You thus should not use it in the homework.

This raises the natural question of whether for all sets X, there exists a set Y whose cardinality is strictly greater than that of X. As we will now see the answer to this is *yes*.

We make the following definition.

Definition 6.8. Let X be a set. The *power set* of X, denoted $\mathcal{P}(X)$, is the set of all subsets of X. □

For the natural numbers $\mathbb{N}$, the power set $\mathcal{P}(\mathbb{N})$ thus contains elements like the following:

- The empty set $\emptyset$.
- One element sets like $\{2\}$ and $\{3\}$.
- Finite sets like $\{3, 231, 7\}$.
- The even numbers, and the odd numbers.
- The whole set $\mathbb{N}$.

We have the following lemma.

Lemma 6.9. *For all sets X, we have $|X| \leq |\mathcal{P}(X)|$.*

Proof. We must exhibit an injection $\iota\colon X \to \mathcal{P}(X)$. We define ι via the following formula:

$$\iota(p) = \{p\} \qquad (p \in X).$$

Here $\{p\} \in \mathcal{P}(X)$ is the one-element set containing p. $\square$

The following theorem of Cantor shows that the cardinality of the power set is strictly greater than the cardinality of X. Its proof is another example of the diagonal argument, though in a more abstract form.

Theorem 6.10 (Cantor)**.** *For all sets X, we have $|X| \neq |\mathcal{P}(X)|$. In other words, there is no bijection $f\colon X \to \mathcal{P}(X)$.*

Proof. In fact, we will show that there is no surjection $f\colon X \to \mathcal{P}(X)$. Consider some function $f\colon X \to \mathcal{P}(X)$. To see that f is not a surjection, it is enough to exhibit some $S \in \mathcal{P}(X)$ such that S is not in the image of f.

This element S should be a subset of X. Define

$$S = \{p \in X \mid p \notin f(p)\}.$$

We claim that S is not in the image of f. Indeed, consider some $a \in X$. We want to show that $f(a) \neq S$. We either have $a \in f(a)$ or $a \notin f(a)$. We consider these two cases separately:

- If $a \in f(a)$, then by definition $a \notin S$. Since $a \in f(a)$ but $a \notin S$, we must have $f(a) \neq S$.
- If $a \notin f(a)$, then by definition $a \in S$. Since $a \notin f(a)$ but $a \in S$, we must have $f(a) \neq S$.

In both cases we have shown that $f(a) \neq S$, as desired. $\square$