Week 1 summary

Wednesday: Discussed course, and honours vrs nonhonours math

Thursday: part of "Language and Grammar of Math" theme

sets: the language of math (see Section 1.A of Bowman's notes)

A *set* is a collection of things (called its *elements*). The things can be numbers, words, animals, other sets, ... Order doesn't matter. Each element occurs only once: e.g. $\{1,2,3\}$ and $\{2,3,2,1,1\}$ are the same set.

"$a \in A$" means "a is an element of the set A"

- for example, if $A = \{1,2,3,4,5\}$, then $2 \in A$

"$B \subseteq A$" means "B is a subset of A"

- for example, if $A = \{1,2,3,4,5\}$, then $\{2,3\} \subseteq A$.

  WARNING: Bowman's Notes and Muldowney's Notes both use $A \subset B$ to denote subset. '$\subseteq$' is more common, and what we will use.

$\{\}$ is called the *empty set*. It has no elements, but it is a subset of every set.

"$A \cup B$" means "the union of A and B": $A \cup B$ is the set containing all elements of A as well as all elements of B

- for example, $\{1,2,3\} \cup \{2,3,4\} = \{1,2,3,4\}$

"$A \cap B$" means "the intersection of A and B": $A \cap B$ is the set containing all elements in A which are also in B

- for example, $\{1,2,3\} \cap \{2,3,4\} = \{2,3\}$

Russell's paradox: Let R be the set of all sets which don't contain themselves. Is $R \in R$?
(Russell's paradox will not be on any exam. The modern explanation of it is that R is not a well defined set. The "set of all sets" doesn't exist)

We build sets using notation sometimes called *set-builder notation*:

- for example, $\{n \in \mathbb{N} \,|\, 5 < n \leq 9\}$ means "all natural numbers n such that n is greater than 5 and less than or equal to 9". This set equals $\{6,7,8,9\}$.
- for another example, $\{x \in \mathbb{R} \,|\, x^2 + 3x + 2 = 0\}$ means "all real numbers x such that x satisfies the equation $x^2+3x+2=0$". This set equals $\{-1,-2\}$.

You must get familiar with sets
(fortunately, this is pretty easy)

Numbers (see Section 1.B of Bowman's Notes)

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

integers $\mathbb{Z} = \{...,-3,-2,-1,0,1,2,3,...\} = \mathbb{N} \cup \{0\} \cup -\mathbb{N}$

rationals $\mathbb{Q} =$ the set of all fractions $= \{a/b \,|\, a \in \mathbb{Z},\ b \in \mathbb{N}\}$

reals $\mathbb{R}$ (all numbers on the number line, e.g. all fractions, as well as numbers like π, $\sqrt{2}$,...)

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$$

(NOTE: infinity ∞ is *not* in any of these sets of numbers)

Friday: More on the "Language and Grammar of Math" theme

We had a **quiz** (Quiz 1 on eclass) on sets. (Quizzes aren't for marks)

Russell's Paradox tells us that Humans are bad at math. Our intuitions lead us astray. Things that look reasonable, can be completely wrong. So we have to be very very careful, very very precise, very very logical. We don't want to be, but we have to be. Or we'll get into all kinds of trouble. So let's describe the grammar of math, which is logic!

A *statement* is any meaningful mathematical expression which is either true or false (If it has variables in it, then it will be true or false once you've substituted the variables for precise numbers or sets or functions or ..., which ever is appropriate)

- for example, "$x \geq 4$" is a statement. Here, the variable x stands for a number. When x is big enough (at least 4), then the statement is true; when x is small enough, it is false.

The logical words

The logical words are ways to build up more complicated statements from simpler ones. You have to be as good with these as you are with the addition table. The mathematical use of these words is more precise than what we normally use in language. I'm writing them in capital letters here, but usually we'll write them in lowercase.

NOT just changes the truth value, from **true** to **false**, and **false** to **true**

- for example, "NOT($x \geq 3$)" is the same as "$x < 3$"

AND: "*Statement1* AND *Statement2*" is true exactly when both *Statement1* is true, and *Statement2* is true.

- for example, "n is a natural number" AND "$n < 5$" AND "$n \geq 3$", is the same as

$$\{n \in \mathbb{N} \,|\, 5 > n \geq 3\}$$

The 'truth table' for AND is

AND	**T**	**F**
T	**T**	**F**
F	**F**	**F**

This means e.g. "**true** AND **false** is **false**, etc.

OR: "*Statement1* OR *Statement2*" is true when either *Statement1* is true, or *Statement2* is true, or both are true.

- for example, "$n < 5$" OR "$n \geq 3$", is satisfied by all natural numbers n, including $n = 3, 4$.

The 'truth table' for OR is

OR	**T**	**F**
T	**T**	**T**
F	**T**	**F**

Week 2 summary

Monday:

IF *Statement1* **THEN** *Statement2*

Statement1 is called the premise, and *Statement2* the conclusion. This has a direction (i.e. is noncommutative), unlike AND and OR. "IF **true** THEN **false**" is false, all other truth combinations are true.

- for example, "IF $x \geq 3$, THEN $x \geq 1$" is **true** for all real numbers x (even e.g. $x = 0$)

Notice that if the premise is **false**, then "IF...THEN" is **true** regardless of the conclusion.

- for example, "IF $2 < 1$ THEN all men are from Mars" is **true**

The conclusion may not have anything to do with the premise, but the IF...THEN is still true.

- for example, "IF $2 \geq 1$ THEN the empty set is a subset of $\{1, 2, 3\}$" is **true**

There are other ways to write "IF *Statement1* THEN *Statement2*":

Statement1 $\Longrightarrow$ *Statement2*
Statement2 if *Statement1*
Statement1 only if *Statement2*
NOT(*Statement1*) OR *Statement2*
Statement2 is necessary for *Statement1*
Statement1 is sufficient for *Statement2*

We will primarily use IF...THEN as well as $\Longrightarrow$

IF AND ONLY IF

"*Statement1* IF AND ONLY IF *Statement2*"
means the same as both
"IF *Statement1* THEN *Statement2*" AND "IF *Statement2* THEN *Statement1*"

It is common to abbreviate "IF AND ONLY IF" to "IFF", and this is what we'll usually do. Another symbol for it is $\Longleftrightarrow$

"*Statement1* IFF *Statement2*" means that *Statement1* and *Statement2* are equals, as far as logic is concerned. When "*Statement1* IFF *Statement2*" is true, we think of *Statement2* as the same as *Statement1*, just written in a different way.

- for example, "$x \in \mathbb{N}$ IFF ($x \in \mathbb{Z}$ AND $x > 0$)" is **true**

The 'truth table' for IFF is

IFF	**T**	**F**
T	**T**	**F**
F	**F**	**T**

IFF has two directions, so there are **TWO** two things to prove when you have to prove a Theorem involving IFF. One direction is "$\Longrightarrow$" (i.e. assume *Statement1* and prove *Statement2*), and the other direction is "$\Longleftarrow$" (i.e. assume *Statement2* and prove *Statement1*).

Definition, Theorem, proof, axiom: part of the grammar of math

A *Theorem* is a statement which is true for any substitution of the variables. When we say "Theorem", we usually mean that we also know how to prove it.

A *Lemma* is a little theorem. It is easy to prove. Lemmas are simple but often very useful.

A *Definition* is not true or false; it is an agreement to make a short-hand notation or terminology.

An *axiom* or *postulate* is a statement which we define to be true.

- For example, we can define that an integer n is even if $n = 2k$ for some integer k, otherwise we call n odd.

- For example, a Theorem is that even plus even is even; and even times even is even.

A *proof* is a sequence of statements. Each statement follows logically from axioms, or the premises of the Theorem we're trying to prove, or previous steps in the proof, or replacing words (like "even" or "odd") with what a definition says they mean. Going from one statement in the proof to the next is supposed to be small and obvious.

A *proof strategy* is a friendly recommendation on how to get started when doing a proof.

Wednesday:

Logic and sets: proof strategies

"$x \in A \cap B$" means "$x \in A$ AND $a \in B$"

Proof strategy: To prove $x \in A \cap B$, you must prove two things: that $x \in A$ and that $x \in B$

"$x \in A \cup B$" means "$x \in A$ OR $a \in B$"

Proof strategy: To prove $x \in A \cup B$, you need to prove that either $x \in A$ or $x \in B$. Hopefully it is obvious that it is in one or the other. Otherwise, you can suppose that $x \notin A$, and then prove it must be in B.

"$A \subseteq B$" means "IF $a \in A$ THEN $a \in B$".

Proof strategy: To prove that $A \subseteq B$, start your proof with "Let $a \in A$." Then write down the properties that a has, thanks to it being an element of A. You want to show that those properties are enough to know that a is also in B, so write down what properties a would need to have in order that it be in B. Then prove those properties one by one.

"$A = B$" means the same as

"$A \subseteq B$ AND $B \subseteq A$"

"$A = B$" also means the same as

"$a \in A$ IFF $a \in B$"

Proof strategy: To prove $A = B$, remember there are **TWO** things to prove.

"$\Longrightarrow$": Assume $a \in A$, and prove $a \in B$.

"$\Longleftarrow$": Assume $b \in B$, and prove $b \in A$.

Lemma 1. $A \subseteq A \cup B$

To prove Lemma 1, first assume $a \in A$. We need to show that that a also lies in $A \cup B$. But certainly $a \in A \cup B$, because $a \in A$. Q.E.D.

Lemma 2. $A \cap B \subseteq A$

To prove Lemma 2, first take any $x \in A \cap B$. That means both $x \in A$ and $x \in B$. In particular, $x \in A$, and we are done. QED

Thursday:

Theorem. $A \subseteq B$ IFF $A = A \cap B$. Also, $A \subseteq B$ IFF $B = A \cup B$.

Proof. We must show 2 things: $A \subseteq B \Rightarrow A = A \cap B$, and $A \subseteq B \Leftarrow A = A \cap B$.

'$\Rightarrow$': To prove $A \subseteq B \Rightarrow A = A \cap B$, we can assume the premise: $A \subseteq B$. We want to use that to prove $A = A \cap B$, in otherwords, to prove both $A \subseteq A \cap B$ and $A \supseteq A \cap B$. To prove $A \subseteq A \cap B$, let $a \in A$. We need to show both $a \in A$ (this is obvious) and $a \in B$ (this is because $A \subseteq B$). Therefore $A \subseteq A \cap B$. Proving $A \subseteq A \cap B$ is trivial: just use Lemma 2. Together those tell us that $A = A \cap B$, and we're done the direction ' $\Rightarrow$'.

'$\Leftarrow$': To prove $A \subseteq B \Leftarrow A = A \cap B$, we can assume the premise: $A = A \cap B$. We want to prove $A \subseteq B$, i.e. that whenever $a \in A$, then $a \in B$. So assume $a \in A$. Since $A = A \cap B$, this means $a \in A \cap B$, hence that both $a \in A$ (which we already knew) and $a \in B$ (which we want to know!). So we're done. QED to theorem

Theme 3: What is Number?

This is harder to answer than you'd think. Let's start with: What are the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$? What is e.g. the 'number' 5?

Well, 5 is the number of x's on the whiteboard when we draw 5 x's on the whiteboard. But this is pretty unsatisfactory, for many reasons: e.g. it is dangerously circular. Also, you can't define some fundamental mathematical concept using the physical world. Math is prior to physics.

An italian mathematician Peano in the 1880s came up with a really nice answer.

A set $\mathcal{N}$ is called a set of natural numbers if there is a function $S : \mathcal{N} \to \mathcal{N}$ satisfying the following properties (called axioms):

Axiom A0. $1 \in \mathbb{N}$. (1 doesn't have any special properties, except in A1 below)

Axiom A1. If $n \in \mathcal{N}$, then $S(n) \neq 1$. In other words, there is no $n \in \mathcal{N}$ such that $S(n) = 1$.

Axiom A2. If $m, n \in \mathcal{N}$, and $m \neq n$, then $S(m) \neq S(n)$. In other words, if $m, n \in \mathcal{N}$ and $S(m) = S(n)$, then $m = n$.

Axiom A3. Let $K \subseteq \mathcal{N}$ have 2 properties:

(i) $1 \in K$

(ii) Whenever $n \in K$, then $S(n) \in K$.

Then $K = \mathcal{N}$.

For example, we can take $\mathcal{N}$ to be $\mathbb{N}$, and take $S(n) = n + 1$.

(By the way, when we write $f : A \to B$, we mean a function with inputs in A and outputs in B.)

S here is called the *successor* or *next* function.

Bertrand Russell (the guy who found Russell's paradox of Week 1) chose 1 to be $\{\}$, and defined $S(\{\}) = \{\{\}\}$, $S(\{\{\}\}) = \{\{\{\}\}\}$ etc, or equivalently $S(A)$ to be $\{A\}$. So natural numbers don't have to look like natural numbers.

Friday:

Define $2 = S(1)$, $3 = S(2)$, $4 = S(3)$, etc etc. So in Russell's version of natural numbers, $2 = \{\{\}\}$, $3 = \{\{\{\}\}\}$ etc.

First note that $2 \neq 1$. After all, $2 = S(1)$, and Axiom A1 says e.g. $S(1) \neq 1$.

More generally, $3 \neq 1$, $4 \neq 1$, etc, for the same reason. E.g. $4 = S(3)$ can't equal 1, by Axiom A1.

Why can't $2 = 4$? Well, $1 \neq 3$ by the previous paragraph, so $S(1) \neq S(3)$ by Axiom A2, i.e. $2 \neq 4$.

Define addition on $\mathcal{N}$ as follows: $m+1$ is defined to be $S(m)$, and $m+S(n)$ is defined to equal $S(m+n)$.

E.g. $2+3 = 2+S(2) = S(2+2) = S(2+S(1)) = S(S(2+1)) = S(S(S(2))) = S(S(3)) = S(4) = 5$.

E.g. $3+2 = 3+S(1) = S(3+1) = S(S(3)) = S(4) = 5$

(The reason Peano knew how to define addition was, he secretly thought about $\mathbb{N}$ and the formula $S(n) = n+1$.)

Theorem. Addition in $\mathcal{N}$ is associative: i.e. for any $\ell, m, n \in \mathcal{N}$, $(\ell+m)+n = \ell+(m+n)$

Proof. Let

$$K = \{n \in \mathcal{N} \mid (\ell+m)+n = \ell+(m+n) \text{ for all } \ell, m \in \mathcal{N}\}$$

In words, K is the set of all natural numbers n such that $(\ell+m)+n = \ell+(m+n)$ for all ℓ, m. We need to show that $K = \mathcal{N}$. This will follow from Axiom A3, if we can show that: (i) $1 \in K$, and (ii) whenever $n \in \mathcal{N}$, then $S(n)$ is also in $\mathcal{N}$.

First, let's show $1 \in K$. We need to show that, for every $\ell, m \in \mathcal{N}$, $(\ell+m)+1 = \ell+(m+1)$. But

$$(\ell+m)+1 = S(\ell+m) = \ell+S(m) = \ell+(m+1)$$

as desired.

Next, let's show (ii). So assume $n \in K$. We need to show that $S(n) \in K$, i.e. that $(\ell+m)+S(n) = \ell+(m+S(n))$ for all $\ell, m \in \mathcal{N}$. But

$$(\ell+m)+S(n) = S((\ell+m)+n) = S(\ell+(m+n)) = \ell+S(m+n) = \ell+(m+S(n))$$

as desired. QED to Theorem

From Week 2, the important things are to understand the logical words, and understand proof strategies.

You must get familiar with Definitions, Theorems, axioms, proofs.

The hard part is proofs. And unfortunately this will take most of you all term...

Week 3 summary

Monday: Peano revisited

Peano's axioms aren't important for us. They are meant to help teach you about the way modern mathematicians think. There is a philosophical side, a conceptual side, to mathematics. Remember the point of Peano's axioms: Because of things like Russell's paradox, we learned that we humans aren't so good with math. We are wrong about sets: things we thought we could do with them, are actually wrong. What else are we wrong about? Are we wrong about calculus? Are we wrong about real numbers? Are we even wrong about counting??

Peano came to the rescue about counting. His axioms show that, with a little work, we can recover all that we knew about counting and natural numbers.

We should say a few words about '='. Whenever we write '=', we need it to satisfy 3 things: $x = x$; $x = y \Rightarrow y = x$, and ($x = y$ AND $y = z$)$\Rightarrow x = z$. So if we use = in an axiom or definition, or if you see it in some question, you can always assume those 3 properties hold.

Mathematicians always ask questions like: "What does number mean?" "What does equality mean?" etc The answer is always obtained by asking instead: "How do numbers behave?" "How does equality behave?" Our answer for "How does equality behave?" is those 3 properties listed in the previous paragraph. Anything satisfying those 3 propoerties deserves the name 'equality'. The official word we use for such relations is "equivalence relation". So equivalence relation is a generalization of the idea of equality. It is really what equality means to a mathematician.

Theorem. Addition in Peano's axioms is commutative: $m+n = n+m$ for all $m, n \in \mathcal{N}$.

Proof. Let $K = \{n \in \mathcal{N} \mid m+n = n+m \text{ for all } m \in \mathcal{N}\}$. We need to show that $K = \mathcal{N}$. To show this, we need to show 2 things: (i) $1 \in K$; (ii) if $n \in K$ then $S(n) \in K$.

To prove $1 \in K$, let's also use Axiom A3. Write $K' = \{n \in \mathcal{N} \mid n+1 = 1+n\}$. First, we need to show $1 \in K'$, but this is trivial: obviously, $1+1 = 1+1$, so $1 \in K'$. Next, assume $n \in K'$, i.e. $n+1 = 1+n$. We need to show that $S(n)$ is also in K', i.e. $S(n)+1 \stackrel{?}{=} 1+S(n)$. Rewriting the left side, we get $S(n)+1 = (n+1)+1$ using the definition of addition. Rewriting the right side, we get $1+S(n) = S(1+n)$ by the definition of addition. But $S(1+n) = S(n+1)$ because $n \in K'$. And $S(n+1) = (n+1)+1$ by definition of addition. Thus we see that the left side of $S(n)+1 \stackrel{?}{=} 1+S(n)$ equals the right side, so indeed $S(n)+1 = 1+S(n)$. By Axiom A3, this concludes the proof that $K' = \mathcal{N}$. In other words, we now know that $1 \in K$.

Now we want to prove (ii). Assume that $n \in K$, i.e. that $m+n = n+m$ for all $m \in \mathcal{N}$. We want to prove that $S(n)$ is also in K, i.e. that $m+S(n) \stackrel{?}{=} S(n)+m$ for all $m \in \mathcal{N}$. The left side says $m+(n+1)$. The right side says $S(n)+m = (n+1)+m$, which equals $n+(1+m)$ by associativity, which equals $n+(m+1)$ since $1 \in K$, which equals $(n+m)+1$ by associativity, which equals $(m+n)+1$ because $n \in K$, which equals $m+(n+1)$ by associativity. Thus the left side of $m+S(n) \stackrel{?}{=} S(n)+m$ equals the right side, so $S(n) \in K$. By Axiom A3, $K = \mathcal{N}$ and we are done. *QED*

Wednesday

Generalized associativity and commutativity

So addition in Peano's axiomatic system is both commutative and associative. Whenever both of these are true, then *Generalized* associativity and commutativity holds. This means that if you have any number of numbers added together and paired off in a certain order, then you can rearrange them however you like, and pair them off in brackets in any new way, and you'll still get the same answer. E.g. $(((a+b)+c)+d)+e = d+((e+(a+c))+b)$.

Let's see how this works. Take the right side and move a to the extreme left, b beside it, etc:

$$d+((e+(a+c))+b) = d+(((a+c)+e)+b) = (((a+c)+e)+b)+d = ((a+c)+(e+b))+d$$

$$= ((a+c)+(b+e))+d = (((a+c)+b)+e)+d = ((a+(c+b))+e)+d = ((a+(b+c))+e)+d$$

$$= (a+(b+c))+(e+d) = (a+(b+c))+(d+e) = ((a+(b+c))+d)+e = (((a+b)+c)+d)+e$$

Knowing that generalized associativity and commutativity is true, is a big deal: it means we don't have to be so careful and slow at rearranging terms and brackets etc. Unless told otherwise, you can assume Generalized associativity and commutativity holds for anything that you know is both associative and commutative.

Induction proof technique

Let $P(n)$ be some statement depending on $n \in \mathbb{N}$. Suppose we want to show that $P(n)$ is true for all n. We can do this, similarly to how we proved associativity and commutativity above, by following Axiom A3.

(i) First show the base case: that $P(1)$ is true.

(ii) Assume the induction hypothesis: that $P(n)$ is true for some n. Then prove that $P(n+1)$ is also true.

If you can do both **(i)** and **(ii)**, then we know (using Axiom A3) that $P(n)$ is true for all $n \in \mathbb{N}$

Induction is a very powerful proof strategy, when you have to prove things for all natural numbers.

Here is an example.

Theorem. Show that $1+2+3+\cdots+n = \frac{n(n+1)}{2}$ for all $n \in \mathbb{N}$.

To prove this, let $P(n)$ be the statement: "$1+2+3+\cdots+n = \frac{n(n+1)}{2}$". Certainly true for $n=1$, so the base case **(i)** holds. Suppose $P(n)$ is true, i.e. $1+2+3+\cdots+n = \frac{n(n+1)}{2}$ (this is the inductive hypothesis). Is $P(n+1)$ true? Of course we have $1+2+3+\cdots+(n+1) = (1+2+3+\cdots+n)+(n+1)$. By the inductive hypothesis, this equals $\frac{n(n+1)}{2}+(n+1)$, which equals $\frac{(n+1)(n+2)}{2}$. Thus $P(n+1)$ is true. Hence by induction, the theorem is true.

Thursday: more induction

We did Quiz 3 today.

Another induction proof is from last year's Midterm 1 (solutions to all questions on that midterm are given in our eclass page). There are hundreds of examples of induction proofs (see the Bowman notes for examples), and you should practice them. Here's another:

Theorem. $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$

The formula for $n = 1$ is true.

If the formula is true for n, i.e. $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$, is it necessarily true for $n + 1$? Well,

$$1^2+2^2+\cdots+n^2+(n+1)^2 = \frac{n(n+1)(2n+1)}{6}+(n+1)^2 = \frac{n(n+1)(2n+1)+6(n+1)^2}{6}$$

$$= \frac{(n+1)\big(n(2n+1)+6(n+1)\big)}{6} = \frac{(n+1)\big(2n^2+7n+6\big)}{6} = \frac{(n+1)(n+2)(2n+3)}{6}$$

so the formula also works for $n + 1$, and induction then tells us it works for all $n \in \mathbb{N}$.

By the way, a convenient notation is $\sum_{k=1}^{n} k^2$ for $1^2+2^2+\cdots+n^2$ or $\sum_{k=1}^{n} k = 1+2+\cdots+n$.

You don't have to start induction at $n = 1$. For example, suppose you are asked to prove that $2^0 + 2^1 + \cdots + 2^n = 2^{n+1} - 1$. This is true for all $n = 0, 1, 2, \ldots$. So to prove that, the base case would be $n = 0$.

Incidentally, $2^0 + 2^1 + \cdots + 2^n = \sum_{k=0}^{n} 2^k$.

Friday: more numbers

Let's return to Peano's axioms of $\mathbb{N}$. We can define multiplication by: $m \cdot 1 = m$, and $m \cdot S(n) = (m \cdot n) + m$. Since any $n \in \mathbb{N}$ is either 1 or the successor of something (and not both), this defines recursively what $m \cdot n$ is for any $m, n \in \mathbb{N}$.

For example, $1 \cdot 3 = 1 \cdot S(2) = 1 \cdot 2 + 1 = 1 \cdot S(1) + 1 = (1 \cdot 1 + 1) + 1 = (1 + 1) + 1 = S(S(1)) = 3$.

We can prove that $\mathbb{N}$ with these two operations '+' and '$\cdot$' satisfies all the usual properties:

Cancellation Law for addition: $k + m = k + n \Rightarrow m = n$
Distributivity: $(k + m) \cdot n = k \cdot n + m \cdot n$ and $k \cdot (m + n) = k \cdot m + k \cdot n$
associativity of multiplication: $(k + m) + n = k + (m + n)$
commutativity of multiplication: $m + n = n + m$
cancellation law for multiplication: $k \cdot m = k \cdot n \Rightarrow m = n$
etc etc etc

We are not so interested in the natural numbers, so I won't bother to prove these.

We can get the integers $\mathbb{Z} = \mathbb{N} \cup -\mathbb{N} \cup \{0\}$ from $\mathbb{N}$, and define addition and multiplication for $\mathbb{Z}$ in terms of that for $\mathbb{N}$: there is a nice way (using an equivalence relation) to do this, but I don't have time.

$\mathbb{Z}$ is mathematically nicer than $\mathbb{N}$, because it has more structure. In particular, there is an additive identity (or neutral element for addition): namely 0. And there are additive inverses: namely $-n$. As always, these are defined by how they behave. Namely, $0+n=n$ and $n+(-n)=0$.

We can get the rational numbers $\mathbb{Q}$ from $\mathbb{Z}$ by bringing in fractions m/n, where $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. We can define how to add and multiply rationals using the addition and multiplication of integers.

The rationals are mathematically nicer than the integers, because they have more structure. In particular, they have a multiplicative identity (also called neutral element for multiplication), namely 1. ($\mathbb{N}$ and $\mathbb{Z}$ also have the multiplicative identity). But $\mathbb{Q}$ has multiplicative inverses: $1/r = r^{-1}$.

$\mathbb{Z}$ doesn't have multiplicative inverses: $2^{-1} = 0.5$ is not an integer.

By the way, we only discuss addition and multiplication, but not subtraction and division, because addition and multiplcation are simpler. E.g. they are associative and commutative (hence generalized associative and commutative). Subtraction and division are neither associative nor commutative. For example $(1-2)-3 = -1-3 = -4$ but $1-(2-3) = 1-(-1) = 2$. Subtraction and division are composite operations: $a-b$ is defined to be $a+(-b)$ and $a \div b$ is defined to be $a \cdot b^{-1}$.

We're primarily interested in real numbers $\mathbb{R}$ (see Section 1.C of Bowman's Notes), not natural numbers $\mathbb{N}$ nor integers $\mathbb{Z}$ nor rationals $\mathbb{Q}$. Our course is about calculus, more precisely all about functions $f(x)$, $f : \mathbb{R} \to \mathbb{R}$. So in order to prove theorems about these functions, we need to understand precisely what it means to be a real number, and precisely which properties we are allowed to assume. **This is hard to do.** For example, we could define real numbers by their decimal expansions, and then define how to add and multiply them by usual school math, but this is pretty messy. And the same number can have different decimal expansions: e.g. $0.99999\ldots = 1$ and $2.36999\ldots = 2.37$.

In modern math, we focus on how things act, not on what they look like. So to capture the real numbers, we write down all the properties of real numbers we need. These basic properties are called *axioms*.

We begin by listing the axioms of real number arithmetic.

Axioms of field

A *field* is a set $\mathbb{F}$ of things (which we'll call 'numbers'), together with two operations we'll call 'addition' and 'multiplication', which we'll write $x+y$ and $x \cdot y$ respectively. So $x+y \in \mathbb{F}$ and $x \cdot y \in \mathbb{F}$, whenever $x, y \in \mathbb{F}$. Addition satisfies these properties:

(AC) $x+y = y+x$ for all $x, y \in \mathbb{F}$. (*commutativity of addition*)

(AA) $(x+y)+z = x+(y+z)$ for all $x, y, z \in \mathbb{F}$. (*associativity of addition*)

(AN) This is a 'number' $0 \in \mathbb{F}$ such that $0+x = x$ for all $x \in \mathbb{F}$. (*additive identity*, or *neutral element for addition*)

(AI) For each $x \in \mathbb{F}$, there is some 'number' $-x \in \mathbb{F}$ such that $x+(-x) = 0$. (*additive inverse*)

'Multiplication' satisfies analogous properties:

(MC) $x \cdot y = y \cdot x$ for all $x, y \in \mathbb{F}$. (*commutativity of multiplication*)

(MA) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ for all $x, y, z \in \mathbb{F}$. (*associativity of multiplication*)

(MN) This is a 'number' $1 \in \mathbb{F}$ such that $1 \cdot x = x$ for all $x \in \mathbb{F}$. (*multiplicative identity*, or *neutral element for multiplication*)

(MI) For each $x \in \mathbb{F}$, except for $x = 0$, there is some 'number' $x^{-1} \in \mathbb{F}$ such that $x \cdot x^{-1} = 1$. (*multiplicative inverse*)

We all know that we're not supposed to divide by 0: this is why we exclude 0 from **(MN)**.

Finally, 'addition' and 'multiplication' satisfy 'distributivity':

(D) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ for all $x, y, z \in \mathbb{F}$.

The name 'field' was chosen for some obscure historical reason, which no longer makes much sense.

As always, do operations inside brackets first. But when there are no brackets, do 'multiplication' before 'addition'. So we can rewrite the Distributivity axiom **(D)** as $x \cdot (y + z) = x \cdot y + x \cdot z$, if we like.

Anything that satisfies these 9 axioms, is called a *field.* Fields are the number systems that you need for linear algebra (they're often called 'scalars' there). So you should see a lot of them in Math 127 and Math 227. We are primarily interested for now in real numbers. Real numbers are a field, but they are much more than merely a field. There are lots of different fields.

- Boring examples of fields: $\mathbb{R}$ and $\mathbb{Q}$ with the usual addition and multiplication.
- The integers $\mathbb{Z}$ (with usual addition and multiplication) are *not* a field. The problem is the **(MI)** axiom: all integers $x \neq 0$ have multiplicative inverses, but these are rational numbers and rarely are integers. **(MI)** requires that the inverses lie in $\mathbb{F}$, which is $\mathbb{Z}$ here. For example, the inverse of 2 is 0.5, which is not an integer. So $x = 2$ is a *counterexample* to **(MI)** for the choice $F = \mathbb{Z}$. All other 8 axioms hold, though.

In these axioms, you see how often we write 'for all'. A convenient short-hand is to write $\forall$ for 'for all'. E.g. $x + y = y + x \ \forall x, y \in \mathbb{F}$.

This week, the most important things are Generalized associativity and commutativity, the induction proof strategy, and the definition of a field.

Week 4 summary

Monday:

NOTE: To prove any of these axioms requires showing it for *all* $x, y, z \in \mathbb{F}$. To show something is a field, you have to prove all 9 axioms, for every possible value of $x, y, z \in \mathbb{F}$. So your proof will almost always have to use variables. To show something is *NOT* a field, it is enough to find one *counterexample* to one axiom. This means that it is enough to find one choice of $x, y, z \in \mathbb{F}$ which, when you substitute it into one of the axioms, you have a false statement.

- A very important example of a field is the complex numbers $\mathbb{C}$, which include $\sqrt{-1}$. We'll discuss them later in our course, and you'll also see them in Linear Algebra.
- A very important example of a field is $F = \{\mathbf{even}, \mathbf{odd}\}$, with the addition and multiplication we've discussed in class. Note that $0 = \mathbf{even}$ and $1 = \mathbf{odd}$. Note that $-\mathbf{odd} = \mathbf{odd}$ and $\mathbf{odd}^{-1} = \mathbf{odd}$. Linear algebra based on this field is used e.g. in cryptography.
- A very important example of a field in geometry is: the 'numbers' are all ratios $p(x)/q(x)$, where $p(x), q(x)$ are polynomials. You add and multiply these function fractions the same way you add and multiply usual fractions. Here, the additive identity is $0/1$ and the multiplicative identity is $1/1$.
- A fun example of something which is *almost* a field is **Tropical Numbers**. Here, $F = \mathbb{Z} \cup \{\infty\}$. Here, ∞ is 'infinity', so it is larger than any integer, and $\infty + n = \infty$ for any $n \in \mathbb{Z}$. 'Addition' is defined by $x \oplus y = \min(x, y)$, and 'multiplication' is defined by $x \odot y = x + y$. This satisfies all axioms except **(AI)**. For example, $0' = \infty$ and $1' = 0$. Multiplicative inverses are 'n^{-1}'$= -n$. To show Tropical numbers are not a field, all we need is one counterexample to one axiom. A counterexample to **(AI)** is: take $x = 0$, then $x \oplus y$ will always be an integer ≤ 0, so it will never equal $0' = \infty$, so '-0' here does not exist.

This doesn't mean Tropical numbers aren't useful – they have applications to optimization problems for example. The integers are also almost but not quite a field, and they are also useful. I like tropical numbers, because they emphasize the philosophy of modern math: it doesn't matter what something looks like, what matters is how it acts. The tropical version of 0 (i.e. the 'additive' identity $0'$) is infinity! What matters is that $\infty \oplus x = x$, i.e. $\min(\infty, x) = x$.

Wednesday:

A useful notation is '$\forall$'. It means *for all*. For example, "$\forall x \in \mathbb{R}$, $x^2 \geq 0$" means "for all $x \in \mathbb{R}$, $x^2 \geq 0$" or "for all real numbers x, the square is greater than or equal to 0", which is the same thing.

Likewise, the notation $\exists$ is short-hand for *there exists*. For example, "$\exists x \in \mathbb{R}$, such that $x > 5$" means "there exists an $x \in \mathbb{R}$ such that $x > 5$"

To show that e.g. commutativity of addition fails, we have to prove NOT($\forall x, y \in \mathbb{F}$, $x + y = y + x$). That amounts to finding one *counterexample*, i.e. one choice of $x, y \in \mathbb{F}$ such that $x+y \neq y+x$. There may be more than one counterexample, but all that matters is that there is at least one. Just find one counterexample – you don't get extra marks for finding 2 or 3. We can think of this more formally as: "NOT($\forall x, y \in \mathbb{F}$, $x + y = y + x$)" is logically identical to "$\exists x, y \in \mathbb{F}$ such that $x + y \neq y + x$." Or another way to say the same thing: "NOT($\forall x, y \in \mathbb{F}$, $x + y = y + x$)" IFF "$\exists x, y \in \mathbb{F}$ such that $x + y \neq y + x$."

Likewise, the negation of Axiom (MN) is: "NOT($\exists 1 \in \mathbb{F}$ such that for all $x \in \mathbb{F}$, $1 \cdot x = x$)" IFF "$\forall y \in \mathbb{F}$, $\exists x \in \mathbb{F}$ such that $y \cdot x \neq x$." After all, if for some y we can find an $x \in \mathbb{F}$ such that $y \cdot x \neq x$, then that y cannot be the multiplicative identity.

Consequences of the field axioms: Section 1.C of **Bowman's Notes**

We are still in the long process of defining precisely what are the real numbers. Last week we defined a *field*, and we've given some examples and nonexamples. This captures the arithmetic of the real numbers, but there are many other examples of fields than merely the reals. Our main interest in this course (unfortunately) are the real numbers.

We'll often write multiplication $x \cdot y$ as xy (without the dot).

In these notes we'll give our Lemmas etc official numbers. You can also use the numbers given in **Bowman's Notes**.

Lemma F.1. Let $\mathbb{F}$ be any field. Then:

(a) $x + 0 = x \ \forall x \in \mathbb{F}$.

(b) $-x + x = 0 \ \forall x \in \mathbb{F}$.

(c) $x \cdot 1 = x \ \forall x \in \mathbb{F}$.

(d) $x^{-1}x = 1 \ \forall x \in \mathbb{F}$, $x \neq 0$.

(e) $(x + y)z = xz + yz \ \forall x, y, z \in \mathbb{F}$.

The proofs of this Lemma are just commutativity (**axioms AC, MC**) applied to some other axiom (namely, **AN**, **AI**, **MN**, **MI**, **D**, respectively). We didn't explicitly write this down in class, because it's pretty trivial.

Note that if we want to prove something, we can't start with it and after a few steps derive something true, and conclude the original statement is true. For example, consider this 'proof':

$$1 \stackrel{?}{=} 2$$

$$0 \cdot 1 \stackrel{?}{=} 0 \cdot 2$$

$$0 \stackrel{?}{=} 0$$

Therefore $1 = 2$.

Thursday:

Lemma F.2. Let $\mathbb{F}$ be any field. Then:

(a) $\forall x \in \mathbb{F}$, $0x = 0$.

(b) $\forall x \in \mathbb{F}$, $-x = (-1) \cdot x$.

(c) $\forall x \in \mathbb{F}$, $-(-x) = x$.

(d) $\forall x \in \mathbb{F}$, $(-1)^2 = 1$.

Part (a) is proved by writing $0x = (0 + 0)x = 0x + 0x$ using Axioms **AN** and **D**, so $0 = 0x + (-0x)$ equals $(0x + 0x) + (-0x) = 0x + (0x + (-0x)) = 0x + 0 = 0x$, using **Axioms AI, AA,AN** and we're done.

The proof of part (b) is: the additive inverse $-x$ equals $(-1) \cdot x$, iff $x + (-1) \cdot x = 0$, since that is the definition of additive inverse. So we have to prove that $x + (-1) \cdot x \stackrel{?}{=} 0$. Rewrite LHS: $x + (-1)x = 1 \cdot x + (-1) \cdot x = (1 + (-1))x = 0x$ using Axioms **MN, D,AI**. Now Lemma F.2(a) tells us that $0x = 0$. Thus, $x + (-1) \cdot x = 0$ and we're done.

The proof of (c) is: the additive inverse $-(-x)$ of $-x$ is x, iff $(-x) + x = 0$. So we need to prove that $(-x) + x \stackrel{?}{=} 0$. But this is just Lemma F.1(b).

Combining $-(-1) = 1$ (Lemma F.2(b)) with Lemma F.2(c) gives $(-1) \cdot (-1) = 1$.

When you prove things, you can use not merely the axioms, but also the Lemmas and Theorems we have already proved. So the more you work, the more tools you can use.

We can define the integer powers x^n for any field, as follows. We define $x^1 = x$, $x^2 = x \cdot x$, $x^3 = x \cdot x \cdot x$, ... Another way to say this is: if $n \in \mathbb{N}$ and we know x^n, then x^{n+1} is defined to be $x \cdot x^n$. This recursively defines x^n for any $n \in \mathbb{N}$. (It is an example of what we call 'induction'.) For $x \neq 0$, we define $x^0 = 1$, and we already know what x^{-1} is; and if for any $n \in \mathbb{N}$ we know what x^{-n} is, then we can define x^{-n-1} as $x^{-1}x^{-n}$. Try now to prove the basic laws of exponents: where they are defined, $x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$ (these are proved in Bowman's Notes, and we'll prove them later in the course.)

By the way, the additive version of powers were discussed on Assignment 3. Of course $1 \in \mathbb{F}$, by Axiom **AN**. We can define $2 = 1 + 1 \in \mathbb{F}$, $3 = 2 + 1 \in \mathbb{F}$, etc. Thus we can think of any $n \in \mathbb{N}$ as being in $\mathbb{F}$. Likewise, we can define $0 \in \mathbb{F}$ by Axiom **AN** and $-1 \in \mathbb{F}$ by Axiom **AI**. Again we can define $-2 = -1 - 1$, $-3 + -2 - 1$, etc. In this way we can regard any $n \in \mathbb{Z}$ as being in any field $\mathbb{F}$.

- For example, in the **even, odd** field, $\mathbf{odd}^n = \mathbf{odd}$ for any $n \in \mathbb{Z}$. Moreover, $n = \mathbf{even}$ when n is even, and $n = \mathbf{odd}$ when n is odd. So although we can put the integers into any field, in some fields different integers can get sent to the same field elements.

- In most fields, raising elements to fractional powers can't be done without making your field bigger. For example, in the rational numbers $\mathbb{Q}$, $2^{1/2}$ doesn't exist (i.e, isn't a rational number) — to make sense of it you need a bigger field, like the real numbers $\mathbb{R}$. In $\mathbb{R}$, $(-1)^{1/2}$ doesn't exist to make sense of it you need a bigger field, like the complex numbers $\mathbb{C}$. Similarly, you can't always define fractional multiples: e.g. in the **even,odd** field, $\frac{1}{2}\mathbf{odd}$ does not exist.

Proof-by-contradiction: a powerful proof strategy

First, let's discuss the *Proof-by-contradiction* proof strategy. This is something you can all easily understand if you give it half a chance. And you better, because we're going to use it a lot! Remember basic logic, that we tested e.g. on Quiz2. Suppose that

we have logically derived something like this:

"IF **statement1** THEN **statement2**"

And suppose we know for some reason that **statement2** must be false. Then what can we conclude about **statement1**? From the definition of "IF...THEN", the only possibility is that **statement1** is also **false**.

This seems a little circular: to show some statement must be false, we need to show some other statement is false. But the idea is that some statements are *obviously* false, and others can be false but for complicated reasons. For example, "$1 = 2$" is obviously **false** (in $\mathbb{N}$ at least!), whereas "There are only finitely many primes" or "$\sqrt{2}$ is a rational number" are both **false** but not obviously so. So the idea is to start with **statement1**, something whose truth value is not obvious, and try to derive something which is *obviously* false. Then that tells us **statement1** must be **false**.

We know how to prove something of the form "IF **statement1** THEN **statement2**": assume **statement1** is **true** and derive **statement2**. Now, contradictions are obviously false. So we want to assume **statement1** is **true**, and then derive a contradiction. By the above reasoning, that is enough to prove **statement1** is **false**.

- For example, suppose we want to prove there are infinitely many prime numbers. The *proof-by-contradiction* proof strategy suggests that we should start with the hypothesis that there are NOT infinitely many primes, and then try to derive a contradiction. So we would start our proof with: "Suppose for contradiction that there are only finitely many primes. Call them $p_1, p_2, p_3, ..., p_n$." Then somehow we'd use that to derive a contraction.
- For example, suppose we want to prove that $\sqrt{2}$ is not in $\mathbb{Q}$. The *proof-by-contradiction* proof strategy suggests that we should start the proof like this: "Suppose for contradiction that $\sqrt{2}$ is in $\mathbb{Q}$. That means $\sqrt{2} = a/b$ where $a \in \mathbb{Z}$ and $b \in \mathbb{N}$." And from that hypothesis we'd try to derive a contradiction.

We've seen other examples of the *proof-by-contradiction* proof strategy earlier in the course, and we'll see others in the future. It's a great strategy, because it often gives you something concrete to hold in your hands. For example, you're given integers a, b such that $\sqrt{2} = a/b$. What properties must a, b have for that equation to hold?

Theorem. $\sqrt{2}$ is not rational.

Proof. Suppose for contradiction that $\sqrt{2} \in \mathbb{Q}$. Then $\sqrt{2} = p/q$ for some $p \in \mathbb{Z}$ and $q \in \mathbb{N}$. We may assume that at least one of p and q is odd (otherwise divide 2 from top and bottom). We have $q\sqrt{2} = p$. Square both sides: $2q^2 = p^2$. So p^2 is even. Then p must be even (since odd $\cdot$ odd is odd). Write $p = 2k$. Then $2q^2 = (2k)^2 = 4k^2$. Divide 2 from each side, to get $q^2 = 2k^2$. Hence q must also be even. But then both p and q are even, a contradiction. Thus $\sqrt{2}$ cannot be rational. QED

This was proved in Section 1.B of Bowman's notes. (Really, to prove that at least one of p, q is odd, we should use induction in the Peano axiom framework.) Try to do the proof that $\sqrt{3}$ is not rational. How about $\sqrt{4}$?

Friday:

We did Quiz 4.

A second proof-by-contradiction:

Theorem. There are infinitely many primes.

Recall that a prime p is a natural number $p > 1$ whose only divisors are 1 and itself. For example, $2, 3, 5, 7, 11$ are the first 5 primes. Prime numbers are important because they are the bricks building up any natural number: any natural number $n > 1$ can be written as a product of primes. For example, $30 = 2 \cdot 3 \cdot 5$ and $100 = 2^2 \cdot 5^2$.

To prove the theorem, suppose for contradiction that there are only finitely many primes. Call the primes $p_1, p_2, \ldots, p_n$. Define $N = p_1 p_2 p_3 \cdots p_n + 1$. Then $N \in \mathbb{N}$. It must have a decomposition into a product of primes. But which prime can divide it? None of them can: p_i divides evenly into $p_1 p_2 \cdots p_i \cdots p_n$, so it will have a remainder of 1 when you try to divide p_i into N. So N cannot be written as a product of any primes. This contradiction means that our initial premise, namely that there are only finitely many primes, must be wrong. Therefore there must be infinitely many primes. QED

We spent the rest of the class addressing questions for the midterm.

From Week 4, the important things are to be familiar with the field axioms (you don't need to memorize these axioms though), and proving simple theorems about fields, as well as the proof-by-contradiction proof strategy. Know about counterexamples and how to negate statements involving phrases like "for all" or "there exists".

Week 5 summary

Monday was the midterm!

Wednesday

Real number inequalities: the definition of an ordered field

Remember: our goal is to understand carefully what means calculus. We focus in this class (just like Math 144, Math 100,... do) on calculus for functions defined on $\mathbb{R}$. So we need to understand what precisely $\mathbb{R}$ is (otherwise, how can we prove theorems about it?). A big part of $\mathbb{R}$ is its arithmetic (addition, subtraction, multiplication, division). We captured this in the axioms of a field, which we've studied the previous week or so. But there are lots of fields. What are additional properties $\mathbb{R}$ has, that other fields don't?

A big part of real numbers are that we can talk about 'bigger' and 'smaller'. Here is one way to do this.

Definition. Let $\mathbb{F}$ be any field. $\mathbb{F}$ is called an ordered field if it has a subset $P \subseteq \mathbb{F}$ (which we think of as the 'positive numbers') satisfying these 3 axioms:

OA If $x, y \in P$, then $x + y \in P$.

OM If $x, y \in P$, then $xy \in P$.

OT For each $x \in \mathbb{F}$, one and only one of the following will occur:

$$x \in P\,, \qquad -x \in P\,, \qquad x = 0$$

If $x \in P$, we call x 'negative'. So any number is either positive, negative, or 0. And no number can be both positive and negative. And 0 is neither positive nor negative. The name **OT** means 'order trichotomy'.

These axioms and their consequences are discussed in section 1.C of Bowman's Notes.

Math 127 also introduced 'ordered field' this week. They used 4 axioms. But it really is the same definition in disguise: using their 4 axioms, they can prove our 3 axioms. And using our 3 axioms, we can prove their 4 axioms. So any field which is ordered in the 127 sense, is ordered in the 117 sense, and vice versa. Every theorem 127 can prove about ordered fields, is also true for us, and vice versa.

Let $\mathbb{F}$ be an ordered field. We write $x > y$ (and say x is "greater than y") when $x - y \in P$. We write $x < y$ when $y > x$, i.e. when $y - x \in P$. We write $x \geq y$ when either $x > y$ or $x = y$. We write $x \leq y$ when $y \geq x$.

Note that $x \in P$ iff $x > 0$. Also, x is negative iff $x < 0$. Note that **OT** now implies that for any $x, y \in \mathbb{F}$, exactly one of the following holds: either $x > y$, or $x < y$, or $x = y$.

Examples of ordered fields

- Boring examples of ordered fields: $\mathbb{R}$ and $\mathbb{Q}$ with the usual notion of positive: e.g. P for $\mathbb{R}$ is $P = \{x \in \mathbb{R} \,|\, x > 0\}$.

- The even,odd field $\mathbb{F} = \{\mathbf{even}, \mathbf{odd}\}$ is NOT an ordered field.

Proof that $\mathbb{F} = \{\mathbf{even}, \mathbf{odd}\}$ is not ordered: Suppose for contradiction that the even,odd field $\mathbb{F}$ is ordered. Since $\mathbf{odd} \neq 0$, Axiom **OT** would then tell us **odd** must be either positive or negative but not both. But **odd** is positive iff $-\mathbf{odd}$ is negative. But $-\mathbf{odd} = \mathbf{odd}$. So **odd** is positive iff **odd** is negative. And this contradicts Axiom **OT**. This contradiction means $\mathbb{F}$ is not ordered. QED

More generally, no finite field is ordered, except for the stupid finite field $\mathbb{F} = \{0\}$ which has exactly 1 element $0 = 1$: it is trivially ordered. (For any prime power p^n, like 8 or 25 or 81, there is a finite field which has exactly p^n elements.)

Thursday

Complex numbers

Complex numbers $z = a + ib \in \mathbb{C}$ are briefly discussed in the Appendix at the end of Bowman's notes. Here, a is called the *real* part; b is called the *imaginary* part. They can be any real number. i can't be simplified – it is $\sqrt{-1}$. But whenever you see i^2, you should replace it with -1.

Addition and subtraction are trivial: just add the real parts, and add the imaginary parts, separately. e.g.

$$(2 + 3i) + (7 - 4i) = (2 + 7) + i(3 - 4) = 9 - i$$

Multiplication is trivial: use distributivity, and remember $i^2 = -1$. E.g.

$$(2 + 3i) \cdot (7 - 4i) = 2 \cdot 7 + 2 \cdot (-4i) + (3i) \cdot 7 - 3i \cdot 4i = 14 - 8i + 21i - 12i^2 = 26 + 13i$$

To do division, rationalise the denominator:

$$\frac{1 + 2i}{3 + 4i} = \frac{1 + 2i}{3 + 4i}\frac{3 - 4i}{3 - 4i} = \frac{3 - 4i + 6i + 8i^2}{9 - 16i^2} = \frac{-5 + 2i}{25}$$

The complex numbers are a field: e.g. unit of addition is $0 = 0 + 0i$, unit of multiplication is $1 = 1 + 0i$, additive inverse of $x + iy$ is $-x - iy$, multiplicative inverse of $x + iy$ is $\frac{x-iy}{x^2+y^2}$. You can check for yourself that addition and multiplication are commutative, associative and satisfy distributivity.

Calculus can be done over the complex numbers just like it can be done over the real numbers, except that complex calculus is far simpler and prettier than real calculus. In algebra, complex numbers are far more important than real numbers: the special property that complex numbers have in algebra is that any polynomial $x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_0 = \sum_{k=0}^{n} a_k x^k$, where each coefficient a_k is allowed to be complex (or real), can be factored completely into linear factors:

$$\sum_{k=0}^{n} a_k x^k = (x - r_1)(x - r_2) \cdots (x - r_n)$$

where these numbers $r_k \in \mathbb{C}$ are called roots or zeros; on the other hand, most polynomials $\sum_{k=0}^{n} a_k x^k$ with real coefficients a_k, will have at least two complex roots, and so cannot be factorized completely over the real numbers. A famous example is $x^2 + 1$: try to factorize that over the reals. In fundamental physics (quantum theories like the Standard model of particle physics, or string theory), complex numbers are more important than real numbers.

It turns out that *the complex numbers $\mathbb{C}$ are not ordered.*

Proof. Suppose for contradiction that $\mathbb{C}$ is ordered. By Axiom **OT**, either $i = \sqrt{-1}$ is 'positive', or $-i$ is 'positive'. Then by **OM**, either $i^2 = -1$ is positive (if $i > 0$) or $(-i)^2 = -1$ is positive (if $-i > 0$) – so no matter what, we know $-1 > 0$ in $\mathbb{C}$. That's bad, because it means (again by Axiom **OM**) that $(-1)^2 = 1$ would also be positive. So putting all this together, we have derived that both -1 and 1 are positive. This contradicts Axiom **OT**. Therefore $\mathbb{C}$ cannot be ordered. QED

• The grooviest example of an ordered field which I know, is the field of rational functions, which we described in Monday of Week4. The 'numbers' in this field are fractions of the form $p(x)/q(x)$, where $p(x)$, $q(x)$ are polynomials with real number coefficients. Note that the real numbers are all in this field (why?).

I need to say which of these fractions $p(x)/q(x)$ are 'positive'. To do this, I first need to explain what 'leading coefficient' of a polynomial is. Any polynomial $p(x)$ can be written in the form $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where all coefficients $a_n, ..., a_0$ are real numbers. If the polynomial is the zero-polynomial, i.e. all coefficients a_k are 0, then we say the leading coefficient is 0. For any other polynomial, at least one of the coefficients will be different from 0. For these polynomials, I can throw away all terms of the form $0x_j$. In particular, I can assume $a_n \neq 0$. So n is the degree of the polynomial, and an is called the leading coefficient of the polynomial. One place the leading coefficient comes in is, for any polynomial $p(x)$, if we substitute into $p(x)$ really large positive values of x, then if these values are big enough, the values of $p(x)$ will be positive iff the leading coefficient is positive. Intuitively, we can say that the 'value' $p(+\infty)$ is positive iff the leading coefficient is positive.

For example, the leading coefficient of $0x^5 - 2x^3 + 7x^2 - \sqrt{2}x + 1$ is -2.

We say $p(x)/q(x)$ is in P, i.e. is 'positive', if the leading coefficient of $p(x)$, divided by the leading coefficient of $q(x)$, is a positive real number. For example, the 'number' $(0x^5 - 2x^3 + 7x^2 - \sqrt{2}x + 1)/(-x + \pi^2)$ has the ratio of leading coefficients $(-2)/(-1) = 2$, which is a positive real number, so $(0x^5 - 2x^3 + 7x^2 - \sqrt{2}x + 1)/(-x + \pi^2) \in P$. Equivalently, $p(x)/q(x)$ is in P iff $p(+\infty)/q(+\infty)$ is positive. You can check that this choice of P satisfies Axioms **OA**, **OM**, **OT**, and so the field of rational functions is ordered.

For example, both 1 and x are in our field. Which is bigger? Well, $x - 1 = (x - 1)/1$ has ratio of leading coefficients equal to $1/1$, and this is a positive real number. So this means $x - 1 \in P$, which means $x > 1$. Similarly, $x > 10000000$. In fact x is larger than any real number. So x is like ∞! However, x^2 is larger than x, as is e.g. $x^3 - 15x^2 + \pi x - \sqrt{2}$.

Also, $1/x$ is in our field. What is bigger: $1/x$ or 1? Well, $1/x - 1 = (1 - x)/x$ has ratio of leading coefficents equal to $(-1)/1$, so $1/x - 1$ is negative, so $1/x < 1$. Similarly, $1/x < 0.000000001$ and in fact $1/x$ is smaller than every positive real number. But $1/x > 0$. So $1/x$ is like an infinitesimal!

Friday

Ordered fields: consequences (see Section 1.C of Bowman's Notes)

Recall the definition of a field, the 9 axioms **AC** to **D** given in the Week3 notes. An ordered field obeys those axioms, together with Axioms **OA**, **OM**, **OT** of this week. The most important example of an ordered field is the real numbers $\mathbb{R}$.

Lemma OF.1. Let $\mathbb{F}$ be any ordered field. Then:

(a) $\forall x, y, z \in \mathbb{F}$, if $x > y$ and $y > z$, then $x > z$.

(b) $\forall x, y \in \mathbb{F}$, if $x \geq y$ and $y \geq x$, then $x = y$.

(c) $\forall x, y, z \in \mathbb{F}$, if $x > y$ then $x + z > y + z$.

(d) $\forall w, x, y, z \in \mathbb{F}$, if $w > x$ and $y > z$, then $w + y > x + z$.

To prove parts (a), (c), and (d), just convert from $x > y$ etc to $x - y \in P$ etc and add, using Axiom **OA**. Our proof of part (b) used a proof-by-contradiction: suppose $x \neq y$; then either $x - y > 0$ (i.e. $x > y$) or $y - x > 0$ (i.e. $x < y$). etc etc etc.

This Lemma and the next should be familiar to you: you've known for years that the reals satisfy them. The point here is that any ordered field satisfies them: to prove them, you only need the 3 axioms of order.

When we write $a > b > c$, we mean both $a > b$ and $b > c$. By Lemma **OF.1(a)**, this implies $a > c$. Define $a \geq b > c$ etc in a similar way.

Lemma OF.2. The following hold in any ordered field.

(a) If both $a > b$ and $c > 0$, then $ca > cb$.

(b) If both $a > b$ and $c < 0$, then $ca < cb$.

(c) If $a \neq 0$, then $a^2 > 0$.

(d) $1 > 0$, $2 > 0$, $3 > 0$ etc.

(e) If $a > 0$, then $a^{-1} > 0$

(f) If $a > b > 0$, then $a^{-1} < b^{-1}$.

The proof of (a) uses the definition of '$>$' and Axiom OM. The proof of (c) uses Axiom OT and $(-1)^2 = 1$. For (d): Use (c) and $1^2 = 1$ to get $1 > 0$. To get $2 > 0$, recall that 2 is defined (on Assignment 3) to be $1 + 1$. Since $1 > 0$, Axiom OA then implies $1 + 1 > 0$.

Part (e) can be done with a proof by contradiction: Assume $a > 0$. Suppose for contradiction that a^{-1} is not positive. Then by Axiom OT, either $a^{-1} = 0$ or $a < 0$. Certainly $a^{-1} \neq 0$ (since $a^{-1}a = 1$ but $0a = 0$). So we must have $a^{-1} < 0$. Then (a) (applied to $0 > a^{-1}$ and $a > 0$) implies $a0 > aa^{-1}$, i.e. $0 > 1$, which contradicts (d). This contradiction means a^{-1} is positive. This concludes the proof of (e).

To prove (f), assume $a > b > 0$. Then $ab > 0$ by Axiom OM. So $(ab)^{-1} > 0$ by (e). But $(ab)^{-1} = a^{-1}b^{-1}$ (why?). Apply (a) to $a > b$ and $a^{-1}b^{-1} > 0$ to get $(a^{-1}b^{-1})a > (a^{-1}b^{-1})b$. This simplifies to $b^{-1} > a^{-1}$. QED to Lemma OF.2.

Recall that in some fields, 2 can equal 0, or 3 can equal 0, ... and 0 can never be positive. Those field obviously cannot be ordered. In particular, any finite field cannot be ordered.

From Week 5, the important things are the definition of ordered field; examples and nonexamples of ordered field; and the two lemmas. You should be comfortable with how they're proved. But you don't have to memorize the exact statements of Lemmas and their proofs.

Week 6 summary

4th theme: inequalities, absolute value

Intervals: (Section 1.G in Bowman's notes)

The following interval notation is convenient: learn it! We uses different brackets to specify whether endpoints are included or excluded.

Definition. Let $\mathbb{F}$ be any ordered field. Choose any $a, b \in \mathbb{F}$, with $a < b$. Then

$$[a, b] = \{x \in \mathbb{F} \,|\, a \leq x \leq b\}$$

$$(a, b) = \{x \in \mathbb{F} \,|\, a < x < b\}$$

$$(a, b] = \{x \in \mathbb{F} \,|\, a < x \leq b\}$$

$$[a, b) = \{x \in \mathbb{F} \,|\, a \leq x < b\}$$

We also write $(-\infty, a) = \{x \in \mathbb{F} \,|\, x < a\}$, $(-\infty, a] = \{x \in \mathbb{F} \,|\, x \leq a\}$, $(a, \infty) = \{x \in \mathbb{F} \,|\, x > a\}$, $[a, \infty) = \{x \in \mathbb{F} \,|\, x \geq a\}$, and $(-\infty, \infty) = \mathbb{F}$.

Notice that $(a, b) \subseteq (a, b]$, $[a, b) \subseteq [a, b]$, etc. We call (a, b) an open interval, because it does not contain the endpoints a and b. We call $[a, b]$ a closed interval, because it contains the endpoints a and b.

Lemma OF.3. Let $\mathbb{F}$ be an ordered field. Let $a, b \in \mathbb{F}$, and $a < b$. Then the open interval (a, b) is not empty. In other words, there always are numbers strictly between a and b.

To prove this, we showed $(a + b)/2 \in (a, b)$, in other words, we showed $(a + b)/2 > a$ and $(a + b)/2 < b$. (First note that $2 \neq 0$ in an ordered field, so $(a + b)/2 = (a + b)2^{-1}$ exists. Also, $(a + b)2^{-1} - a = a2^{-1} + b2^{-1} - a2\,2^{-1} = b2^{-1} - a2^{-1} = (b - a)2^{-1} \in P$ since $b - a \in P$ (why?) and $2^{-1} \in P$, so $(a + b)/2 > a$. The proof of $(a + b)/2 < b$ is similar.

Absolute value (Section 1.D in Bowman's notes)

Let $\mathbb{F}$ be an ordered field. We're most interested in the case $\mathbb{F} = \mathbb{R}$.

For $x \in \mathbb{F}$, define the absolute value $|x|$ to be

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

- For example, in $\mathbb{R}$, $|2| = 2$, $|0| = 0$, $|-3| = 3$.

Absolute value will play a big role in our course. We can think of absolute value as meaning how far we are from 0.

$|x - c| < r$ means all x which are closer than r to c. In interval language, it means $(c - r, c + r)$. By comparison, $|x - c| = r$ means the two points $\{c - r, c + r\}$ – that is what a 1-dimensional circle looks like!

When you're trying to solve inequalities like $|x - 1| + 2|x + 2| < 9$, break it up into cases. Here there will be $2^2 = 4$ cases (2 cases for each absolute value):

case 1. $x - 1 \geq 0$ and $x + 2 \geq 0$

case 2. $x - 1 \geq 0$ and $x + 2 < 0$

case 3. $x - 1 < 0$ and $x + 2 \geq 0$

case 4. $x - 1 < 0$ and $x + 2 < 0$

In case 1, x lies in $[1, \infty) \cap [-2, \infty) = [1, \infty)$. For such x, our inequality $|x-1|+2|x+2| < 9$ simplifies to $(x-1)+2(x+2) < 9$, i.e. $3x+3 < 9$, i.e. $x < 2$. So the x in case 1 that survive our inequality are the x in $[1, \infty) \cap (-\infty, 2) = [1, 2)$.

In case 2, x lies in $[1, \infty) \cap (-\infty, -2) = \{\}$. So we can ignore case 2.

In case 3, x lies in $(-\infty, 1) \cap [-2, \infty) = [-2, 1)$. For such x, our inequality $|x-1| + 2|x+2| < 9$ simplifies to $-(x-1)+2(x+2) < 9$, i.e. $x+5 < 9$, i.e. $x < 4$. So the x in case 1 that survive our inequality are the x in $[-2, 1) \cap (-\infty, 4) = [-2, 1)$.

In case 4, x lies in $(-\infty, 1) \cap (-\infty, -2) = (-\infty, -2)$. For such x, our inequality $|x-1|+2|x+2| < 9$ simplifies to $-(x-1)-2(x+2) < 9$, i.e. $-3x-3 < 9$, i.e. $x > -4$ (dividing by -3 reverses the inequality). So the x in case 1 that survive our inequality are the x in $(-\infty, -2) \cap (-4, \infty) = (-4, -2)$.

Putting all this together, the set of all x which satisfy $|x-1|+2|x+2| < 9$ is the union of all x surviving case 1, case 2, case 3, case 4, i.e. $[1, 2) \cup \{\} \cup [-2, 1) \cup (-4, -2) = (-4, 2)$.

Using cases is a pretty common (but pretty tedious) strategy for proving something. It is very common with absolute value.

Here are the basic properties of absolute value:

Properties of Absolute Value. Let $\mathbb{F}$ be any ordered field, and let $x.y \in \mathbb{F}$.

(AV1) $|x| \geq 0$.

(AV2) $|x| = 0$ iff $x = 0$

(AV3) $|-x| = |x|$

(AV4) $|xy| = |x|\,|y|$

(AV5) $|x| - |y| \leq |x+y| \leq |x| + |y|$ and $|x| - |y| \leq |x-y| \leq |x| + |y|$

All of these are pretty easy to prove. The most important of these is (AV5), which is called the *triangle inequality*. To prove it, note that we always have $x \leq |x|$ and $-x \leq |x|$ and $y \leq |y|$ and $-y \leq |y|$ (why?), so adding these inequalities, we get $x - y \leq |x| + |y|$ and $-x + y \leq |x| + |y|$, which together imply $|x-y| \leq |x| + |y|$ (why?). Replacing y with $-y$ also gives $|x+y| \leq |x| + |y|$. Applying this to $x = (x-y)+y$ gives $|x| \leq |x-y| + |y|$; rearranging gives $|x| - |y| \leq |x-y|$.

Example. If $|a-x| \leq 0.01$ and $|b-y| \leq 0.02$, what the biggest that $|(2a-3b)-(2x-3y)|$ can be?

Solution: $|(2a-3b)-(2x-3y)| = |2(a-x)-3(b-y)| \leq |2(a-x)| + |3(b-y)|$ by triangle inequality. This equals $2|a-x|+3|b-y|$ by Lemma AV4, which is $\leq 2 \cdot 0.01 + 3 \cdot 0.02 = 0.08$

Bounds: (see Section 1.H of Bowman's notes)

Remember, our goal is to capture the real numbers axiomatically. We're almost there: the reals are an ordered field. We took 12 axioms to define an ordered field. The reals are an ordered field, but so are the rational numbers, and the field of rational functions

(=fractions of polynomials). We only need one more axiom, which we're slowly building up to. Keep in mind this question: What is a way to distinguish $\mathbb{Q}$ from $\mathbb{R}$?

Let $\mathbb{F}$ be an ordered field. Let $S \subseteq \mathbb{F}$ be any set of numbers in $\mathbb{F}$. By an *upper bound* for S, we mean any number $u \in \mathbb{F}$ which is $\geq$ any number in S: $u \geq x \ \forall x \in \mathbb{F}$. By a *lower bound* for S, we mean any number $l \in \mathbb{F}$ which is $\leq$ any number in S: $l \leq x \ \forall x \in \mathbb{F}$. A set may or may not have an upper bound, and may or may not have a lower bound. If S has at least one upper bound, then it has infinitely many and we say S is *bounded above*. If S has at least one lower bound, then it has infinitely many and we say S is *bounded below*. If S is both bounded above and bounded below, we say S is *bounded*.

- For example, take $\mathbb{F} = \mathbb{R}$ and consider $S = \{1, 2, 3, 4\}$. This is bounded: e.g. $u = 4$ and $u = 100$ are upper bounds, and $l = 1$ and $l = -234$ are lower bounds. $S = \mathbb{N}$ is unbounded: it is bounded below (e.g. $l = 1$) but unbounded above.

Note that a set S is bounded iff there exists $a, b \in \mathbb{F}$ such that $S \subseteq [a, b]$. In particular, a can be any lower bound, and b will be an upper bound.

From Week 6, the important things are interval notation, real comfort with absolute value, and comfort with upper and lower bounds.

Week 7 summary

Maximum, minimum, supremum, infimum: (see Section 1.I of Bowman's notes)

Again, let $\mathbb{F}$ be an ordered field. If a set S in $\mathbb{F}$ is bounded above, then it will have infinitely many upper bounds. It may or may not happen that it has a *least upper bound*, also called a *supremum*, denoted $\sup S$. So the supremum (if it exists) is an upper bound of S, and it is less than or equal to any other upper bound. Similarly, the *greatest lower bound*, also called the *infimum* and denoted $\inf S$, is a lower bound of S which is greater than or equal to any lower bound of S. If S is unbounded above, then we write $\sup S = +\infty$. If S is unbounded below, then we write $\inf S = -\infty$.

So $\inf S$ and $\sup S$ will either be $-\infty$ or $+\infty$, or will be elements of $\mathbb{F}$. We are *not* saying e.g. $\infty \in \mathbb{F}$. It is a convenient convention to say that e.g. $\sup S = \infty$ is S is unbounded above, as we'll see. We sometimes write ∞ instead of $+\infty$.

If $\sup S$ exists and is in S, then we call it the maximum of S, and denote it $\max S$. And if $\inf S$ exists and is in S, then we call it the minimum of S, and denote it $\min S$.

• For example, take $\mathbb{F} = \mathbb{Q}$. Then $S = (0, 1]$ is bounded; $\sup S = \max S = 1$; $\inf S = 0$ but $\min S$ does not exist. Another example: $\inf \mathbb{Z} = -\infty$ and $\sup \mathbb{Z} = +\infty$, and max and min of $\mathbb{Z}$ don't exist.

When sup exists, it will be unique (a set has at most one supremum). Similarly, a set has at most one infimum, at most one maximum, at most one minimum.

To prove $s = \sup S$, you must show **two** things: first, that s is an upper bound of S, and secondly, that if ℓ is any other upper bound of S, then $\ell \geq s$. Similarly to prove that something is $\inf S$, you must prove two things.

Any finite set has a maximum and a minimum.

Supremum is more important than maximum (because it exists much more often). Likewise, infimum is more important than minimum.

Completeness: The final axiom! (Bowman's notes 1.J)

Let $\mathbb{F}$ be an ordered field (so obeys the 9 field axioms, and the 3 order axioms). $\mathbb{F}$ is called *complete* if it obeys this axiom:

Axiom C: Every nonempty subset $S \subseteq \mathbb{F}$ has a supremum.

If S is not bounded above, then as mentioned above $\sup S = +\infty$. 'Nonempty' means not the empty set.

If $\mathbb{F}$ is complete, then also any nonempty set $S \subseteq \mathbb{F}$ has an infimum (which will be $-\infty$ if S is unbounded below).

• For example, $\mathbb{Q}$ is not complete. It turns out that the set $S = \{x \in \mathbb{Q} \,|\, x^2 < 2\}$ does not have a supremum (in $\mathbb{Q}$). To prove this, suppose for contradiction that it does have a supremum $s \in \mathbb{Q}$. If $s^2 < 2$, show that s couldn't be an upper bound of S. If $s^2 > 2$, show that we could find an upper bound of S which is smaller than s. Thus $s^2 = 2$. But we proved long ago that any s satisfying $s^2 = 2$ cannot be rational.

• For example, the ordered field of rational functions is not complete. If you take S to be the set of all negative real numbers, then S doesn't have a least upper bound. To see that, use a proof by contradiction: assume $\sup S$ exists and equals s. Well, 0 is an upper bound, so $s \neq \infty$. But $-1/x$ is also an upper bound, and $0 > -1/x$, so we know $s < 0$. We know that $1 < 2$, so $s > 2s$. So $2s$ cannot also be an upper bound. Therefore $2s < x$ for some negative number $x \in \mathbb{R}$. But then $s < x/2$, which contradicts the fact that s is an upper bound of S.

Recall that in any field we can define $2 = 1 + 1, 3 = 2 + 1, \ldots$, i.e. we can speak of $\mathbb{N} \subset \mathbb{F}$, i.e. the natural numbers $\mathbb{N}$ fit inside $\mathbb{F}$. In an ordered field, we know $1 < 2 < 3 < 4 < \cdots$. Here is a useful fact:

Archimedian property theorem. Let $\mathbb{F}$ be a complete ordered field. Then for any $x \in \mathbb{F}$, there is an $n \in \mathbb{N}$ such that $n > x$.

To prove this, use a proof by contradiction: suppose for contradiction that some $x \in \mathbb{F}$ is greater than any $n \in \mathbb{N}$. Let $y = \sup \mathbb{N}$. Then $y \in \mathbb{F}$ (because $\mathbb{N}$ is bounded). Then $y/2$ (which is less than y because $2^{-1} < 1$) is not an upper bound for $\mathbb{N}$, so $y/2 < n$ for some $n \in \mathbb{N}$, so $y < 2n$, a contradiction.

An ordered field $\mathbb{F}$ is called *Archimedian* if for every $x \in \mathbb{F}$, there is an $n \in \mathbb{N}$ such that $x < n$. The Archimedian Property is discussed in Bowman's notes section 1.J. Another way to think about Archimedian fields: if $\mathbb{F}$ is Archimedian, then for any $\epsilon > 0$, there exists an $n \in \mathbb{N}$ such that $0 < 1/n < \epsilon$. (Indeed, take $x = 1/\epsilon$)

For example, $\mathbb{Q}$ is Archimedian: if $x = p/q$, then $x < 1 + |p|$. This shows that an Archimedian ordered field may not be complete. On the other hand, the ordered field of rational functions is not Archimedian: the polynomial x

is bigger than any $n \in \mathbb{N}$. This gives another proof that the rational functions are not complete.

Theorem C.1. There is one and only one complete ordered field.

This unique ordered field is called the real numbers $\mathbb{R}$. We won't prove this theorem, though the Archimedian property is a start: it can be used to show that any complete ordered field lies inside $\mathbb{R}$. In Math 217 we return to these sorts of questions.

This theorem tells us that, just as any number system obeying Peano's axioms is really $\mathbb{N}$, any complete ordered field is really $\mathbb{R}$. So we finally know what $\mathbb{R}$ is: it satisfies the 9 axioms of a field, the 3 axioms of order, and the 1 axiom of completeness.

The complex numbers are not complete in this sense, because they are not ordered. There is a better, more general notion of completeness, which describes it as not having any missing points. The line of rational numbers misses points like $\sqrt{2}$. We can see this by considering *sequences* of rational numbers: 1, 1.4, 1.41, 1.414, 1.4142, ... This sequence gets closer and closer to one of the (infinitely many) holes in the line of rational numbers. The reals and the complex numbers do not have any holes in this sense. Completeness is really about sequences and limits, as we will soon see.

We're now ready to discuss calculus! A derivative is really a limit: $\frac{d}{dx}f(x) = \lim\limits_{h\to 0} \frac{f(x+h)-f(x)}{h}$. An integral is a limit of Riemann sums. So we're going to study now limits, which is a new theme.

Sequences: The definition (Section 2.1 in Bowman's notes)

A *sequence* of real numbers is a function $f : \mathbb{N} \to \mathbb{R}$. We usually write these with subscripts: e.g. $a_n = f(n)$. So a sequence looks like $a_1, a_2, a_3, a_4, \ldots$. In our course we're most interested in sequences with $a_n \in \mathbb{R}$, but you can also consider sequences of complex numbers, or sequences of matrices or sequences of functions etc.

Limit of a sequence: Section 2.A of **Bowman's Notes**

Consider the sequence $a_n = 1/n$. So the first few terms are:

$$1, 0.5, 0.333..., 0.25, 0.2, 0.1666...$$

etc. As n gets bigger, this sequence gets closer and closer to 0. For each n, $1/n$ is always positive, so it never equals 0, but intuitively we can think that ?as n reaches infinity, $1/n$ finally reaches 0'.

A number L is called the *limit* of a sequence a_n if, for any $\epsilon > 0$, there is a number N depending on ϵ, such that $|L - a_n| < \epsilon$ for all $n > N$.

We know what $|L-x| < \epsilon$ means: it means the interval $L-\epsilon < x < L+\epsilon$. We think of ϵ being a super-small number. It is the greek letter 'e', and we think of it as 'error'. There might be noise, maybe lots of noise, where the sequence does all kinds of silly things. But eventually (and this is where N comes in), every term in the sequence gets super-close to L. ϵ quantifies what 'super-close' means; N tells which n's does the sequence gets superclose to L.

This is a really important definition. It tells you how we think of infinity in math: we approximate it with finite $n \in \mathbb{N}$ which get bigger and bigger. We secretly think of the limit (when it exists) as a_∞.

We can also talk about sequences in $\mathbb{R}^2$. Then limit has the same definition, but now $L \in \mathbb{R}^2$ and $|L - x| < \epsilon$ means all points in a disc of radius ϵ with centre L. We can also talk about sequences in $\mathbb{R}^3$, but then $L \in \mathbb{R}^3$ and $|L - x| < \epsilon$ means all points in a ball of radius ϵ with centre L. And so on.

We say $L = \lim\limits_{n\to\infty} a_n$, or that $a_n \to L$ or that a_n converges to L. We say a sequence *diverges* if it doesn't converge to some limit.

For example, the sequence $a_n = 1/n$ converges to 0. To see this, choose any $\epsilon > 0$. Take $N = 1/\epsilon$. Then for any $n > N$, $|0 - a_n| = 1/n < 1/N = \epsilon$. This is what it means to have 0 as its limit.

Another example of limit: take $a_n = 1/\sqrt{n}$. This also converges to $L = 0$. To see this, choose any $\epsilon > 0$. Take $N = 1/\epsilon^2$. Then for any $n > N$, $|0 - a_n| = 1/\sqrt{n} < 1/\sqrt{N} = \epsilon$, and we are done.

You may wonder how we find N as a function of ϵ. How would we guess that we should take $N = 1/\epsilon^2$ here? To do this, work backwards. On a scrap sheet of paper, do the calculation $|0 - a_n| = 1/\sqrt{n} < 1/\sqrt{N}$. We want this to be $< \epsilon$, so we want $1/\sqrt{N} = \epsilon$, so solve that for N.

On the other hand, $b_n = (-1)^n$ doesn't converge anywhere. To see it doesn't converge to 1, take $\epsilon = .5$. Then any even n obeys $|1-b_n| = |1-1| = 0 < \epsilon$, which is good, but any odd n obeys $|1 - b_n| = |1 - (-1)| = 2 > \epsilon$. No matter how big you choose N, I can take n to be any odd number $> N$, and the inequality $|1 - b_n| < \epsilon$ will be violated.

From Week 4, the important things are the notion of maximum and minimum, and especially sup and inf, and the notion of completeness and that the reals are complete. You'll need to know about sequences and limits.

Week 8 summary

We will often need to approximate fractions: i.e. simplify both the top (called numerator) and bottom (called denominator), and know whether our simplifications make the fraction bigger or smaller.

Golden Rule of Fraction Inequalities: Suppose you have a fraction $\frac{top}{bottom}$, where both top and bottom are positive. If you make *top* bigger, and *bottom* smaller, then the fraction gets bigger: if $top \leq top'$ and $bottom \geq bottom'$, then $\frac{top}{bottom} \leq \frac{top'}{bottom'}$.

Example. Prove from the definition that $a_n = \frac{2n^2+n+1}{n^2-1} \to 2$. We compute

$$|a_n - 2| = \left|\frac{2n^2+n+1}{n^2-1} - 2\frac{n^2-1}{n^2-1}\right| = \left|\frac{2n^2+n+1-2n^2+2}{n^2-1}\right| = \left|\frac{n+3}{n^2-1}\right| \leq \left|\frac{4n}{n^2/2}\right|$$

using the Golden rule twice: $n+3 \leq 4n$ and $n^2-1 \geq n^2/2$. So the error $|a_n - 2|$ is always less then $4n/(n^2/2) = 8/n$. We want the error to be $< \epsilon$. We will get that if n is big enough that $8/n < \epsilon$.

That's the secret work. Now let's write this in a fancy way: Choose any $\epsilon > 0$. Choose $N = 8/\epsilon$. Then for any $n > N$,

$$|a_n - 2| = \left|\frac{2n^2+n+1}{n^2-1} - 2\frac{n^2-1}{n^2-1}\right| = \left|\frac{2n^2+n+1-2n^2+2}{n^2-1}\right| = \left|\frac{n+3}{n^2-1}\right| \leq \left|\frac{4n}{n^2/2}\right| = \frac{8}{n} < \epsilon$$

as desired.

For our next result, we need to know about subsequences. A subsequence of a sequence just means you cross out some terms in the sequence: e.g. the boldface terms in

$$1, \mathbf{2}, \mathbf{3}, 4, \mathbf{5}, 6, \mathbf{7}, 8, 9, 10, \mathbf{11}, 12, ...$$

are 2,3,5,7,11,..., which is a subsequence (the primes) of the sequence 1,2,3,4,5,6,... In fancy talk, a subsequence of a sequence $a_1, a_2, a_3, ...$ is $b_1 = a_{n_1}, b_2 = a_{n_2}, ...$, where the subscripts n_k is a choice of numbers $1 \leq n_1 < n_2 < n_3 < n_4 < \cdots$. E.g. in the prime number example given above, $n_1 = 2, n_2 = 3, n_3 = 5, n_4 = 7, ..., n_k =$ the kth prime number.

Subsequence Theorem. Let a_n be a sequence. Then $\lim_{n\to\infty} a_n = L$ iff, for every subsequence b_k, $\lim_{k\to\infty} b_k = L$.

Proof. This is an 'iff', so there are two directions.
$\Longrightarrow$: Assume that a_n converges to L. Let $b_k = a_{n_k}$ be any subsequence. We want to show that b_k also converges to L.

Choose any $\epsilon > 0$. Then there is some N such that $|a_n - L| < \epsilon$ whenever $n > N$. Now choose any $k > N$. Note that $n_1 \geq 1$, $n_2 \geq 2$, $n_3 \geq 3$, ... $n_k \geq k$ for all k. (You can prove this by induction.) Since $k > N$, then $n_k \geq k > N$, so $|b_k - L| = |a_{n_k} - L| < \epsilon$, and we are done: $b_k \to L$.

$\Longleftarrow$: This is trivial. Suppose we know that any subsequence of a_n converges to L. We want to prove that a_n itself converges to L. But $b_k = a_k$ (i.e. $n_k = k$) is a subsequence of a_n (a pretty silly subsequence, but that's OK). We were told that any subsequence converges to L, so this b_k must also converge to L. QED

The main use of the Subsequence Theorem is to show that some sequences won't converge. The idea is given in this Corollary ('corollary' means 'consequence').

Nonconvergence Corollary. If a sequence a_n has subsequences b_k and c_ℓ, which converge to different limits, then a_n does not converge to anything.

This is just a rephrasing of the Subsequence Theorem.

Example. The sequence $a_n = (-1)^n \frac{n+1}{n}$ does not converge. Take one subsequence to be $b_k = \frac{2k+1}{2k}$ (this uses the choice $n_k = 2k$): it clearly has limit 1. Take another subsequence to be $c_\ell = -\frac{2\ell}{2\ell-1}$ (this uses the choice $n_\ell = 2\ell - 1$): it clearly has limit -1. These two limits are different, so a_n can't converge to anything, by the Nonconvergence Corollary.

Example. Consider the sequence

$$1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1..... \tag{1}$$

(so keep increasing by 1 the number of 0's between the 1's. Take one subsequence to be the 1's:

$$\mathbf{1}, 0, \mathbf{1}, 0, 0, \mathbf{1}, 0, 0, 0, \mathbf{1}, 0, 0, 0, 0, \mathbf{1}, 0, 0, 0, 0, 0, \mathbf{1}.....$$

so this subsequence looks like 1,1,1,... Certainly it converges to 1. Next, take the subsequence

$$1, \mathbf{0}, 1, \mathbf{0}, \mathbf{0}, 1, \mathbf{0}, \mathbf{0}, \mathbf{0}, 1, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, 1, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, 1.....$$

So this sequence is 0,0,0,0,... which clearly converges to 0. Because these two subsequences don't converge to the same thing, the original sequence in (1) does not converge.

A very useful result to prove convergence is:

Squeeze Theorem. Let a_n, b_n, c_n be sequences, and suppose $a_n \leq b_n \leq c_n$. Suppose both a_n and c_n converge to L. Then b_n also converges to L.

To prove this, choose any $\epsilon > 0$. Because $a_n \to L$, there is a N' such that $|a_n - L| < \epsilon$, for all $n > N'$. Because $c_n \to L$, there is an N'' such that $|c_n - L| < \epsilon$ for all $n > N''$.

Let $N = \max\{N', N''\}$. Choose any $n > N$. Then $n > N'$,so $L - a_n < \epsilon$. And $n > N''$, so $c_n - L < \epsilon$. We want to show $|b_n - L| < \epsilon$. There are two possibilities:

case 1: $L \geq b_n$. Then $|b_n - L| = L - b_n \leq L - a_n$ since $a_n \leq b_n$. But $L - a_n < \epsilon$, by the previous paragraph. So putting this together, we get $|b_n - L| < \epsilon$ in this case.

case 2: $L < b_n$. Then $|b_n - L| = b_n - L \leq c_n - L$ since $c_n \geq b_n$. But $c_n - L < \epsilon$, by a previous paragraph. So putting this together, we get $|b_n - L| < \epsilon$ in this case, also.

So in all cases, we get $|b_n - L| < \epsilon$. This proves that $\lim_{n\to\infty} b_n = L$. QED

It is a useful way to prove sequences converge:

Example: Let d_n be the nth digit of π, so $d_1 = 3$, $d_2 = 1$, $d_3 = 4$ etc. Consider the sequence $d_n/\sqrt{n}$. The complicated part of this sequence is the d_n's, which bounce around pretty randomly between 0 and 9. So the biggest they get is 9, and the smallest they get is 0. So the sequence $d_n/\sqrt{n}$ is never more than $9/\sqrt{n}$, and never less than $0/\sqrt{n} = 0$: we have $0 \leq d_n/\sqrt{n} \leq 9/\sqrt{n}$. Certainly the sequence 0,0,0,... has a limit: its limit is 0. And the sequence $9/\sqrt{n}$ has a limit, namely 0. So by the Squeeze Thm, the limit of $d_n/\sqrt{n}$ is 0.

Example: Consider the sequence $a_n = \frac{1}{n}\cos(n)$. The complicated part of this sequence is $\cos(n)$, which oscillates randomly between -1 and 1:

$$-1 \leq \cos(\theta) \leq 1\,,\ -1 \leq \sin(\theta) \leq 1 \qquad \forall \theta \in \mathbb{R}$$

So we get

$$\frac{-1}{n} \leq \frac{1}{n}\cos(n) \leq \frac{1}{n}$$

But both $-1/n$ and $1/n$ converge to 0. So the Squeeze Thm says $\frac{1}{n}\cos(n)$ also converges to 0.

By the way, here's an interesting sequence: $a_n = \sqrt{n+1} - \sqrt{n}$. What does it converge to? Think of it this way:

$$\sqrt{n+1} - \sqrt{n} = (\sqrt{n+1} - \sqrt{n})\frac{\sqrt{n+1} + \sqrt{n}}{\sqrt{n+1} + \sqrt{n}} = \frac{n+1-n}{\sqrt{n+1} + \sqrt{n}}$$

which tends to 0.

Here's a little technical result we'll need occassionally. We call a sequence a_n *bounded*, if the set $\{a_1, a_2, a_3, ...\}$ is bounded. In other words, there is some M, M' such that $M' \leq a_n \leq M$ for all n. In other words, a_n only gets so big and only so negative.

Bounded Lemma. Suppose a_n is a convergent sequence. Then it is bounded.

Proof. Call L the limit of a_n. Choose $\epsilon = 1$. Then there is some N such that $|a_n - L| < 1$ for all $n > N$. This means $a_n \in (L-1, L+1)$, for all $n > N$. Let M be the maximum of the numbers $\{a_1, a_2, ..., a_N, L-1\}$, and let M' be the minimum of the numbers $\{a_1, a_2, ..., a_N, L-1\}$. Both are finite sets, so have a maximum and a minimum. Then obviously $M' \leq a_n \leq M$ for all n: for $n \leq N$ this is clear from the definition of M, M', and for all $n > N$ this is also clear from the definition of N. So a_n is bounded. QED

This is used in the proof of the following Theorem.

Pretty useful Theorem. Suppose $\lim\limits_{n\to\infty} a_n = L$ and $\lim\limits_{n\to\infty} b_n = L'$. Then:

(a) For any constant $\alpha \in \mathbb{R}$, the sequence αa_n converges to αL.

(b) $a_n + b_n$ converges to $L + L'$, and $a_n - b_n$ converges to $L - L'$. More generally, for any constants $\alpha, \beta \in \mathbb{R}$, $\alpha a_n + \beta b_n = \alpha L + \beta L'$.

(c) $a_n b_n$ converges to LL'. If no $b_n = 0$, and if $L' \neq 0$, then a_n/b_n converges to L/L'.

Proof. Let's prove that $a_n + b_n$ converges to $L + L'$. Choose any $\epsilon > 0$. Then there is an N' such that for all $n > N'$, $|a_n - L| < \epsilon/2$. And there is an N'' such that for any $n > N''$, $|b_n - L'| < \epsilon/2$. Now let $N = \max\{N', N''\}$. Then for any $n > N''$, $|(an + bn) - (L + L')| \leq |a_n - L| + |b_n - L'|$ by the triangle inequality. We know $n > N'$, so $|a_n - L| < \epsilon/2$. And we know $n > N''$, so $|b_n - L'| < \epsilon/2$. Putting these together, we get $|(a_n + b_n) - (L + L')| < \epsilon/2 + \epsilon/2 = \epsilon$. This means $a_n + b_n$ converges to $L + L'$.

The proof of the others are in Bowman's Notes, section 2.A. QED

When using the Pretty Useful Theorem, the following facts are useful:

Claim. For any $r > 0$, $\lim_{n\to\infty} 1/n^r$ equals 0.

For example, we know $1/n \to 0$, so using the Pretty Useful Theorem 21 times, we see that $1/n^{22} \to 0$. Likewise, thanks to an old assignment question, we know $1/\sqrt{n} \to 0$, etc. And $\sqrt{n} = n^{1/2}$. On an Assignment 7 question, we'll prove this Claim. You may assume it on Assignment 6 and your Midterm 2, if you want.

Example.

$$\frac{3n^4 + 5n^3 - 4n^2 - 9}{-n^4 + \sqrt{n}} = \frac{3 + 5/n - 4/n^2 - 9/n^4}{-1 + 1/n^{3.5}} \to \frac{3 + 5\cdot 0 - 4\cdot 0 - 9\cdot 0}{-1 + 1\cdot 0} = -3$$

Infinite limits:

Just as it is convenient to say that sup and inf are ∞ or $-\infty$, it is convenient to say that $a_n \to \infty$ and $a_n \to -\infty$. Plugging say $L = \infty$ into the limit definition is nonsense: you get $|a_n - \infty| < \epsilon$. So a different way to say we "approach infinity" is needed. Here it is:

Definition. We say that a_n tends to ∞, or has limit ∞, or $\lim_{n\to\infty} a_n = \infty$, or $a_n \to \infty$, if for any $M > 0$, there is an N such that for all $n > N$, we have $a_n > M$.
We say that a_n tends to $-\infty$, or has limit $-\infty$, or $\lim_{n\to\infty} a_n = -\infty$, or $a_n \to -\infty$, if for any $M > 0$, there is an N such that for all $n > N$, we have $a_n < -M$.

We don't say that a_n converges to ∞ or $-\infty$ though. We reserve the word 'convergence' for finite limits.

So there are two changes. One is, we use $M > 0$ instead of $\epsilon > 0$, but this is pretty trivial. The serious change is that "$|a_n - L| < \epsilon$" has been replaced with "$a_n > M$. We think of ϵ as a very small positive error; as we move it closer and closer to 0, i.e. as a_n gets closer and closer to L, N and hence n gets bigger and bigger. On the other hand, we think of M as large; as M gets larger and larger, so a_n gets closer and closer to ∞, N and hence n gets closer and closer to ∞.

Examples. The limit of $a_n = n^2$ is ∞. To see this, choose any $M > 0$. Let $N = \sqrt{M}$. Then for any $n > N$, $a_n = n^2 > N^2 = M$, and we're done.

The limit of $a_n = -\sqrt{n}$ is $-\infty$. To see this, choose any $M > 0$ and let $N = M^2$. Then for any $n > M$, $a_n = -\sqrt{n} < -\sqrt{N} = -M$ and we're done.

The sequence $a_n = (-1)^n n$ does not have a limit. Certainly it can't have a finite limit, because $|a_n|$ grows to ∞. It can't tend to ∞ because the odd a_n's are negative, and it can't tend to $-\infty$ since the even a_n's are positive.

Most of our theorems have an analogue for sequences with infinite limits, as we'll see on Monday. For example:

Theorem. A sequence a_n converges to $L = \infty$ iff all subsequences converge to ∞. (same for $L = -\infty$).

Theorem. Suppose $a_n \to L$ and $b_n \to L'$, where $L, L' \in \mathbb{R} \cup \{\pm\infty\}$. Then:

(a) As long as we don't have $L = \infty = -L'$ or $L = -\infty = -L'$, then $a_n + b_n \to L + L'$.

(b) As long as we don't have $L = \pm\infty$ and $L' = 0$ or $L = 0$ and $L' = \pm\infty$, then $a_n b_n \to LL'$.

The proofs are as before. Similar results hold for $\alpha a_n + \beta b_n$ and a_n/b_n – write these statements out yourselves.

Example.

$$\frac{3n^4 - \sqrt{n}}{4 - n^3} = n\frac{3n^3 - 1/\sqrt{n}}{-n^3 - 4} \to -\infty$$

From Week 8, the important things are to be familiar with sequences and their limits. Know to use the Nonconvergence Corollary to prove a sequence doesn't converge. The squeeze theorem is a nice way to show things converge. The Pretty Useful Theorem is very useful. Be familiar with infinite limits too. The proofs of these theorems is not important.

Week 9 summary

Monday: Infinite limits

We defined infinite limits: what it means for $\lim_{n\to\infty} a_n = \infty$ or $\lim_{n\to\infty} a_n = \infty$. This was given at the end of Week8 so there is no need to repeat it here.

Wednesday: Monotone sequences (Sections 2.B,2.C of Bowman's Notes)

We know from last week that a convergent sequence is bounded. Is the converse true: Is a bounded sequence convergent? No, of course not, $(-1)^n$ is bounded but not convergent. But the following is true.

A *monotone increasing sequence* means $a_1 \le a_2 \le a_3 \le \cdots$. A *monotone decreasing sequence* means $a_1 \ge a_2 \ge a_3 \ge \cdots$. A *monotone* or *monotonic sequence* is one which is either monotone increasing or monotone decreasing. Recall that a_n being bounded means that there is some M such that $|a_n| < M$.

Example: For example, $a_n = n$ is monotone increasing. $a_n = 1/n$ is monotone decreasing. $a_n = 1$ is both monotone increasing and monotone decreasing.

Monotone Convergence Theorem. (a) If a_n is bounded and monotone, then it converges.
(b) If a_n is monotone, then $\lim_{n\to\infty} a_n$ exists.

The limit in (a) will be finite; the limit in (b) may be $\pm\infty$. When the sequence is monotone increasing, the limit will be $\sup a_n$. When the limit is monotone decreasing, the limit will be $\inf a_n$.

This is the first place we needed completeness. Note that an increasing sequence a_n of rational numbers often won't converge to a rational number. For example 1, 1.4, 1.41, 1.414, 1.4142, ... is a sequence of rational numbers converging to $\sqrt{2}$. In this sense, the "number line" of rational numbers has small holes in it. Those holes are at the irrational numbers.

Material for Midterm 2 stops here

No questions on the midterm will require you to know the following material (lim sup, lim inf, nor the Power Corollary). But you are free to use these on the midterm if they help.

Let a_n be a bounded sequence. Consider the sequence s_n defined by $s_n = \sup\{a_n, a_{n+1}, a_{n+2}, ...\}$ for each n. Then s_n is a bounded sequence (bounded by same thing as a_n). Moreover, s_n is monotone increasing. Hence, s_n is convergent (by Monotone convergence Theorem). The limit is called lim sup a_n. If a_n is not bounded above, then lim sup a_n will be infinity.

Similarly, define the sequence i_n by $i_n = \inf\{a_n, a_{n+1}, a_{n+2}, ...\}$ for each n. Then i_n is a bounded sequence (bounded by same thing as a_n). Moreover, i_n is monotone idecreasing. Hence, i_n is convergent (by Monotone convergence Theorem). Its limit is called lim sup a_n. If a_n is not bounded below, then lim inf a_n will be $-\infty$.

What's special about lim sup and lim inf is that they always exist; on the other hand, most sequences don't have limits. Note that $\lim\limits_{n\to\infty} a_n = L$ iff lim sup $a_n = L =$ lim inf a_n.

For example, when $a_n = (-1)^n(1+1/n)$, lim sup $a_n = +1$ and lim inf $a_n = -1$.

Power Corollary. Let c be any constant.

(a) If $c \leq -1$, then c^n diverges (has sup sup $\neq$ lim inf)
(b) If $|c| < 1$, then c^n converges to 0.
(c) If $c = 1$ then c^n converges to 1.
(d) If $c > 1$, then $c^n \to \infty$.

For $c > 1$, c^n is monotonic increasing; for $0 < c < 1$, c^n is monotonic decreasing. When $c < 0$, c^n alternates in sign, but c^{2n} is either monotonic increasing or decreasing depending on whether $|c|$ is greater or less than 1, respectively.

Thursday: Review class for midterm 2

Friday: Midterm 2!

From Week 9, the important things are to be comfortable with infinite limits, know the Monotone convergence theorem, know whether c^n converges, know the ratio test. Only infinite limits and Monotone convergence are required for the midterm.

Week 10 summary

Consequences of Monotone convergence:

We've already seen last week two consequences of the Monotone Convergence Theorem. One was $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$. You work out the basic properties of lim sup and lim inf on Assn.7. It is not so important for our course: the main purpose for us is it yields a fast proof of Bolzano-Weierstrass Theorem, later this week.

Another consequence was the Power Corollary, which describes how sequences of powers r^n behave.

One consequence of the Power Corollary is the *geometric series* $1 + r + r^2 + r^3 + \cdots$. Recall that $\sum_{k=0}^{n} r^k = 1 + r + r^2 + \cdots + r^n$:

Claim. For $|r| < 1$, $\lim_{n\to\infty} \sum_{k=0}^{n} r^k = \frac{1}{1-r}$.

To see this, first observe that, for any $r \neq 1$,

$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r} \tag{1}$$

To see that, either use induction on n, or multiply both sides by $1 - r$ and simplify. The claim then follows from eq.(1) by taking the limit as $n \to \infty$, using the fact that $r^n \to 0$ for $|r| < 1$.

It is common to write $\sum_{k=0}^{\infty} r^k = 1 + r + r^2 + \cdots$ for $\lim_{n\to\infty} \sum_{k=0}^{n} r^k$.

Another consequence of the Monotone Convergence Theorem is decimal expansions:

Example: Decimal expansion. What does $\pi = 3.1415926...$ mean? It means

$$3 \cdot 10^0 + 1 \cdot 10^{-1} + 4 \cdot 10^{-2} + 1 \cdot 10^{-3} + 5 \cdot 10^{-4} + 9 \cdot 10^{-5} + 2 \cdot 10^{-6} + 6 \cdot 10^{-7} + \cdots$$

More generally, choose any 'digits' $d_0, d_1, d_2, \ldots \in \{0, 1, 2, \ldots, 9\}$. Then the decimal expansion $d_0.d_1d_2d_3\ldots$ equals the infinite sum (called a *series*) $\sum_{n=0}^{\infty} d_n 10^{-n}$. How do we make sense of this infinite sum? With a sequence! Let $a_1 = 3\cdot 10^0 = 3$, $a_2 = 3\cdot 10^0 + 1\cdot 10^{-1} = 3.1$, $a_2 = 3\cdot 10^0 + 1\cdot 10^{-1} + 4\cdot 10^{-2} = 3.14$, etc. Then $a_1 \le a_2 \le a_3 \le \cdots$ is a monotone increasing sequence. It is certainly bounded above by $3.999\cdots$ (which below we'll see equals 4). So it converges to a finite limit; that limit is π. More generally, the sequence

$a_n = \sum_{k=0}^{n-1} d_k 10^{-k}$ is also monotone increasing and bounded, so converges to some number which we denote by the decimal expansion $d_0.d_1d_2d_3....$

Of course, there is nothing special about 10 here. We use 10 only because we have 10 fingers. Martians have 3 arms, each with 3 fingers, so they use base 9: for them, $d_k \in \{0, 1, 2, ..., 8\}$ and for them $d_0.d_1d_2... = \sum_{n=0}^{\infty} d_n 9^{-n}$. E.g. for them, $\pi = 3.1241881...$

So what does 0.99999.... equal? Clearly,

$$0.9999.... = \sum_{k=1}^{\infty} 9 \cdot 10^{-k} = \frac{9}{10} \sum_{k=0}^{\infty} \left(\frac{1}{10}\right)^k = \frac{9}{10} \frac{1}{1 - \frac{1}{10}} = 1$$

Similarly, $-1.9999... = -2$, $2.12499999... = 2.125$, etc. Lots of numbers have two different decimal expansions.

Ratio Test. Let a_n be a sequence, with no $a_n = 0$, and suppose that the limit of the ratios, $\lim_{n\to\infty} a_{n+1}/a_n$, exists and equals R.

(a) If $R < -1$, then a_n diverges (has sup$\neq$inf)

(b) If $|R| < 1$, then a_n converges to 0.

(c) If $R > 1$, then $a_n \to \infty$ or $a_n \to -\infty$.

This follows from the c^n Corollary given above. If $a_{n+1}/a_n \to 1$, then you can't say anything. E.g. $a_n = -n$ has $R = 1$ but tends to $-\infty$, while $b_n = 8.3$ has $R = 1$ but limit 8.3, and the sequence

$$1, 2, 2^{1/2}, 1, 2^{1/3}, 2^{2/3}, 2, 2^{3/4}, 2^{2/4}, 2^{1/4}, 1, 2^{1/5}, 2^{2/5}, ...$$

has $R = 1$ but diverges (one subsequence is all 1's and another subsequence is all 2's).

If $R = -1$, then sequence can be divergent (e.g. $a_n = (-1)^n$), or converge to 0 (e.g. $a_n = (-1)^n/n$).

Part (a) resp. (c) also applies if $R = -\infty$ resp. $R = \infty$. In (c), it should be easy to tell if the limit of a_n is $+\infty$ or $-\infty$: eventually, all the numbers in the sequence will either be positive (in which case $L = +\infty$), or all negative (in which case $L = -\infty$).

To see how to prove the ratio test, let's do the hardest one: part (b). Let's assume $0 \le R < 1$. Fix $\epsilon > 0$ small enough so that $R + \epsilon < 1$. Then $\frac{a_{n+1}}{a_n} \to R$ implies that there will exist an N such that $|\frac{a_{n+1}}{a_n} - R| < \epsilon$ for all $n > N$. In other words, for all $n > N$, $\left|\frac{a_{n+1}}{a_n}\right| < R + \epsilon < 1$. Now, for any

$n > N$, $a_n = \frac{a_n}{a_{n-1}} \frac{a_{n-1}}{a_{n-2}} \cdots \frac{a_{N+1}}{a_N} a_N$, so

$$|a_n| = \left|\frac{a_n}{a_{n-1}}\right| \left|\frac{a_{n-1}}{a_{n-2}}\right| \cdots \left|\frac{a_{N+1}}{a_N}\right| |a_N| < (R+\epsilon)\cdot(R+\epsilon)\cdots(R+\epsilon)|a_N| = |a_N|(R+\epsilon)^{n-N}$$

In other words, for $n > N$,

$$-\frac{|a_N|}{(R+\epsilon)^N} r^n < a_n < \frac{|a_N|}{(R+\epsilon)^N} r^n$$

where $r = R+\epsilon \in (-1, 1)$. By the Power corollary, both $-\frac{|a_N|}{(R+\epsilon)^N} r^n \to 0$ and $\frac{|a_N|}{(R+\epsilon)^N} r^n \to 0$. So by the Squeeze thm, $a_n \to 0$.

The ratio test is a very powerful tool. We avoid $a_n = 0$ in the ratio test in order to avoid division by 0.

Example 1: Let $c \in \mathbb{N}$ and $d > 1$. Consider $a_n = n^c/d^n$. We compute

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)^c/d^{n+1}}{n^c/d^n} = (1 + \frac{1}{n})^c/d$$

Now, $1 + \frac{1}{n} \to 1$ so $(1 + \frac{1}{n})^c \to 1$ by the Product of Limits Theorem of Week 8. So $a_{n+1}/a_n \to 1/d$. But $d > 1$, so $a_n \to 0$ by the Ratio test.

More generally, if $a_n = P(n)/d^n$ where $P(n)$ is any fixed polynomial in n, then the Ratio Test again gets a limit of ratios equal to $1/d$, so again $a_n \to 0$. So exponentials grow faster than any polynomial.

Example 2: By the factorial $n!$ we mean $1 \cdot 2 \cdot 3 \cdots n$. E.g. $1! = 1, 2! = 2, 3! = 6, 4! = 24, 5! = 120$ etc. In high school you may have seen factorials in probability; next term we'll see factorials in Taylor series. Note that the sequence $a_n = \frac{2^n}{n!}$ has limit 0:

$$\frac{a_{n+1}}{a_n} = \frac{2^{n+1}/(n+1)!}{2^n/n!} = \frac{2^{n+1}}{2^n} \frac{n!}{(n+1)!} = \frac{2}{n+1} \to 0$$

so by the ratio test, $\frac{2^n}{n!} \to 0$.

Recall that every convergent sequence is bounded, but some bounded sequences diverge. However, any bounded monotonic sequence converges. The next result tells us what we can say when a bounded sequence is not monotonic.

Bolzano-Weierstrass Theorem. Every bounded sequence has a convergent subsequence.

The analogue of this holds in any dimension $\mathbb{R}^d$: if you have a sequence $\vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots$ of points inside our classroom, then there will be a subsequence $\vec{x}_{n_k}$ which converges to some point in our classroom.

Bolzano-Weierstrass is a pretty nice theorem, but we won't see many consequences in our course. It says that you can't stuff infinitely many points into a bounded sequence without at least one spot where some points 'pile up' ='accumulate', i.e. get closer and closer together. There may be one, two, ..., infinitely many such spots. Any such spot will have its own subsequence converging to it. If there is exactly one such spot, then the sequence will converge to that spot; but if there are more than one accumulation spot, then the sequence won't converge. (Of course each subsequence will converge).

Example: For example, the sequence $a_n = (-1)^n(1 + \frac{1}{n})$ is a sequence that never gets bigger than 2 nor smaller than -2, so it's bounded. There are has two spots where the points in this sequence accumulate: ± 1. So there is a subsequence (e.g. $b_k = a_{2k}$) converging to 1, and a subsequence (e.g. $c_k = a_{2k-1}$) converging to -1. Since a_n has two subsequences converging to different points, a_n must diverge.

The proof of Bolzano-Weierstrass is easy for us: the sequence is bounded, so $\limsup\limits_{n\to\infty} a_n$ exists and is finite. By an Assn.7 question, there will be a subsequence a_{n_k} which converges to lim sup. That is a convergent subsequence. Similarly, there will be a subsequence converging to lim inf, and there may be several others which converge to points in between the lim sup and the lim inf.

From Week 10, the most important things are the Ratio test and (to a lesser extent) Bolzano-Weierstrass. Next term the geometric series will be important.

Week 11 summary

Cauchy sequences (Section 2.E of Bowman's notes)

Our old definition of convergent sequence from Week 8 is fine, except it requires that you already know what the limit L is. There is another way to guarantee that a sequence converges, without necessarily knowing what the limit is.

Consider this sequence $s_n = \sum_{k=1}^{n} \frac{\cos(\sqrt{2}k)}{k^2}$. So $s_1 = \cos(\sqrt{2})$, $s_2 = \cos(\sqrt{2}) + \frac{\cos(2\sqrt{2})}{4}$, $s_3 = \cos(\sqrt{2}) + \frac{\cos(2\sqrt{2})}{4} + \frac{\cos(3\sqrt{2})}{9}$, etc. Sequences like this (based on sums) are called *series*. We write $\lim\limits_{n\to\infty} s_n = \sum_{k=1}^{\infty} \frac{\cos(\sqrt{2}n)}{n^2}$. But how can we try to prove that s_n converges to some finite number, when we have no idea what that limit would be?

This is how:

Definition. A *Cauchy sequence* a_n is a sequence with this property: for any $\epsilon > 0$, there is an N such that, whenever $m > n > N$, $|a_n - a_m| < \epsilon$.

It doesn't matter whether we require $m > n > N$ (as we did) or $m, n > N$ (as Bowman did): the two definitions would be the same. (i.e. a sequence is Cauchy with the $m > n > N$ in the definition, iff it is Cauchy with the $m, n > N$ definition). We chose to use $m > n > N$, because it is more convenient with series (and later, with integrals).

Cauchy sequence Theorem. A sequence a_n converges to a finite number iff it is a Cauchy sequence.

This is proved using Bolzano–Weierstrass (see also Bowman's notes). Implicit in the proof is the Completeness Axiom. To see this, note that the decimal approximations to $\sqrt{2}$, namely 1, 1.4, 1.41, 1.414, 1.4142,..., is a Cauchy sequence in $\mathbb{Q}$ which does not converge to a rational number. So not all Cauchy sequences in $\mathbb{Q}$ converge (to something in $\mathbb{Q}$). The only difference between $\mathbb{Q}$ and $\mathbb{R}$ is the Completeness Axiom. You can turn this around, and assume the Cauchy sequence Theorem as an axiom, and prove the Completeness Axiom.

So Cauchy is the same as convergence, with the advantage (and also disadvantage) that it doesn't say anything about what the limit is.

Example 1. $a_n = (-1)^n/n^2$ is Cauchy. To see that, choose any $\epsilon > 0$ and let $N = \sqrt{2/\epsilon}$. If $n, m > N$, then

$$|a_n - a_m| = n^{-2} + m^{-2} < N^{-2} + N^{-2} = 2N^{-2} = 2(\sqrt{2/\epsilon})^{-2} = 2/(2/\epsilon) = \epsilon$$

This is consistent with the Cauchy sequence Theorem, since we already knew that $a_n = (-1)^n/n^2$ converged (in fact it converges to 0).

Example 2. Consider the series $s_n = \sum_{k=1}^{n} \frac{\cos(\sqrt{2}k)}{k^2}$. Let's prove it is Cauchy. Write $a_k = \frac{\cos(\sqrt{2}k)}{k^2}$ for convenience, so we don't have to keep writing that fraction over and over again. Choose any $\epsilon > 0$. We need to find an N such that, for all $m > n > N$, $|s_m - s_n| < \epsilon$. First, note that because $m > n$,

$$\begin{aligned} s_m - s_n &= (a_1 + a_2 + \cdots + a_n + a_{n+1} + \cdots + a_m) - (a_1 - a_2 - \cdots - a_n) \\ &= a_{n+1} + a_{n+2} + \cdots + a_m = \sum_{k=n+1}^{m} a_k = \sum_{k=n+1}^{m} \frac{\cos(\sqrt{2}k)}{k^2} \end{aligned}$$

(The same calculation works for any series.) So

$$|s_m - s_n| = \left| \sum_{k=n+1}^{m} \frac{\cos(\sqrt{2}k)}{k^2} \right| \leq \sum_{k=n+1}^{m} \left| \frac{\cos(\sqrt{2}k)}{k^2} \right| = \sum_{k=n+1}^{m} \frac{|\cos(\sqrt{2}k)|}{|k^2|} \leq \sum_{k=n+1}^{m} \frac{1}{k^2}$$

The first '$\leq$' is because of the mighty Triangle Inequality: the absolute value of a finite sum is less than or equal to the sum of the absolute values. The second '$\leq$' is because of the inequality $|\cos(\theta)| \leq 1$.

To finish off the proof that our series s_n in Example 2 is Cauchy, we use the fact that $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges and therefore is also Cauchy. This is a question on your Assn.8. So there is some N such that for all $m > n > N$, $|\sum_{k=n+1}^{m} \frac{1}{k^2}| < \epsilon$ (the absolute values here can be dropped because everything is positive). Plugging this into our inequality for $|s_m - s_n|$, we get that $|s_m - s_n| < \epsilon$. Thus our series is Cauchy. Thus, by the Cauchy sequence theorem, $s_n = \sum_{k=1}^{n} \frac{\cos(\sqrt{2}k)}{k^2}$ converges to some finite number. I personally have absolutely no clue what the limit actually is, but it exists!

Example 3. Prove the series $s_n = \sum_{k=1}^{n} \frac{1}{k}$ is not Cauchy. So we have to prove that:

$$\text{NOT}\,(\forall \epsilon > 0,\ \exists N \text{ such that } \forall m > n > N,\ |s_m - s_n| < \epsilon)\,.$$

You are experts on simplifying these sorts of expressions: $\forall$ becomes $\exists$; $\exists$ becomes $\forall$; don't change the orders of anything; change the final statement (here, it is $|s_m - s_n| < \epsilon$) to its negation. So the result is: to prove something is not Cauchy, we have to prove:

$$\exists \epsilon > 0 \text{ such that } \forall N,\ \exists m > n > N \text{ such that } |s_m - s_n| \geq \epsilon$$

So we must fix ϵ: take $\epsilon = 1/2$ (other choices also work). Our enemy gets to choose N. So we must leave it as a variable. But we now get to select $m > n > N$ which work: select $n = N+1$ and $m = 2N+4$. Then

$$|s_m - s_n| = \sum_{k=n+1}^{m} \frac{1}{k} = \frac{1}{N+2} + \frac{1}{N+3} + \cdots + \frac{1}{2N+4} \geq \frac{1}{2N+4} + \frac{1}{2N+4} + \cdots + \frac{1}{2N+4}$$

We've used the Golden Rule of fraction inequalities to get smaller fractions by making the denominator(=bottom) bigger. So how many identical copies of $\frac{1}{2N+4}$ do we now have? In other words, how many different terms are there in a sum from $N+2$ to $2N+4$? The way I do this is to subtract $N+1$ from all those terms, so the sum actually goes from $(N+2)-(N+1) = 1$ to $(2N+4)-(N+1) = N+3$, so there are $N+3$ terms. So our inequality becomes

$$|s_m - s_n| \geq \frac{N+3}{2N+4} > \frac{N+2}{2N+4} = \frac{1}{2} = \epsilon$$

This then is how you show the series $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges (to ∞).

The best way to show series (infinite sums) converge or diverge, is through Cauchy series.

New Theme: Functions and their limits

Functions (Section 3.A of Bowman's Notes)

Definition. Let D, Y be sets. A *function* $f : D \to Y$ is a choice of $y \in Y$ for each $x \in D$. We write $f(x)$ for this y. The set D is called the *domain* (it is the inputs), the set Y is called the *codomain* (it contains all the outputs).

In this course, we are mainly interested in D being unions of finitely many intervals in $\mathbb{R}$, and Y being $\mathbb{R}$.

Example. The constant functions have $f(x)$ a constant value, e.g. $f(x) = 3$ for all $x \in \mathbb{R}$. The polynomials look like $f(x) = \sum_{k=0}^{n} a_k x^k$ for constants $a_0, \ldots, a_n$. If $a_n \neq 0$, we call n the degree of the polynomial. The *rational functions* (as we know) are of the form $poly(x)/poly(x)$. The polynomial $f(x) = x^2 + 3$ has domain $\mathbb{R}$, and codomain $\mathbb{R}$, as does any polynomial.

We can always make the domain smaller, by throwing away healthy points, but this is a bit strange. By the *natural domain* of a function, we mean the largest subset of $\mathbb{R}$ where the function still makes sense. So any polynomial has natural domain $\mathbb{R}$; $f(x) = 1/x$ has natural domain $\{x \in \mathbb{R} \,|\, x \neq 0\} = (-\infty, 0) \cup (0, \infty)$.

Similarly, we can often make the codomain smaller by throwing away points that the function can never equal. The smallest possible choice of codomain we can make, is called the]it range: it is the set of all possible outputs. Sometimes knowing the range is important, but usually it is very difficult to find. The codomain is usually chosen as simple as possible, which is why we take it to be $\mathbb{R}$. E.g. $f(x) = x^2$ has range $[0, \infty)$, while $f(x) = 1/x$ has range $(-\infty, 0) \cup (0, \infty)$. The rational function $f(x) = (x^2+1)/(x-1)$ has natural domain $(-\infty, 1) \cup (1, \infty)$, codomain $\mathbb{R}$, and range $(-\infty, 2-2\sqrt{2}] \cup [2+2\sqrt{2}, \infty)$. We will mostly ignore range in this course; if we don't give the domain and range, then take the natural domain for the domain, and take the range to be $\mathbb{R}$.

Function limits (Section 3.C of Bowman's Notes)

Calculus is all about limits. For example, the definition of derivative we're going to give shortly is: $\frac{df}{dx}(a) = \lim\limits_{h \to 0} \frac{f(a+h)-f(a)}{h}$. But calculus is about limits of functions, not limits of sequences. Instead of sending $n \to \infty$ along the natural numbers, we want to send e.g. $h \to 0$ along the real numbers.

So we need to define these real number or function limits. The point of wasting so much time on the sequence limits, is that they are a baby version of the function limits. Sequence limits are easier, but if you understand them, then it should be a piece of cake to understand function limits.

Recall what $\lim\limits_{n \to \infty} a_n = L$ means:

for any $\epsilon > 0$, there is an N such that $|a_n - L| < \epsilon$ whenever $n > N$.

The $n > N$ thing is how we say that n gets closer and closer to ∞. The $|a_n - L| < \epsilon$ thing is how we say that a_n gets closer and closer to L. So the definition says: as n gets closer to ∞, a_n gets closer to L. So the following definition should seem pretty natural:

Definition (infinite-finite limit). Let $f(x)$ be any function *and assume the interval* (c, ∞) *is in the domain of* $f(x)$*, for some* $c \in \mathbb{R}$. We say $\lim\limits_{x \to \infty} f(x) = L$ if, for all $\epsilon > 0$, there is an $N > c$ such that $|f(x) - L| < \epsilon$ whenever $x > N$. We also say that the limit *converges* to L.

Define $\lim\limits_{x \to -\infty} f(x) = L$ similarly: $f(x)$ is a function *with* $(-\infty, c)$ *in its domain for some* $c \in \mathbb{R}$; then for any $\epsilon > 0$, there is an $N < c$ such that $|f(x) - L| < \epsilon$ whenever $x < N$.

In words, we say that as x gets closer to $+\infty$ (or $-\infty$), $f(x)$ gets closer to L. The weird stuff in italics about (c, ∞) or $(-\infty, c)$ being in the domain of $f(x)$, and $N > c$ respectively $N < c$, is just to make sure that $f(x)$ is defined. Otherwise $|f(x) - L| < \epsilon$ is meaningless. So just ignore c at first, it is just legal-ese. For $x \to \infty$, we're interested in N being bigger and bigger, $N = 100$, $N = 10000$, $N = 10000000$ etc etc etc. For $x \to -\infty$, we're interested in N getting more and more negative, i.e. $N = -100$, $N = -10000$, $N = -1000000$ etc etc etc.

Example of infinite-finite: Show $\lim\limits_{x\to\infty} \frac{2x^2+x-1}{x^2+x+1} = 2$

OK, given any $\epsilon > 0$, we want an N such that whenever $n > N$, $|\frac{2x^2+x-1}{x^2+x+1} - 2| < \epsilon$. So the big question is, what is N?

Secret work:

$$\left|\frac{2x^2+x-1}{x^2+x+1} - 2\right| = \left|\frac{2x^2+x-1-2x^2-2x-2}{x^2+x+1}\right| = \left|\frac{-x-3}{x^2+x+1}\right| = \frac{x+3}{x^2+x+1}$$

at least when $x > -3$ (we want to get rid of the absolute values, to simplify things). We want to get upper bounds for this, so that means we want the top (numerator) to get bigger and/or the bottom (denominator) to get smaller, and we want both to get simpler and simpler. Well, for $x > 3$, $x + 3 < 2x$. And $x^2 + x + 1 > x^2$ when $x > -1$. So, as long as $x > 3$, we get

$$\frac{x+3}{x^2+x+1} < \frac{2x}{x^2} = \frac{2}{x} < \frac{2}{N}$$

when $x > N$. We are done simplifying, so we can set this equal to ϵ: $\epsilon = 2/N$, which means $N = 2/\epsilon$, and we're done!

Official answer:

Choose any $\epsilon > 0$. Let $N = \max\{3, 2/\epsilon\}$. Then for any $x > N$,

$$\left|\frac{2x^2+x-1}{x^2+x+1} - 2\right| = \left|\frac{2x^2+x-1-2x^2-2x-2}{x^2+x+1}\right| = \left|\frac{-x-3}{x^2+x+1}\right| = \frac{x+3}{x^2+x+1}$$

$$< \frac{2x}{x^2} = \frac{2}{x} < \frac{2}{N} \leq \epsilon$$

and we're done.

Definition (infinite-infinite limit). Let $f(x)$ be any function *and assume that the interval* (c, ∞) *is in the domain of* $f(x)$. We say $\lim\limits_{x\to\infty} f(x) = \infty$ if, for all M, there is an $N > c$ such that $f(x) > M$ whenever $x > N$. Likewise, we say $\lim_{x\to\infty} f(x) = -\infty$ if, for all M, there is an $N > c$ such that $f(x) < M$ whenever $x > N$.

Similarly, let $f(x)$ be any function *and assume that the interval* $(-\infty, c)$ *is in the domain of* $f(x)$. We say $\lim_{x\to-\infty} f(x) = \infty$ if, for all M, there is an $N < c$ such that $f(x) > M$ whenever $x < N$. Likewise, we say $\lim_{x\to-\infty} f(x) = -\infty$ if, for all M, there is an $N < c$ such that $f(x) < M$ whenever $x < N$.

Again, the italics is silly legal-ese, and you can ignore it for now. These 4 mini-definitions of infinite-infinite limits are meant to mimic what is done for sequences when we wrote e.g. $\lim_{n\to\infty} a_n = \infty$.

Example of infinite-infinite: Show $\lim_{x\to-\infty} \frac{-2x^2+x-1}{2x+1} = \infty$

OK, given any M (we're especially interested in one very positive), we want an N (which will get more and more negative as M gets more and more positive) such that whenever $x < N$, $\frac{2x^2-x+1}{-2x-1} > M$. So the big question is, what is N?

Secret work: For $x < 0$ (guaranteed if $N < 0$), $2x^2 - x + 1 > 2x^2$ and $-2x - 1 < -2x$ so

$$\frac{2x^2 - x + 1}{-2x - 1} > \frac{2x^2}{-2x} = -x > -N$$

So we should take $N = -M$.

Official answer: Choose any M. Let $N = \min\{0, -M\}$. Then for any $x < N$,

$$\frac{2x^2 - x + 1}{-2x - 1} > \frac{2x^2}{-2x} = -x > -N \geq M$$

and we're done.

The most important class of limits is:

Definition (finite-finite limit). Let $f(x)$ be any function *and assume both the intervals* (c, a) *and* (a, d) *are in the domain of* $f(x)$*, for some* $c < a < d$. We say $\lim_{x\to a} f(x) = L$ if, for all $\epsilon > 0$, there is a $\delta > 0$ such that $a - \delta > c$ *and* $a + \delta < d$*, and* $|f(x) - L| < \epsilon$ whenever $0 < |x - a| < \delta$. We also say that the limit *converges* to L.

So δ refers to the x-axis, and ϵ to the y-axis. This definition says that, as x gets closer to a, $f(x)$ gets closer to L. All that weird italics stuff about

(c, a) and (d, a) being in the domain of $f(x)$, and that $a-\delta > c$ and $a+\delta < d$, is legal-ese just making sure that $f(x)$ will always be defined. You can ignore it.

Note another piece of legal-ese: we insist that $0 < |x - a|$, so we deliberately avoid $x = a$. The reason is that we want to probe what happens *near* $x = a$; later we'll compare it to what happens *at* $x = a$. We do the same thing in the Definition for the infinite-finite limit: we don't ask what $f(\infty)$ equals. I didn't emphasise this in class this week, because I wanted to focus on the eseentials of the definition, but we'll discuss it in Week 12. The decision to avoid $x = a$ in the definition is just a convention; some books probably allow $x = a$. But considering only the x satisfying $0 < |x - a| < \delta$, is a pretty standard one, and one we will subscribe to in Math 117 and 118.

Your enemy chooses ϵ. You have to respond to the challenge by finding a corresponding δ. You're saying that, as long as x is within δ of a, then $f(x)$ will be within ϵ of L.

Note a difference between tending toward $\pm\infty$, and tending toward a finite number: the limits to $\pm\infty$ are what we'll later call "one-sided limits". This reflects the fact that, really, the number line is a circle and $+\infty = -\infty$. But from our shallow finite perspective that we have in our course, we think of these $\pm\infty$ as different. Nevertheless, in this course let's continue to be shallow, and always think of $+\infty$ as different from $-\infty$. It isn't "wrong" to think of them as different, it is a decision that we are free to make (like considering real numbers instead of complex numbers). From the perspective of modern geometry, it looks rather old-fashioned to treat $\pm\infty$ as different (and to consider real numbers instead of complex numbers).

Example: Let $f(x) = 2x + 3$. Show $\lim\limits_{x\to 1} f(x) = 5$.

Secret work: So that means $a = 1$ and $L = 5$, and we're in the finite-finite world. We need to find explicitly how close x must be to 1, in order that $f(x)$ be close to 5 (i.e. we must find δ explicitly, given ϵ). We compute

$$|f(x) - L| = |(2x + 3) - 5| = |2x - 2| = 2|x - 1|$$

so in order to have $|f(x) - 5| < \epsilon$, we need $|x - 1| < \epsilon/2$. In other words, we should take $\delta = \epsilon/2$.

Official answer: The domain of $f(x) = 2x + 3$ is all of $\mathbb{R}$, so we can ignore c, d here. Choose any $\epsilon > 0$. Take $\delta = \epsilon/2$. Then when $|x - 1| < \delta$,

$$|f(x) - L| = |(2x + 3) - 5| = |2x - 2| = 2|x - 1| < 2\delta = \epsilon$$

and we're done.

Example: Show $f(x) = 1/x$ is continuous at $x = 4$.

Secret work: So that means $a = 4$ and $L = f(4) = 1/4$, and we're in the finite-finite world. We need to find explicitly how close x must be to 4, in order that $f(x)$ be close to $f(4)$. We compute

$$|f(x) - L| = |1/x - 1/4| = |\frac{4-x}{4x}| = \frac{|x-4|}{4|x|}$$

We want to bound both terms $|x - 4|$ and $|x|$. We can always force $x > 2$ here (it just amounts to insisting that $\delta < 2$, since $|x-4| < \delta$ means $4-\delta < x < 4+\delta$). So $|x| > 2$, and we get $|f(x)-L| < |x-2|/8$. So in order to have $|f(x) - 4| < \epsilon$, we need $|x - 2| < 8\epsilon$. In other words, we should take $\delta = 8\epsilon$ (and also $\delta < 2$).

Official answer: The domain of $f(x) = 1/x$ is $(-\infty, 0) \cup (0, \infty)$, so we have $c = 0, d = \infty$ here. Choose any $\epsilon > 0$. Take $\delta = \min\{2, 8\epsilon\}$. Then when $|x - 4| < \delta$, $x > 2$ (since $\delta \leq 2$, so $x > 4 - \delta \geq 2$), and

$$|f(x) - L| = |1/x - 1/4| = |\frac{4-x}{4x}| = \frac{|x-4|}{4|x|} < \frac{|x-4|}{8}$$

since $\delta < 2$ (and hence again $x > a - \delta \leq 2$). But $|x - 4|/8 < \delta/8 \leq \epsilon$, and we're done.

The final possibility:

Definition (finite-infinite limit). Let $f(x)$ be any function *and assume both the intervals* (c, a) *and* (a, d) *are in the domain of* $f(x)$*, for some* $c < a < d$. We say $\lim_{x\to a} f(x) = \infty$ if, for all M, there is a $\delta > 0$ such that $a-\delta > c$ *and* $a + \delta < d$, *and* $f(x) > M$ whenever $0 < |x - a| < \delta$. Likewise, we say $\lim_{x\to a} f(x) = -\infty$ if, for all M, there is a $\delta > 0$ such that $a - \delta > c$ *and* $a + \delta < d$, *and* $f(x) < M$ whenever $0 < |x - a| < \delta$.

Again, ignore the italics. Again we see the decision to use $0 < |x-a| < \delta$ and not the simpler $|x - a| < \delta$.

Example of finite-infinite: Show $\lim_{x\to 1} \frac{-2x^2+x-1}{x^2-2x+1} = -\infty$

OK, given any M (we're interested in M being really negative, like $M = -1000000$ and worse), we want a $\delta > 0$ (we're interested in δ like $\delta = 0.00001$) such that whenever $0 < |x - 1| < \delta$, $\frac{-2x^2+x-1}{x^2-2x+1} < -M$. So the big question is, what is δ?

Secret work: For $0 < x < 2$ (guaranteed if $\delta < 1$), $-2x^2 + x - 1 = -2(x - 1/4)^2 - 7/8 \leq -7/8$ so $\frac{-2x^2+x-1}{x^2-2x+1} < \frac{-7}{8(x-1)^2} < \frac{-7}{8\delta^2}$. So we should take $M = \frac{-7}{8\delta^2}$, i.e. $\delta = \sqrt{7}/\sqrt{8|M|}$.

Official answer: Choose any M. Let $\delta = \min\{1, \sqrt{7}/\sqrt{8|M|}\}$. Then for any $0 < |x - 1| < \delta$,

$$\frac{-2x^2 + x - 1}{x^2 - 2x + 1} = \frac{-2(x - 1/4)^2 - 7/8}{(x - 1)^2} \leq \frac{-7}{8(x - 1)^2} < \frac{-7}{8\delta^2} \leq M$$

and we're done.

The statement that $\lim_{x \to a} f(x) \neq L$ is:

NOT $(\forall \epsilon > 0,\ \exists \delta > 0$ such that $\forall x$ satisfying $0 < |x - a| < \delta,\ |f(x) - L| < \epsilon)$.

(I'm discarding the silly italics stuff to keep things simple) By the usual rules, this simplifies to:

$\exists \epsilon > 0$, such that $\forall \delta > 0\ \exists x$ satisfying $0 < |x - a| < \delta$, such that $|f(x) - L| < \epsilon$.

So to prove that $\lim_{x \to a} f(x) \neq L$, we have to come up with a specific value for ϵ (it can depend on L, we leave δ as a variable, and we find a specific x that will work.

From Week 11, the important things are to be really comfortable with limits of functions. The most important of these are finite-finite ones. Don't worry so much about the italics stuff. Cauchy sequences are not so important in this course – they are more useful next term when we discuss integration.

Week 12 summary

Non-example of finite-infinite: Let's explain why, from the definitions, that $\lim_{x\to 0} 1/x$ does not equal ∞. Take 'NOT' of the definition of limit for finite-infinite:

$$\text{NOT}(\forall M,\ \exists \delta > 0 \text{ such that } \forall x \text{ satisfying } 0 < |x - 0| < \delta,\ 1/x > M)$$

which is the same as

$$\exists M,\ \text{such that } \forall \delta > 0,\ \exists x \text{ satisfying } 0 < |x| < \delta,\ \text{and } 1/x \le M$$

Take $M = 0$. Choose any $\delta > 0$. Take $x = -\delta/2$. Then $0 < |x| = \delta/2 < \delta$, and $1/x = -2/\delta < 0 = M$ and we're done.

The Bridge: Section 3.D,3.E in Bowman's notes

We have seen two main topics since the first midterm: sequences and their limits, and functions and their limits. It's time to connect them:

Bridge Theorem. $\lim_{x\to a} f(x) = L$ iff for all sequences $a_n \to a$ (with $a_n \neq a$), $f(a_n) \to L$.

This is Theorem 3.1 in Bowman's notes. This applies to all 4 limit types: finite-finite, infinite-finite, etc. The reason for requiring each $a_n \neq a$, is the same as requiring $0 < |x - a|$ in the definition of $\lim_{x\to a}$. I'm being sloppy regarding domains: for now, I'm requiring $x = a$ to be an interior point of the domain of f, which means there are $c < a < d$ such that $(c, a) \cup (a, d) \subseteq$ domain(f), and then we require all $a_n \in (c, a) \cup (a, d)$ — if this domain condition confuses you, ignore it.

The Bridge Theorem doesn't look that wonderful, but it allows us to carry over ('bridge') everything we know about limits of sequences, which is quite a bit. We'll prove it shortly. Here's one immediate application:

Nonconvergence Corollary. Choose any function $f(x)$. $\lim_{x\to a} f(x)$ doesn't exist, iff there are sequences $a_n, b_n \to a$ (where $a_n, b_n \neq a$) and $\lim_{n\to\infty} f(a_n) \neq \lim_{n\to\infty} f(b_n)$.

For example, to prove that $\lim_{x\to 0} 1/x$ doesn't exist, just consider these two sequences: $a_n = 1/n$, and $b_n = -1/n$. Both converge to 0. But $1/a_n \to$

$+\infty$ and $1/b_n \to -\infty$, and $+\infty \neq -\infty$, so we're done! Another example: $\lim_{x\to 0} \operatorname{sgn}(x)$ does not exist. Our two sequences are $a_n = 1/n$ and $b_n = -1/n$. Then $\operatorname{sgn}(a_n) = +1$ which tends to $+1$ (obviously!), while $\operatorname{sgn}(b_n) = -1$ which tends to -1. So that means $\lim_{x\to 0} \operatorname{sgn}(x)$ does not exist.

Another consequence:

Pretty Useful Theorem for function limits. Assume $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = L'$. Then:

- **(a)** $\lim_{x\to a}(\alpha f(x) + \beta g(x)) = \alpha L + \beta L'$, for any constants $\alpha, \beta \in \mathbb{R}$;
- **(b)** $\lim_{x\to a} f(x)g(x) = LL'$;
- **(c)** If $g(x) \neq 0$ in some interval (c, d) containing a, and $L' \neq 0$, then $\lim_{x\to a} f(x)/g(x) = L/L'$.

This is the best way to compute most function limits. Of course, the proof is a combination of the Bridge Theorem with the Pretty Useful Theorem on sequence limits of sums, products, quotients of sequences.

We've just seen the usefulness of the Bridge Theorem. Now we'll prove them. We'll just prove one of them (finite-finite), as they're all proved the same way. Assignment 8 Question 5 deals with the infinite-infinite case.

Bridge Theorem (finite-finite). Suppose that $(b, a) \cup (a, c)$ lies in the domain of f, for some b, c with $b < a < c$. Then $\lim_{x\to a} f(x) = L$ **iff** all sequences $a_n \to a$, which have all a_n in the domain of f and $a_n \neq a$ for all n, also have $f(a_n) \to L$.

The condition that "$(b, a) \cup (a, c)$ is in the domain of f", just says that all points close to a are in the domain of f. So we are excluding a being an endpoint of the domain (this is what one-sided limits are all about), and we're also excluding weird things like the domain just consisting of rational numbers but not irrational numbers. The condition that "all a_n [are] in the domain of f" just makes sure that each $f(a_n)$ is defined. Ignore it if you want.

Proof. This is an "iff", so there are two directions to prove.
$\Longrightarrow$ For this direction, we assume that $\lim_{x\to a} f(x) = L$. Let a_n be any sequence entirely in the domain of f, and we assume a_n converges to a and that each $a_n \neq a$. We want to prove that the sequence $f(a_n)$ converges to L.

$\lim_{x\to a} f(x) = L$ tells that when x is close to a, then $f(x)$ will be close to L. $a_n \to a$ tells us that for all big enough n, a_n will be close to a. So it shouldn't be too hard to see how to make your proof:

Choose any $\varepsilon > 0$. Because $\lim_{x\to a} f(x) = L$, there exists a $\delta > 0$ such that, whenever $0 < |x - a| < \delta$, then $|f(x) - L| < \epsilon$. Because $a_n \to a$, there is an N such that, whenever $n > N$, $|a_n - a| < \delta$. Putting these together, we see that for all $n > N$, $|f(a_n) - L| < \epsilon$. QED to $\Rightarrow$

$\Longleftarrow$ Assume that for any sequence $a_n \to a$, where all terms a_n lie in the domain of f and $a_n \neq a$, we have $f(a_n) \to L$. We want to prove that $\lim_{x\to a} f(x) = L$.

Suppose for contradiction that $\lim_{x\to a} f(x) \neq L$. Then that means that there is some $\varepsilon_0 > 0$ such that, for any $\delta > 0$, there is an $x_\delta \neq a$ in the domain of f such that $|x_\delta - a| < \delta$ but $|f(x_\delta) - L| \geq \varepsilon_0$. Let's use this to create a sequence. The idea is to use $\delta = 1/n$ for each n.

Let $a_n = x_{1/n}$. Then by the properties of the x_δ, we know $0 < |a_n - a| < 1/n$, so $a_n \to a$ and $a_n \neq a$. Also, $|f(a_n) - L| \geq \varepsilon_0$. So certainly $f(a_n)$ doesn't converge to L.

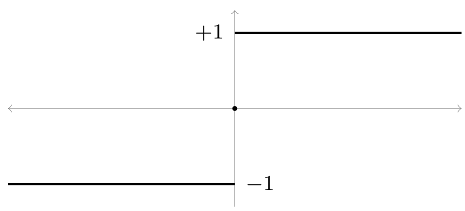
But this cannot happen: We assumed at the start that *any time* a sequence a_n tends to a, then $f(a_n)$ *must* tend to L. This terrible contradiction means that our most recent assumption (or supposition) was wrong. So it is false that "$\lim_{x\to a} f(x) \neq L$", which means that it is true that $\lim_{x\to a} f(x) = L$, and we're done. QED to $\Leftarrow$, hence QED to Bridge Thm

Continuity: Section 3.D,3.E in Bowman's notes

A *continuous function* is a function that can be graphed without your pencil leaving the paper. So $\sin(x)$ and any polynomial, are continuous functions. A function *discontinuous* at some point $x = c$ has a jump there. For example,

$$\operatorname{sgn}(x) = \begin{cases} +1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

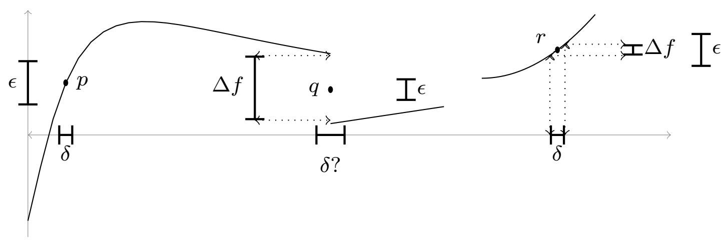
is discontinuous at $x = 0$. This is its graph, with its jump at $x = 0$:



Fancy definition of continuity: We say that $f(x)$ is *continuous at* $x = a$ if the interval (c, d) is in the domain of $f(x)$ for some $c < a < d$, and $\lim_{x \to a} f(x) = f(a)$. Otherwise we say $f(x)$ is *discontinuous at* $x = a$.

Note that continuity only applies to finite-finite limits. Note that, when $f(x)$ is continuous, we can write $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

This picture may help:



The basic idea is borrowed from sequences. Look at the point $p = (a, b)$ in the graph, so $b = f(a)$: If someone tells us they want $f(x)$ to be within ϵ of $f(a)$, we just have to make x within δ of a. So ϵ refers to the error in the y-direction (i.e. the error in the values of f), and δ refers to how close x has to be to a in order that the range of values of f be within ϵ. The smaller ϵ is chosen, the smaller we have to make δ. This is shown in more detail on the right side at the point r: if we force x to be in the δ interval shown, the range of $f(x)$ will be in the Δf interval shown, and this is within the ϵ interval chosen. If ϵ is made much smaller, we'll have to make δ much smaller, but we'll always be able to do that. So the graph is continuous at both p and r. It isn't continuous at q because no matter how small you make δ, you'll get a Δf much bigger than the ϵ chosen.

So continuity means that, no matter how small your enemy makes ϵ, you'll be able to constrain x (using δ) so that the range of f over that δ-interval, what we call $\Delta f(x)$ in the picture, is smaller than ϵ.

How about absolute value $f(x) = |x|$: is it continuous? The answer is yes. One way to show this is the old way, directly from the definition. Choose any $a \in \mathbb{R}$, and $\epsilon > 0$. Let $\delta = \epsilon$. Then for any x satisfying $0 < |x - a| < \delta$, the triangle inequality says $\big||x| - |a|\big| \le |x-a| \le \delta = \epsilon$ and we're done. That was surprisingly easy!

Another example: let's prove discontinuity using δ-ϵ. Consider again the sign function $\mathrm{sgn}(x)$. From the graph given earlier, we see the jump at $x = 0$. To get that $\lim_{x\to 0} \mathrm{sgn}(x)$ cannot equal 0 (in fact the limit doesn't exist), we need to find an $\epsilon > 0$ such that, for all $\delta > 0$, there is an x satisfying $0 < |x - 0| < \delta$ such that $|\mathrm{sgn}(x) - 0| > \epsilon$.

$\epsilon = 0.5$ works (as does any $\epsilon < 1$). For any $\delta > 0$, take $x = \delta/2$. Then $0 < |x| < \delta$, and $|\mathrm{sgn}(x)| = 1 > \epsilon$. So $\mathrm{sgn}(x)$ is discontinuous at $x = 0$.

Similarly, we can show $\lim_{x\to 0} \mathrm{sgn}(x)$ does not exist. Suppose for contradiction that $\lim_{x\to 0} \mathrm{sgn}(x) = L$. Take $\epsilon = 0.5$. Then there is some $\delta > 0$ such that for any x with $0 < |x| < \delta$, $|\mathrm{sgn}(x) - L| < 0.5$. Well, compare $x = \delta/2$ with $x = -\delta/2$: both satisfy $0 < |x| < \delta$, so $x = \delta/2$ gives $|+1 - L| < 0.5$ and $x = -\delta/2$ gives $|-1 - L| < 0.5$. In other words, the first says $0.5 < L < 1.5$ and the second says $-1.5 < L < -0.5$, but both can't be satisfied. So no such limit L can exist.

The Bridge Theorem together with the Pretty useful Theorem tells us:

Pretty Useful Continuity Thm (a) f, g continuous at $x = a$ implies $\alpha f(x) + \beta g(x)$ is also continuous at $x = a$, for any constants $\alpha, \beta \in \mathbb{R}$.

(b) f, g continuous at $x = a$ implies $f(x)g(x)$ is also continuous at $x = a$.

(c) If f, g are continuous at $x = a$, and $g(a) \neq 0$, then $f(x)/g(x)$ is also continuous at $x = a$.

This is pretty incredible! It is trivial that the constant function $f(x) = 1$, and the function $g(x) = x$, are both continuous everywhere. Hence by (a), any linear function $x \mapsto \alpha + \beta x = \alpha f(x) + \beta g(x)$ is continuous everywhere. But by (b), the powers $x \mapsto x^n = g(x)\, g(x) \cdots g(x)$ will be continuous, so together (a) and (b) say that any polynomial is continuous everywhere. Hence (a),(b),(c) say that any rational function $p(x)/q(x)$ will be continuous anywhere the denominator $q(x)$ doesn't vanish.

Because $g(x)$ is continuous at $x = a$, the condition that $g(a) \neq 0$ actually implies that $g(x) \neq 0$ for all $x \in (c, d)$, for some $c, d \in \mathbb{R}$ satisfying $c < a < d$. The reason: take $\epsilon = |g(a)|$, then there is some $\delta > 0$ such that, for all $|x - a| < \delta$, $|f(x) - f(c)| < \epsilon$. Take $c = a - \delta, d = a + \delta$; then for all $x \in (c, d)$, $f(c) - \epsilon < f(x) < f(c) + \epsilon$, i.e. $f(x) \neq 0$.

The Bridge Thm gives a fast way to prove that a function is or is not continuous at some point.

Theorem. Choose any function $f(x)$.

(a) $f(x)$ is continuous at $x = a$, iff for all sequences $a_n \to a$ we have $\lim_{n\to\infty} f(a_n) = f(a)$.

(b) $f(x)$ is discontinuous at $x = a$, iff we can find a sequence $a_n \to a$ such that $\lim_{n\to\infty} f(a_n) \neq f(a)$.

For example, we can use this to show that $\operatorname{sgn}(x)$ is not continuous at $x = 0$. Take $a_n = 1/n$. Then $a_n \to 0$, but $\operatorname{sgn}(a_n) = +1 \to 1$ which doesn't equal $\operatorname{sgn}(0)$, and we're done! (by (b)).

A useful source of nice examples (and counterexamples) exploits rational and irrational numbers. For this purpose, here is a description of which numbers are rational or irrational. Let x be a real number. Then it has a decimal expansion $...d_2d_1d_0.d_{-1}d_{-2}\ldots$ where all but finitely many digits to the left of the decimal point are 0. (some numbers have 2 decimal expansions, but that isn't important here). Our number x is rational, iff eventually the decimal expansion to the right starts to repeat. So 1426.78431212121212121... is rational, but 0.1001000100001000001... is irrational. There is nothing special about base 10; the same works for any base.

In any case, the following is true: for any real number x, there is a sequence a_n of rational numbers tending to x, and a sequence b_n of irrational numbers tending to x. For example, if x is rational, then $a_n = x + 10^{-n}$ and $b_n = x + \sqrt{2}10^{-n}$ work. For x irrational, $a_n = ...d_2d_1d_0.d_{-1}...d_{-n}$ and $b_n = x + 10^{-n}$ work. (e.g. for $x = \pi$, $a_1 = 3$, $a_2 = 3.1$, $a_3 = 3.14$, etc).

OK, let's use the theorem to show that the function

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

is discontinuous everywhere. Choose any $a \in \mathbb{R}$, and let $a_n, b_n \to a$ where a_n are rational and b_n are irrational. Then $f(a_n) = 1 \to 1$ and $f(b_n) = 0 \to 0$, so $\lim_{x\to a} f(x)$ doesn't exist, so certainly $f(x)$ can never be continuous.

Another example: let's show $g(x) = xf(x)$ is continuous at $x = 0$, where $f(x)$ is that rational/irrational function. Choose any sequence $x_n \to 0$ (it doesn't matter whether the x_n are rational or irrational or both). We get that $g(x_n)$ equals either x_n or 0, depending on whether x_n is rational or irrational. So we get the inequalities $-|x_n| \leq g(x_n) \leq |x_n|$. Now use the Squeeze Theorem: absolute value is continuous, so $\pm|x_n| \to \pm|0| = 0$, so we must have $g(x_n) \to 0$ as well. This means $g(x)$ is continuous at $x = 0$. (It will be discontinuous everywhere else)

We can use that theorem to give a fast proof of the fact that composition (nesting) of continuous functions is continuous:

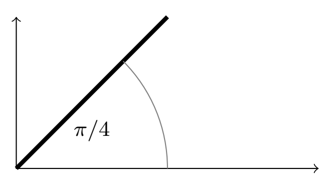
Theorem. Suppose $f(x)$ is continuous at $x = a$, and $g(y)$ is continuous at $y = f(a)$. Then the function $(g \circ f)(x) = g(f(x))$ is continuous at $x = a$.

To prove this, let a_n be any sequence tending to a. Then $b_n = f(a_n)$ is a sequence tending to $f(a)$, since f is continuous at a. So that means that $g(b_n) = g(f(a_n))$ tends to $g(f(a))$, since $g(y)$ is continuous at $y = f(a)$. Hence by the above theorem, $g \circ f$ is continuous at $x = a$.

Nesting functions is a great way to make more complicated functions from simpler ones.

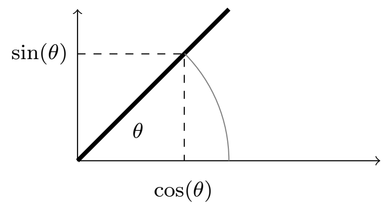
Trigonometry: (Section 3.B of Bowman's Notes)

It turns out that there is a preferred unit to measure angles. (There is not a preferred unit to measure length.) The preferred angle measurement is called *radians*. Draw a circle with radius 1; draw two lines meeting at the origin; the angle between the lines is measured by the arc-length between where those lines cross the circle. So a 90° angle is a quarter of the full circle, so a quarter of 2π, so $\pi/2$. 180° equals π radians, etc. Equivalently, you can also say angle θ radians is a pizza-slie of area $\theta/2$ (e.g. a full circle, i.e. 2π radians, has area π).



It's when we finally get to calculus, that we'll really feel the benefit of this choice of units.

All of you are familiar with the trignometric functions. The most important of these are sine and cosine. Choose any angle, draw a line at that angle with respect to the x-axis. The coordinates of the point where your line hits the circle is $(\cos(\theta), \sin(\theta))$.



You can use that as the way you define sine and cosine, or equivalently you can use the usual triangle way. Then $\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$ is the slope of the line.

You can ignore secant, cosecant, cotangent, etc.

The most important formula is $\cos^2(\theta) + \sin^2(\theta) = 1$. The notation is a little ambiguous: what this means is $(\cos(\theta))^2 + (\sin(\theta))^2 = 1$. This holds because these are the coordinates of a point on the unit circle: $x^2 + y^2 = 1$.

$\cos(\theta)$ and $\sin(\theta)$ both have domain $\mathbb{R}$ and range $[-1, 1]$. So $-1 \leq \sin(\theta) \leq 1$ and $-1 \leq \cos(\theta) \leq 1$. They both are periodic with period 2π: $\cos(\theta + 2\pi) = \cos(\theta)$ and $\sin(\theta + 2\pi) = \sin(\theta)$. $\cos(\theta)$ is even and $\sin(\theta)$ is odd: $\cos(-\theta) = \cos(\theta)$ and $\sin(-\theta) = -\sin(\theta)$.

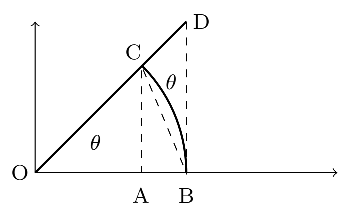
The angle sum formulas are:

$$\cos(\theta+\phi) = \cos(\theta)\ \cos(\phi) - \sin(\theta)\ \sin(\phi) \ \text{ and } \ \sin(\theta+\phi) = \sin(\theta)\ \cos(\phi) + \cos(\theta)\ \sin(\phi)$$

In particular, $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) = 2\cos^2(\theta) - 1$ and $\sin(2\theta) = 2\sin(\theta)\cos(\theta)$.

Most important for us is an additional Trig inequality:

$$|x|\,|\cos(x)| \leq |\sin(x)| \leq |x|\,, \ \forall x \in [-\pi/2, \pi/2]$$



To see that $\sin(\theta) \leq \theta$ for $0 \leq \theta \leq \pi/2$, note that the length of segment AC is $\sin(\theta)$. We know segment CB is longer, since it is the hypothenuse of right-triangle ABC, and the hypothenuse is always the longest side of a right-triangle (this follows from Pythagorus' Thm). We know the circular arc-length of BC must be longer than that hypothenuse BC, since the shortest distance between two points is the straight line. That arc-length is precisely θ. So we see that $\sin(\theta)$, which is the length of AC, must be less than θ, which is the arc-length BC: i.e. $\sin(\theta) \leq \theta$.

Now consider the big triangle OBD. It has area $\frac{1}{2} \cdot \text{base} \cdot \text{height} = \frac{1}{2} \cdot 1 \cdot \tan(\theta)$, since BD=BD/OB=AC/OA=$\sin(\theta)/\cos(\theta)$. Certainly, that big triangle has bigger area than the area of the pizza slice OCB, which is $\frac{1}{2}\theta$. Thus $\frac{1}{2}\tan(\theta) \geq \frac{1}{2}\theta$, i.e. $\sin(\theta) \geq \theta \cos(\theta)$.

Comment on arbitrariness in Mathematics & Science

Mathematics strives for the Universal, much more than any other science. So smart dolphins, or smart 3-armed 3-fingered martians, or smart Cloud Creatures in the atmosphere of Jupiter, should have the same math as us.

An example of arbitrariness is our choice of Base 10 and the decimal expansion of real numbers. We chose that because we have 10 fingers, so 10 is our 'unit' of counting. There is no mathematical reason that I know of, that tells us one base is preferred over another. Certainly, base 10 is *not* special.

Is it clear that all these smart beings know about the real numbers? Not obvious to me that $\mathbb{R}$ is part of Universal Math.

We have 360° in a circle, because 360 is close to the number of days in a year, and because 360 is a highly divisible number. It is possible no other intelligent beings measure angles like that. There is a universal way to measure angles, however: *radians.*

One-sided limits: (Section 3.F of Bowman's Notes)

The final loose-end on limits is 1-sided limits. This was discussed on your Assn.9. We only approach a from one side, either from the left (the "negative" side, so we call it $\lim_{x\to a^-} f(x)$), or from the right (the "positive" side, so we call it $\lim_{x\to a^+} f(x)$).

Definition (finite-finite$^+$). Let $f(x)$ be any function, and assume that the domain of f contains the interval (a, b) for some $b > a$. We say that $\lim_{x\to a^+} f(x)$ exists and equals $L \in \mathbb{R}$, if for any $\varepsilon > 0$ there is a $\delta > 0$ such that for any x in the domain of f with $0 < x - a < \delta$, $|f(x) - L| < \varepsilon$.

Definition (finite-finite$^-$). Let $f(x)$ be any function, and assume that the domain of f contains the interval (b, a) for some $b < a$. We say that

$\lim_{x \to a^-} f(x)$ exists and equals $L \in \mathbb{R}$, if for any $\varepsilon > 0$ there is a $\delta > 0$ such that for any x in the domain of f with $0 < a - x < \delta$, $|f(x) - L| < \varepsilon$.

For example, the $\mathrm{sgn}(x)$ function obeys Then $\lim_{x \to 0^+} \mathrm{sgn}(x) = 1$, and $\lim_{x \to 0^-} \mathrm{sgn}(x) = -1$.

Note that $\lim_{x \to a} f(x) = L$, iff both $\lim_{x \to a^+} f(x) = L$ and $\lim_{x \to a^-} f(x) = L$.

We say $f(x)$ is *continuous from the right* at $x = a$, if $\lim_{x \to a^+} f(x) = f(a)$. We say $f(x)$ is *continuous from the left* at $x = a$, if $\lim_{x \to a^-} f(x) = f(a)$. We say $f(x)$ is continuous on the interval $[a, b]$, if $f(x)$ is continuous at all $x = c$ for $a < c < b$, and also $f(x)$ is continuous from the right at $x = a$, and continuous from the left at $x = b$.

Note that $f(x)$ is continuous at $x = a$, iff it is continuous at a from the left and continuous at a from the right.

For example, $\mathrm{sgn}(x)$ is continuous on the intervals $(0, \infty)$ and $(-\infty, 0)$, but it is not continuous at $x = 0$, nor continuous there from the right nor from the left. However, if we change sgn(0) so that sgn(0)=1 (instead of sgn(0)=0), then $\mathrm{sgn}(x)$ would be continuous at $x = 0$ from the right. This newly defined $\mathrm{sgn}(x)$ would be continuous on $[0, \infty)$. It is still not continuous at $x = 0$ from the left.

The usual definitions can be made for infinite limits:

Definition (finite$^+$-infinite). Let $f(x)$ be any function, and assume that the domain of f contains the interval (a, b) for some $b > a$. We say that $\lim_{x \to a^+} f(x)$ exists and equals ∞, if for any $M > 0$ there is a $\delta > 0$ such that for any x in the domain of f with $0 < x - a < \delta$, $f(x) > M$. We say that $\lim_{x \to a^+} f(x)$ exists and equals $-\infty$, if for any $M < 0$ there is a $\delta > 0$ such that for any x in the domain of f with $0 < x - a < \delta$, $f(x) < M$.

Definition (finite$^-$-infinite). Let $f(x)$ be any function, and assume that the domain of f contains the interval (a, b) for some $b > a$. We say that $\lim_{x \to a^-} f(x)$ exists and equals ∞, if for any $M > 0$ there is a $\delta > 0$ such that for any x in the domain of f with $0 < a - x < \delta$, $f(x) > M$. We say that $\lim_{x \to a^-} f(x)$ exists and equals $-\infty$, if for any $M < 0$ there is a $\delta > 0$ such that for any x in the domain of f with $0 < a - x < \delta$, $f(x) < M$.

As before, $\lim_{x \to a} f(x) = \pm\infty$, iff $\lim_{x \to a} f(x) = \lim_{x \to a} f(x) = \pm\infty$.

For example, $\lim_{x \to 0^+} 1/x = \infty$, and $\lim_{x \to 0^-} 1/x = -\infty$.

If the limit as x approaches a is infinite, then the function can't be continuous there. So continuity only applies to finite-finite limits.

How about infinite-finite or infinite-infinite one-sided limits? We've already done those! Recall that we can think of the real number circle (where $+\infty$ and $-\infty$ are identified). When we write $\lim_{x\to\infty} f(x)$, we really mean $\lim_{x\to\infty^-} f(x)$. When we write $\lim_{x\to-\infty} f(x)$, we really mean $\lim_{x\to\infty^+} f(x)$, or if you prefer $\lim_{x\to-\infty^+} f(x)$.

The bridge theorems work for one-sided limits. For example: $\lim_{x\to a^+} f(x) = L$ iff for all sequences $a_n \to a$, where all a_n are in domain of f and $a_n > a$, $\lim_{n\to\infty} f(a_n) = L$. Another example: $f(x)$ is continuous at $x = a$ from the left, iff for all sequences $a_n \to a$, where all a_n are in domain of f and $a_n < a$, $\lim_{n\to\infty} f(a_n) = f(a)$.

From Week 12, the important things are the Bridge Theorem and its consequences, and what continuity means. Most important as always is finite-finite limits. Be comfortable with sine and cosine; know the inequality $|\theta|\,|\cos(\theta)| \leq |\sin(\theta)| \leq |\theta|$ (we'll see what it's good for next week).

Week 13 summary

Example: Prove that $\sin(x)$ is continuous for all $x \in \mathbb{R}$.

The trick is to use the trig identity $\sin(x) - \sin(y) = 2\cos(\frac{x+y}{2})\ \sin(\frac{x-y}{2})$. This identity is proved by the angle-sum formulas:

$$\sin(x) - \sin(y) = \sin(\frac{x+y}{2} + \frac{x-y}{2}) - \sin(\frac{x+y}{2} - \frac{x-y}{2})$$

$$= \sin(\frac{x+y}{2})\ \cos(\frac{x-y}{2}) + \cos(\frac{x+y}{2})\ \sin(\frac{x-y}{2})$$

$$-(\sin(\frac{x+y}{2})\ \cos(\frac{x-y}{2}) + \cos(\frac{x+y}{2})\ \sin(\frac{y-x}{2}))$$

$$= 2\cos(\frac{x+y}{2})\ \sin(\frac{x-y}{2})$$

Secret work: Fix any $a \in \mathbb{R}$. We're again in the finite-finite world. We need to find explicitly how close x must be to a, in order that $\sin(x)$ be close to $\sin(a)$. Using that trig identity, we compute

$$|\sin(x) - \sin(a)| = |2\cos(\frac{x+a}{2})\ \sin(\frac{x-a}{2})| \leq 2|\frac{x-a}{2}|$$

using the bounds $|\cos(\theta)| \leq 1$ (hence $|\cos(\frac{x+a}{2})| \leq 1$) and $|\sin(\theta)| \leq |x|$ (hence $|\sin(\frac{x-a}{2})| \leq \frac{|x-a|}{2}$). So $|\sin(x) - \sin(a)| \leq \epsilon$ will happen if $|x-a| < \epsilon$, so we should take $\delta = \epsilon$.

Official answer: The domain of $f(x) = \sin(x)$ is all of $\mathbb{R}$, so we can ignore c, d here. Fix any $a \in \mathbb{R}$: we want to prove $\lim\limits_{x \to a} \sin(x) = \sin(a)$. So choose any $\epsilon > 0$. Take $\delta = \epsilon$. Then

$$|\sin(x) - \sin(a)| = |2\cos(\frac{x+a}{2})\ \sin(\frac{x-a}{2})| \leq 2|\frac{x-a}{2}|$$

using the bounds $|\cos(\theta)| \leq 1$ (hence $|\cos(\frac{x+a}{2})| \leq 1$) and $|\sin(\theta)| \leq |x|$ (hence $|\sin(\frac{x-a}{2})| \leq \frac{|x-a|}{2}$). So $|\sin(x) - \sin(a)| < \delta = \epsilon$, and we're done.

Work out for yourself the proof that $\cos(x)$ is also continuous for all $x \in \mathbb{R}$. Use the identity $\cos(x) - \cos(y) = 2\sin(\frac{x+y}{2})\ \sin(\frac{y-x}{2})$.

Example: Prove that $\lim\limits_{x \to 0} \frac{\sin(x)}{x} = 1$.

For $0 < x < \pi/2$, we know from last week that $x\cos(x) \le \sin(x) \le x$, hence $\cos(x) \le \frac{\sin(x)}{x} \le 1$. For $-\pi/2 < x < 0$ we have $x\cos(x) \ge \sin(x) \ge x$ and hence $\cos(x) \le \frac{\sin(x)}{x} \le 1$. So $\cos(x) \le \frac{\sin(x)}{x} \le 1$ holds for all $0 < |x| < \pi/2$.

Let x_n be any sequence tending to 0, where all $x_n \ne 0$. We can assume all $|x_n| < \pi/2$, by throwing away the first few terms if necessary. Then

$$\cos(x_n) \le \frac{\sin(x_n)}{x_n} \le 1$$

holds for all n. Now $\cos(x_n) \to \cos(0) = 1$ thanks to the continuity of $\cos(x)$. So by the squeeze thm for sequences, we get that $\lim_{n\to\infty} \frac{\sin(x_n)}{x_n} = 1$. But x_n was arbitrary. So by the Bridge Thm, $\lim_{x\to 0} \frac{\sin(x)}{x}$ exists and equals 1.

Intermediate value Theorem (Section 3.G of Bowman's notes)

Here's a pretty cool theorem:

Intermediate Value Theorem. Let $f(x)$ be continuous on the interval $[a, b]$. Choose any L between $f(a)$ and $f(b)$. Then there is some c between a and b such that $f(c) = L$.

The idea is that a continuous function can be graphed without your pen leaving the paper. If you start drawing the graph at $x = a$ at $f(a)$, and finish drawing the graph at $x = b$ at $f(b)$, then somewhere in between a and b you're going to have to cross the value L, at least once. You don't know exactly where, but it will be somewhere between a and b.

For example, consider a polynomial like $p(x) = x^{45} + 23x^{40} - 200x^{33} + \pi x^{28} - \sqrt{101}x^{16} + 13$. Then as $x \to \infty$, $p(x) \to \infty$, so for sufficiently large $x > 0$, $p(x)$ is positive. And as $x \to -\infty$, $p(x) \to -\infty$, so for sufficiently negative $x < 0$, $p(x)$ is negative. So take a to be really negative, like -10000000 or something, and take b to be really positive, like $+10000000$ or something. Then $p(a) < 0 < p(b)$. Now, any polynomial is continuous everywhere, so in particular it will be continuous on $[a, b]$. So by the Intermediate Value Theorem, $p(x)$ must equal 0 somewhere between $x = a$ and $x = b$. In fact we can do even better: note that $p(0) = 13$, $p(1)$ is a number around -200 (more precisely, $p(1) = -169.9082829...$), $p(2)$ is massively positive (more precisely, $p(2)$ is around 10^{14} or so) and $p(-2)$ is massively negative (more precisely, $p(-2)$ is around -10^{13} or so). So the Intermediate Value Theorem tells us $p(x)$ has a root somewhere between -2 and 0 (since $L = 0$ is between $p(-2) = -10^{13}$ and $p(0) = 13$), and another root somewhere between 0 and 1 (since $L = 0$ is between $p(0) = 13$ and $p(1) = -169$), and another root somewhere between 1 and 2 (since $L = 0$ is between $p(1) = -169$ and $p(2) = 10^{14}$). Estimates of these roots are $-1.9064800422...$, $.9171041...$, and $1.32946314...$

We see that the only thing special about this polynomial is that it has odd degree. So the Intermediate Value Theorem tells us that any polynomial of odd degree will have at least 1 root. ("root" means a value of x where the polynomial equals 0). Note that this isn't true for even degree polynomials: some can have no roots (e.g. $x^2 + 1$).

Proof of Intermediate Value Theorem. If $f(a) = L$ or $f(b) = L$, then we're done (take $c = a$ or $c = b$). So *we can assume without loss of generality* that $f(a) < L < f(b)$ (the other possibility, namely $f(a) > L > f(b)$, is proved the same way).

Put $m_1 = (a + b)/2$. If $f(m_1) = L$, then we're done (take $c = m_1$), so we can ignore this possibility. There are two other possibilities:

(i) $f(m_1) > L$: put $a_1 = a$, $b_1 = m_1$;

(ii) $f(m_1) < L$: put $a_1 = m_1$, $b_1 = b$.

In either case we have $f(a_1) < L < f(b_1)$ and $b_1 - a_1 = (b - a)/2$.

Now put $m_2 = (a_1 + b_1)/2$. If $f(m_2) = L$, then we're done (take $c = m_2$), so we can ignore this possibility. There are two other possibilities:

(i) $f(m_2) > L$: put $a_2 = a$, $b_2 = m_2$;

(ii) $f(m_2) < L$: put $a_2 = m_2$, $b_2 = b$.

In either case we have $f(a_2) < L < f(b_2)$ and $b_2 - a_2 = (b_1 - a_1)/2 = (b - a)/4$.

Continue like this: we get sequences a_n and b_n such that both $f(a_n) < L < f(b_n)$ and $b_n - a_n = (b - a)/2^n$. Now, a_n is monotonic increasing, and bounded above by b. So $\lim_{n\to\infty} a_n$ is some finite number $c \leq b$. And b_n is monotonic decreasing, and bounded below by a. So $\lim_{n\to\infty} b_n$ is some finite number $c' \geq a$. But $b_n - a_n \to 0$, so $c = c'$, and $a \leq c \leq b$. By continuity, $f(a_n), f(b_n) \to f(c)$. By the Squeeze Theorem applied to $f(a_n) < L < f(b_n)$, we get $f(c) = L$, and we're done. QED

Note that this fails over $\mathbb{Q}$. To see this, consider $f(x) = x^2 - 2$. This is continuous, and $f(1) = -1$ and $f(2) = 2$, so the Intermediate Value Theorem tells us it should have a root somewhere between 1 and 2, and over the reals we know what it is: $\sqrt{2} = 1.4142....$ But if we just worked over the rational numbers, the graph of $f(x)$ would cross the x-axis without ever touching it. It would squeeze between a gap in the rational numbers. The Completeness Axiom says that $\mathbb{R}$, unlike $\mathbb{Q}$, has no gaps.

Another application of Intermediate Value Theorem are antipodal points on the earth: these are the points on the opposite side, so draw a line through the point and centre of earth, and where it intersects the surface of the earth on the other side is called the antipodal point. E.g. the antipodal point to Edmonton is in the southern Indian ocean nearish to Antarctica.

The Intermediate Value Theorem says that there are antipodal points on the Equator with exactly the same temperature. To see this, define a function $f(x)$ to be the temperature on equator at x° longitude, minus the tempurature at the equator at $(x + 180)^\circ$ longitude (its antipodal point). E.g. 0° longitude on equator is in the Atlantic ocean just south of Ghana (which at this moment has a temperature of 29° C), and the antipodal point is somewhere near Fiji in Pacific ocean (which at this moment has a temperature of 24° C), so $f(0) = 29 - 24 = 5$. Likewise, Quito is at longitude $x = -78$ and temperature 16°, and its antipodal point is somewhere in Indonesia with temperature 26°, so $f(-78) = 16 - 26 = -10$. So somewhere between Quito ($x = -78$) and Ghana ($x = 0$) will be a point with $f(x) = 0$. Even if we didn't know that $f(x)$ is positive in Ghana and negative at Quito, we would know that f at say Quito will be the negative of f at Indonesia, so somewhere between $x = -78$ and $x = 180 - 78$ will have $f(x) = 0$.

Maximum principle (Section 3.G in Bowman's Notes)

We say a function is *bounded* on a subset S of its domain, if there is some number $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in S$. So the same M works for all $x \in S$. For example, $\sin(x)$ is bounded by $M = 1$ on $S = \mathbb{R}$. $f(x) = x^2$ is bounded on $S = (-10, 10)$ by $M = 100$ (it's also bounded there by $M = 200$ or any other number $M > 100$). If f is bounded on S, then both the supremum $\sup\{f(x) : x \in S\}$ and infimum $\inf\{f(x) : x \in S\}$ are finite.

We say f *attains its supremum* on S, if there is some $c \in S$ such that $f(c) = \sup\{f(x) : x \in S\}$. We say f *attains its infimum* on S, if there is some $d \in S$ such that $f(d) = \inf\{f(x) : x \in S\}$.

Maximum Principle. Every continuous function $f(x)$ on a closed interval $[a, b]$ is bounded. $f(x)$ will attain both its supremum and infimum on $[a, b]$.

Proof. Let $L = \sup\{f(x) : x \in S\}$ (one thing we need to show is that $L < \infty$). Let $z_n < L$ be any sequence tending to L (from the left). Then by definition of supremum, there will be $x_n \in [a, b]$ such that $z_n < x_n \leq L$. Hence $f(x_n) \to L$ by the Squeeze Theorem. But the Bolzano–Weierstrass Theorem tells us there must be a subsequence x_{n_k} which converges to some point $c \in [a, b]$. Then $f(x_{n_k}) \to L$ (because $f(x_n) \to L$), but continuity of f means that $f(x_{n_k}) \to f(c)$ (since $x_{n_k} \to c$). Hence $f(c) = L$. This means two things: first of all, $L < \infty$ (since it equals the number $f(c)$), and $f(x)$ attains its supremum.

The proof for infimum is identical. QED

This is a combination of Theorems 3.3 and 3.4 in Bowman's notes. We need a closed interval here. For example, $f(x) = 1/x$ is continuous on $(0, 1]$, but it is not bounded there (the supremum is infinite).

Continuity of inverse functions (Section 4.I in Bowman's Notes)

Recall: Let $f : D \to R$ be a function. D is called the domain, and R the codomain. The range is the set of all images $\{f(x) : x \in D\}$. Officially, a function consists of 3 things: a domain, a codomain, and a formula (or a rule for assigning to each point x in the domain, a point $f(x)$ in the codomain.

We say $f : D \to R$ is *one-to-one* (or *injective*) if whenever $a, b \in D$ and $a \neq b$, then $f(a) \neq f(b)$. So different points get sent to different points. You can also say this like this: every point in codomain comes from at most one point in domain.

For example, $f(x) = x^2$ is not one-to-one if we take its domain to be $\mathbb{R}$: e.g. $f(-2) = 4 = f(2)$. However, if we restrict the domain to be e.g. $[0, \infty)$, then it is one-to-one.

We say $f : D \to R$ is *onto* (or *surjective*) if range=codomain. So any point in the codomain comes from at least one point in the domain.

You can always make a function one-to-one, by restricting to an interval where it is increasing (or decreasing). You can always make a function onto by replacing its codomain with its range.

For example, $f(x) = x^2$, when $f : \mathbb{R} \to \mathbb{R}$, is not onto (because x^2 can never be negative). However, changing the codomain to $[0, \infty)$, this f becomes onto. It is not one-to-one, if the domain is all of $\mathbb{R}$, since e.g. $f(-1) = f(1)$. However if we restrict the domain to e.g. $[0, \infty)$, it becomes one-to-one.

An *inverse* to $f : D \to R$ is a function $g : R \to D$ with the property that

$f \circ g : R \to R$ is the identity $y \mapsto y$ on R, and $g \circ f : D \to D$ is the identity $x \mapsto x$ on D. Most functions don't have an inverse; if it does, the function is called *invertible*. A function has at most one inverse; we denote the inverse by f^{-1} when it exists.

Don't confuse f^{-1} with the reciprocal $1/f$.

For example, $f : [0, \infty) \to [0, \infty)$, where $f(x) = x^2$, is invertible with inverse $f^{-1}(y) = \sqrt{y}$.

A function is invertible, iff it is both one-to-one and onto. Try to see why this is. (One direction: If a function $f : D \to R$ is both one-to-one and onto, then every point $y \in R$ comes from one and only one $x \in D$, so we write $f^{-1}(y) = x$).

Assume $f(x)$ is invertible. Then the graph of f^{-1} is just the graph of f, with the x-axis reflected to the y-axis: the graph of f^{-1} is the mirror image through the diagonal line $y = x$ of the graph of f. Now, intuitively, a continuous function has a graph that has no jumps, so it can be drawn without your pencil leaving the paper; if this holds for the graph of f, then it should also hold for the graph of f^{-1}. In other words, if f is continuous, so should be f^{-1}:

Continuity of inverse Theorem. Suppose $f : [a, b] \to [c, d]$ is continuous and invertible. Then $f^{-1} : [c, d] \to [a, b]$ is continuous.

Proof. Let $y_n \to L$ in $[c, d]$. Put $x_n = f^{-1}(y_n)$. We want to show $x_n \to f^{-1}(L)$ (by the Happy Corollary, this would show f^{-1} is continuous, as desired).

Suppose for contradiction that x_n does not converge to $f^{-1}(L)$. That means there is some subsequence x_{n_k} which stays far from $f^{-1}(L)$: that is, some $\varepsilon_0 > 0$ such that for all k, $|x_{n_k} - f^{-1}(L)| > \varepsilon_0$. Now, Bolzano-Weierstrass says that this subsequence x_{n_k} has itself a subsequence $x_{n_{k_l}}$ which converges to some $w \in [a, b]$. Now, continuity of f tells us that $f(x_{n_{k_l}}) \to f(w)$. But $x_{n_{k_l}} = f^{-1}(y_{n_{k_l}})$ by definition of the x_n, so $f(x_{n_{k_l}}) = f(f^{-1}(y_{n_{k_l}})) = y_{n_{k_l}}$. And $y_n \to L$, so $y_{n_{k_l}} \to L$. Hence $f(x_{n_{k_l}}) \to L$. Therefore $f(w) = L$, or equivalently $w = f^{-1}(L)$.

But all x_{n_k}, hence all $x_{n_{k_l}}$, are at least ε_0 from $f^{-1}(L)$, so $x_{n_{k_l}}$ cannot tend to $f^{-1}(L)$. This contradiction means that x_n must converge to $f^{-1}(L)$. QED

There is nothing special about the domain being a closed interval. It also holds if it is say an open interval like (a, b), or half-open, half-closed, like

$(a, b]$. The easiest way to see this is to approximate say (a, b) with closed intervals $[a', b'] \subset (a, b)$ for $a' \to a$, $b' \to b$. The fact that the codomain is an interval is because of the Intermediate Value Theorem: if the domain of a continuous invertible function is e.g. an open interval, then so is the range.

There is a standard way to get invertible functions. We'll show shortly that if f is a strictly increasing (or strictly decreasing) function, then it is one-to-one and hence is invertible if the codomain is chosen to be the range. ($f(x)$ *strictly increasing* means $x < x'$ implies $f(x) < f(x')$; $f(x)$*strictly decreasing* means $x < x'$ implies $f(x) > f(x')$.)

For example, $f(x) = x^2$ is strictly increasing on $[0, \infty)$. Then the inverse $f^{-1}(y) = \sqrt{y}$ is continuous. We already knew this.

But $f(x) = x^n$ is also strictly increasing on $[0, \infty)$ for any $n \in \mathbb{N}$, so its inverse $f^{-1}(y) = y^{1/n}$ will be continuous.

$f(x) = \sin(x)$ is increasing on $[-\pi/2, \pi/2]$ with range $[-1, 1]$. So the inverse $\sin^{-1}(y) = \arcsin(y)$, from $[-1, 1] \to [-\pi/2, \pi/2]$, is continuous.

Theorem. Let $f(x)$ be a continuous function with domain $[a, b]$. Write $M = \sup_{x \in [a,b]}\{f(x)\}$ and $m = \inf_{x \in [a,b]}\{f(x)\}$. Then $f : [a, b] \to [m, M]$. If $f(x)$ has an inverse, then $f(x)$ is either strictly increasing (so $m = f(a)$ and $M = f(b)$), or $f(x)$ is strictly decreasing (so $m = f(a)$ and $m = f(b)$).

Proof. By the Maximum Principle, we know that $f : [a, b] \to [m, M]$.

Now assume f^{-1} exists. Suppose for contradiction that $f(x)$ is neither strictly increasing nor strictly decreasing. Then there are $x_1, x_2, x_3 \in [a, b]$ such that $x_1 < x_2 < x_3$ and either $f(x_1) < f(x_2) > f(x_3)$, or $f(x_1) > f(x_2) < f(x_3)$ (why is this true?)

Without loss of generality, assume $f(x_1) < f(x_2) > f(x_3)$. f must be one-to-one, so $f(x_1) \neq f(x_3)$, so either $f(x_1) > f(x_3)$ or $f(x_1) < f(x_3)$. If $f(x_1) > f(x_3)$ then we have $f(x_3) < f(x_1) < f(x_2)$. By the Intermediate Value Theorem, there must exist some $c \in (x_2, x_3)$ such that $f(c) = f(x_1)$. This contradicts f being one-to-one. If instead we have $f(x_1) < f(x_3)$, you work out how to modify this argument to get a similar contradiction. These contradictions mean that f must be strictly increasing or strictly decreasing. QED

Theorem. Let $f(x)$ be a continuous function with domain $[a, b]$, which is strictly increasing. Then $f : [a, b] \to [f(a), f(b)]$ is invertible.

Proof. Choose any $x, x' \in [a, b]$, $x \neq x'$. Assume without loss of generality that $x < x'$. Then $f(x) < f(x')$. In particular, $f(x) \neq f(x')$. So f is one-to-one.

For any $a < x < b$, $f(a) < f(x) < f(b)$. This means that the range of f is a subset of $[f(a), f(b)]$. Now choose any $y \in [f(a), f(b)]$. By the Intermediate Value Theorem, there is some $x \in [a, b]$ such that $f(x) = y$. This means that $[f(a), f(b)]$ is contained in the range of f. Together, we get that the range of f equals $[f(a), f(b)]$.

But a one-to-one onto function is necessarily invertible. So $f : [a, b] \to [f(a), f(b)]$ is invertible. QED

There is nothing special about a closed interval here: e.g. (a, b) would also work, and the range will also be an open interval, though we can't write it as $(f(a), f(b))$.

There is nothing special about strictly increasing in this last theorem; strictly decreasing would work as well.

If the function is neither strictly increasing nor strictly decreasing, say it increases to a point and then decreases, then it won't be one-to-one so won't be invertible.

Some sci-fi

We're now at the end of Chapter 3 in Bowman's Notes, and the end of our course. Here are two indications that continuity is weirder than you'd think.

By a *curve* in $\mathbb{R}^2$, we mean a map $t \mapsto (x(t), y(t))$ from $[0, 1]$ to $\mathbb{R}^2$, where $x(t)$ and $y(t)$ are continuous. We think of t as the time, and $x(t)$ and $y(t)$ are the x, y-coordinates of a particle that's zipping around. There's nothing special about curves in $\mathbb{R}^2$; e.g. curves in $\mathbb{R}^3$ would also have a $z(t)$ coordinate. There is nothing special about $[0, 1]$; the time t can come from any interval in $\mathbb{R}$.

E.g., the graph of any continuous function $f(x)$ is a curve $t \mapsto (t, f(t))$.

The weird thing is that it is possible to have *space-filling curves*. E.g. there are continuous curves $[0, 1]$ *onto* the unit square $[0, 1] \times [0, 1]$. So the (solid) unit square consists of all points (x, y) where $0 \leq x \leq 1$, $0 \leq y \leq 1$.

The first person to do this was Peano. Here is his curve, now called the *Peano curve*. Write $t \in [0, 1]$ in base 3: so $t = 0.t_1t_2t_3...$ where each $t_i \in \{0, 1, 2\}$, so $t = \sum_{i=1}^{\infty} 3^{-i} t_i$. For example, $7/9 = 0.21$ and $1/2 = 0.11111....$

Define

$$x(t) = 0.t_1\, s^{t_2}(t_3)\, s^{t_2+t_4}(t_5)... , \quad y(t) = 0.s^{t_1}(t_2)\, s^{t_1+t_3}(t_4)....$$

where $s^{odd}(t_j) = 2 - t_j$ and $s^{even}(t_j) = t_j$. Then the curve is $t \mapsto (x(t), y(t))$.

For example, $x(7/9) = 0.222... = 1, y(7/9) = 0.1 = 1/3$ so $7/9 \mapsto (1, 1/3)$. Likewise, $5/9 \mapsto (1/3, 1/3)$ and $1/2 \mapsto (1/2, 1/2)$.

- The interval $0 \leq t \leq 1/9$ maps *onto* the subsquare $[0, 1/3] \times [0, 1/3]$,
- the interval $1/9 \leq t \leq 2/9$ maps *onto* the subsquare $[0, 1/3] \times [1/3, 2/3]$,
- the interval $2/9 \leq t \leq 3/9$ maps *onto* the subsquare $[0, 1/3] \times [2/3, 1]$,
- the interval $3/9 \leq t \leq 4/9$ maps *onto* the subsquare $[1/3, 2/3] \times [2/3, 1]$,
- the interval $4/9 \leq t \leq 5/9$ maps *onto* the subsquare $[1/3, 2/3] \times [1/3, 2/3]$,
- the interval $5/9 \leq t \leq 6/9$ maps *onto* the subsquare $[1/3, 2/3] \times [0, 1/3]$,
- the interval $6/9 \leq t \leq 7/9$ maps *onto* the subsquare $[2/3, 1] \times [0, 1/3]$,
- the interval $7/9 \leq t \leq 8/9$ maps *onto* the subsquare $[2/3, 1] \times [1/3, 2/3]$,
- and the interval $8/9 \leq t \leq 1$ maps *onto* the subsquare $[2/3, 1] \times [2/3, 1]$.

Here's a continuous function $f : [0, 1] \to [0, 1]$, called the *Devil's staircase.* It also uses base 3. To find out what $f(x)$ equals, follow these steps:

Step 1: Write $x = 0.x_1x_2x_3...$, the base 3 expansion.

Step 2: Let x_i be the first digit to equal 1 (if no digits equal 1, proceed to step 3). Change x_i to 2, and make $x_{i+1}, x_{i+2}, ...$ all equal 0.

Step 3: Now every digit equals 0's or 2's. Change all 2's to 1's. So now all digits equal 0's or 1's. Interpret this number as a base 2 expansion. The value of that base 2 expansion is $f(x)$.

For example, for $1/3 \leq x \leq 2/3$, $x_1 = 1$, so Step 2 changes x to 0.2, and Step 3 changes this to 0.1, which in base 2 equals 1/2. So $f(x) = 1/2$.

Likewise, for $1/9 \leq x \leq 2/9$, $f(x) = 1/4$. For $7/9 \leq x \leq 8/9$, $f(x) = 3/4$.

It isn't hard to show that this is continuous: if x and x' are close, then they'll agree to the first several base-3 digits x_i, so Step 2 will also agree to the first several base-3 digits, so the final answer will also agree to the first several base-2 digits, so $f(x)$ and $f(x')$ will also be close.

Note that $f(0) = 0$ and $f(1) = 1$. Note that lots of x's are on flat parts (=constant parts) of the graph. As we said before, the graph is flat for $x \in [1/3, 2/3]$ ($f(x) = 1/2$ there). And its flat on [1/9,2/9] and [7/9,8/9]. And it's flat on [1/27,2/27], [7/27,8/27], [19/27,20/27] and [25/27,26/27]. The total length of intervals its flat on is

$$1/3+(1/9+1/9)+(1/27+1/27+1/27+1/27)+\cdots = 1/3+2/9+4/27+8/81+\cdots$$

$$= \frac{1}{3}\sum_{i=0}^{\infty}(2/3)^i = 1$$

So 100% of the $x \in [0,1]$ lie on parts of the graph which is flat. Indeed, the graph is flat (horizontal) at any x whose base-3 expansion has a 1. And almost every x will eventually have an $x_i = 1$.

So the Devil's staircase is an increasing continuous function from 0 to 1, which is horizontal 100% of the time. If you randomly choose a spot on the function, it will be horizontal there. Nevetheless it somehow gets from 0 to 1, without any gaps.

Infinity in geometry

Write a fraction a/b as $[a,b]$. We know that $(an)/(bn) = a/b$, so we say $[an, bn] = [a,b]$ (at least when $n \neq 0$). Then $[a,b] = [a/b, 1]$ when $b \neq 0$. We say $\infty = [1,0]$. Note that $[-1,0] = [1,0]$, so $-\infty = \infty$.

We can visualize this by saying that we attach 1 point, at infinity, to the real number line. So the real number line is a circle. If you go far enough in the positive direction, you wrap around and eventually become negative!

(For those of you who know about complex numbers, there is similarly one point at infinity that you add to the complex plane. You can get there by heading off in any direction. So the complex plane is really a sphere!)

Now how does this look in 2-dimensions, for $\mathbb{R}^2$? Well, think of points as $[x,y,z]$, where we identify $[x,y,z] = [xn, yn, zn]$. So as long as $z \neq 0$, $[x,y,z] = [x/z, y/z, 1]$. We think of this as the usual boring finite point $(x/z, y/z)$. The infinite points have $z = 0$. So they are $[x,y,0]$. These are either $[x/y, 1, 0]$ (when $y \neq 0$), or $[1, 0, 0]$. So the infinite points for $\mathbb{R}^2$ form a real number circle! You get a different infinite point for each direction.

You can see curves at infinity. Consider the parabola $y = x^2$. How many infinite points does it have? To do this, add a z variable, so that each term has degree 2. The way to do this is $yz = x^2$. The boring finite points are $z = 1$. The infinite points have $z = 0$, i.e. $y0 = x^2$, i.e. $x = 0$, i.e. $[0,y,0] = [0,1,0]$. So there is 1 infinite point, in the direction of the y-axis. The parabola is a stretched out ellipse.

Consider next the hyperbola $y = 1/x$. This becomes $xy = z^2$. As you can see, this has 2 infinite points, so this is also just a stretched out ellipse.

From Week 13, the important things are the continuity of $\sin(x), \cos(x)$; $\lim\limits_{x \to 0} \frac{\sin(x)}{x} = 1$; Intermediate Value Thm; Maximum Principle; and inverse functions. The geometry at infinity, and Peano curves, and Devil's staircase, are included just for fun.