Week 1 summary

Wednesday: Discussed course, and honours vrs nonhonours math

Thursday:

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 $\{2,3\} \subseteq A$

$\{\}$ 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 Y be the set of all sets which don't contain themselves. Is $Y \in Y$?
(Russell's paradox will not be on any exam)

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:

Definition, Theorem, proof: the grammar of math

A *definition* makes precise some terminology or notation.

- For example, we defined even and odd.

A *theorem* is an unambiguous statement we are declaring to be true

- For example, we stated the Theorem: even plus even is even; even times even is even; odd plus odd is even; odd times odd is odd.

A *proof* of a theorem is a sequence of statements. Each statement either is something we are given in the theorem (e.g. two even numbers), uses definitions to reduce terms to simpler ones, uses the basic properties we are allowed to assume (e.g. properties of natural numbers such as $a(b + c) = ab + ac$), or uses theorems we have already proved.

- For example, we gave the proof of our theorem.

You must get familiar with Definitions, Theorems, proofs.

(Unfortunately this will take most of you all term...)

Week 2 summary

Monday:

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

Elementary logic revisited: more grammar of math

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.

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,
  "products mn and sums $m + n$ of natural numbers are natural numbers"
  is an axiom for us.

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.

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.

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**

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

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

Wednesday:

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** regardless of

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*

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

Thursday:

IF AND ONLY IF

"*Statement1* IF AND ONLY IF *Statement2*"
means the same as
"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*).

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 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.

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

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

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

Real numbers (see Section 1.C of Bowman's Notes)

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.

Friday:

Real number arithmetic: The axioms of a field

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

(AC) $x \oplus y = y \oplus x$ for all $x, y \in F$. (*commutativity of addition*)
(AA) $(x \oplus y) \oplus z = x \oplus (y \oplus z)$ for all $x, y, z \in F$. (*associativity of addition*)
(AN) This is a 'number' $0' \in F$ such that $0' + x = x$ for all $x \in F$. (*additive identity*, or *neutral element for addition*)
(AI) For each $x \in F$, there is some 'number' '$-x \in F$' such that $x \oplus$ '$-x$'$= 0'$. (*additive inverse*)

'Subtraction' is defined to be $x \ominus y = x \oplus$ '$-y$'. Note that ordinary subtraction of numbers is usually neither commutative nor associative.

'Multiplication' satisfies analogous properties:

(MC) $x \odot y = y \odot x$ for all $x, y \in F$. (*commutativity of multiplication*)
(MA) $(x \odot y) \odot z = x \odot (y \odot z)$ for all $x, y, z \in F$. (*associativity of multiplication*)
(MN) This is a 'number' $1' \in F$ such that $1' \odot x = x$ for all $x \in F$. (*multiplicative identity*, or *neutral element for multiplication*)
(MI) For each $x \in F$, except for $x = 0'$, there is some 'number' '$x^{-1} \in F$' such that $x \odot$ 'x^{-1}'$= 1'$. (*multiplicative inverse*)

'Division' is defined to be x'/'$y = x \odot$ 'y^{-1}'. Note that ordinary division of numbers is usually neither commutative nor associative. 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) $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$ for all $x, y, z \in F$.

In class we let N (for **N**umber) denote a field; but it is more common to let F (for **F**ield) denote a field. 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 $x \odot (y \oplus z) = x \odot y \oplus x \odot z$, 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. Of course, for those boring fields, $0' = 0$ and $1' = 1$.

- 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 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.

NOTE: To prove any of these axioms requires showing it for *all* $x, y, z \in F$. To show something is a field, you have to prove all 9 axioms, for every possible value of $x, y, z \in F$. 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 $x, y, z \in F$ 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, $0' = 0/1$ and $1' = 1/1$.

- 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. $\min(\infty, x) = x$.

From Week 2, the important things are to understand the logical words, and understand proof strategies. You don't need to memorize the Field axioms (though these are pretty easy).

Week 3 summary

Monday:

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 gave some examples. 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.

From this point on, we'll use the familiar notation from real number arithmetic. So 'addition' is denoted $x + y$, not $x \oplus y$, and 'multiplication' is denoted xy or $x \cdot y$, not $x \odot y$. The neutral (a.k.a. identity) element for addition will be denoted 0, not $0'$, and the additive inverse will be denoted $-x$, not '$-x$'. The neutral (a.k.a. identity) element for multiplication will be denoted 1, not $1'$, and the multiplicative inverse will be denoted x^{-1}, not 'x^{-1}'. This makes our formulas and Lemmas much more readable! But all of the following hold for any field F. So a useful exercise is to rewrite these in the Week2 notation.

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

A useful notation is '$\forall$'. It means *for all.* For example, $\forall x \in \mathbb{R}$ means "for all $x \in \mathbb{R}$" or "for all real numbers x", which is the same thing.

Lemma F.1. Let F be any field. Then:

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

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

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

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

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

The proofs of this Lemma are just commutativity (**axioms AC, MC**) applied to some other axiom (namely, **AN**, **AI**, **MN**, **MI**, **D**, respectively).

Cancellation Laws. Let F be any field. Then:

(a) $\forall x, y, z \in F,\ x + z = y + z \implies x = y$.

(b) $\forall x, y, z \in F,\ z + x = z + y \implies x = y$.

(c) $\forall x, y, z \in F$ where $z \neq 0$, $xz = yz \implies x = y$.

(d) $\forall x, y, z \in F$ where $z \neq 0$, $zx = zy \implies x = y$.

To prove these, just use the inverse axioms **AI**, **MI** and associativity.

Lemma F.2. Let F be any field. Then:

(a) $\forall x \in F,\ -(-x) = x$.

(b) $\forall x \in F,\ 0x = 0$.

(c) $\forall x \in F$, $-x = -1 \cdot x$.

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

Parts (a) and (d) are proved using **Lemma F.1(b),(d)**, as well as **Axioms AI,MI** which defines what inverses mean (and tell you they exist). Parts (b) and (c) are proved using the trick of replacing x with $1x$ and using distributivity (**axiom D**). In all these proofs, it is convenient to use the Cancellation Laws. (Part (d) was actually stated and proved in Wednesday's class, but it fits best in this Lemma.)

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

Wednesday:

We had a **quiz** (Quiz 2 on eclass) on logic. (Quizzes aren't for marks)

Consequences of the field axioms (continued)

Corollary F.3. Let F be any field. Then:

(a) $\forall x, y \in F$, $(-1)^2 = 1$.

(b) $\forall x, y \in F$, $-(x+y) = -x - y$.

Both of these are proved by using **Lemma F.2(c)**. For part (a), apply this to $-(-1) = 1$ (**Lemma F.2(a)**). For part (b), also use distributivity **Axiom D**: $-1 \cdot (x+y) = -1 \cdot x + -1 \cdot y$.

In class, we called this a **Lemma** (which means 'Little Theorem'), but a better name for this would be **Corollary**, which means 'Consequence'. Both these parts are pretty easy consequences of **Lemma F.2(c)**. But whether you call something 'Lemma', 'Theorem', 'Corollary', or something else, isn't important. In class this Lemma was slightly different: class included what we now call **Lemma F.2(d)**, and didn't include part (b). I'm including part (b) here because it was convenient later in the week.

In **Lemma F.3(a)**, we are using the square-notation. More generally, 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 the assignment due this week (Assignment 2). Using **Lemma F.2(b),(c)**, we can define *integer multiples* nx for any field as follows. Notation here is a little ambiguous, so we'll go back to the old field notation of Week2. We write $1x = x$ (this is reminiscent of, but different from, **Axiom AN** which says $1' \cdot x = x$); then for any $n \in \mathbb{N}$, $(n+1)x = x + nx$. For instance we have $2x = x + x$, which is reminiscent of, but different from, the Assignment 2 question which asks you to prove $2' \cdot x = x + x$. We define $0x = 0'$ and $-1x = -x$ (these are reminiscent of, but different from, **Lemma F.2(b),(c)** which in this more careful notation actually say $0' \cdot x = 0'$ and $(-1') \cdot x = -x$); for any $n \in \mathbb{N}$, we define $(-n-1)x = -x + (-n)x$. But as on Assignment 2, we can sort of fit the integers inside any field by defining n', for any integer n, to be $n1'$. This matches our old notation for $0', 1', -1'$, and now $2' = 1' + 1'$ as on Assignment 2, etc. You can try to prove $(mx) + (nx) = (m+n)x$ and $m(nx) = (mn)x$.

- For example, in the **even, odd** field, $\mathbf{odd}^n = \mathbf{odd}$ for any $n \in \mathbb{Z}$. Moreover, $n\,\mathbf{odd} = \mathbf{even}$ when n is even, and $n\,\mathbf{odd} = \mathbf{odd}$ when n is odd. Note that in this **even,odd** field, $2' = 0' = \mathbf{even}$. 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.

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 114, 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 for the past 3 classes 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 F be any field. F is called an *ordered field* if it has a subset $P \subseteq F$ (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 $x \in F$, one and only one of the following will occur:

$$x \in P \quad -x \in P \quad x = 0$$

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.

Really, we should write these using the weird notation of Week2: **OA** says that if $x, y \in P$, then $x \oplus y \in P$; **OM** says that if $x, y \in P$, then $x \odot y \in P$; **OT** says that either $x \in P$, '$-x$' $\in P$, or $x = 0'$. But whenever we can, we'll continue to use the notation you're familiar with, from the real numbers.

Thursday:

Examples of ordered fields

Definition. Let F be an ordered field. We write $x > y$ when $x - y \in P$. We write $x < y$ when $y > x$. We write $x \geq y$ when either $x > y$ or $x = y$. We write $x \leq y$ when either $y \geq x$.

Note that $x \in P$ iff $x > 0$. Also, x is negative iff $x < 0$.

Really, we should write these using weird notation $x \prec y$, $x \succeq y$ etc, because order might be weird. But as much as possible, we'll use real number notation.

- Boring examples of ordered fields: $\mathbb{R}$ and $\mathbb{Q}$ with the usual notion of positive.
- The **even,odd** field $F = \{\mathbf{even}, \mathbf{odd}\}$ is **NOT** an ordered field.

Proof that $F = \{\mathbf{even}, \mathbf{odd}\}$ *is not ordered:* The **even,odd** field F is either ordered or it is not. Let's investigate the possibility that it 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**. So assuming that F is ordered got us into trouble: we got a contradiction. This contradiction eliminates the possibility that F is ordered. The only other possibility is that F is not ordered. So we know that F cannot be ordered. *QED*

This illustrates an important proof strategy, that's given the sexy name *proof-by-contradiction.* It's a bit convoluted, but you'll get used to it. Actually, the basic idea is really simple. We discussed it in detail in the Friday class (even though we used it earlier), so look below in the Friday section below to see it explained.

- The complex numbers $\mathbb{C}$ are a very important field which includes the real numbers, as well as $\sqrt{-1}$, and all combinations thereof. We'll discuss $\mathbb{C}$ a little later in the course. For those of you who already know about $\mathbb{C}$, it is **NOT** an ordered field. This is the only way $\mathbb{R}$ is better than $\mathbb{C}$. You should be able to come up with your own proof. Here is one.

proof that $\mathbb{C}$ *is not ordered:* Suppose for contradiction that $\mathbb{C}$ is ordered. In $\mathbb{C}$, -1 has two square-roots, namely $\pm\sqrt{-1}$. By **Axiom OT**, one of these is 'positive' and the other is 'negative'. Let $s\sqrt{-1}$ be the positive one, for some choice $s = \pm$ of sign. By **OM**, $(s\sqrt{-1})^2$ would also be positive, but $(s\sqrt{-1})^2 = -1$, so that means -1 would be positive. But then (again by **Axiom OM**) $(-1)^2$ would also be positive. But $(-1)^2 = 1$. So putting all this together, we have derived that both -1 is positive, and that 1 is 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 Friday of Week2. 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_i 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 a_n 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 $x \succ 1$. Similarly, $x \succ 10000000$. 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 coeffients equal to $(-1)/1$, so $1/x - 1$ is negative, so $1/x \prec 1$. Similarly, $1/x \prec 0.000000001$ and in fact $1/x$ is smaller than every positive real number. But $1/x \succ 0$. So $1/x$ is like an infinitesimal!

Friday:

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 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 examples of the *proof-by-contradiction* proof strategy earlier in the course, and we'll see lots more 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 $\sqrt{2} = a/b$. So try to see what properties a, b must have for that equation to hold.

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

Lemma OF.1. Let F be any ordered field. Then:

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

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

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

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

To prove parts (a), (c), and (d), just convert from the definition of $x > y$ etc to $x - y \in P$ etc and add, using **Axiom OA**. Our proof of part (b) used a proof-by-contradiction.

From Week 3, the important things are to be familiar with the field axioms and the order axioms, as well as the proof-by-contradiction proof strategy. You should also be familiar with our Lemmas, and how they're proved. But you don't have to memorize the exact statements of Lemmas and their proofs.

Week 4 summary

Monday:

Consequences of the ordered field axioms (continued): Section 1.C of **Bowman's Notes**

Recall the definition of a field, the 9 axioms **AC** to **D** given in the Week2 notes. An **ordered field** obeys those axioms, together with Axioms **OA, OM, OT** of Week3. We will use the familiar notation for addition, multiplication, neutral elements for addition and multiplication, additive and multiplicative inverses, and inequalities. The most important example of an ordered field is the real numbers $\mathbb{R}$.

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 $a > b > 0$ and $c < 0$, then $ca < cb$.
- **(c)** If $a \neq 0$, then $a^2 > 0$.
- **(d)** $1 > 0$
- **(e)** If $a > 0$, then $a^{-1} > 0$
- **(f)** If $a > b > 0$, then $a^{-1} < b^{-1}$.
- **(g)** $2 > 0$.

The proof of **(a)** uses the definition of '$>$' and Axiom **OM**. The proof of **(c)** uses Axiom **OT** and $(-1)^2 = 1$. **(d)** uses **(c)** and $1^2 = 1$.

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 $b^{-1} > a^{-1}$.

To prove **(g)**, recall that 2 is defined (on Assignment 2) to be $1 + 1$. We know from **(d)** that $1 > 0$, so Axiom **OA** implies $1 + 1 > 0$, and we're done. QED to **Lemma OF.2**.

In class we didn't put part **(g)** into **Lemma OF.2**, but rather included it in the proof of **Lemma OF.3**. Recall that in some fields, 2 can equal 0, which can never be positive. Those field obviously cannot be ordered.

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

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

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

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

$$(a, b) = \{x \in F \,|\, a < x < b$$
$$(a, b] = \{x \in F \,|\, a < x \leq b$$
$$[a, b) = \{x \in F \,|\, a \leq x < b$$

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

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. The intervals $[a, b)$ and $(a, b]$ are called *half-open*, for obvious reasons.

Lemma OF.3. Let F be an ordered field. Let $a, b \in F$, 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, $(a+b)/2 > a$ and $(a+b)/2 < b$.

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

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

For $x \in F$, 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 $\mathbb{R}$, $|2| = 2$, $|0| = 0$, $|-3| = 3$. In the ordered field of rational functions $p(x)/q(x)$ discussed in Week3, $|-x^2 + 1/x| = x^2 - 1/x$. **But be careful with the order on the rational function field – we will use it only in silly examples. The order on $\mathbb{R}$ is the one we will use throughout our course.**

We can think of absolute value as meaning how far we are from 0. Here are its basic properties:

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

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

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

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

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

(AV5) Assume $y \geq 0$. Then $|x| \leq y$ iff $x \in [-y, y]$. Also, $|x| \leq y$ iff both $x \leq y$ and $-x \leq y$. Also, $|x| < y$ iff $x \in (-y, y)$. Also, $|x| < y$ iff $x < y$ and $-x < y$.

(AV6) $x \in [-|x|, |x|]$.

(AV7) $\big||x| - |y|\big| \leq |x + y| \leq |x| + |y|$ and $\big||x| - |y|\big| \leq |x - y| \leq |x| + |y|$

All of these are pretty easy to prove, except **(AV7)**, which is called the **triangle inequality**. It is a bit tedious to prove. For example, we always have $x \leq |x|$ and $-x \leq |x|$ and $y \leq |y|$ and $-y \leq |y|$, so adding these inequalities, we get $x - y \leq |x| + |y|$ and $-x + y \leq |x| + |y|$, which together imply $|x - y| \leq |x| + |y|$ (we're using use **AV5)** here).

Applying this to $x = (x - y) + y$ and $y = (y - x) + x$ gives $|x| \leq |x - y| + |y|$ and $|y| \leq |y - x| + |x|$; rearranging the first gives $|x| - |y| \leq |x - y|$ and rearranging the second gives $-|x| + |y| \leq |x - y|$, so $\big||x| - |y|\big| \leq |x - y|$ (again using **AV5)**). The other inequalities come by replacing y with $-y$ in the ones we just proved.

Wednesday:

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 F be an ordered field. Let $S \subseteq F$ be any set of numbers in F. By an *upper bound* for S, we mean any number $u \in F$ which is $\geq$ any number in S: $u \geq x \ \forall x \in F$. By a *lower bound* for S, we mean any number $\ell \in F$ which is $\leq$ any number in S: $\ell \leq x \ \forall x \in F$. 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, we say S is *bounded above*; otherwise we say S is *unbounded above*. If S has at least one lower bound, we say S is *bounded below*; otherwise we say S is *unbounded below*. If S is both bounded above and bounded below, we say S is *bounded*. If S is not bounded, we say it is *unbounded*.

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

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

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

Again, let F be an ordered field. If a set S in 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 $\leq$ 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 $\geq$ any lower bound of S. If S is unbounded above, then we write $\sup S = +\infty$. If S is unbounded below, then we write $\inf S = -\infty$.

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 $F = \mathbb{Q}$. 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 $\mathbb{Z}$ doesn't exist. We will see on friday's class that the inf of $\{x \in \mathbb{Q} \,|\, x^2 > 2\}$ does not exist (in $\mathbb{Q}$).

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.

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

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

If S is not bounded above, then as mentioned above $\sup S = +\infty$.

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

- For example, $\mathbb{Q}$ is not complete. We will see on friday that $S = \{x \in \mathbb{Q} \,|\, x^2 < 2\}$ does not have a supremum (in $\mathbb{Q}$).
- For example, the ordered field of rational functions is not complete. Take S to be the set of all negative real numbers. So 0 is an upper bound. But so is $-1/x$, and $-2/x$, and $-3/x$, etc etc. And $0 > -1/x > -1/x^2 > -1/x^3 > \cdots$. You can convince yourself that S doesn't have a least upper bound.

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

We won't prove this theorem, though on Friday we prove half of it: that any complete ordered field lies inside $\mathbb{R}$. In Math 217 we return to these sorts of questions.

This unique ordered field is called the real numbers $\mathbb{R}$.

Archimedian Property (Bowman's notes 1.J)

Let F be an ordered field. $\mathbb{N}$ fits inside F as follows: certainly $1 \in F$; define $2 = 1+1$ as on Assignment 2; define $3 = 2+1$; $4 = 4+1$; etc. Then by Axiom **OA**, $0 < 1 < 2 < 3 < 4 < \cdots$.

An ordered field F is called *Archimedian* if for every $x \in F$, there is an $n \in \mathbb{N}$ such that $x \leq n$.

- For example, $\mathbb{Q}$ is Archimedian: if $x = p/q$, then $x \leq |p|$.
- For example, the ordered field of rational functions is not Archimedian: x is bigger than any $n \in \mathbb{N}$.

Another way to think about Archimedian fields: if F is Archimedian, then for any $\epsilon > 0$, there exists an $n \in \mathbb{N}$ such that $0 < 1/n < \epsilon$. (Indeed, take $x = 1/\epsilon$)

Lemma C.1. A complete ordered field (in other words, $\mathbb{R}$) is Archimedian.

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

Thursday:

Induction (Bowman's notes 1.E)

Over the past couple weeks, we've stated the 13 axioms defining the real numbers $\mathbb{R}$. Really, we should define carefully everything. Most important now are the axioms of the natural numbers $\mathbb{N}$. Once we know $\mathbb{N}$, we can define the integers $\mathbb{Z}$ using $\mathbb{N}$, and define the rationals $\mathbb{Q}$ using $\mathbb{Z}$.

So what are the axioms for $\mathbb{N}$? Here is one list:

Axiom N.1 $1 \in \mathbb{N}$.

Axiom N.2 If $n \in \mathbb{N}$, so is $n+1$

Axiom N.3 If $m, n \in \mathbb{N}$, and $m+1 = n+1$, then $m = n$

Axiom N.4 There is no $n \in \mathbb{N}$ which has $n+1 = 1$

Axiom N.5: Let $P(n)$ be some statement depending on $n \in \mathbb{N}$.

IP.1 (base case) Suppose that $P(1)$ is true.

IP.2 (induction hypothesis) Suppose that, whenever $P(n)$ is true, then $P(n+1)$ is also true.

Then $\forall n \in \mathbb{N}$, $P(n)$ is true.

Axiom **A.5** is called the **Induction Principle**. Then one defines addition recursively by: if we know what $n+m$ is, then $n+(m+1)$ is defied to be $(n+m)+1$. You can also define multiplication: $n \cdot 1 = n$, and if we know what $n \cdot m$ is, then $n \cdot (m+1) = (n \cdot m) + n$. We can define inequalities by: $n > m$ iff there exists some $a \in \mathbb{N}$ such that $n = m + a$.

There are other choices for axioms; Professor Troitsky likes to include instead

Every nonempty subset $S \subseteq \mathbb{N}$ has a minimum in $\mathbb{N}$.

and then he must prove the Induction Principle. If instead we assume the Induction principle, then you'd have to assume every nonempty $S \subseteq \mathbb{N}$ has a minimum. This isn't hard: Suppose for contradiction that $S \subseteq \mathbb{N}$ is nonempty and has no minimum. Let $P(n)$ be the statement "No $m \leq n$ is in S". Then **IP.1** is true (otherwise 1 would be the minimum). Assume $P(n)$ is true, so no $1 \leq m \leq n$ is in S. Then $P(n+1)$ must also be true (otherwise $n+1$ would be the minimum). Thus **IP.2** is also true. Hence all $P(n)$ are true, hence no n is in S, hence S is empty, a contradiction.

The axioms of $\mathbb{N}$, and how to define addition, multiplication and order, are not important for us. The important thing is the induction principle. It doesn't matter whether you call it an axiom, or prove it from other axioms. The important thing from this class: **The Induction Principle is a very powerful proof strategy, when you have to prove things for all natural numbers.**

Theorem. Let F be any ordered field. Let S be a finite subset of F. Then S has both a maximum and a minimum.

To prove this, let $P(n)$ be the statement that "Any subset $S \subseteq \mathbb{N}$ with size n has a minimum and a maximum."

Theorem. $4^n - 1$ is divisible by three

To prove this, let $P(n)$ be the statement: "$4^n - 1$ is divisible by 3". Certainly true for $n = 1$, so **IP.1** holds. Suppose $P(n)$ is true. Write $4^n - 1 = 3k$. Is $P(n+1)$ true? In other words, is $4^{n+1}-1$ divisible by 3? Well, $4^{n+1}-1 = 4(4^n-1)+3 = 4\cdot 3k+3 = 3(4k+1)$, so indeed 3 divides $3(4k+1) = 4^{n+1}-1$, so $P(n+1)$ is true. Hence by induction, the theorem is true.

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 $1^2+2^2+\cdots+n^2$.

Your midterm 1 exam covers all material up to here
(so up to and including thursday, but not friday)

Friday:

Square-roots in $\mathbb{R}$ and $\mathbb{Q}$

For now on, we can assume $\mathbb{R}$ exists and satisfies all 13 axioms.

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.) Try to do the proof that $\sqrt{3}$ is not rational. How about $\sqrt{4}$?

Corollary. $\mathbb{Q}$ is not complete.

Let $S=\{x\in\mathbb{Q}\,|\,x^2<2\}$. Does S have a supremum in $\mathbb{Q}$? You can show that if it does have a supremum, call it s, then $s^2=2$ (if $s^2<2$, then it wouldn't be an upper bound; if $s^2>2$ then you could find a smaller upper bound). But the Theorem says that there is no rational number s such that $s^2=2$.

The same argument shows:

Theorem. $\sqrt{2} \in \mathbb{R}$.

Of course we 'know' this, but it is nice to know we can prove this from our 13 axioms. The set S given above must have a supremum s, by completeness axiom **C**; it must satisfy $s^2 = 2$, so $s = \pm\sqrt{2}$.

Again, there is nothing special about $\sqrt{2}$. E.g. $\sqrt{3}$ also is a real number. How about $\sqrt{-1}$, why doesn't the same proof say $\sqrt{-1} \in \mathbb{R}$?

More truths about Archimedian fields

(this isn't so important for our course)

This section sketches part of the argument how the only complete field is really $\mathbb{R}$ (though maybe in disguise). Any complete field is Archimedian; **Corollary A.2** below show it fits inside $\mathbb{R}$. More is true: it must equal $\mathbb{R}$, but that is a topic for a different course.

We know that any ordered field contains a copy of the natural numbers. That means it contains a copy if the integers $\mathbb{Z}$ (by including 0 and the additive inverses of $\mathbb{N}$, and hence a copy of $\mathbb{Q}$ (since the multiplicative inverses n^{-1} must exist, so so must the products $n^{-1}\mathbb{Z}$).

Lemma A.1. Let F be an Archimedian (hence ordered) field. Then:

(i) for any $a \in F$ where $a > 0$, the interval $[a, a+1]$ contains a natural number.

(ii) If $a < b$, then the interval (a, b) contains a rational number.

Let S be the set of all $n \in \mathbb{N}$ such that $a \leq n$. By the Archimedian property, S is nonempty. But every nonempty subset of $\mathbb{N}$ has a minimum (see the Thursday discussion). Let m be the minimum of our set S. If $a+1 < m$, then $m-1$ would be in S, contradicting the minimality of m. So $a+1 \geq m$, hence $m \in [a, a+1]$ and we're done.

- Both **(i)**,**(ii)** fail for the ordered field of rational functions. There is no natural number in $[x, x+1]$, and there is no rational number in $(1/x, 2/x)$.

Corollary A.2. Let F be an Archimedian field. Choose any $a \in F$. Then $a = \sup\{x \in \mathbb{Q} \,|\, x \leq a\}$

This shows that the rationals are 'close' to anything in F, i.e. that you can approximate any $a \in F$ arbitrarily closely with rationals. It fails for rational functions: e.g. take $a = 1/x$, then $\{x \in \mathbb{Q} \,|\, x \leq a\} = \{r \in \mathbb{Q} \,|\, r \leq 0\}$ has supremum 0, not $1/x$. So there is a gap between the rationals and the 'infinitesimal' $1/x$.

Corollary A.3. Any Archimedian field sits inside $\mathbb{R}$. Hence any complete ordered field lies in $\mathbb{R}$.

As we know, more is true: any complete ordered field *equals* $\mathbb{R}$.

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. $a_n = f(n)$. Sometimes sequences are also denoted as $\{a_n\}_{n \in \mathbb{N}}$ or $(a_n)_{n \in \mathbb{N}}$.

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 is is always positive, so it never equals 0, but intuitively we can think that 'as n reaches infinity, $1/n$ finally reaches 0'. Next week we make this notion precise: we'll say the limit of $(1/n)_{n\in\mathbb{N}}$ equals 0, and write $\lim_{n\to\infty} 1/n = 0$.

> **From Week 4, the important things** are to be familiar with LemmasOF.2, the interval notation, the definition and basic properties of absolute value, the notion of upper and lower bounds, maximum and minimum, and especially sup and inf, the notion of completeness and that the reals are complete, and the induction proof strategy. It is good to understand the proof that $\sqrt{2}$ is irrational (but not for Midterm 1). You'll need to know about sequences (but not for the midterm). As always, you don't need to memorize the exact statements of Lemmas and their proofs.

Week 5 summary

Monday had Quiz 3 and a midterm review

Wednesday was the midterm

Thursday

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

A *sequence* looks like $a_1, a_2, a_3, a_4, \ldots$. In our course we're most interested in sequences with $a_n \in \mathbb{R}$

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. 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 ns the sequence gets superclose to L.

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 $L = \lim_{n\to\infty} a_n$, 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.

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.

Friday

More limits

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.

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 $|L - a_n| < \epsilon$, for all $n > N'$. Because $c_n \to L$, there is an N'' such that $|L - c_n| < \epsilon$ for all $n > N''$. This implies 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 $|L - b_n| < \epsilon$. There are two possibilities:

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

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

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

From Week 5, the important things are to be familiar with sequences and their limits. Know the definition, what it means, how to prove something converges, how something diverges. The squeeze theorem is a nice way to show things converge. The proof of the squeeze theorem is not important.

Week 6 summary

Monday: Stuff about limits

Last friday we gave the Squeeze Thm. 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_ns, 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. 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.

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 $a_1, a_2, a_3, ...$ is $b_1 = a_{n_1}, b_2 = a_{n_2}, \ldots$, 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 thee are two directions.

$\Longrightarrow$: Assume that a_n converges to L. Let $b_k = n_{n_k}$ 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 $|L - a_n| < \epsilon$ whenever $n > N$.

Now choose any $k > N$. Note that $b_1 \geq 1$, $b_2 \geq 2$, $b_3 \geq 3$,... $b_k \geq k$ for all k. (You can prove this by induction.)

Since $k > N$, then $n_k \geq k > N$, so

$$|L - b_k| = |L - a_{n_k}| < \epsilon$$

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

Wednesday Even more limits

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

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

This is just a rephrasing of the Squeeze 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 $b_k = -\frac{2k+1}{2k}$ (this uses the choice $n_k = 2k - 1$): it clearly has limit -1. These two limits are different, so a_n can't converge to anything, by the Squeeze Corollary.

Example. Consider the sequence

$$1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1.....$$

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

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

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.

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:

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

$$= \frac{(n+1) - n}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}} < \frac{2}{\sqrt{n}}$$

which tends to 0.

Thursday: even more limits

Here's a little technical result we'll need later today. 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 such that $|a_n| < M$ 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 $|L - a_n| < 1$ for all $n > N$. This means $a_n \in (L-1, L+1)$, for all $n > N$. So $|a_n| < |L| + 1$ for all $n > N$.

Let M be the maximum of the numbers $\{|a_1|, |a_2|, \ldots, |a_N|, |L|+1\}$. This is a finite set, so it has a maximum (this is a Theorem in the Week 4 (Thursday) notes. Then obviously $|a_n| < M$ for all n: for $n \leq N$ this is clear from the definition of 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 $\lim_{n\to\infty} a_n = L$ and $\lim_{n\to\infty} b_n = L'$. Then:
(a) $a_n + b_n$ converges to $L + L'$, and $a_n - b_n$ converges to $L - L'$. More generally, for any constants c, d, $c\,a_n + d\,b_n = cL + dL'$.
(b) $a_n b_n$ 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 $n > N$, $L - a_n| < \epsilon/2$. And there is an N' such that for any $n > N'$, $|L' - b_n| < \epsilon/2$. Now let $N'' = \max\{N, N'\}$. Then for any $n > N''$,

$$|(L + L') - (a_n + b_n)| = |(L - a_n) + (L' - b_n)| \leq |L - a_n| + |L' - b_n|$$

by the triangle inequality. We know $n > N$, so $L - a_n| < \epsilon/2$. And we know $n > N'$, so $|L' - b_n| < \epsilon/2$. Putting these together, we get $|(L + L') - (a_n + b_n)| < \epsilon/2 + \epsilon/2 = \epsilon$. 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:

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

For example, $1/n \to 0$, $1/\sqrt{n} \to 0$, etc.

Example.

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

Friday

Infinite limits

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 $n > N$, $a_n > M$.
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 $n > N$, $a_n < -M$.

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

Examples. The limit of n^2 is ∞. The limit of $-\sqrt{n}$ is $-\infty$. The sequence $(-1)n$ does not have a limit.

Most of our theorems have an analogue for sequences with infinite limits. For example:

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

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 $ca_n + db_n$ and a_n/b_n — write these out yourselves.

Example.

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

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 $b = 1$. 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 becomes the complex 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 are $z = 0$, 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 $xy = z^2$. Work it out for yourself: this has 2 infinite points, and this is also just a sgtretched out ellipse.

From Week 6, the important things are to get good at finding limits.

Week 7 summary

Monday: Have I got more limits for you (Sections 2.B,2.C of Bowman's Notes)

We know from last Thursday 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 $a_1 \leq a_2 \leq a_3 \leq \cdots$. A *monotone decreasing sequence* means $a_1 \geq a_2 \geq a_3 \geq \cdots$. 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.

Bounded monotone 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 $\inf a_n$.

Corollary. Let c be any constant.
(a) If $c \leq -1$, then c^n diverges (has $\sup \neq \inf$)
(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.

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 $R \leq -1$, 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 (inf 1 and sup 2).

This also applies if $R = \infty$. In (c), it should be easy to tell if the limit of a_n is $+\infty$ or $-\infty$. We avoid $a_n = 0$ in order to avoid division by 0.

Example: 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 6. 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.

Wednesday: Have I got even more limits for you (Sections 2.D, 2.E of Bowman's Notes)

Recall from last class: A bounded monotone sequence converges.

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 $3.999 \cdots = 4$. 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, $\pi = 3.1241881...$

Recall from Week 6 that every convergent sequence is bounded, but some bounded sequences diverge. However, last class we learned that any bounded monotonic sequence converges.

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

The analogue of this holds in any dimension $\mathbb{R}^d$. 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.

Our old definition of convergent sequence from Week 5 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:

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

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

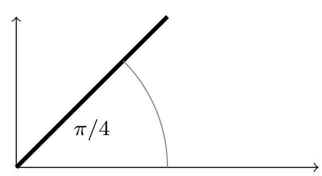
$$|a_n - a_m| = n^{-2} + m^{-2} < N^{-2} + N^{-2} = 2N^{-2} = 2(\sqrt{2/\epsilon})^{-2} = 2/(2/\epsilon) = \epsilon$$

Cauchy sequence Theorem. The sequence a_n converges iff it is a Cauchy sequence.

This is proved using Bolzano–Weierstrass. 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. You check this on the assignment.

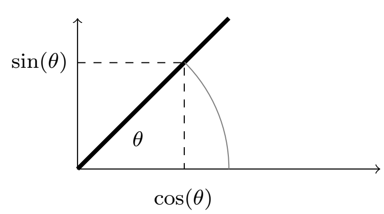
Thursday: trig (Section 3.B of Bowman's Notes)

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:

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

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

Friday: Limits and functions

Have I got a little more on limits for you (Section 2.E of Bowman's Notes)

In the definition of Cauchy sequences, we see that $m, n > N$. It is not enough to have $n > N$ and $m = n + 1$. For example, $a_n = \sqrt{n}$ diverges and it is not Cauchy. But it is the case that $a_{n+1} - a_n = \sqrt{n+1} - \sqrt{n}$ tends to 0.

Another example is the 'harmonic series' $a_n = \sum_{k=1}^{n} 1/k$. This diverges and is not Cauchy, but $a_{n+1} - a_n = 1/(n+1) \to 0$. (To see a_n isn't Cauchy, for any n, $a_{2n} - a_n = \sum_{k=n+1}^{2n} 1/k \leq \frac{n}{2n} = \frac{1}{2}$.)

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

What is a *function*?

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*, the set Y is called the *codomain*. The set of all values $f(x)$ is called the *range*.

In this course, we are mainly interested in D and Y being subsets of $\mathbb{R}$. Almost always they'll be intervals, or unions of finitely many intervals.

A function $f : \mathbb{R} \to \mathbb{R}$ is *even* if $f(-x) = f(x)$ for all $x \in \mathbb{R}$. It is called *odd* if $f(-x) = -f(x)$ for all x.

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* are of the form $poly(x)/poly(x)$. $f(x) = x^4 - 3x^2 + 1$ is even; $f(x) = 6x^3 + 14x$ is odd. The polynomial $f(x) = x^2 + 3$ has domain $\mathbb{R}$, codomain $\mathbb{R}$, and range $[3, \infty)$. The polynomial $f(x) = x^3 + 3$ has domain, codomain and range $\mathbb{R}$. The rational function $f(x) = (x^2 + 1)/(x - 1)$ has domain $(-\infty, 1) \cup (1, \infty)$ and range $(-\infty, 2 - 2\sqrt{2}] \cup [2 + 2\sqrt{2}, \infty)$.

From Week 7, the important things are to be really comfortable with limits of sequences. Be comfortable with the basics of sine and cosine.

Week 8 summary

Monday: Miscillaneous

We did Quiz 4, on sequences and their limits.

Additional Trig inequality: Section 3.B in Bowman's notes

$$\sin(x) \leq x \leq \tan(x)$$

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. We chose that because we have 10 fingers, so 10 is our 'unit' of counting. We discuss this a bit in the Week 7 notes.

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*. This is discussed in Thursday of Week7.

Comment on the Calculus Revolution vrs the Next Revolution. Our calculus course is based on the revolution in math 200 years ago which followed the development of calculus by Newton and Leibniz. Newton was interested in an intuitive tool for his physics and geometry. It took mathematicians a century to make precise mathematical sense of Newton's (and Leibniz's) ideas. The resulting branch of math is called *Analysis*, and this course (Math 117) is really a first course in Analysis. (Math 114 etc is really just a course describing Newton's tools, not making much effort at explaining why they work.)

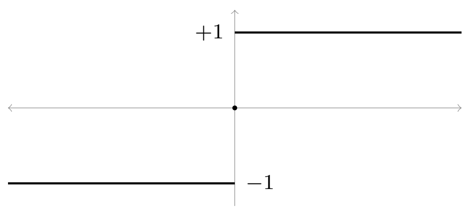
In the 20th century a much deeper physical theory was developed, called Quantum Field Theory. And it makes even less mathematical sense than calculus did. It will take all of the 21st century to make mathematical sense of Quantum Field Theory. And the implications to math (and hence to all of the mathematical sciences, including especially physics) will be far more profound than the calculus revolution was.

Intuition concerning continuity

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 $x = 0$:



We'll give the precise definition of continuity and discontinuity next class. We'll define next class what it means to take the limit of $f(x)$ as x gets closer and closer to a (but never equals a). We call this $\lim\limits_{x\to a} f(x)$. Then we say f is continuous at $x = a$ if $\lim\limits_{x\to a} f(x) = f(a)$.

Wednesday: δ's and ϵ's (Sections 3.C, 3.E of Bowman's Notes)

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

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

More important 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 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. Note that 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.

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.

There are two additional possibilities:

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 $0 < |x - a| < \delta$.

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 $\lim_{x \to \infty} f(x) = \infty$ 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 $x > N$.

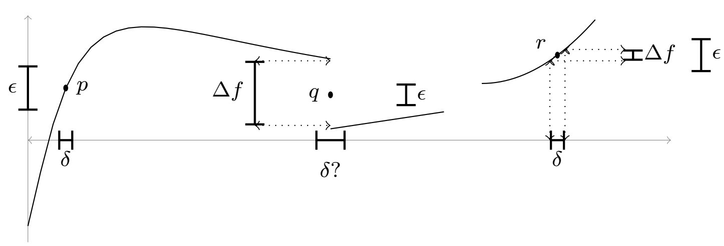
Similarly, let $f(x)$ be any function and assume that the interval $(-\infty, c)$ is in the domain of $f(x)$. We say $\lim_{x \to -\infty} f(x) = \infty$ if, for all M, there is an $N < c$ such that $f(x) > M$ whenever $x < N$. Likewise, we say $\lim_{x \to -\infty} f(x) = -\infty$ if, for all M, there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

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 perspective 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).

The final definition:

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. Of all these limit definitions, the most important are the finite-finite limit, and continuity. 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.

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

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 $f(x) = x^2$ is continuous at $x = 2$.

Secret work: So that means $a = 2$ and $L = f(2) = 4$, and we're in the finite-finite world. We need to find explicitly how close x must be to 2, in order that $f(x)$ be close to $f(2)$ (i.e. we must find δ explicitly, given ϵ). We compute

$$|f(x) - L| = |x^2 - 4| = |(x - 2)(x + 2)| = |x - 2|\,|x + 2|$$

We want to bound both factors $|x - 2|$ and $|x + 2|$. We can always force $x < 3$ here (it just amounts to insisting that $\delta < 1$, since $|x - 2| < \delta$ means $2 - \delta < x < 2 + \delta$). So $|x + 2| < 5$, and we get $|f(x) - L| < 5|x - 2|$. So in order to have $|f(x) - 4| < \epsilon$, we need $|x - 2| < \epsilon/5$. In other words, we should take $\delta = \epsilon/5$ (and also $\delta < 1$).

Official answer: The domain of $f(x) = x^2$ is all of $\mathbb{R}$, so again we can ignore c, d here. Choose any $\epsilon > 0$. Take $\delta = \min\{1, \epsilon/5\}$. Then when $|x - 2| < \delta$,

$$|f(x) - L| = |x^2 - 4| = |x - 2|\,|x + 2| < |x - 2| 5$$

since $\delta < 1$ (and hence $x + 2 < a + \delta + 2 \leq 5$). But $|x - 2|5 < 5\delta \leq \epsilon$, and we're done.

Thursday: More δ-ϵ examples

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.

Example: Show $f(x) = \sqrt{x}$ is continuous at $x = 4$.

Secret work: So that means $a = 4$ and $L = f(4) = 2$, and we're again 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

$$|f(x) - L| = |\sqrt{x} - 2| = |(\sqrt{x} - 2)\frac{\sqrt{x} + 2}{\sqrt{x} + 2}| = |\frac{x - 4}{\sqrt{x} + 2} = \frac{|x - 4|}{|\sqrt{x} + 2|}$$

We want to bound both factors $|x - 4|$ and $|\sqrt{x} + 2|$. We want an upper bound for that fraction, so that means an upper bound for the top and a lower bound for the bottom. We can always force $\sqrt{x} > 1$ here (it just amounts to insisting that $\delta < 3$, since $|x - 4| < \delta$ means $4 - \delta < x < 4 + \delta$).

So $|\sqrt{x}+2| > 3$, and we get $|f(x)-L| < 3|x-4|$. So in order to have $|f(x)-2| < \epsilon$, we need $|x-4| < \epsilon/5$. In other words, we should take $\delta = \epsilon/3$ (and also $\delta < 3$).

Official answer: The domain of $f(x) = \sqrt{x}$ is all $x \geq 0$, so $c = 0, d = \infty$ here. Choose any $\epsilon > 0$. Take $\delta = \min\{3, \epsilon/3\}$. Then when $|x-4| < \delta$, $x > 1 > 0$ so x is in the domain of $f(x)$, and

$$|f(x)-L| = |\sqrt{x}-2| = |(\sqrt{x}-2)\frac{\sqrt{x}+2}{\sqrt{x}+2}| = |\frac{x-4}{\sqrt{x}+2} = \frac{|x-4|}{|\sqrt{x}+2|}$$

since $\delta < 3$ (and hence $\sqrt{x}+2 > \sqrt{a-\delta}+2 \leq 3$). But $|x-4|3 < 3\delta \leq \epsilon$, and we're done.

If instead $f(x) = 3/x - 7\sqrt{x}$, at say $a = 4$, then use the triangle inequality to treat $3/x$ and $-7\sqrt{x}$ separately. You can see for yourself that $\delta = \min\{12\epsilon/37, 3\}$ works. ($\delta \leq 3$ means $x > 1$; use $|f(x) - f(4)| \leq 3|x-4|/(4|x|) + 7|x-4|/(\sqrt{x}+2) < ((3/4) + (7/3))|x-4| < \epsilon$).

Friday: Have I got more examples for you

There are two other pictures of limits:

(a) the neighbourhood method. Call U a neighbourhood of a if it contains an interval (c,d) with $c < a < d$. Call U a neighbourhood of $+\infty$ (respectively $-\infty$) if it contains an interval (N,∞) (resp. $(-\infty, N)$) for some $N \in \mathbb{R}$. We say that $f(x)$ is continuous at $x = a$ if, for each neighbourhood V of $f(a)$, there is a neighbourhood U of a contained in the domain of $f(x)$ such that $f(U) \subseteq V$. Here, $f(U) = \{f(x) \,|\, x \in U\}$. Here, U takes the place of the interval $|x-a| < \delta$ and V takes the place of the interval $|f(x) - f(a)| < \epsilon$.

The neighbourhood method is the way we'll define continuity in much more general contexts, starting in 2nd year. We won't use it in this course, but maybe it'll help some of you see what's going on.

(b) the sequence method. We say that $\lim_{x\to a} f(x) = L$ if, for every sequence $x_n \to a$, where all $x_n \neq a$, $f(x_n) \to L$. This works for all 4 limits (finite-finite, infinite-finite, etc).

In Week9, we will prove the equivalence of the sequence method with the δ-ϵ definition. This is a nice reformulation of limit which we will use.

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.

From Week 8, the important thing is to be comfortable with limits of functions and what continuity means. Most important is finite-finite limits.

Week 9 summary

Monday: More limits of functions

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.

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

OK, given any M, we want a $\delta > 0$ 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. $\delta = \sqrt{7}/\sqrt{8M}$.

Official answer: Choose any M. Let $\delta = \min\{1, \sqrt{7}/\sqrt{8M}\}$. 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.

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

OK, given any M, we want an N 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 $M = N$.

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

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

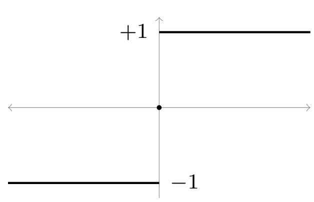
and we're done.

Wednesday: Happy Hallowe'en! (Sections 3.C, 3.E of Bowman's Notes)

First, let's prove discontinuity using δ-ϵ. Consider the sign function

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

If we graph this function, we get



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.

Happiness: 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:

Happy 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 $(c, a) \cup (a, d) \subseteq \mathrm{domain}(f)$, and then we require all $a_n \in (c, a) \cup (a, d)$ — if this domain condition confuses you, ignore it.

The Happy Theorem doesn't look that wonderful, but it is a 'bridge theorem' allowing us to carry over everything we know about limits of sequences, which is quite a bit. We'll prove it next week.

Happy Corollary. f continuous at $x = a$ iff for all $a_n \to a$, $f(a_n) \to f(a)$.

This is just applying the Happy Theorem to the formula $\lim_{x \to a} f(x) = f(a)$.

Theorem H.1(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. Likewise, we know $\sin(x)$ and $\cos(x)$ are continuous everywhere, so e.g. $\tan(x) = \sin(x)/\cos(x)$ is continuous at any x where $\cos(x) \neq 0$, i.e. at any $x \neq (2n+1)\pi/2$. (for $n \in \mathbb{Z}$)

Of course, the proof is a combination of the Happy Corollary with the theorem on limits of sums, products, quotients of sequences.

Theorem H.2. If $g(x)$ is continuous at $x = a$ and $f(y)$ is continuous at $y = g(a)$, then the composition $(f \circ g)(x) = f(g(x))$ is continuous at $x = a$.

So for example, $\sin(\frac{x^3+5}{x^2-x-6})$ is continuous everywhere, except possibly at $x = 3, -2$ (where the denominator $x^4 - x - 6$ vanishes).

To see this, let a_n be any sequence tending to a. Then by continuity of $g(x)$ at $x = a$, $g(a_n) \to g(a)$. In other words, $b_n = g(a_n)$ is a sequence converging to $b = g(a)$, so the continuity of $f(y)$ at $y = b$ means $f(b_n) \to f(b)$, which gives Theorem H.2.

Thursday: More Happiness (Sections 3.D, 3.E of Bowman)

The same argument that gave us Theorem H.2, tells us that, if $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = L'$, then

$$\lim_{x\to a}(f(x)+g(x)) = L + L'\,, \quad \lim_{x\to a}(f(x)g(x)) = LL'\,,$$

If $L' \neq 0$, then $\lim_{x\to a}(f(x)/g(x)) = L/L'$.

We can also write these as $\lim_{x\to a}(f(x)+g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$ etc, if you prefer. Note that, when $f(x)$ is continuous, we can write $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

Note that the notion of continuity helps evaluate limits quickly. For example, suppose you want to compute $\lim_{x\to\infty} \sin(\pi \frac{x^3-x^2+5}{4x^3-9x^2+5x-17}) = \sin(\pi/4) = \sqrt{2}/2$, because $\sin(x)$ is continuous.

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 $\big||x| - |a|\big| \leq |x - a| \leq \delta = \epsilon$ and we're done. That was surprisingly easy!

You can also prove continuity using Happiness. Recall that $|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$ We have continuity at any $a > 0$, because all that matters for those points is the branch of $|x|$ for positive x, which is just the line $y = x$. This is certainly continuous. Similarly, $|x|$ is continuous for any $a < 0$. The only question is: what happens where the 2 branches meet, i.e. at $x = 0$? (This is typically what continuity reduces to.) So let $a_n \to 0$. Then for any $\epsilon > 0$, there is an N such that for all $n > N$, $|a_n - 0| < \epsilon$. Then for any $n > N$, $\big||a_n| - 0\big| = \big||a_n|\big| = |a_n| < \epsilon$. This is the statement that $|a_n| \to 0$, which is what we wanted. So $|x|$ is also continuous at $x = 0$, as we knew.

Happiness gives a fast way to prove that limits don't exist, and that a function is not continuous at some point.

Theorem H.3. Choose any function $f(x)$.
(a) $\lim\limits_{x\to a} f(x)$ doesn't exist if there are sequences $a_n, b_n \to a$ (where $a_n, b_n \neq a$) and $\lim\limits_{n\to\infty} f(a_n) \neq \lim\limits_{n\to\infty} f(b_n)$.
(b) $f(x)$ is discontinuous at $x = a$, if we can find a sequence $a_n \to a$ such that $\lim\limits_{n\to\infty} f(a_n) \neq f(a)$.

For example, we can use this to show that $\text{sgn}(x)$ is not continuous at $x = 0$. Take $a_n = 1/n$. Then $a_n \to 0$, but $\text{sgn}(a_n) = +1 \to 1$ which doesn't equal $\text{sgn}(0)$, and we're done! (by H.3(b)). The same argument shows that $\lim\limits_{x\to 0} \text{sgn}(x)$ does not exist. Our two sequences are $a_n = 1/n$ and $b_n = -1/n$. Then $\text{sgn}(a_n) = +1$ which tends to $+1$ (obviously!), while $\text{sgn}(b_n) = -1$ which tends to -1. So that means $\lim\limits_{x\to 0} \text{sgn}(x)$ does not exist (using Thm.H.3(a)).

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, $a_n = ...d_2 d_1 d_0.d_{-1}...d_{-n}$ 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 Happiness (i.e. Thm.H.3) 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 $\lim\limits_{x\to a} f(x)$ 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 $-|x_n| \le g(x_n) \le |x_n|$. 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 $x = 0$.

Friday: Quiz and midterm review

From Week 9, the important thing is Happiness.

Week 10 summary

Monday: Midterm 2!!

Wednesday: Midterm discussion; Proof of Happy Theorem (Section 3.D of Bowman's Notes)

In Week 9 we saw the usefulness of the Happy Theorems. Now we'll prove them. We'll just prove one of them, as they're all proved the same way.

Happy Corollary. Suppose the interval (b, c) is in the domain of f, and that $b < a < c$. Then $f(x)$ is continuous at $x = a$ **iff** all sequences $a_n \to a$, which have all a_n in the domain of f, also have $f(a_n) \to f(a)$.

The conditions that "(b, c) is in the domain of f, and that $b < a < c$", just say that a is in the domain of f, as are all points close to a. So we are excluding a being an endpoint of the domain (this will be discussed next class), 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 $f(x)$ is continuous at $x = a$. (As always with continuity, this is a "finite-finite" limit.) Let a_n be any sequence entirely in the domain of f, and we assume a_n converges to a. We want to prove that the sequence $f(a_n)$ converges to $f(a)$.

Continuity of f tells that when x is close to a, then $f(x)$ will be close to $f(a)$. $a_n \to a$ 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 $f(x)$ is continuous at $x = a$, there exists a $\delta > 0$ such that, whenever $|x - a| < \delta$, then $|f(x) - f(a)| < \epsilon$. 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 $n > N$, $|f(a_n) - f(a)| < \epsilon$. QED to $\Rightarrow$

Thursday: Finishing Happiness proof; 1-sided limits (Sections 3.D, 3.F of Bowman)

Let's finish off the proof of the Happy Corollary, by proving the other direction:

$\Longleftarrow$ Assume that for any sequence $a_n \to a$, where all terms a_n lie in the domain of f, we have $f(a_n) \to f(a)$. We want to prove that, in spite of that, $f(x)$ is not continuous at $x = a$.

Suppose for contradiction that $f(x)$ is not continuous at $x = a$. Then that means that there is some $\varepsilon_0 > 0$ such that, for any $\delta > 0$, there is an x_δ in the domain of f such that $|x_\delta - a| < \delta$ but $|f(x_\delta) - f(a)| > \varepsilon_0$. 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 $|a_n - a| < 1/n$, so $a_n \to a$. Also, $|f(a_n) - f(a)| > \varepsilon_0$. So certainly $f(a_n)$ doesn't converge to $f(a)$.

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 $f(a)$. This terrible contradiction means that our most recent assumption (or supposition) was wrong. So it is false that "$f(x)$ is not continuous at $x = a$", which means that it is true that $f(x)$ is continuous at $x = a$, and we're done. QED to $\Leftarrow$, hence QED to Happy Corollary

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

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 $\lim\limits_{x\to a^+} f(x)$ 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 $\lim\limits_{x\to a^-} f(x)$ 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, consider the function

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

Then $\lim\limits_{x\to 0^+} \operatorname{sgn}(x) = 1$, and $\lim\limits_{x\to 0^-} \operatorname{sgn}(x) = -1$.

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

We say $f(x)$ is *continuous from the right* at $x = a$, if $\lim\limits_{x\to a^+} f(x) = f(a)$. We say $f(x)$ is *continuous from the left* at $x = a$, if $\lim\limits_{x\to a^-} f(x) = f(a)$. 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 $\lim\limits_{x\to a^+} f(x)$ 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 $\lim\limits_{x\to a^+} f(x)$ 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 $\lim\limits_{x\to a^-} f(x)$ 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 $\lim\limits_{x\to a^-} f(x)$ 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, $\lim\limits_{x\to a} f(x) = \pm\infty$, iff $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} f(x) = \pm\infty$.

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

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 $\lim\limits_{x\to\infty} f(x)$, we really mean $\lim\limits_{x\to\infty^-} f(x)$. When we write $\lim\limits_{x\to-\infty} f(x)$, we really mean $\lim\limits_{x\to\infty^+} f(x)$, or if you prefer $\lim\limits_{x \to -\infty^+} f(x)$.

The happiness theorems work for one-sided limits. For example: $\lim\limits_{x \to a^+} f(x) = L$ 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 $a_n < a$, $\lim_{n\to\infty} f(a_n) = f(a)$.

Friday: 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 $\lim\limits_{n\to\infty} a_n$ is some finite number $c \le b$. And b_n is monotonic decreasing, and bounded below by a. So $\lim\limits_{n\to\infty} b_n$ is some finite number $c' \ge a$. But $b_n - a_n \to 0$, so $c = c'$, and $a \le c \le b$. 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.

From Week 10, the important thing is one-sided limits, continuity on closed intervals, Intermediate Value Theorem.

Week 11 summary

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

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 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 $f(x) = 0$.

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 S.

Proof. Let $L = \sup\{f(x) : x \in S\}$ (we need to show 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 $z_n < x_n \le L$. 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).

Wednesday: Continuity of inverse functions (Section 3.G 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.

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.

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.

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 prove next class 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.

Thursday: Some sci-fi

Let's begin with a little result we mentioned at the end of last class:

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

(a) $f(x)$ is one-to-one.

(b) $f(x)$ has range $[f(a), f(b)]$.

(c) $f : [a, b] \to [f(a), f(b)]$ 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; 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.

We're now at the end of Chapter 3 in Bowman's Notes. 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 $0 \le x \le 1$, $0 \le y \le 1$.

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

In class I described a related curve, by Hilbert, which uses base 4.

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 $1/3 \le x \le 2/3$, $x_1 = 1$, 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 $1/9 \le x \le 2/9$, $f(x) = 1/4$. For $7/9 \le x \le 8/9$, $f(x) = 3/4$.

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.

Friday: Definition of derivative (Section 4.A of Bowman's notes)

Definition. Let (a, b) be in the domain of f, and choose any $c \in (a, b)$. We say $f(x)$ is *differentiable* at $x = c$, if $\lim\limits_{h\to 0} \frac{f(c+h)-f(c)}{h}$ converges to a finite number. We call that limit the *derivative* $f'(c)$ or $\frac{df}{dx}(c)$.

For example, consider $f(x) = ax + b$. Then for any $c \in \mathbb{R}$,

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = \lim_{h\to 0} \frac{a(c+h) + b - (ac+b)}{h} = \lim_{h\to 0} \frac{ah}{h} = a$$

so $f(x)$ is differentiable at any $x = c$, with derivative $f'(c) = a$.

For example, consider $f(x) = x^2$. Then for any $c \in \mathbb{R}$,

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = \lim_{h\to 0} \frac{(c+h)^2 - c^2}{h} = \lim_{h\to 0} \frac{2ch + h^2}{h} = 2c$$

so $f(x)$ is differentiable at any $x = c$, with derivative $f'(c) = 2c$.

For example, consider $f(x) = x^n$. Then for any $c \in \mathbb{R}$,

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = \lim_{h\to 0} \frac{(c+h)^n - c^n}{h} = \lim_{h\to 0} \frac{nc^{n-1}h + n(n-1)c^{n-2}h^2/2 + \cdots + h^n}{h} = nc^{n-1}$$

so $f(x)$ is differentiable at any $x = c$, with derivative $f'(c) = nc^{n-1}$.

From Week 11, the important things are the Maximum Principle, the definition of invertibility, the Continuity of Inverse Theorem, the definition of derivative, and the ability to differentiate from the definition.

Week 12 summary

Monday: Derivatives of trig functions; continuity and derivative (Section 4.A in Bowman's Notes)

A more interesting example to consider is the derivative of $\sin(x)$. For this purpose, use the inequality

$$\sin(x) \leq x \leq \tan(x) \quad \text{for } 0 \leq x \leq \pi/2 \tag{1}$$

which we proved on Monday of Week 8. From the first inequality in eq.(1) we get $\sin(x)/x \leq 1$, and from the second we get $x \leq \sin(x)/\cos(x)$, i.e. $\cos(x) \leq \sin(x)/x$ (both x and $\cos(x)$ are positive for $0 < x < \pi/2$ so the inequalities don't change when we multiply or divide by them). Putting these together, we get $\cos(x) \leq \sin(x)/x \leq 1$ for $0 < x > \pi/2$. We also need this inequality for negative x: for $-\pi/2 \leq x \leq 0$, replace x with $-x$ in eq.(1) to get

$$\sin(x) \geq x \geq \tan(x) \quad \text{for } 0 \geq x \geq pi/2$$

This gives $\sin(x)/x \leq 1$ and $\cos(x) \leq \sin(x)/x$ (now $\cos(x) > 0$ but $x < 0$, so inequalities reverse). The net result is that

$$\cos(x) \leq \frac{\sin(x)}{x} \leq 1\,, \quad \text{for } x \in (-\frac{\pi}{2}, 0) \cup (0, \frac{\pi}{2}) \tag{2}$$

OK, let's try to differentiate $\sin(x)$ at $x = 0$. We need to evaluate $\lim\limits_{h\to 0} \frac{\sin(h)-\sin(0)}{h} = \lim\limits_{h\to 0} \frac{\sin(h)}{h}$. But for $h \neq 0$ but close to 0, we have the iniequalities in eq.(1) above, so:

$$\lim_{h\to 0} \cos(h) \leq \lim_{h\to 0} \frac{\sin(h)}{h} \leq \lim_{h\to 0} 1$$

Now, the right limit is obviously equal to 1. And $\cos(x)$ is continuous everywhere, so the left limit equals $\cos(0) = 1$. So by the Squeeze theorem, we get that $\lim\limits_{h\to 0} \sin(h)/h = 1$. But this means that the derivative of $\sin(x)$ exists at $x = 0$, and equals 1.

We can use this to compute the derivative of $\cos(x)$ at $x = 0$. We need to evaluate $\lim\limits_{h\to 0} \frac{\cos(h)-\cos(0)}{h} = \lim\limits_{h\to 0} \frac{\cos(h)-1}{h}$. But there's a trick: write $h' = h/2$, then $\cos(h) - 1 = \cos(2h') - 1 = \cos^2(h') - \sin^2(h') - 1 = 1 - 2\sin^2(h') - 1 = -2\sin^2(h')$. So

$$\lim_{h\to 0} \frac{\cos(h) - 1}{h} = \lim_{h\to 0} \frac{-2\sin^2(h')}{2h'} = -\lim_{h\to 0} \frac{\sin^2(h')}{h'^2} h' = -\left(\lim_{h'\to 0} \frac{\sin(h')}{h'}\right)\left(\lim_{h'\to 0} h'\right) = -1^2 \cdot 0 = 0$$

Thus the derivative of $\cos(x)$ at $x = 0$ exists, and equals 0.

How about the derivative of $\sin(x)$ at any point x?

$$\lim_{h\to 0} \frac{\sin(x+h) - \sin(x)}{h} = \lim_{h\to 0} \frac{\sin(x)\cos(h) + \cos(x)\sin(h) - \sin(x)}{h}$$

$$= \lim_{h\to 0} \sin(x)\frac{\cos(h) - 1}{h} + \lim_{h\to 0} \cos(x)\frac{\sin(h)}{h} = \sin(x)\cdot 0 + \cos(x)\cdot 1 = \cos(x)$$

using the above calculations of the $\sin(x)$ and $\cos(x)$ derivatives at $x = 0$. So $\sin(x)$ is differentiable everywhere, and $\frac{d}{dx}\sin(x) = \cos(x)$.

How about the derivative of $\cos(x)$ at any point x?

$$\lim_{h\to 0} \frac{\cos(x+h) - \cos(x)}{h} = \lim_{h\to 0} \frac{\cos(x)\cos(h) - \sin(x)\sin(h) - \cos(x)}{h}$$

$$= \lim_{h\to 0} \cos(x)\frac{\cos(h) - 1}{h} - \lim_{h\to 0} \sin(x)\frac{\sin(h)}{h} = \cos(x)\cdot 0 - \sin(x)\cdot 1 = -\sin(x)$$

using the above calculations of the $\sin(x)$ and $\cos(x)$ derivatives at $x = 0$. So $\cos(x)$ is differentiable everywhere, and $\frac{d}{dx}\cos(x) = -\sin(x)$.

How about $\tan(x)$ and the other trig functions? They're all built from $\sin(x)$ and $\cos(x)$, so you just have to use quotient rule etc to calculate them.

So are all functions differentiable everywhere? No!

Diff$\subset$Cont Theorem. If $f(x)$ is differentiable at $x = c$, then it must be continuous at $x = c$.

This is Theorem 4.1 in Bowman's Notes. The easiest way to see this is to consider the definition of derivative at $x = c$:

$$f'(c) = \lim_{x\to c} \frac{f(x) - f(c)}{x - c}$$

As $x \to c$, the denominator tends to 0. So the only chance to get a finite number (namely $f'(c)$) from this limit, is if the numerator also tends to 0. So $f(x) \to f(c)$. This is exactly what it means for $f(x)$ to be continuous at $x = c$. Bowman's Notes give an alternate argument.

So it is easy to come up with examples of functions which aren't differentiable at a point. E.g. $\operatorname{sgn}(x)$ is discontinuous at $x = 0$, so it is not differentiable at $x = 0$. Another example: $f(x) = 1$ when $x \in \mathbb{Q}$, otherwise $f(x) = 0$; this function is discontinuous everywhere, so it isn't differentiable anywhere.

Is the Converse to this theorem true? In other words, if $f(x)$ is continuous at $x = c$, is it necessarily differentiable there?

No! The easiest example is $f(x) = |x|$. This is continuous everywhere. Does $|x|'$ exist everywhere? For $x > 0$,

$$\frac{d}{dx}|x| = \lim_{h\to 0}\frac{|x+h| - |x|}{h} = \lim_{h\to 0}\frac{x+h-x}{h} = 1$$

so $|x|' = 1$ when $x > 0$. For $x < 0$,

$$\frac{d}{dx}|x| = \lim_{h\to 0}\frac{|x+h| - |x|}{h} = \lim_{h\to 0}\frac{-(x+h)-(-x)}{h} = -1$$

so $|x|' = -1$ when $x < 0$. For $x > 0$, we compute the one-sided limits: $\lim_{h\to 0^+}\frac{|h|-|0|}{h} = +1$ and $\lim_{h\to 0^-}\frac{|h|-|0|}{h} = -1$. So the two 1-sided limits don't agree, so $|x|$ is not differentiable at $x = 0$.

Functions which are continuous but not differentiable at a point, have a corner there.

Wednesday: Product, quotient, and chain rules (Section 4.A in Bowman's Notes)

A very very common tactic in math is to build up the complicated from the simple. The simple are things like $f(x) = x$ and trig functions. We can build new functions by adding, multiplying, dividing, nesting,... We already know that these operations send continuous functions to continuous functions. So we want to know what happens to the derivative.

Properties of Derivatives. Suppose $f(x), g(x)$ are differentiable at $x = c$. Then:

(a) *sum rule:* for any constants α, β, the function $\alpha f(x) + \beta g(x)$ is differentiable at $x = c$, with derivative $\alpha f'(c) + \beta g'(c)$;
(b) *product rule:* the function $f(x)\, g(x)$ is differentiable at $x = c$, with derivative $f'(c)\, g(x) + f(x)\, g'(c)$;

(c) *quotient rule:* provided $g(c) \neq 0$, the function $f(x)/g(x)$ is differentiable at $x = c$, with derivative $\frac{f'(c)\, g(x) - f(x)\, g'(c)}{g^2(x)}$.

This is Theorem 4.2 in Bowman's notes, where it is proved.

The sum rule (a) tells us that e.g. the set of all functions, say $f : (0,1) \to \mathbb{R}$, which are differentiable at some point c, forms a *vector space*. If we consider say the set of all functions $f : (0,1) \to \mathbb{R}$ which are differentiable everywhere, then not only is that a vector space, but the derivative $\frac{d}{dx}$ is a linear transformation from that vector space (the domain) to the space of *all* functions $g : (0,1) \to \mathbb{R}$ (which may or may not be differentiable or even continuous).

We learned on Friday of Week 11 that $f(x) = x$ is differentiable everywhere, as is the constant function $f(x) = 1$. Then by the product rule, so is $x \cdot x = x^2$, so so is $x^2 \cdot x = x^3$ etc. The sum rule tells us that $\alpha + \beta x + \gamma x^2$ etc will also be differentiable everywhere. We thus obtain:

Corollary. Every polynomial is differentiable everywhere. Any rational function (i.e. function of the form $\mathrm{poly}(x)/\mathrm{poly}(x)$) is differentiable everywhere it is defined (i.e. everywhere the denominator doesn't vanish).

We can easily prove by induction, using the product rule, that $\frac{d}{dx}x^n = nx^{n-1}$ for all $n \in \mathbb{N}$. So we can differentiate any polynomial.

For example, let V be the set of all polynomials. This is an infinite-dimensional vector space. The derivative of any polynomial is another polynomial. So $\frac{d}{dx}$ is a linear transformation $V \to V$. It corresponds to a matrix, as soon as we choose a basis for V. The obvious basis is $1, x, x^2, x^3, ...$ The the matrix $[\frac{d}{dx}]$ corresponding to this basis is $\infty \times \infty$, but the northwest corner of it looks like

$$\left[\frac{d}{dx}\right] = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 3 & \cdots \\ \vdots & \vdots & & & \ddots \end{pmatrix}$$

Recall the composition $f \circ g$ of functions: the basic formula is $(f \circ g)(x) = f(g(x))$. An example is $\cos(x^2 + 1)$: here $f(y)$ is $\cos(y)$ and $g(x) = x^2 + 1$. So composition is how we *nest* functions. Don't confuse it with product $f(x)g(x)$ of functions.

Chain rule. Let $g(x)$ be differentiable at $x = c$, and $f(y)$ be differentiable at $y = g(c)$. Then $(f \circ g)(x)$ is differentiable at $x = c$, and the derivative is $\frac{d}{dx} f(g(x)) = f'(d)\, g'(c)$ where we write $d = g(c)$.

We write $f'(d)$ here, but we could also write it $f'(g(c))$, but in that case make sure you remember that you're differentiating $f(y)$ as usual, and then evaluating it at $y = g(c)$.

So $\cos(x^2+1)$ will be differentiable everywhere, with derivative $(-\sin(x^2+1))\ (2x) = -2x\sin(x^2+1)$. Note that $f'(y) = -\sin(y)$ here (since $f(x) = \cos(x)$), but we substitute in $y = g(x) = x^2 + 1$, to get $-\sin(x^2+1)$.

The proof of the Chain rule amounts to a change of variables, and games with sequences. The full proof is given in Thm.4.3 of Bowman's notes.

Thursday: Inverse function rule (Section 4.I of Bowman's notes)

The final "rule" for derivatives we'll give is the formula for differentiating inverse functions. By the *inverse function* $f^{-1}(y)$ we mean the function (if it exists) such that $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$. So $f^{-1}(y)$ undoes $f(x)$. We discussed inverse on Wednesday of Week 11. f^{-1} exists iff f is one-to-one and onto. We learned in Week 11 (and Assignment 9 Question 2) that a continuous function can be invertible only if it is increasing or decreasing. Don't confuse $f^{-1}(y)$ here with $1/f(x)$. the inverse function rule is related to the Chain rule, not the quotient rule.

Inverse function rule. Suppose $f(x)$ is continuous and invertible on the interval (a, b), and differentiable at some point $c \in (a, b)$. Write $d = f(c)$. Then $f^{-1}(y)$ is differentiable at $y = d$, iff $f'(c) \neq 0$. In this case, $\frac{d}{dx} f^{-1}(d) = \frac{1}{f'(c)}$.

WARNING: Like the Chain rule, this formula mixes c's and d's. We want to know the derivative of f^{-1} at $y = d$, but the formula gives it to us using $x = c$. So a better way to write the inverse function rule is: $\frac{d}{dx} f^{-1}(d) = \frac{1}{f'(f^{-1}(d))}$. Again, this is just the ordinary derivative of f, but evaluated at $c = f^{-1}(d)$. We'll give examples shortly.

This is Corollary 4.6.2 of Bowman's notes, where it is proved.

First example: $f(x) = \sin(x)$. This increases from -1 at $x = -\pi/2$ to 1 at $x = \pi/2$, so this is an invertible function on the interval $[-\pi/2, \pi/2]$. We write $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ for this inverse. Note that the derivative of $\sin(x)$, namely $\cos(x)$, vanishes at the endpoints $x = \pm\pi/2$, so $\arcsin(y)$ is only differentiable in the interval $(-1, 1)$. Write $y = \sin(x)$. The derivative $\frac{d}{dy} \arcsin(y)$ equals $1/\cos(x)$, by the Inverse function rule, but this involves both variables x, y. We just want the answer in terms of y. So we have to evaluate $\cos(\arcsin(y))$. The way to do this is to express $\cos(x)$ as a function of $\sin(x)$. We know $\sin^2(x) + \cos^2(x) = 1$, so $\cos(x) = \pm\sqrt{1 - \sin^2(x)}$. So $\cos(\arcsin(y)) = \pm\sqrt{1 - y^2}$. The only question is which sign to take. But remember we are on the interval $(-\pi/2, \pi/2)$. $\cos(x)$ is positive on this interval (it starts at 0, rises to 1, then decreases back to 0). So on this interval, $\cos(x) = \sqrt{1 - \sin^2(x)}$, so we should take the $+$-sign. (Remember that $\sqrt{y}$ is always taken to be the positive root.) So the final answer:

$$\frac{d}{dy} \arcsin(y) = \frac{1}{\sqrt{1 - y^2}} \text{ on the interval } y \in (-1, 1)$$

Second example: $f(x) = \cos(x)$. This decreases from 1 at $x = 0$ to -1 at $x = \pi$, so this is an invertible function on the interval $[0, \pi]$. We write $\arccos : [-1, 1] \to [0, \pi]$ for this inverse. Note that the derivative of $\cos(x)$, namely $-\sin(x)$, vanishes at the endpoints $x = 0, \pi$, so $\arccos(y)$ is only differentiable in the interval $(-1, 1)$. Write $y = \cos(x)$. The derivative $\frac{d}{dy} \arccos(y)$ equals $-1/\sin(x)$, by the Inverse function rule, but this involves both variables x, y. We just want the answer in terms of y. So we have to evaluate $\sin(\arccos(y))$. The way to do this is to express $\sin(x)$ as a function of $\cos(x)$. We know $\sin^2(x) + \cos^2(x) = 1$, so $\sin(x) = \pm\sqrt{1 - \cos^2(x)}$. So $\sin(\arccos(y)) = \pm\sqrt{1 - y^2}$. The only question is which sign to take. But remember we are on the interval $(0, \pi)$. $\sin(x)$ is positive on this interval (it starts at 0, rises to 1, then decreases back to 0). So on this interval, $\sin(x) = \sqrt{1 - \cos^2(x)}$, so we should take the $+$-sign. So the final answer:

$$\frac{d}{dy} \arccos(y) = \frac{-1}{\sqrt{1 - y^2}} \text{ on the interval } y \in (-1, 1)$$

The minus sign comes from the minus sign when we differentiated $\cos(x)$.

Third example: $f(x) = x^2$. This in increasing on $[0, \infty)$, so it has an inverse $f^{-1} : [0, \infty) \to [0, \infty)$, namely the square-root $f^{-1}(y) = \sqrt{y} = y^{1/2}$. $f'(x) = 2x$ vanishes at $x = 0$, so the square-root is not differentiable at $y = 0$. At any point $y > 0$ though, the derivative $\frac{d}{dy} f^{-1}(y)$ exists and equals $1/(2x)$. Again, we want to replace x with some function of y, but this is easy: $x = \sqrt{y}$. So we get the final answer:

$$\frac{d}{dy}\sqrt{y} = \frac{1}{2\sqrt{y}} \text{ on the interval } y \in (0, \infty)$$

Writing $\sqrt{y}$ as $y^{1/2}$, this is consistent with the formula $\frac{d}{dx}x^n = nx^{n-1}$ where now $n = 1/2$.

The analysis is identical for $f(x) = x^n$ when $n \in \mathbb{N}$ is even: The inverse $f^{-1}(y) = y^{1/n}$ exists on $[0, \infty)$, is differentiable for $y > 0$, with derivative $\frac{d}{dy}y^{1/n} = \frac{1}{n}y^{\frac{1}{n}-1}$.

Forth example: $f(x) = x^3$. This is increasing on all of $\mathbb{R}$, so it has an inverse $f^{-1} : \mathbb{R} \to \mathbb{R}$, namely the cube-root $f^{-1}(y) = \sqrt{y} = y^{1/3}$. $f'(x) = 3x^2$ vanishes at $x = 0$, so the cube-root is not differentiable at $y = 0$. At any point $y \neq 0$ though, the derivative $\frac{d}{dy}f^{-1}(y)$ exists and equals $1/(3x^2)$. Again, use $x = y^{1/3}$, to get the final answer:

$$\frac{d}{dy}y^{1/3} = \frac{1}{3y^{2/3}} \text{ on } y \neq 0$$

As before, this is consistent with the formula $\frac{d}{dx}x^n = nx^{n-1}$ where now $n = 1/3$.

The analysis is identical for $f(x) = x^n$ when $n \in \mathbb{N}$ is odd: The inverse $f^{-1}(y) = y^{1/n}$ exists on $\mathbb{R}$, is differentiable for $y \neq 0$, with derivative $\frac{d}{dy}y^{1/n} = \frac{1}{n}y^{\frac{1}{n}-1}$.

Friday: Quiz, proof of inverse rule

We gave Quiz 6, which took up a lot of time. It is given on our e-class page, with solutions.

We then gave the proof of the Inverse function rule. The formula comes from the Chain rule, but before we can use the Chain rule, we need to know f^{-1} is differentiable. This occupies most of the proof. We use the Happy theorem to convert it to a question about sequences, and then use f^{-1} to make a change-of-variables.

From Week 12, the important things are being comfortable with differentiating functions built from polynomials and trig functions; the product, chain, inverse rules, differentiable imples continuous.

Week 13 summary

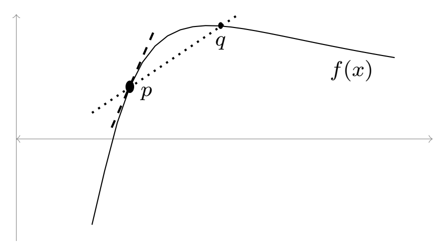
Because of the proximity of the Final Exam, I'll focus here only on the material that has a bearing on your Final exam. Before Math 118 begins, i'll supplement the Week 13 summary with the other material we covered.

Monday: The meaning of derivative (Section 4.A of Bowman)

The real meaning of derivative is the closest linear approximation to your (probably nonlinear!) function, at a given point. (In class we discussed what derivatives of higher dimensional functions, like $f : \mathbb{R}^n \to \mathbb{R}^m$, but ignore for now,)

What this means for our functions, like $f : \mathbb{R} \to \mathbb{R}$, is that derivative gives the slope of the line tangent to the curve $y = f(x)$, at some point $x = c$.

In this picture, the dotted line is the line passing through the point $p = (c, f(c))$ on the graph at $x = c$, and some other point $q = (x, f(x))$ on the graph. It's not such a good approximation to our function. But as x gets closer and closer to c, it hugs the graph (near p) closer and closer. In the limit, this becomes the tangent line, drawn here with a dashed line.



The slope of a line is $m = \Delta y/\Delta x$: change in y coordinates divided by change in x coordinates. The slope of the dotted line will be $m = \frac{f(x)-f(c)}{x-c}$. The slope of the tangent line is the limit of this as $x \to c$: $\lim\limits_{x\to c} \frac{f(x)-f(c)}{x-c}$. This is the derivative $f'(c)$.

If we have to determine the equation of the line tangent to the graph at $x = c$, well, we know the slope (namely $f'(c)$) and a point on the line (namely $p = (c, f(c))$). So we can use the formula $m = \frac{y-y_0}{x-x_0}$, which here becomes

$$f'(c) = \frac{y - f(c)}{x - c}$$

For example, the equation of the tangent line at $x = 1$ to $f(x) = x^2 + 2$ is $2 = \frac{y-3}{x-1}$, i.e. $y = 2x + 1$. Check: it has the right slope (namely $f'(1) = 2$), and it passes through $(1, f(1)) = (1, 3)$.

(We also discussed derivative as rate of change: velocity is rate of change of position; acceleration is rate of change of velocity; force is rate of change of momentum. But it's not on exam, so ignore for now)

Wednesday: Derivative tests Sections 4.B, 4.C in Bowman's Notes

We say $f(x)$ is *increasing* on an interval $[a,b]$ if $f(c) \leq f(d)$ for any $c,d \in [a,b]$ with $c<d$. We say $f(x)$ is *strictly increasing* on an interval $[a,b]$ if $f(c) < f(d)$ for any $c,d \in [a,b]$ with $c<d$. The definitions for *decreasing* and *strictly decreasing* are similar. The definitions for open or half-open intervals (a,b) or $(a,b]$ etc are the same.

We say $f(x)$ is *monotonic* on the interval $[a,b]$ if it is either increasing there, or decreasing there.

Increasing/decreasing Theorem. Suppose $f(x), g(x)$ are differentiable at $x=c$. Then: **(a)** If $f'(c)>0$, then there exists a $\delta>0$ such that $f(x)>f(c)$ when $c<x<c+\delta$, and $f(x)<f(c)$ when $c-\delta<x<c$.
(b) If $f'(c)<0$, then there exists a $\delta>0$ such that $f(x)<f(c)$ when $c<x<c+\delta$, and $f(x)>f(c)$ when $c-\delta<x<c$.

The proof starts with the observation that if say $\lim\limits_{x\to c}\frac{f(x)-f(c)}{x-c}>0$, then there must be a $\delta>0$ such that $\frac{f(x)-f(c)}{x-c}>0$ for all $x\in(c-\delta,c+\delta)$. You can prove this by contradiction, using the Happy Theorem: if we can't find such a δ, then each possible choice $\delta=1/n$ won't work, which means there must be an $x_n \in (c-1/n, c+1/n)$ such that $\frac{f(x_n)-f(c)}{x_n-c}\leq 0$. But then $\lim\limits_{n\to\infty}\frac{f(x_n)-f(c)}{x_n-c}\leq 0$. But $\lim\limits_{n\to\infty}\frac{f(x_n)-f(c)}{x_n-c}=\lim\limits_{x\to c}\frac{f(x)-f(c)}{x-c}>0$ by the Happy Theorem, and we get a contradiction.

Once we know $\frac{f(x)-f(c)}{x-c}>0$ for all $x\in(c-\delta,c+\delta)$, it is easy to get from this the inequalities of (a). The proof of (b) is the same.

We say $f(x)$ has a global maxima at c if $f(x)\leq f(c)$ for all x in the domain of f. We say $f(x)$ has an interior local maxima at c if there is a $\delta>0$ such that $(c-\delta,c+\delta)$ is in the domain of f, and $f(x)\leq f(c)$ for all $x\in(c-\delta,c+\delta)$. Global minima and interior local minima are defined the same way.

Corollary. Suppose $f(x)$ has an interior local minima or maxima at $x=c$, and that $f'(c)$ exists. Then $f'(c)=0$.

This is an immediate corollary of the Increasing/decreasing theorem, because $f'(c) > 0$ means $f(x)$ is increasing next to c, and $f'(c) < 0$ means its decreasing, so all that's left is $f'(c) = 0$.

Note that differentiability at c is essential. For example, $f(x) = |x|$ has an internal local minima at $x = 0$, but the derivative doesn't equal 0 there. In fat, the derivative doesn't exist there.

Note that it is not an iff: just because the derivative equals 0 at a point, doesn't mean it is a local maximal or minima. An easy example is $f(x) = x^3$: at $x = 0$ its derivative equals 0, but the function strictly increases everywhere.

We can use this to determine all local maxima and minima of e.g. $f(x) = 2x^3+3x^2-12x+1$ on the interval $[-3, 2]$ say. The derivative is $6x^2+6x-12$, which has zeros at $x = 1$ and $x = -2$. We compute $f(-2) = 21$ and $f(1) = -6$. The endpoints have values $f(-3) = 10$ and $f(2) = 5$. So is there a local maxima or minimal at $x = -2$? Well, we know $f'(x)$ exists and is continuous everywhere.$f(-3) < f(-2)$, so $f(x)$ must increase on the interval $(-3, -2)$ (one way to see this is compute $f'(-3) = 24$, and if $f'(x)$ ever switched sign on $(-3, -2)$, then by Intermediate Value Theorem it would have to equal 0 somewhere in that interval, and we know it doesn't, so $f'(x)$ must remain positive on that interval, which means it increases. Similarly, $f'(0) = -12$ (I chose $x = 0$ since it's inside the interval $(-2, 1)$) so by the same reasoning $f'(x)$ must stay negative in that interval $(-2, 1)$. This means $f(x)$ decreases on $(-2, 1)$. This means there is a local maxima at $x = -2$. Finally, $f'(2) = 24$, so $f'(x)$ must remain positive on the interval (1,2), so $f(x)$ increases there. So there is a local minima at $x = 1$. $x = -2$ is the only internal local maxima; $x = 1$ is the only internal local minima. Both are global maxima/minima, comparing the values at the endpoints.

(Later we'll get the 2nd derivative test for local max/min, but we don't have it yet.)

(Irrelevant for now, but Devil's staircase is differentiable almost everywhere (i.e. 100% of points); the derivative when it exists is 0; but it increases from $f(0) = 0$ to $f(1) = 1$.)

Thursday: Integration

This is a rapid fire intro to integration, for those of you who are switching to Math 115 or Math 101 or whatever. Also, you may need integration in other courses, like physics. We'll treat it properly a month or so into Math 118.

There are two kinds of integrals: indefinite and definite.

By the antiderivative $F(x)$ of $f(x)$, we mean any function with $\frac{d}{dx}F(x) = f(x)$. For example, the antiderivative of $2x$ is x^2. But x^2+3 is another antiderivative: $\frac{d}{dx}(x^2+3) = 2x$. So there are lots of antiderivatives. If you have one antiderivative, then adding a constant to it will be another antiderivative.

A more accurate phrase for antiderivative is "inverse of derivative". As we discussed in Week 12, the derivative is a function: it is a linear map sending the vector space of differentiable functions to the vector space of functions. It is not invertible, since it is not one-to-one (different functions can have the same derivative). It's like x^2: it's not one-to-one on $\mathbb{R}$, so it's not invertible, but we can still write the multi-valued 'inverse' as $\pm\sqrt{y}$. Here, we can write the multi-valued inverse of derivative as $F(x)+c$, where $F(x)$ is any function satisfying $\frac{d}{dx}F(x) = f(x)$.

The *indefinite integral* $\int f(x)\,dx$ is defined to be all antiderivatives of $f(x)$, i.e. $F(x)+c$. Don't forget the $+c$. There is *never* a single antiderivative. See p.139 Bowman's Notes. The antiderivative may not exist.

The other kind of integral is the *definite integral*. Suppose $f(x)$ is defined on (a,b), and $f(x) \geq 0$ there. We ask: what is the area between $f(x)$ and the real axis, and between $x = a$ and $x = c$. We call that region a wavy rectangle.

One way to try to define what that area would mean, is to approximate it with things we do know how to find the area of. The easiest are rectangles. So we can try to cover our wavy rectangle as closely as possible with disjoint unions of rectangles.

One way to make sense of this idea is as follows. Choose any points a_i satisfying $a = a_0 < a_1 < a_2 < \cdots < a_n = n$. This partitions the interval $[a,b]$ into n parts. For each interval $[a_i, a_{i+1}]$ let $Y_i = \sup\{f(x)\}$ be the supremum of $f(x)$ for $a_i \leq x \leq a_{i+1}$. Let y_i the the infimum over the same interval. Then the rectangles with base $[a_i, a_{i+1}]$ and sides $[0, Y_i]$ cover all of the wavy rectangle, and a bit more. And the rectangles with base $[a_i, a_{i+1}]$ and sides $[0, y_i]$ fall inside the wavy rectangle. So the area of the tall rectangles will overestimate the area of the wavy rectangle, and the total area of the short rectangles will underestimate the area of the wavy rectangle. So we get

$$\sum_{i=0}^{n-1}(a_{i+1}-a_i)y_i \leq A \leq \sum_{i=0}^{n-1}(a_{i+1}-a_i)Y_i$$

where A is the area of wavy rectangle. If we take the supremum of all lower estimates, as the rectangles get finer and finer, we still get a lower estimate for A (but it'll be a pretty good one). Similarly, the infimum of all upper estimates, will be a pretty good upper estimate for A.

If this infimum of upper estimates equals the supremum of lower estimates, then we say the definite integral $\int_a^b f(x)\,dx$ exists and equals that common value. This will be the area of the wavy rectangle. This definition of definite integral is called the Riemann integral. A better name would be clunky integral since it approximates area with chunks.

(You don't need to make much sense of this definition of Riemann integral – we'll do it properly next term.)

This notion of area and integral is discussed in Section 5.A in Bowman's Notes.

Friday: Integration (cont)

Not all functions are integrable. For example, consider $f(x) = 0$ or 1 depending on whether x is rational or not. Take the interval $[a, b] = [0, 1]$. The lower estimates will always be 0. The upper estimates will always be ≥ 1. The sup of lower estimates will be 0; the inf of upper estimates will be 1. $0 \neq 1$, so $\int_0^1 f(x)\,dx$ does not exist.

However, when a function is continuous on $[a, b]$, then $\int_a^b f(x)\,dx$ will exist. (Theorem 5.6 in Bowman's Notes)

The relation between the two integrals is Fundamental Theorem of Calculus (Thm 5.9 of Bowman's Note). Suppose $f(x)$ is integrable and has an antiderivative $F(x)$ on the interval $[a, b]$. So $\frac{d}{dx}F(x) = f(x)$. Then $\int_a^b f(x)\,dx = F(b) - F(a)$.

That is how you evaluate definite integrals.

From Week 12, the important things are being comfortable with tangent line to graph, derivative and increasing/decreasing, max/min, indefinite integral, and the Fundamental Theorem of Calculus. Be able to calculate easy definite integrals. (We don't need trig substitutions or integration by parts, etc – all that is for next term.