MAT157 Lecture Notes

Arkyter

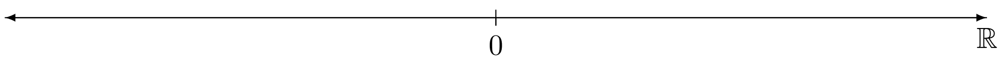
'23-'24 Fall & Winter Semesters

§1 Day 1: Introduction to Class (Sep. 8, 2023)

MAT157 is meant to focus on the more theoretical side of things, while 137 is more applied and focuses on computations more. This class covers foundations, and serves as an introduction to modern math, including logic, set theory, and so on.

§1.1 Informal Exploration of Properties of Numbers

In general, when we talk about arithmetic, we think of basic operators such as $+$ or $\cdot$ (addition and multiplication), or relations such as order, $a < b$, of which are defined on the real number line,



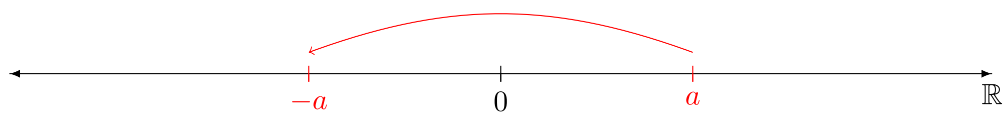
Of these operations, there are certain properties that come along with them (which are also defined in Spivak's book)[1]:

P1. (Additive Associativity) For any three numbers a, b, c, we have

$$a + (b + c) = (a + b) + c.$$

P2. (Additive Identity) There is a number, which we'll call "0", where $a + 0 = a = 0 + a$ for all a.

P3. (Additive Inverse) For every number a, there exists a number $-a$ where $a + (-a) = 0$.



P4. (Additive Commutativity) For any two numbers a, b, we have $a + b = b + a$. [2]

P5. (Multiplicative Associativity) For any three numbers a, b, c, we have

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

[1] avoiding formal logic and set notation for now

[2] i'm feeling like this is gonna lead into group theory lmao

P6. (Multiplicative Identity) There exists a neutral element[3], "1", such that

$$a \cdot 1 = a = 1 \cdot a$$

for all choices of a.

P7. (Multiplicative Inverse) There exists a number a^{-1} such that $a^{-1}a = 1 = aa^{-1}$.

P8. (Multiplicative Commutativity) For any two numbers a, b, we have $a \cdot b = b \cdot a$.

P9. (Distributive) For any three numbers a, b, c, we have $a \cdot (b + c) = ab + ac$.

To wrap up, we have a final question: can there be more than 1 identity?[4] Notice that, suppose we have $b + x = b$. By adding $-b$ to both sides, we obtain

$$-b + (b + x) = -b + b \qquad \text{(P1)}$$
$$(-b + b) + x = (-b + b) \qquad \text{(P3)}$$
$$x = 0.$$

I'm not sure where Prof. Burchard is going with this proof, so let me include mine instead.

Let $(S, \circ)$ be an algebraic structure. Suppose m, n are both identity elements of $(S, \circ)$. Then, $\forall k \in S$, we have $k \circ m = k = n \circ k$, where upon substituting $k = m, n$, we obtain $m = m \circ n = n \Rightarrow m = n$. Therefore, identities must be unique.

[3] confused about this wording, wouldn't it be called an identity? also note to self, left / right identity in the future might need to be differentiated

[4] ext. more than 1 two-sided identity in any algebraic structure?

§2 Day 2: Rest of Axioms (Sep. 11, 2023)

Today, we introduce the rest of the axioms (trichotomy and all that stuff). I was *not there* at lecture, so I'm just really going over what people sent me. Sorry!

§2.1 Fields

Formally, a field is any set of elements that satisfies the field axioms (aka, **P1** to **P9**) for both addition and multiplication. In a sense, fields are simply commutative division rings (in particular, all multiplicative elements have a multiplicative inverse). An example (provided in-class) of a non-commutative division ring would be the Quaternions ($\mathbb{H}$). Some examples of fields include, $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, but we also have $\mathbb{F}_2 = \{0, 1\}$, where we show the Cayley table:

$\mathbb{F}_2, +$	0	1
0	0	1
1	1	0

$\mathbb{F}_2, \times$	0	1
0	0	0
1	0	1

or $\text{GF}(4) = \{0, 1, \omega, \omega + 1\}$ with ω being a root of $x^2 + x + 1$.

$\text{GF}(4), +$	0	1	ω	$\omega+1$
0	0	1	ω	$\omega+1$
1	1	0	$\omega+1$	ω
ω	ω	$\omega+1$	0	1
$\omega+1$	$\omega+1$	ω	1	0

$\text{GF}(4), \times$	0	1	ω	$\omega+1$
0	0	0	0	0
1	0	1	ω	$\omega+1$
ω	0	ω	$\omega+1$	1
$\omega+1$	0	$\omega+1$	1	ω

§2.2 P10-12

P10. (Trichotomy) For every number a, one and only one of the following holds,

$$\begin{cases} a & = 0 \\ a & \in P \\ -a & \in P. \end{cases}$$

where P represents the set of all positive reals.

P11. (Additive Closure) If a and b are in P, then $a + b$ is in P.

P12. (Multiplicative Closure) If a and b are in P, then $a \cdot b$ is in P.

Now, we have a final practice problem.

Problem 2.1
Show that if a is not 0, then $a^2 \in P$.

The proof is left up to the reader, I am tired and I want to go hug my blåhaj.

§3 Day 3: Positive Sets in Fields and Ordering (Sep. 13, 2023)

Office hours are now open for MAT157 at Bahen Center (Floor 6) in Prof. Burchard's office, 12pm-1pm on Wednesdays. Aside from that, seems like something is happening with Tut 501 and Pra 501 so don't sign up for those.

§3.1 Positive Sets

Suppose there is a set P in a field K such that for all $a \in K$, exactly one of the 3 below holds (call this property **P10**, trichotomy):

$$\begin{cases} a & \in P, \\ -a & \in P, \\ a & = 0. \end{cases}$$

Suppose further that this set P has additive and multiplicative closure (call it **P11** and **P12**) as well. Now, there is a property we wish to show:

Theorem 3.1

Given an $a \in K$, we have $a \in P$ if and only if $a + a \in P$.

We can prove this by elimination. The forward implication ($a \in P \implies a + a \in P$) is resolved directly through additive closure. However, for the other way around, we first start by showing that $a \neq 0$. This follows directly from the fact that $0 + 0 = 0$, which is not in P. Now, for the case $-a \notin P$. Suppose otherwise that indeed $-a \in P$. Then, by **P11**, we see that $-(a + a)$ must also be in P, which is a contradiction with our initial assumption that $a + a \in P$ (violating **P10**). Since there is no other possibility, a must be in P. □

Problem 3.2 (Further Proposition)

Can a finite field F contain such a positive set P?

Given the characteristic p of F where[5]

$$\underbrace{a + a + \cdots + a}_{p \text{ summands}} = 0,$$

for all $a \in F$, we see that P cannot be closed, which would be a contradiction.

[5] if you don't know what a characteristic is, no worries, this wasn't discussed in class. figured i'd include it for a bit of funsies. also, the characteristic of $GF(4)$ from last lecture is 2.

§3.2 Ordering

Given a field K and subset P endowed with **P10-P12**, we wish to define the notion of order. For a, b in our field K, define the binary relations $<, \leq$ where we say

- $a > b$ if and only if $a - b \in P$.
- $a \geq b$ if and only if $a = b$, or $a - b \in P$.
- $a < b$ if and only if $b - a \in P$.
- $a \leq b$ if and only if $b = a$, or $b - a \in P$.

Where in the above, $<$, $\leq$, etc. are symbols representing order. In general, we say there is a total order on a certain set if all elements are comparable with each other, such as the binary relation $\leq$ defining a total order on $\mathbb{R}$. However, if there exist mutually incomparable elements, we say there is a *partial* order instead.[6]

> **Remark 3.3.** A field is called an *ordered field* if it is equipped with a strict total order '$<$' compatible with field operations, such as the rationals and reals.

Once again, we have another trichotomy for the binary relation $<$, where, if and only if K is totally ordered, we have, for all pairs $(a, b) \in K$, exactly one of the following is true:

$$\begin{cases} a < b, \\ a = b, \\ a > b. \end{cases}$$

We now prove a handful of the properties associated with it.

Theorem 3.4 (Transitivity)
Given $a, b, c \in K$, if $a < b$ and $b < c$, then $a < c$.

Given $a < b \implies b - a \in P$ and $b < c \implies c - b \in P$, we can write

$$c - a = (c - b) + (b - a) \in P. \tag{P11}$$

The sum of two elements of P is also in P. Therefore, $c - a \in P \implies a < c$.

Theorem 3.5
If $a < b$ and $c \in P$, then $ac < bc$.

We wish to show $bc - ac \in P$ (as it implies $ac < bc$). Following the distributive property, **P9**, we see that

$$bc - ac = \underbrace{(b - a)}_{\text{in } P} c \in P.$$

By multiplicative closure, our implication is complete.

[6] we didn't really cover this in class, just in office hours, so i won't write more.

§4 Day 4: Induction (Sep. 15, 2023)

The natural numbers, $\mathbb{N}$, is the set containing 1 and all successors of n for all $n \in \mathbb{N}$; in particular, this implication yields to $\mathbb{N} = \{1, 2, 3, \ldots, \}$[7]. Formally, we may write:

$$\mathbb{N} : \begin{cases} 1 \in \mathbb{N} \\ \forall n \in \mathbb{N} & \implies n + 1 \in \mathbb{N}, \end{cases}$$

in which equality of successors $n + 1 = m + 1$ implies $n = m$, and that 1 is not the successor of anyone.

§4.1 Induction

The induction principle is that for any subset $A \subseteq \mathbb{N}$, if $1 \in A$ and for all $n \in A$, its successor belongs to A as well, we know $A = \mathbb{N}$. In general, when inducting, we want to prove statements by first demonstrating a base case, then generalizing onwards; here's an example:

Problem 4.1 (Gauss Sum)

For all $n \in \mathbb{N}$, show that

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

We claim that

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

In the base case, let $n = 1$. It is immediately verifiable that

$$1 = \frac{1 \cdot 2}{2}.$$

For the inductive step, suppose we are given $n = k$, and we wish to prove the equation for case $n = k + 1$.

$$\sum_{i=1}^{k+1} i = \left(\sum_{i=1}^{k} i\right) + (k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{(k+1)(k+2)}{2}.$$

Since we know case $n = k \in \mathbb{N}$ implies case $n = k + 1$, we have proven the equation.

[7] does this class consider 0 a natural?

§5 Day 5: Properties of Sets; Summation and Product; Binomial Coefficients (Sep. 18, 2023)

[8]A *set* is defined as a collection of objections; for example,

$$\underbrace{\{1,2,3\} = \{1,3,2\} = \{3,1,2,1,1,1\}}_{\text{order/repetition do not matter}}.$$

as mentioned above, order and repetition do not matter in sets: we are describing a collection of distinct objects. For example, the naturals $\mathbb{N}$ (counting numbers) are defined as an infinite set

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

There's a handful of notation that comes with the notion of sets, of which are listed below (also covered in **MAT240** notes):

$$\begin{aligned}
x \in S &: x \text{ lies in } S.\\
A \subset B &: A \text{ is a subset of } B; \text{ formally, } \forall a \in A, a \in B.\\
A \subseteq B &: A \text{ is a subset, or is equal to } B;\\
A \subsetneq B &: \text{denotes the notion of strict/proper subset, is equivalent to } \subset.\\
x \in A \cap B &: \text{if and only if } x \in A \text{ and } x \in B.\\
x \in A \cup B &: \text{if and only if } x \in A \text{ or } x \in B.\\
A \times B &= \{(a,b) \mid a \in A, b \in B\}.
\end{aligned}$$

Given $A \subset S$, A^C is the set complement, and denotes the set $S \setminus A = \{x \in S \mid x \notin A\}$.

And finally, $\mathcal{P}(A)$ denotes the *Power Set* of A, which is the set of all subsets of A. An example of the power set is $\mathcal{P}(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$.

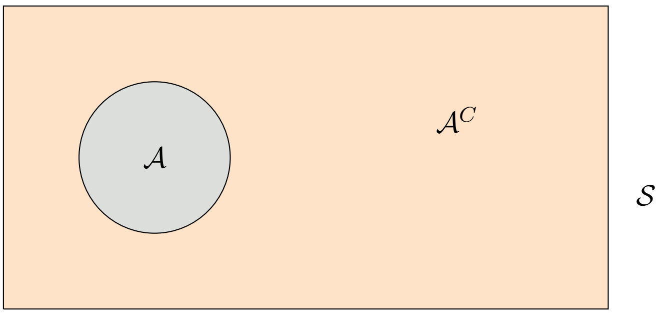


Figure 1: Set Complement

> **Remark 5.1.** Jimmy poked me about it in class but $\mathcal{P}(\mathbb{N})$ is uncountable by Cantor's Theorem; there is no surjection $A \to \mathcal{P}(A)$ for any given set A.

§5.1 Summation and Product, Binomial Coefficients (???)

Summation and product operate in an inductive nature;

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n = a_n + \sum_{i=1}^{n-1} a_i,$$

[8]god i'm addicted to semicolons

$$\prod_{i=1}^{n} a_i = a_1 a_2 \dots a_n = a_n \prod_{i=1}^{n-1} a_i.$$

As for binomial coefficients, given $1 \leq k \leq n$[9], we denote

$$\binom{n}{k} \underbrace{:=}_{defines} \frac{n!}{(n-k)!k!}$$

with the conventions that $\binom{n}{0} = 1$, and $\binom{n}{k} = 0$ if $k > n$. As an exercise, show that

Problem 5.2 (Pascal's Identity)

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

The proof is left up to the reader; try it combinatorially (done in **MAT344** notes day 2 or algebraically).

[9]these can be reals, but for now are naturals

§6 Tutorial 1: Induction Paradox; Binomial Coefficients and Field Review (Sep. 19, 2023)

We begin by providing a prototypical example of a *paradoxical proof*:

Theorem 6.1 (All Horses are the Same Color)

Suppose we have a collection of n horses and we wish to show that they are all of the same color. We proceed to demonstrate this by inducting on n:

- For the base case, we see that a single horse must have the same color as itself.
- For the inductive step, assume that any k horses are of the same color; in a collection of $k+1$ horses, we may select the first k or the last k horses (omitting the last and first horses respectively); in this manner, we see that all the horses in our $k+1$ collection must be of the same color.

However, we see that the proof itself is fallacious; our inductive step does not cover the base case or the case $n = 2$; it only works for $n \geq 3$, which implies that our proof is invalid.

§6.1 Binomial Coefficients and Combinatorial Arguments

Honestly, not too sure why we're going over this, but sure:

Problem 6.2 (Pascal's Rule)

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Problem 6.3 (Choose Function Inversion)

$$\binom{n}{k} = \binom{n}{n-k}.$$

Problem 6.4 (Stars and Bars)

How many ways are there to choose a positive integer 4-tuple (x_1, x_2, x_3, x_4) such that

$$x_1 + x_2 + x_3 + x_4 = 10?$$

These are left up to the reader as exercise.

§6.2 Field Exercises

Some more!

1. Show that

$$\underbrace{(1+1+\cdots+1)}_{m \text{ summands}}\underbrace{(1+1+\cdots+1)}_{n \text{ summands}} = \underbrace{(1+1+\cdots+1)}_{mn \text{ summands}}$$

2. Show that the field characteristic is necessarily prime or 0. (Additionally, give an example of a non-finite field with nonzero characteristic[10]).

3. Show that in a finite field, there exists $m, n \in \mathbb{N}$ such that

$$\underbrace{(1 + 1 + \cdots + 1)}_{m \text{ summands}} = \underbrace{(1 + 1 + \cdots + 1)}_{n \text{ summands}}$$

even though $m \neq n$.

4. Show that finite fields necessarily have positive characteristic.

[10] i just added this in

§7 Day 6: More Binomial Coefficients (Sep. 20, 2023)

Today we go over more binomial coefficients. A quick comment made in class today, but we will proceed soon to construct $\mathbb{R}$ with Dedekind cuts.

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

where we take the first four for granted, and we seek to properly construct $\mathbb{R}$.

§7.1 More Binomial Coefficients

Last lecture, we went over the definition of factorials and choice; recursively defined as

$$n! : 1! = 1$$
$$(n+1)! = n! \cdot (n+1)$$

or alternatively $\prod_{j=1}^{n} j$. Note that i, j may be used interchangeably as iterators; they may be thought of as the variables in "for" loops from CS, if that's your thing. For the choose function or binomial coefficient, given $0 \leq k \leq n$, we have

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Which is also equipped with the handy identity

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

for $0 \leq k \leq n-1$.

Now, let's go over the propositions from Lecture 5:

1. $\binom{n}{k}$ equals the number of subsets of $\{1, 2, \ldots, n\}$ with k elements.
2. It is also equal to the number of up-right (Manhattan metric) paths from $(0, 0)$ to $(k, n-k)$.
3. It is also equal to the number of binary strings $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ where $a_i \in \{0, 1\}$, with $i \in \{1, 2, \ldots, n\}$. If you're in **MAT344**, this should be familiar (read: $\{0, 1\}$-strings).
4. The coefficient of x^k in $(1+x)^k = \sum a_k x^k$.

These propositions serve as elementary exercises, and can all be quickly proven with combinatorial arguments; take it as practice to argue them rigorously. Alternatively, see Karim's Notes for the solutions.

§8 Day 7: Bounded Set, Infimum / Supremum; Dedekind Cuts (Sept. 22, 2023)

More reminders on how binomial coefficients work! Remember that

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$
$$\binom{0}{0} = 1$$
$$\binom{n}{n} = 1$$

Another reminder on set ordering, that any subset of a totally ordered set $\mathbb{Q}, \mathbb{R}$, etc. is also totally ordered given that they follow the same order relation. We also have, for the order relation $\leq$, for an ordered field F,

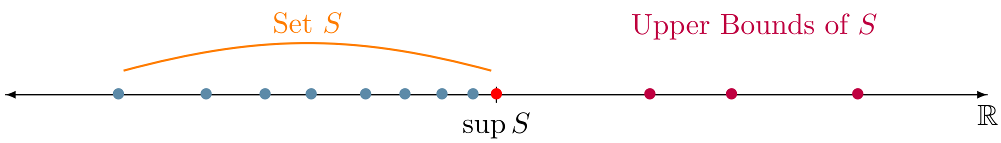
- $a \leq a$ for all $a \in F$,
- Transitivity holds! If $a \leq b$ and $b \leq c$, then $a \leq c$.
- *Something else, reminder to add!*

§8.1 Bounded Sets

First, we start with a bit of notation. Let P be a (partially) ordered set, and let A be a subset of P.

- We say $x \in P$ is an upper bound for A if and only if $a \leq x$ for all $a \in A$.
- We say $s \in P$ is the *supremum* of A if and only if x is the *smallest* upper bound for A; that is, for every $a \in A$, $a \leq s$, and if $a \leq x$ for some upper bound x for all $a \in A$, then $s \leq x$. We denote this as $\sup(A)$.
- We say $\sup(A)$ is a *maximum* of A if $\sup(A) \in A$. [11]
- We say $y \in P$ is a lower bound for A if and only if $a \geq y$ for all $a \in A$.
- We say $i \in P$ is the *infimum* of A if and only if i is the *largest* lower bound for A; that is, for every $a \in A$, $a \geq i$, and if $a \geq y$ for some lower bound y for all $a \in A$, then $i \geq y$. We denote this as $\inf(A)$.
- We say $\inf(A)$ is a *minimum* of A if $\inf(A) \in A$. [12]

Naturally, the supremum and infimum are unique if they exist. Moreover, we also say that a set is bounded if it is bounded above and below: i.e., it has an upper and lower bound. (In the diagram below, Infimum works the other way around... use your imagination)



[11] Note that the supremum doesn't have to actually be an element of A, just that if it is, we say that it is the maximum.

[12] naturally, the same goes for infimum. it don't gotta be in A, just, yk, "hella neat" if it is

Example 8.1

Show that the set $A = \{x \in \mathbb{R} \mid x^2 < 2\}$ is bounded.

Example 8.2

Is the set of all prime numbers bounded above?

Now, we move to more properties of the subsets of the reals:

P13. (Archimedean Property) Given any positive x, y in an ordered field F, there is an integer $n > 0$ such that $nx > y$.[13]

P14. (Existence of Supremum) Every non-empty set of real numbers that is bounded above has a least upper bound / supremum.

With these postulates, we may now define a Dedekind Cut. A Dedekind Cut is a partition of the rationals into a subset α such that

- α is nonempty,
- $\alpha \neq \mathbb{Q}$ (it must be a proper subset),
- α is bounded above,
- If $x \in \alpha$ and $y \leq x$ (with both $x, y \in \mathbb{Q}$, then $y \in \alpha$,
- If $x \in \alpha$, then there exists a $y \in \alpha$ such that $y > x$ (no greatest element).

Given a cut $\alpha \subset \mathbb{Q}$, we know α is bounded above, which means we can find a rational $\frac{p}{q}$ ($p \in \mathbb{Z}, q \in \mathbb{N}$) that is an upper bound. Then, $|p|$ is an upper bound. Define a new set $\beta = \{x \mid x < |p| + 1\}$; clearly, $\alpha \subset \beta$ (we say $\alpha < \beta$ if $\alpha \subset \beta$). Now, we may identify the cut

$$\alpha q = \{x \in \mathbb{Q} \mid x < \alpha\} \text{ with } q.$$

Not sure what happened, I think Almut got rushed at the end. I'll revise this section so it makes sense later.

[13] as an exercise, use this to prove that $\mathbb{Q}$ is dense in $\mathbb{R}$.

§9 Day 8: Dedekind Cuts Pt. II (Sept. 25, 2023)

Today we formally go over Dedekind Cuts.[14]

§9.1 Dedekind Cuts

The idea here is that we want to construct the real numbers from $\mathbb{Q}$ (which we shall take as given for now). Intuitively, a *Dedekind cut* is a subset of the rationals that represents a single real number; we may define any real entirely in terms of sets in $\mathbb{Q}$ with basic set operations.[15] Define a *Dedekind cut* $\alpha \subset \mathbb{Q}$ as a set satisfying,

- α is non-empty.
- α is bounded above. In particular, $\alpha \subsetneq \mathbb{Q}$ (it cannot be the set of all rationals).
- Let $x, y \in \mathbb{Q}$. If $x \in \alpha$ and $y < x$, then $y \in \alpha$. (α is "closed downwards")
- α has no maximum. That is, if $x \in \alpha$, then we may always find $y \in \alpha$ such that $y > x$.

For the set theory nerds (you know who you are), we say a Dedekind cut L is a nonempty proper subset of $\mathbb{Q}$ that has no maximal element, and satisfies

$$(\forall a, b \in \mathbb{Q})[a \in L \wedge b < a] \implies [b \in L].$$

We may illustrate it as such,

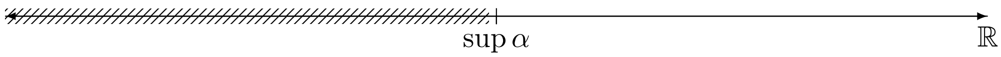


Figure 2: The shaded region is our Dedekind Cut (over $\mathbb{Q}$).

> **Remark 9.1** (Dedekind Cuts are Unique for all Reals)**.** As a bit of intuition that was left out of lecture today, remember that each unique real corresponds to a unique Dedekind cut! Suppose otherwise; say we have distinct reals $a, b \in \mathbb{R}$. Then, there exists a rational $a < p/q < b$; by definition of Dedekind cuts, p/q is in the cut of b, but is not in the cut for a, and thus they are necessarily distinct. □

Having defined our cuts, we will now define our usual operations and order relation, addition, multiplication, and "$\geq$" to show that $\mathbb{R}$ is an ordered field satisfying **P1** to **P9** (field axioms of addition and multiplication), **P10** to **P12** (propositions for order), is archimedean (**P13**), and satisfies the supremum axiom (**P14**). We will do this by extending the operations on the rationals (which we know already is an ordered field) onto the reals:

$\mathbb{Q}$	$\mathbb{R}$
$q \in \mathbb{Q}$	Dedekind cuts $\alpha \subset \mathbb{Q}$; we may write it as $\alpha_q = \{x \in \mathbb{Q} \mid x < q\}$.[16]
$+$	$\alpha + \beta = \{x + y \mid x \in \alpha, y \in \beta\}$
"$\leq$"	$\alpha \leq \beta$ if $\alpha \subseteq \beta$.
$\cdot$	Complicated; see below!

[14] funny story, Almut says she regrets todays lecture lol

[15] note that this is possible because $\mathbb{R}$ is archimedean and $\mathbb{Q}$ is dense in $\mathbb{R}$.

Before we may move onto multiplication, let's define a few more properties regarding the reals. Let the neutral element be $0 = \alpha_0 = \{x \in \mathbb{Q} \mid x < 0\}$, and we claim that for all $\alpha \in \mathbb{R}$, we have $\alpha + \alpha_0 = \alpha$. We see that this follows from the fact that given any $z \in \alpha + \alpha_0$, we have $z = x + y$ with $x \in \alpha$ and $y < 0$, hence $z < x \implies x + y \in \alpha \implies \alpha + \alpha_0 \subseteq \alpha$. In the converse direction, we see that for all $z \in \alpha$, we want to find $x \in \alpha, y < 0$ such that $z = x + y$. Since z is not maximal, we may pick $x > z$ such that

$$\underbrace{z}_{z \in \alpha} = \underbrace{x}_{x \in \alpha} + \underbrace{(z - x)}_{y < 0} \in \alpha + \alpha_0,$$

which resolves our claim. Now, we proceed onto additive inverse: given $\alpha \in \mathbb{R}$, define an *anti-cut* $\bar{\alpha}$ by

$$\bar{\alpha} = \begin{cases} \mathbb{Q} \setminus \alpha & \text{if } \alpha \text{ irrational,} \\ (\mathbb{Q} \setminus \alpha) \setminus \min\{\mathbb{Q} \setminus \alpha\} & \text{if } \alpha \text{ rational.} \end{cases}$$

Realistically, in both cases, when we take the set complement of α with respect to the rationals, we're obtaining a case of "everything greater than or equal to α"; this "equal to" only exists of α represents a rational, which we must remove. In this manner, we

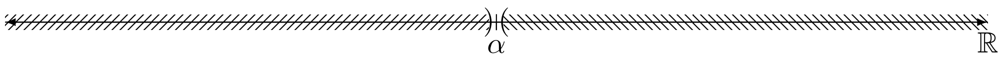


Figure 3: I apologize for the terrible diagram; this is meant to illustrate cuts and anticuts.

may define $-\alpha \in \mathbb{R}$ as $\{-x \mid x \in \bar{\alpha}\}$. With this, we may write

$$\begin{aligned} \alpha \cdot \beta &= \beta \cdot \alpha \text{ if } \alpha, \beta > 0 \\ \alpha_0 \cdot \beta &= \alpha_0 \\ \alpha \cdot \beta &= -(-\alpha \cdot \beta) \text{ if } \alpha < 0. \end{aligned}$$

and class ended here! I'll add multiplication stuff later if I have time.

§10 Tutorial 2: Definition of Symbols; Infimum and Supremum (Sep. 26, 2023)

If you were following my notes for **MAT240** as well, all symbols have been defined in there; supremum and infimum have been defined on **Day 7** of my **MAT157** notes (page 12). Nothing really happened in my tutorial, it was kinda just a review.

§11 Day 9: Functions (Sep. 27, 2023)

The formal definition of a function is, a function $f : X \to Y$ is determined by a subset $\Gamma \subset X \times Y$ such that

- For every entry $x \in X$, there exists a $y \in Y$ such that $(x, y) \in \Gamma$.
- If $(a, b) \in \Gamma$ and $(a, c) \in \Gamma$, then $b = c$ (functions have unique outputs).

For example, a permutation is a bijective function $\sigma : X \to X$; for example, let $X = [5] = \{1, 2, 3, 4, 5\}$, we may let

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

Recall from day 1 of **MAT240** that injectivity, surjectivity, and bijectivity are important properties of functions, as well as the notion of codomain vs. range/image (most of class today was focused on these notions, so I'll omit the repetitive details).

§12 Day 10: Functions Cont.; Introduction to Limits (Sep. 29, 2023)

Previously, we defined $f \circ g(x) = f(g(x))$ for all $x \in \mathrm{dom}(g)$, provided that the image of g was a subset of the domain of f. Function composition is associative by definition (exercise: check by expanding $(f \circ g) \circ h(x) = f \circ (g \circ h)(x)$).

§12.1 Fractional Linear Transformation / Matrix Transformation

We may visualize fractions as matrices; for example, let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a function where $f(x, y) = (ax + b, cx + d)$ for $a, b, c, d \in \mathbb{R}$. Let's first establish the below,

$$f(x,y) = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix} = \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}}_{\text{matrix multiplication}} .$$

That is,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \left(\; A \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} \; \middle| \; B \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} \; \right) \implies f_A \circ f_B = f_{AB}.$$

[17] For the fractional linear transformation[18], write

$$f(x) = \frac{ax+b}{cx+d},$$

where $ad - bc = 0$[19]; the domain of f would be $\mathbb{R}$ excluding the zeroes of $cx + d$. We may write in terms of matrices just as above. Now, for the purposes of function composition with matrix transformation, set f, g such that $f = f_A, g = f_B$ as above,

$$f(x) = \frac{a_{11}x + a_{12}}{a_{21}x + a_{22}}, \qquad g(x) = \frac{b_{11}x + b_{12}}{b_{21}x + b_{22}},$$

such that $a_{11}a_{22} - a_{12}a_{21} \neq 0$ and $b_{11}b_{22} - b_{12}b_{21} \neq 0$. Then we see $f \circ g$ may be expressed as

$$\begin{aligned} \frac{a_{11}\left(\frac{b_{11}x+b_{12}}{b_{21}x+b_{22}}\right) + a_{12}}{a_{21}\left(\frac{b_{11}x+b_{12}}{b_{21}x+b_{22}}\right) + a_{22}} &= \frac{a_{11}(b_{11}x + b_{12}) + a_{12}(b_{21}x + b_{22})}{a_{21}(b_{11}x + b_{12}) + a_{22}(b_{21}x + b_{22})} \\ &= \frac{(a_{11}b_{11} + a_{12}b_{21})x + (a_{11}b_{12} + a_{12}b_{22})}{(a_{21}b_{11} + a_{22}b_{21})x + (a_{21}b_{12} + a_{22}b_{22})}, \end{aligned}$$

which indeed is just as we showed above.

§12.2 Sequences and Limits

Let X be a sequence: a sequence is formally defined as a function

$$\begin{aligned} a &: \mathbb{N} \to X \\ n &\mapsto a_n \in X, \end{aligned}$$

We may write the above as $\{a_n\}$. Suppose that $\{a_n\}$ is in $\mathbb{R}$. Then, supposing that this sequence converges to some real L, we say that $\lim_{n\to\infty} a_n = L$ if, for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that if $n > N$, we have $|a_n - L| < \epsilon$. (note that N depends on the choice of ϵ!)[20]

[17]matrices dun goof up can't insert some vertical bar ;w;
[18]time to study about riemann spheres and meromorphic functions ehe
[19]condition such that we don't end up with trivial constants or absurd edge cases. see: determinant
[20]see day 11 for intuition

§13 Day 11: Formal Definition of a Limit Cont. (Oct. 2, 2023)

Let's start with some definitions (just as a recap)!

(a) A sequence $\{a_n\}_{n\in\mathbb{N}}$ in a set X is a function $f:\mathbb{N}\to X$ where $n\mapsto a_n\in X$.

a) We say $\{a_n\}$ in $\mathbb{R}$ is bounded above if its range $\{a_n \mid n\in\mathbb{N}\} < M$ for some upper bound $M\in\mathbb{R}$ (see 8.1 if confuzzled).

(b) A sequence is *monotone* if either

$$\forall n\in\mathbb{N},\ \begin{cases} a_n \le a_{n+1} & \text{(monotonically increasing / non-decreasing)} \\ a_n \ge a_{n+1} & \text{(monotonically decreasing / non-increasing)} \end{cases}.$$

§13.1 Limits and Convergence

A sequence $\{a_n\}\in\mathbb{R}$ converges to some number $L\in\mathbb{R}$ if for all $\varepsilon>0$, there exists $N\in\mathbb{N}$ such that for all $n>N$, we have $|a_n-L|<\varepsilon$. Then, we may write $\lim_{n\to\infty} a_n = L$ (it is the limit of the sequence). We also write $a_n\to L$ as $n\to\infty$.

Remark 13.1 (Intuition for Limits)**.** Treat it like a game, where I tell you how close the elements of your sequence should be (ε) to the limit, and you tell me when those elements are close enough (N). Your choice of N depends on how tight I want my ε to be, but you should probably have some way to give me a response no matter whatever ε I pick.

Remark 13.2 (Random thought I had in class)**.** Given the sequence $\{a_n\}\in\mathbb{R}$, when you pick your N for some choice of ε, there *is no* "what if my sequence decided to stray away from the limit at the last second" kinda thing. If your choice of N allows the sequence to do that, then you just have to modify what you chose to be N; if there is no suitable N *at all*, then your sequence doesn't converge.

Now for a list of examples from class:

1. Let us have a sequence defined by $a_n=\frac{1}{n}$. We wish to show that $\lim a_n = 0$. For any $\varepsilon>0$, we have (by the archimedean property on $\mathbb{R}$) that there exists some $N>\frac{1}{\varepsilon}$. Then for all $n>N$, we have

$$|a_n-L| = \frac{1}{n} < \frac{1}{N} < \varepsilon.$$

where $L=0$ as desired. Since ε is arbitrary by definition, we have $\frac{1}{n}\to 0$ as $n\to\infty$. □

2. Let us have the sequence $(a_n): a_n = (-1)^n$, with $n\ge 1$. To show that it diverges, we first prove the lemma,

Lemma 13.3 (Sequence convergence implies subseqeunce convergence)

First, we define a subsequence; a sequence $\{b_n\}$ is a subsequence of sequence $\{a_n\}$ if there exists a "re-indexing" sequence $\{n_k\}$ in $\mathbb{N}$ such that $k\mapsto n_k$ with all $n_{k+1}>n_k$ (as a result, $n_k\ge k$ for all k) such that $b_k = a_{n_k}$.

To prove the lemma, take a sequence $\{a_n\}$ with limit L and an arbitrary subsequence $\{b_n\}$. Using the reindexing from above, let $\varepsilon>0$; then there exists N such that $n>N \implies |a_n-L|<\varepsilon$; for whatever choice $k>N$ (since $n_k\ge k$) we have $|b_k-L| = |a_{n_k}-L|<\varepsilon$ as well. □

Using the lemma, we may pick constant subsequences $(b_k) : b_k = a_{2k} = 1$ (then $b_k \to 1$) and $(c_k) : c_k = a_{2k-1} = -1$ (then $c_k \to -1$). Since there exists subsequences with conflicting limits we know $(-1)^n$ diverges.

Theorem 13.4 (Monotone Convergence Theorem)

If a sequence is non-decreasing and bounded above (by a supremum), then said sequence will converge to the supremum. Likewise, if a sequence is non-increasing and bounded below (by an infimum), then said sequence will converge to the infimum.

We prove the bounded above version of this theorem; the other will be left as an exercise to the reader. Let $\{a_n\}_{n\geq 1}$ be a bounded sequence; suppose $\lim a_n = \sup\{a_n \mid n \leq N\}$ (call the supremum s, which exists because $\{a_n\}$ is bounded above). Let $\varepsilon > 0$; we know $s - \varepsilon$ cannot be an upper bound (since s is a supremum; *least* upper bound), this means there must exist some N where $a_N > s - \varepsilon$; since $\{a_n\}$ is monotonically increasing, we have that all $a_{n\geq N} > s - \varepsilon$. Then we have

$$s - \varepsilon < \underbrace{a_n < s + \varepsilon}_{s+\varepsilon \text{ is bound}} \implies |a_n - s| < \varepsilon$$

for sufficiently large n.[21]

[21] i know almut used $\varepsilon/2$ but i don't really see why i would use it for this particular example; this is my proof but I think she proved it differently. well, both work i guess-

§14 Tutorial 3: Review of Dedekind Cuts and Function Properties (Oct. 3, 2023)

I'm just going to go over the quick details that haven't been covered already:

- The *supremum* of a set (closed downwards) of Dedekind cuts $A = \{\alpha_{a_1}, \alpha_{a_2}, \dots, \alpha_{a_n}\}$ (where (a_n) is a sequence of reals[22]) is given by the union; as in,

$$\bigcup_{i=1}^{n} \alpha_{a_i} = \sup(A).$$

- In a similar manner, the *infimum* of a set of Dedekind cuts $A = \{\alpha_{a_1}, \alpha_{a_2}, \dots, \alpha_{a_n}\}$ (where (a_n) is just as above) is given by the intersection; as in,[23]

$$\bigcap_{i=1}^{n} \alpha_{a_i} = \inf(A).$$

- Intuition: really, just see it as $x \cup y = \max\{x, y\}$ and $x \cap y = \min\{x, y\}$.

Problem 14.1 (AAAA)
Let us have a set of Dedekind cuts given by $A = \{\alpha_{1/n} \mid n \in \mathbb{N}\}$; explain why the following is not true,

$$\bigcap_{i=1}^{n} \alpha_{1/n} = \inf(A) = \alpha_0.$$

§14.1 Functions

A word on piecewise functions: they're not really a description of a function, it's just how we interpret what's going on. They *are* very convenient though, for example, the Dirichlet function

$$1_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

Another thing went over in class: functions can be labeled even or odd based on the property,

- A function f is even if $f(-x) = f(x)$ for all x in the domain of f.
- A function f is odd if $f(-x) = -f(x)$ for all x in the domain of f.

Remark 14.2. The notion of even/odd here isn't like the naturals; we can have functions such is the zero function, which is even and odd (exercise: verify that this is indeed the only such function); some functions can also be neither even nor odd. Ex: $f(x) = x^2 + x^3$.

Another interesting remark from class: any function can be written as the unique sum of an even and odd function in this manner,

$$f(x) = \underbrace{\frac{f(x) + f(-x)}{2}}_{\text{even}} + \underbrace{\frac{f(x) - f(-x)}{2}}_{\text{odd}}.$$

[22] hey, not too sure about this. someone check for me?
[23] now that i'm doing the homework i think this is wrong i'm gonna cry

§15 Day 12: Sequence Convergence (Oct. 4, 2023)

Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$; remember from last time that if such a sequence is bounded and monotone, then it is convergent. As a recap, let's prove that if the sequence is monotone *down*, then it is also convergent. In the monotone up case, we've shown that $\lim a_n = \sup a_n$; so now, we want to show that $\lim a_n = \inf a_n$. Consider defining a second sequence $b_n := -a_n$ for $n \in \mathbb{N}$ first. Then we have

$$\lim a_n = \lim(-b_n) = -\lim b_n = -\sup b_n = \inf -b_n = \inf a_n,$$

which we may write since b_n is a monotone up sequence. However, to make things a little more concrete, we also invoke the following lemmas for the above,

Lemma 15.1 (Negation of Limit)

If $\lim x_n = L$, then $\lim(-x_n) = -L$. Given $\varepsilon > 0$, since $\lim x_n = L$, there exists some $N \in \mathbb{N}$ such that $n > N \implies |x_n - L| < \varepsilon$. But also, at the same time,

$$|(-x_n) - (-L)| = |-(x_n - L)| = |x_n - L| < \varepsilon.$$

Since our choice of ε was arbitrary, we see $\lim(-x_n) = -L$.

Lemma 15.2 (Infimum/Supremum)

Let $A \in \mathbb{R}$ be bounded above. and define a set $B = \{-a \mid a \in A\}$. Then

1. If M is an upper bound for A, then $-M$ is a lower bound for B.
2. $\sup(A) = -\inf(B)$.

We start by proving the first statement; the fact that M is an upper bound for A means we have $a < M$ for all $a \in A$, which we may reverse the inequality into $-a > -M$ for all $a \in A$ as well. This means we have $b \geq -M$ for all $b \in B$, and thus $-M$ is a lower bound for B. Similarly, we may argue for point (2) that if S is the supremum for A, then $-S$ naturally is also the infimum for B.

§15.1 Newton-Heron Method

Define a recurring sequence where $x_1 = 1$, and $x_{n+1} = \frac{1}{2}(x_n + \frac{2}{x_n})$ for all $n \geq 1$. We claim that $\lim x_n = \sqrt{2}$; start by writing function $f(x) = \frac{1}{2}(x + \frac{2}{x})$ (such that $f(x) = x_{n+1}$), and notice that the minimum of f lies at $f(\sqrt{2}) = \sqrt{2}$ (by the AM-GM inequality, we have $f(x) \geq \sqrt{x - \frac{2}{x}}$). We claim that for $n \geq 2$, we have $x_{n+1} < x_n$:

$$\begin{aligned}(x_n - x_{n+1}) &= x_n - \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) \\ &= \frac{1}{2}\left(x_n - \frac{2}{x_n}\right) \\ &= \underbrace{\frac{1}{x_n}}_{>\sqrt{2}}(x_n^2 - 2) > 0.\end{aligned}$$

By our earlier established lemmas, the limit $L = \lim x_n = \inf x_n \mid n \geq 2$ exists; and since x_n is decreasing for $n \geq 2$, and x_n is bounded below; clearly, $L \geq 2$ since $x_n \geq 2$ for all $n \geq 2$. Then we may write

$$
\begin{aligned}
(x_{n+1} - \sqrt{2}) &= \frac{1}{2}\left(x_n + \frac{\varepsilon}{x_n} - 2\sqrt{2}\right) \\
&= \frac{1}{2x_n}(x_n^2 - 2\sqrt{2}x_n + 2) \\
&= \frac{2}{x_n}(x_n - \sqrt{2})^2 \to 0
\end{aligned}
$$

Therefore, all lower bounds are less than or equal to $\sqrt{2}$, and so we are done.

§16 Day 13: Behavior of Sequences: Subsequences and Divergence (Oct. 6, 2023)

Recall from last time that given a sequence $\{a_n\}_{n\in\mathbb{N}}$, we say $a_n \to L$ if, for all $\varepsilon > 0$, there exists some natural N such that for all $n > N$, we have $|a_n - L| < \varepsilon$. To express this, we write $\lim a_n = L$. With this, we say $\{a_n\}$ is convergent if $a_n \to L$ for some $L \in \mathbb{R}$; otherwise, it is divergent. Furthermore, we also define a subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$ where $\{n_k\}$ is a strictly increasing sequence of naturals;

Example 16.1 (Subsequence of Prime Squares)

Take the sequence $a_n = n^2$. We may define a subsequence (p_k) where $p_k = a_{n_k}$, with n_k being the kth prime. The first few terms of (p_k) are then

$$2^2, 3^2, 5^2, 7^2, 11^2, \dots.$$

Visualization (subsequences are just a reindexing of the original sequence, really):

$$\underbrace{\mathbb{N} \overbrace{\to}^{\{n_k\}_{k\in\mathbb{N}}} \mathbb{N} \overbrace{\to}^{\{a_n\}} \mathbb{R}}_{\text{subsequence of } \{a_{n_k}\}}.$$

§16.1 Sequence Divergence to Infinity

Divergence to infinity: suppose we have some sequence $\{a_n\}_{n\in\mathbb{N}}$ where $a_n = q^n$ for some given $q > 1$. Since $a_n < a_{n+1}$, by Bernoulli's inequality,

$$a_n = (1 + (q-1))^n \geq 1 + n(q-1).$$

But first, let's establish our condition for divergence to infinity: given a sequence $\{a_n\}_{n\in\mathbb{N}}$, we say it diverges to infinity if, for all m, we may find an $N \in \mathbb{N}$ such that for all $n > N$, we have $a_n > m$. Back to the problem; we may establish that by the Archimedean property, we may always find some $N > \frac{m}{q-1}$,[24] and therefore, $a_n \to \infty$. $\square$

Remark 16.2. If $a_n \to -\infty$, then $-a_n \to \infty$.

Lemma 16.3

Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. $\{a_n\}_{n\in\mathbb{N}}$ is unbounded above if and only if there exists a subsequence $a_{n_k} \to \infty$.

First, suppose that $a_{n_k} \to \infty$. Then $\{a_{n_k} \mid k \in \mathbb{N}\} \subset \{a_n \mid n \in \mathbb{N}\}$, which has no upper bound by definition. For the forward implication, suppose that $\{a_n\}$ is not bounded above; then for all $k \in \mathbb{N}$, we may find an $n \in \mathbb{N}$ such that $a_n > k$. Set $n_k := n + n_{k-1}$, so that by construction, $n_{k+1} > n_k$ for all $k \in \mathbb{N}$ and $a_{n_k} > k$. Therefore, $\{a_{n_k}\}_{k\in\mathbb{N}}$ is indeed such a subsequence with $\lim a_{n_k} \to \infty$. $\square$

To illustrate this, we may take sequence $a_n = (-2)^n$. Since it has a subsequence $a_{2k} = 2^{2k} \to \infty$, we say a_n diverges. Another example would be $b_n = 2^n + (-2)^n$, which has subsequence $b_{2k} = 2 \cdot 2^{2k} \to \infty$, which implies b_n diverges as well by our lemma.

[24] my brain is not working and i can't see why $\frac{m}{q-1}$ works, will revisit and rethink...

§16.2 Classification of Divergent Subsequences

We now write down two important properties of divergent subsequences:

- Every unbounded sequence diverges (in the same way, every convergent sequence is bounded).
- Moreover, if a sequence has two subsequences that converge to two different limits, then the sequence diverges; as in, we may find subsequences $\{b_k\}_{k\in\mathbb{N}}$ and $\{c_k\}_{n\in\mathbb{N}}$ such that $b_k \to L_1$ and $c_k \to L_2$.

Suppose we have a sequence $\{a_n\}_{n\in\mathbb{N}}$, and write the subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$ where, given any $\varepsilon > 0$, we may choose N such that $|a_{n_k} - L| < \varepsilon$ for all $k > N$. Then $a_{n_k} \to L$. This is only possible if a_n was convergent in the first place. In a way, if we may write $\lim a_n = L$ and L', then take $\varepsilon > 0$:

$$\exists N \in \mathbb{N} \text{ such that } |a_n - L| < \frac{\varepsilon}{2} \ \forall n > N,$$
$$\exists N' \in \mathbb{N} \text{ such that } \left|a_n - L'\right| < \frac{\varepsilon}{2} \ \forall n > N'.$$

Then for $n > \max\{N, N'\}$, we have

$$\left|L - L'\right| \le |L - a_n| + |a_n - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \square$$

which also means $L = L_1 = L_2$ if L exists; otherwise, divergent sequences don't have such a thing.

§17 Tutorial 4: Midterm 1 Review (Oct. 10, 2023)

There will be **6 questions**[25] on the midterm; half straightforward, worth 10 marks each, covering:

- Definitions (free marks)
- Recall things from class (I'm going to assume proofs, techniques, etc)
- New proofs (as expected)
- **1 question** on Dedekind cuts, no more no less.

There will also be at least one question on induction and proof by contradiction. Spivak problems, past midterms (posted on course schedule), ASSU test bank (downstairs of Sidney Smith...?), and the final test bank (UofT Website) are all good resources.

§17.1 Past Exam Review

Example 17.1 (Inductively defining the naturals)

If A is a subset of $\mathbb{N}$ such that it contains 1 and, for all $a \in A$, we have $a+1 \in A$, we have $A = \mathbb{N}$. Indeed, by the well ordering principle, if we had a set $B := \mathbb{N} \setminus A = \{x \in \mathbb{N} \mid x \notin A\}$, then let $b \in B$ be the smallest element not in A; then $b - 1 \in A \implies b - 1 + 1 = b \in A$, which is contradictory, so B is necessarily empty. This affirms $A = \mathbb{N}$.

Problem 17.2 (2018 Midterm 1, Q3)

Show that $\alpha_3 \cdot \alpha = \{3x \mid x \in \alpha\}$. (Double inclusion works here).

Let $A = \alpha_3 \cdot \alpha$, and let $B = \{3x \mid x \in \alpha\}$.

- Suppose $\alpha > 0$. Then $A = \{z \in \mathbb{Q} \mid z \leq 0 \text{ or } z = xy \text{ with } x \in \alpha_3, y \in \alpha; x, y > 0\}$, and we may write; for all $x \in \alpha$, pick $x' > x$ such that $x' \in \alpha$ as well (α is a cut, so it has no maximum). Then

$$\frac{x}{x'} < 1 \implies \frac{3x}{x'} < 3 \in \alpha_3,$$

and so we may write

$$\frac{3x}{x'} \cdot x' = 3x \in \alpha_3 \cdot \alpha.$$

- For the other direction, let $z \in \alpha_3 \cdot \alpha$ so that z is of the form $3x$ for some $x \in \alpha$. If $z \leq 0$, we are automatically done by definition; if $z > 0$, write

$$z = 3 \cdot \frac{xy}{3}, \quad 0 < x < 3, \quad y \in \alpha.$$

Since $\frac{x}{3} < 1$ we have $\frac{xy}{3} \in \alpha$.

- Finally, suppose $\alpha < 0$. We can simply use the same argument by rewriting,

$$\begin{aligned}\alpha_3 \cdot \alpha &= -\alpha_3 \cdot |\alpha| \\ &= -\{3x \mid x \in |\alpha|\} \\ &= \{3x \mid x \in -|\alpha|\},\end{aligned}$$

which is equivalent to $\alpha > 0$. The case $\alpha = 0$ has been proven in class and is easy to verify.

[25] "which is sort of a lot i think" - oliver trevett

§17.2 Some Extra Notes

Remember that a Dedekind cut $\alpha \in \mathbb{R}$ is a set, and that $\mathbb{R}$ is a set of Dedekind cuts, which is a set of particular subsets of $\mathbb{Q}$. The Dedekind cut

$$\alpha = \{x \in \mathbb{Q} \mid x^2 < 2 \text{ or } x < 0\}$$

has a supremum of $\sqrt{2}$ in $\mathbb{R}$, but has no supremum in $\mathbb{Q}$. Check 9.1 for what constitutes a Dedekind cut. Moreover, given a collection A (assume it is bounded above and nonempty) of Dedekind cuts, we write

$$\sup(A) = \bigcup_{\alpha \in A} \alpha,$$

and we know this is the supremum because for any upper bound β, we have

$$\beta \supset \bigcup_{\alpha \in A} \alpha = \sup(A).$$

§18 Day 14: Bolzano-Weierstrass Theorem (Oct. 11, 2023)

Correction from last class on subsequences: recall that a sequence $\{a_n\}_{n\in\mathbb{N}}$ has tail $\{a_n\}_{n>N}$ for some $N \in \mathbb{N}$, of which we say convergence and boundedness are properties of the tail.

> **Remark 18.1.** If $\{a_n \mid n \in \mathbb{N}\}$ is bounded above, then so is the tail $\{a_n \mid n > N\}$ (fix some natural N). The forward implication is clear, and for the other direction, we simply fix upper bound $b \geq a_n$ for all $n > N$, as in, "b is an upper bound for the tail."

Anyways, we now re-prove a lemma from last class, that a sequence $\{a_n\}_{n\in\mathbb{N}}$ is unbounded above if and only if there exists a subsequence $b_k = a_{n_k}$ for $k \in \mathbb{N}$ such that $\lim b_n = \infty$. We proved the converse last class, so now we prove the forward implication. Start by constructing n_k recursively such that $n_{k+1} > n_k$ for all $k \in \mathbb{N}$ (which defines the subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$), $a_{n_k} > k$ for all k, and $a_{n_{k+1}} > a_{n_k}$.

- For case $k = 1$, since $\{a_n\}_{n\in\mathbb{N}}$ is unbounded above, we may find n such that $a_{n_1} > 1$ (if we couldn't, then $\{a_n\}$ has an upper bound of 1, which is contradictory)

- We may recursively define indices $n_1 < \cdots < n_k$ for some $k \in \mathbb{N}$ such that $a_j > j$ for $j = 1, \ldots, k$; and so, $a_{n_1} < \cdots < a_{n_k}$. Now, consider $\{a_n \mid n > k\}$; this is a tail of the sequence, hence unbounded above; thus, we may always find $n_{k+1} > n_k$ such that $a_{n_k+1} > \max\{k+1, a_{n_k}\}$. By induction, this defines an increasing subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$ with $a_{n_k} > k$ diverging to infinity. And so we have $\lim a_{n_k} = \infty$.

§18.1 Bolzano-Weierstrass (Not Heine-Borel)

The theorem states that every bounded sequence of real numbers has a convergent subsequence (i.e., there exists $\{x_{n_k}\}_{n\in\mathbb{N}}$ that converges to some limit L). We will prove this by constructing two sequences, $\{a_n\}_{n\in\mathbb{N}}$ which is monotonically increasing, and $\{b_n\}_{n\in\mathbb{N}}$ which is monotonically decreasing, with $a_n < b_n$ for all n such that the set $S_n = \{m \in \mathbb{N} \mid a_n \leq x_m \leq b_n\}$ is infinite (as in, we're bounding the elements of $\{x_n\}$ between two sequences).

- Start out with a_1, a lower bound for $\{x_n\}_{n\in\mathbb{N}}$, and b_1 as an upper bound. Clearly, $a_1 < b_1$; since S_1 contains infinitely many elements, we may divide it into two halves by writing $c_1 = \frac{a_1+b_1}{2}$, and defining

$$S_1^+ = \{m \in \mathbb{N} \mid c_n \leq x_m \leq b_n\}, \qquad S_1^- = \{m \in \mathbb{N} \mid a_n \leq x_m \leq c_n\},$$

  of which $S_1 = S_1^- \cup S_1^+$, where one of these two subsets must have infinitely many elements (otherwise, S_1 is finite, contradiction). If S_1^- has infinitely many elements, define $a_2 = \inf S_1^-$, $b_2 = \sup S_1^-$, and the same goes with S_1^+, then let whichever respective set (the one containing infinitely many indices of x_i) be S_2.

- We may repeat the above procedure recursively. With this, we see, for any $n \in \mathbb{N}$,

$$S_1 \supset S_2 \supset \cdots \supset S_n \implies |b_n - a_n| \leq 2^{-(n-1)} |b_1 - a_1|.$$

  For any $\varepsilon > 0$, we may find $n \in \mathbb{N}$ such that $2^{-(n-1)} |b_1 - a_1| < \varepsilon$, so $|b_n - a_n| \to 0$; thus, for all $i, j > n$, we have $\left|x_{n_i} - x_{n_j}\right| < \varepsilon$, which means $\{x_{n_k}\}$ is Cauchy and necessarily converges. □

§19 Day 15: Bolzano-Weierstrass Theorem (Review) (Oct. 11, 2023)

Today we went over a more in-depth proof of last time's Bolzano-Weierstrass theorem.

§20 Day 16: Epsilon-Delta Limits and Continuity (Section 5) (Oct. 16, 2023)

Quick reminder that we have extra office hours on Oct. 20, 11am to 2pm. Book your calendars! c:

§20.1 Continuity, and Some Definitions

An open interval (on $\mathbb{R}$) is a subset of $\mathbb{R}$ given by $(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$, where $a < b$ are reals. From here on a, b refer to reals (just for convenience). A closed interval is a subset of $\mathbb{R}$ given by $[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$. A half-open interval is given by $[a, b)$ or $(a, b]$ (i.e., one end has $\leq$, the other has $<$), then we have semi-infinite intervals, where we write $(a, \infty), [a, \infty), (-\infty, a), (-\infty, a]$. Note that $\pm\infty$ is never included in our interval; infinity is a concept, not a literal number.

With these done, we may now proceed to define limits and continuity. Let f be a real-valued function defined on some open interval $I \subset \mathbb{R}$[26]. Fix $a \in I$; we say $f(x)$ converges to some number $\ell \in \mathbb{R}$ if, for all $\varepsilon > 0$, there exists $\delta > 0$ such that $0 < |x - a| < \delta \implies |f(x) - \ell| < \varepsilon$. If such a δ exists, we say $\lim_{x \to a} f(x) = \ell$. Note that we don't particularly look at the case where $x = a$; we look at what happens around our chosen a. Moreover, f is continuous in I if

$$\lim_{x \to a} f(x) = f(a)$$

for all $a \in I$ (i.e., f is continuous for all $a \in I$). Now for some examples.

Example 20.1

Show $f(x) = c$ is continuous over $\mathbb{R}$.

For any $a \in \mathbb{R}$, let $\varepsilon > 0$ and pick $\delta = 17.9$ (really a big dumb arbitrary value, it can be whatever you like, you'll see why); now, whenever $|x - a| < \delta$, we have $|f(x) - f(a)| = \underbrace{|f(x) - f(x)|}_{f(a)=f(x)} = 0 < \varepsilon$.

Example 20.2

Show $f(x) = 3x$ is continuous over $\mathbb{R}$.

For any $a \in \mathbb{R}$, let $\varepsilon > 0$ and pick $\delta = \frac{\varepsilon}{10}$ (again, we can pick whatever we like; there are infinitely many suitable choices); we claim that $\lim_{x \to a} f(x) = f(a) = 3a$. We see this from the fact that whenever $|x - a| < \frac{\varepsilon}{10}$, we have

$$|f(x) - f(a)| = |3x - 3a| = 3|x - a| < \frac{3}{10}\varepsilon < \varepsilon$$

as desired.

[26] u don't gotta use an open interval, just convenient for now

Example 20.3

Show $f(x) = x^2$ is continuous over $\mathbb{R}$.

For whatever δ that works, notice that

$$\begin{aligned} |x-a| < \delta \implies |f(x) - f(a)| &= |x+a|\,|x-a| \\ &= |x-a+2a|\,|x-a| \\ &\le \underbrace{(|x-a| + |2a|)}_{<\delta+2|a|}|x-a| \\ &< (\delta + 2|a|)\delta, \end{aligned}$$

implying $(\delta + 2|a|)\delta < \varepsilon$ is sufficient. Now all we have to do is pick a suitable δ, and complete the proof; suppose we wish to pick a $\delta < 1$ (the 1 is arbitrary, it can be whatever). Then $(\delta + 2|a|)\delta < (1 + 2|a|)\delta$ giving

$$\delta < \frac{\varepsilon}{1 + 2|a|},$$

where we may now write

$$\delta = \min\left\{1, \frac{\varepsilon}{1 + 2|a|}\right\}.$$

Example 20.4

Show that $f(x) = \begin{cases} \frac{1}{x}, & x \neq 0, 1 \\ 5, & x = 0 \\ 7, & x = 1 \end{cases}$ is continuous except at $x = 0$ and 1.

We see $\lim_{x\to 1} f(x) = 1 \neq 7$ and $\lim_{x\to 0^+} = \infty, \lim_{x\to 0^-} = -\infty$, both of which are not equal to 5; therefore f is discontinuous at these points. It is left to the reader as an exercise to check for the continuity of f everywhere else.

§21 Tutorial 5: Definition of Limit (Oct. 17, 2023)

Basically just went over the epsilon-delta definition of a limit and what continuity is. See page above!

§22 Day 17: Limits and Continuity Day 2 (Oct. 18, 2023)

Today we go over some more tools / examples.

§22.1 Dirichlet Function

Let $1_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$. We wish to show that it is nowhere continuous.

We proceed to show contradiction against the epsilon-delta definition; let $\varepsilon = 1/2$, and observe the interval $(x-\delta, x+\delta)$. From homework 2, we know any interval of reals necessarily contains rational and irrational numbers (let's pick a rational p and irrational q), so one of the following must occur:

- If x is rational, $|1_{\mathbb{Q}}(x) - 1_{\mathbb{Q}}(q)| = |1-0| = 1 \not< 1/2$,
- If x is irrational, $|1_{\mathbb{Q}}(x) - 1_{\mathbb{Q}}(p)| = |0-1| = 1 \not< 1/2$,

and so we are done. Now, we prove the squeeze lemma:

Lemma 22.1 (Squeeze Lemma)

Let I be an interval with $a \in I$, and let f, g, h be defined on I (not necessarily at a). If, for all $x \in I$ we have

$$g(x) \le f(x) \le h(x)$$

and $\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = L$, then $\lim_{x\to a} f(x) = L$.

Let g, h satisfy

$$\forall\varepsilon > 0, \exists\delta_1 > 0 \text{ such that } |x-a| < \delta_1 \implies |g(x) - L| < \varepsilon,$$
$$\forall\varepsilon > 0, \exists\delta_2 > 0 \text{ such that } |x-a| < \delta_2 \implies |h(x) - L| < \varepsilon.$$

Now, pick $\delta = \min\{\delta_1, \delta_2\}$. Thus, if $|x-a| < \delta$, we have

$$\begin{gathered} g(x) \le f(x) \le h(x) \\ g(x) - L \le f(x) - L \le h(x) - L \\ -\varepsilon < g(x) - L \le f(x) - L \le h(x) - L < \varepsilon \\ -\varepsilon < f(x) - L < \varepsilon, \end{gathered}$$

and so we are done. □

Example 22.2 (Example Duo)

Let $g(x) = xI_{\mathbb{Q}}(x)$ be defined over $\mathbb{R}$; show that g is continuous at $x = 0$ and discontinuous everywhere else.

Notice g is lower bounded by $-|x|$ and upper bounded by $|x|$; we see $\lim_{x\to 0} |x| = \lim_{x\to 0} -|x| = 0$ from picking $\delta = \varepsilon$. Then $0 < |x-a| < \delta$ yields

$$||x| - |a|| \le |x-a| < \delta = \varepsilon$$

from the triangle inequality. The same idea holds for $-|x|$; thus

$$\lim_{x\to 0} g(x) = 0 = 0(I_{\mathbb{Q}}(0)),$$

and we see g is continuous at $x = 0$. For $x \neq 0$, a similar argument can be used (from earlier when we proved $1_{\mathbb{Q}}(x)$ is discontinuous); simply pick a rational and an irrational in the interval and we are done.

§23 Day 18: More Tools for Limits (Oct. 20, 2023)

Today we will go over more tools for limits; next week, we will go over zeroes and extrema of continuous functions, as well as the "3 hard theorems" from Spivak. Derivatives will be the week after.

§23.1 Addition, Multiplication, Division

Let us have f, g where $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Then

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

Let us have δ_1 and δ_2 where

$$0 < |x-a| < \delta_1 \implies |f(x) - L| < \frac{\varepsilon}{2}$$
$$0 < |x-a| < \delta_2 \implies |g(x) - M| < \frac{\varepsilon}{2}$$

Now, let $\delta = \min\{\delta_1, \delta_2\}$. Then if $0 < |x-a| < \delta$, we see both $|f(x) - L|$ and $|g(x) - M|$ must be less than $\frac{\varepsilon}{2}$. We may now sum these two inequalities to obtain

$$|(f(x)+g(x)) - (L+M)| \leq |f(x) - L| + |g(x) - M| < \varepsilon$$

from the triangle inequality. Thus $\lim_{x\to a}(f+g)(x) = L + M$. □

(b) $\lim_{x\to a} f(x)g(x) = L \cdot M$,

We prove this in a similar fashion, except we'll choose some more exotic epsilons. Notice from

$$\begin{aligned} |a_1b_1 - a_2b_2| &= |a_1(b_1 - b_2) + (a_1 - a_2)b_2| \\ &\leq |a_1|\,|b_2 - b_2| + |a_1 - a_2|\,|b_2|, \end{aligned}$$

we're going to be doing some casework so the final expression is $< \varepsilon$. Let us have

$$0 < |x-a| < \delta_1 \implies |f(x) - L| < \frac{\varepsilon}{2(1+|L|)}$$
$$0 < |x-a| < \delta_2 \implies |g(x) - M| < \frac{\varepsilon}{2(1+|M|)}$$
$$0 < |x-a| < \delta_3 \implies |g(x) - M| < 1$$

Given $\varepsilon > 0$, now pick $\delta = \min\{\frac{\varepsilon}{2(M+1)}, \frac{\varepsilon}{2(L+1)}, \delta_3\}$. Then we have

$$\begin{aligned} |f(x)g(x) - LM| &= |f(x)g(x) - Lg(x) + Lg(x) - LM| \\ &\leq |f(x)g(x) - Lg(x)| + |Lg(x) - LM| \\ &= |g(x)|\,|f(x) - L| + |L|\,|g(x) - M| \\ &< (1+|M|)\left(\frac{\varepsilon}{2(1+|M|)}\right) + (1+|L|)\left(\frac{\varepsilon}{2(1+|L|)}\right) \\ &= \varepsilon. \end{aligned}$$

And so we now may conclude $\lim_{x\to a} f(x)g(x) = LM$.[27] □

[27] sry didn't catch almut's proof in lec

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

It suffices to prove that $\lim_{x\to a} \frac{1}{g(x)} = \frac{1}{M}$ (by the product of limits from earlier). First, write

$$\left|\frac{1}{g(x)} - \frac{1}{M}\right| = \frac{|M - g(x)|}{|g(x)M|},$$

and so given $\varepsilon > 0$ (where we may assume $\varepsilon < \frac{M}{2}$ because $M \neq 0$), we have

$$\begin{aligned} |g(x)| &= |g(x) - M - (-M)| \\ &= |M - (M - g(x)| \\ &\geq |M| - |M - g(x)| \\ &\geq \frac{M}{2}. \end{aligned}$$

Now, pick $\delta = \varepsilon \cdot \frac{M^2}{2}$. For $|x - a| < \delta$ we necessarily have

$$\left|\frac{1}{g(x)} - \frac{1}{M}\right| \leq \frac{|M - g(x)|}{|g(x)|\, M} \leq \frac{\varepsilon M^2}{2} \cdot \left(\frac{M}{2} \cdot M\right)^{-1} = \varepsilon,$$

and so we are done.

(d) If $f(x) \leq g(x)$, then $L \leq M$.

Let us have δ_1, δ_2 where

$$\begin{aligned} 0 < |x - a| < \delta_1 &\implies |f(x) - L| < \varepsilon, \\ 0 < |x - a| < \delta_2 &\implies |g(x) - M| < \varepsilon, \end{aligned}$$

which gives $L - \varepsilon < f(x) < g(x) < M + \varepsilon$ from rearrangement; if, whenever $L > M$, let $\varepsilon = \frac{L-M}{3}$. Then let us have $\delta = \min\{\delta_1, \delta_2\}$. We now have

$$(M + \varepsilon) - (L - \varepsilon) = M - L + \frac{2(L - M)}{3} = \frac{M - L}{3} > 0 \implies M > L,$$

which is contradictory. Thus, $L \leq M$, and we are done.

Theorem 23.1 (Composition of Limit)
Let $\lim_{x\to a} g(x) = b$ and $\lim_{y\to b} f(y) = L$. Then $\lim_{x\to a}(f \circ g)(x) = L$.

Given $\varepsilon > 0$, choose δ such that $0 < |y - b| < \delta$ then $|f(x) - L| < \varepsilon$. Then choose $\eta > 0$ such that if $0 < |x - a| > \eta$, we have $|g(x) - b| < \delta$. But with this η, we now have $|f(g(x)) - L| < \varepsilon$, and so we are done. $\square$

§24 Day 19: Intermediate Value Theorem (Oct. 23, 2023)

We start with the function composition from last time; let f, g be real-valued functions, defined on suitable subsets of $\mathbb{R}$, such that $f \circ g$ makes sense. Fix an $a \in \text{dom}(g)$; assuming we have $\lim_{x\to a} g(x) = b$ and f is continuous at b, then

$$\lim_{x\to a}(f \circ g)(x) = f(\lim_{x\to a} g(x)) = f(b).$$

To prove this, take any $\varepsilon > 0$; we want to find $\delta > 0$ such that whenever $0 < |x - a| < \delta$, we have $|f(g(x)) - f(b)| < \varepsilon$. To do this, we will first find $\eta > 0$ such that $0 < |y - b| < \eta$ gives $|f(y) - f(b)| < \varepsilon$ (since $\lim_{y\to b} f(y) = f(b)$). Then, find $\delta > 0$ such that $0 < |x - a| < \delta \implies |g(x) - b| < \eta$. Thus we have $0 < |x - a| < \delta \implies |f(g(x)) - f(b)| < \varepsilon$ as claimed. Note this implies that if f, g are continuous, then so is $f \circ g$.

§24.1 Intermediate Value Theorem

Let $[a, b]$ be a non-empty closed interval, and define a continuous function $f : [a, b] \to \mathbb{R}$. If $f(a) < 0 < f(b)$, then there is a point c with $a < c < b$ such that $f(c) = 0$. To prove

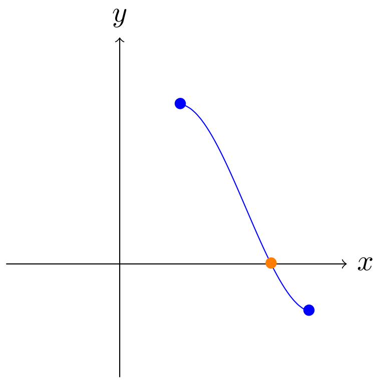


Figure 4: i'm very skill issued at tikz, just imagine this as a fancy Wikipedia graph

this, we will recursively construct a sequence of intervals $[a_n, b_n]$. Let $a_1 = a$, $b_1 = b$, and consider $c = \frac{a+b}{2}$.

- If $f(c) = 0$, then we have found the zero of f,
- If $f(c) < 0$, set $a_2 = c$, $b_2 = b_1$,
- If $f(c) > 0$, set $a_2 = a$, $b_2 = c$.

By construction, $f(a_2) < 0 < f(b_2)$ (unless we've found the zero already). Recursively, we will obtain two sequences

$$\begin{cases} a_1 \le a_2 \le \dots \\ b_1 \ge b_2 \ge \dots \end{cases}$$

of which either the sequence terminates at some n; then $f(\frac{a_n+b_n}{2}) = 0$, or it is infinite, and $f(a_n) < f(c) < f(b_n)$ for all n. Then we may take the limit of both sequences, and since $|b_n - a_n| = \frac{|b_1-a_1|}{2^{n-1}} \to 0$ as $n \to \infty$, we clearly have $c = \lim a_n = \lim b_n$. Since $f(\lim a_n) = \lim f(a_n) \le 0$ and $f(\lim b_n) = \lim f(b_n) \ge 0$ by continuity, we know $0 \le f(c) \le 0 \implies f(c) = 0$ and we are done.

Likewise, if f is continuous on $[a, b]$ and $f(a) > 0 > f(b)$, we can also find $f(c) = 0$; just pick $g = -f$. Moreover, for all $t \in \text{Im}(T)$, by considering $g(x) = f(x) - t$, we may also show $\exists x \in [a, b]$ such that $f(x) = t$.

§25 Tutorial 6: If someone was there please let me know because I didn't wake up in time (Oct. 24, 2023)

§26 Day 20: Extreme Value Theorem (Oct. 25, 2023)

Impending doom has been announced (midterm 2 will be on Nov. 16, after our reading week). Anyways, so we had an alternative proof of IVT come up today! Let a continuous f takes positive and negative values on an interval, then it has a zero. Let $A = \{x \in [a, b] \mid f(x) < 0\}$ (i.e. the preimage of all negative values of f in $[a, b]$), and consider $c = \sup A$. Then we have $f(c) = 0$ (if $f(c) > 0$, then c is not a supremum; if $f(c) < 0$, then it is not an upper bound).

§26.1 Extreme Value Theorem

Let us have $f : [a, b] \to \mathbb{R}$ be continuous; then there exists a $c \in [a, b]$ such that

$$c = \max_{[a,b]} f \iff c = \max\{f(x) \mid a \leq x \leq b\} \iff f(c) \geq f(x)\ \forall x \in [a, b].$$

Any of the three above definitions are equivalent; just depends on how far you're willing to go to be pedantic. With this, we now wish to prove the extreme value theorem (numero duo of Spivak's three hard theorems). The general idea is to construct the pre-image of the supremum of the image of f. To do this, we must first prove[28]

Theorem 26.1 (Boundedness Theorem)

If f is continuous on the interval $[a, b]$, then it is bounded on $[a, b]$ as well.

Suppose that the function is not bounded above on the interval $[a, b]$. By definition, for all $n \in \mathbb{N}$, there must exist $x_n \in [a, b]$ such that $f(x_n) > n$. This gives us the sequence $\{x_n\}_{n\in\mathbb{N}}$, which, by the Bolzano-Weierstrass theorem, necessarily has a convergent subsequence; let it be $\{x_{n_k}\}_{k\in\mathbb{N}}$, and suppose its limit is x^*. By the continuity of f,

$$\lim_{k\to\infty} f(x_{n_k}) = f(x^*),$$

but since $f(x_{n_k}) > n_k \geq k$ for any choice of k, $f(x_{n_k}) \to +\infty$, contradicting the existence of such a limit. Thus, the image of f must be bounded above on $[a, b]$ (a similar argument follows for bounded below). □

Now that we have that the image of f is bounded, there must exist a supremum of f (let it be M). Thus, we wish to find some $c \in [a, b]$ such that $f(c) = M$. Since M is the least upper bound of $\operatorname{Im}(f)$, $M - \frac{1}{n}$ is not an upper bound for any natural $n \in \mathbb{N}$; then there must be some $c_n \in [a, b]$ such that $M - \frac{1}{n} < f(c_n)$. Construct such a sequence $\{c_n\}_{n\in\mathbb{N}}$ defined by such a property, and notice $f(c_n) \to M$ as $n \to \infty$ (note that $\{c_n\}_{n\in\mathbb{N}}$ itself is not necessarily convergent...). By the Bolzano-Weierstrass theorem, we may find a convergent subsequence $\{c_{n_k}\}_{k\in\mathbb{N}}$ with limit c^*. Since $\{f(c_{n_k})\}_{k\in\mathbb{N}}$ is a subsequence of $\{f(c_n)\}_{n\in\mathbb{N}}$ and convergent sequences have the same limits, we see $f(c^*) = M$, and since c^* is indeed in $[a, b]$ by construction, we are done.

[28] this feels like one of those "ta-da" look it's the boundedness theorem look at how awesome and cool it is omg

§27 Day 21 & 22: Derivatives (Oct. 27 & 30, 2023)

Quick recap of the extreme value theorem: let $a \leq b$, then define $f : [a, b] \to \mathbb{R}$ be a continuous function. Then there must exist $c, d \in [a, b]$ such that f attains its maximum at c, and f attains its minimum at d; i.e. $\max_{[a,b]} f = c$ and $\min_{[a,b]} f = d$. Construct a maximizing sequence $\{x_n\}_{n\in\mathbb{N}}$ such that $f(x_n) \to \sup f$; then by the Bolzano-Weierstrass theorem, we may find a convergent subseqeunce $\{x_{n_k}\}_{k\in\mathbb{N}}$ where $x_{n_k} \to x^*$; then we may write

$$f(x^*) = f(\lim x_{n_k}) = \lim f(x_{n_k}) = \sup f$$

by continuity. The same argument comes for infimum. Moreover, if we combine the intermediate value theorem and the extreme value theorem, we see that the image of any continuous function $f : [a, b] \to \mathbb{R}$ is necessarily an interval $[\min_{[a,b]} f, \max_{[a,b]} f]$. If you wish to be silly, we may prove this by noting $\mathbb{R}$ is compact, and so the image of any continuous function f is also compact. By the Heine-Borel theorem, this compact set is also closed and bounded, and so by EVT we have our desired interval. (Not sure if this works lol... figured I'd take a shot at it)

§27.1 Derivatives

Let $f : A \to \mathbb{R}$ where $\operatorname{dom} f = A \subset \mathbb{R}$. Fix an $a \in A$. By doing this, we automatically assume that f is defined in a small open interval around a, i.e., let $I_n = (b, c)$ containing a (as in, $b \leq a \leq c$). This is called a neighborhood of a.

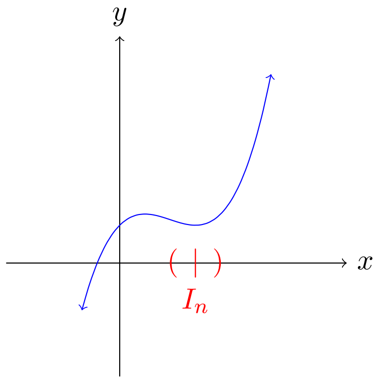


Figure 5: Neighborhood of a

We say that f is differentiable at a given a if

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} \iff \lim_{h\to 0} \frac{f(a + h) - f(a)}{h}$$

exists. Note that both of the definitions above are equivalent (let $x = a + h$); they're just conventions. In the same manner, we may say that there exists a k such that $f(x) = f(a) + k(x - a) + r_1(x, a)$, where

$$\lim_{x\to a} \frac{|r_1(x, a)|}{|x - a|} = 0.$$

If such a k does exist, then we may write

$$k = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} = f'(a).$$

Moreover, if f is differentiable at all $a \in A$, we say f is differentiable. It is called a continuously differentiable function if $a \mapsto f'(a)$ is a continuous map.

Lemma 27.1 (Differentiability implies Continuity)

If f is differentiable at a, then it is continuous at a. We may prove this by writing

$$\begin{aligned}\lim_{x\to a} f(x) &= \lim_{x\to a}\big(f(a) + f'(a)(x-a) + r_1(x,a)\big)\\ &= f(a) + \lim_{x\to a} f'(a)(x-a) + \lim_{x\to a}\frac{|r_1(x,a)|}{|x-a|}|x-a|\\ &= f(a),\end{aligned}$$

since the latter two terms on the right hand side both vanish by definition. $\square$

Moreover, we say that f attains a maximum at a (or minimum) if $f(a) \geq f(x)$ for all $x \in \operatorname{dom} f$ (opposite definition for minimum, i.e. $f(a) \leq f(x)$). We also say a is a *local* maxima or minima if there exists some open interval such that $a \in I \subset \operatorname{dom} f$ with $f(a) \geq f(x)$ for all $x \in I$.

Lemma 27.2 (Local Minima/Maxima if $f'(a) = 0$.)

Check Day 23!

We also went over some Mean Value Theorem things, but I assume that will be covered Wednesday.

§28 Tutorial 7: If someone was there, let me know (Oct. 31, 2023)

§29 Day 23: Derivatives Pt. II (Nov. 1, 2023)

Today, we went over Leibniz v. Newton notation, local extrema, and the mean value theorem.

§29.1 Local Extrema and Derivatives (Part 2)

Given a function $f : A \to \mathbb{R}$ that is differentiable at a given point $a \in A$, we automatically assume to be talking about an open interval $I = (b, c)$ where $a \in I = (b, c) \subset A$; given these conditions, a is said to be a local extrema if $f'(a) = 0$. Today we reprove a lemma from yesterday that I didn't get around to tex-ing:

Lemma 29.1 (Local Minima/Maxima if $f'(a) = 0$.)

If f has such a local extrema at a (suppose it's a maxima, just for the sake of simplicity; consider $-f$ for minima), then there must exist some $\delta > 0$ such that $|x - a| < \delta$ implies $f(x) \leq f(a)$, by the definition of such an extreme point. Note that here, our interval is $(x - \delta, x + \delta)$. Notice that

$$f'(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \geq 0$$

since the numerator and denominator are both positive. In the same manner,

$$f'(a) = \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} \leq 0$$

since the numerator is positive (as above) and the denominator is negative. Thus, we see $f'(a) = 0$, and we are done.

Anyways, the derivative $f'(x)$ can be thought of as the map $x \mapsto f'(x)$. Usually this notation works for single variable derivatives, but when we have more than one variable (as in functions like $f(x, y)$, etc...), then its easier to switch to Leibniz notation, $\frac{df}{dx}$ instead of Newton, $f'(x)$. If we wanted to rewrite our definition of derivatives in terms of Leibniz notation, we would have

$$\frac{df}{dx}(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$
$$\frac{df}{dx} = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

where the latter is where we take the derivative over the given domain. Generally, the idea comes from comparing $\Delta f(x) = f(x + h) - f(x)$ versus $\Delta x = x + h - x = h$ to determine the derivative, but we do get some problems if we think of these as individual elements of a fraction, since

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \frac{0}{0},$$

which is kind of... broken. We'll poke more into these two as differential forms in the second half of **MAT257**! So basically, $f'(x)$ is good, $\frac{df}{dx}$ is good, df, dx on their own is bad (for now).

§30 Day 24: Derivatives Pt. III; Mean Value Theorem, Product/Quotient/Chain Rule (Nov. 3, 2023)

Today we will go over the proof of the Mean Value Theorem (along with applications on monotonicity), as well as the product, quotient, and chain rule. Next class, we will go over higher derivatives; exam on Nov. 16! Check Piazza or Quercus for the announcement.

§30.1 Mean Value Theorem

I'll just pick up from Day 23's notes.

Theorem 30.1 (Mean Value Theorem)

If $f : [a, b] \to \mathbb{R}$ is a continuous on the closed interval $[a, b]$ and differentiable on (a, b) (note that we automatically assume $a < b$), then there exists a point $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

We will prove this by reducing our case to Rolle's Theorem; if $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$. Start by considering this case. By EVT, f attains its maximum and minimum on the interval $[a, b]$; if both the minimum and maximum are at the endpoints, then f is constant and we are done (as all $c \in (a, b)$ satisfy $f'(c) = 0$). Otherwise, either the minimum or maximum of f lie within our interval, of which it must have a derivative of zero (as per definition, see **Lemma 29.1**).

Now, in the more general condition that $f(a) \neq f(b)$, take the function g defined as

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Notice that $g(a) = g(b)$, and so we may apply Rolle's theorem here; pick any $c \in (a, b)$ such that $g'(c) = 0$. Then we have

$$g'(c) = 0 \implies f'(c) = \frac{f(b) - f(a)}{b - a}$$

as desired. □

§30.2 Differentiation Rules

(a) Linearity: $(f + g)' = f' + g'$, and $(\alpha f)' = \alpha f'$ for any real α and $f, g \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ (such that their derivatives exist)

(b) $f'(x) = 0$ for all x on the interval (a, b) if and only if f is constant on (a, b).

(c) Product Rule: $(fg)' = f'g + fg'$; we see this from

$$\begin{aligned}(fg)'(x) &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= \lim_{h\to 0} f(x+h)\left(\frac{g(x+h) - g(x)}{h}\right) + \lim_{h\to 0}\left(\frac{f(x+h) - f(x)}{h}\right) g(x) \\ &= f(x)g'(x) + f'(x)g(x)\end{aligned}$$

by limit laws.

(d) Quotient Rule: Assuming $g(a) \neq 0$, we have

$$\left(\frac{f}{g}\right)'(a) = \left(f \cdot \frac{1}{g}\right)'(a) = \frac{f'(a)}{g(a)} + f(a)\left(\frac{1}{g}\right)'(a),$$

which means we only need to solve for $(\frac{1}{g})'(a)$ now. Write

$$\begin{aligned}\left(\frac{1}{g}\right)(a) &= \lim_{h\to 0} \frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} \\ &= \lim_{h\to 0} \frac{g(a) - g(a+h)}{hg(a+h)g(a)} \\ &= \lim_{h\to 0} \frac{g(a) - g(a+h)}{h} \cdot \frac{1}{g(a+h)g(a)} \\ &= -\frac{g'(a)}{g(a)^2},\end{aligned}$$

plugging it back in, we have

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \quad \square$$

(e) Power Rule: $(x^n)' = nx^{n-1}$, which we may verify from

$$\begin{aligned}\lim_{h\to 0} \frac{(x+h)^n - x^n}{h} &= \lim_{h\to 0} \frac{1}{h}\left(x^n + \binom{n}{1}x^{n-1}h + \cdots + \binom{n}{n}h^n - x^n\right) \\ &= \lim_{h\to 0} \binom{n}{1}x^{n-1}h^0 + \cdots + \binom{n}{n}h^{n-1} \\ &= nx^{n-1}.\end{aligned}$$

Exercise: generalize this to negative integer exponents (quotient rule) and rational exponents (chain rule).

(f) Chain Rule: $(f \circ g)'(x) = f'(g(x))g'(x)$. Ideally, we would have,

$$\lim_{x\to a} \frac{f(g(x)) - f(g(a))}{x - a} = \lim_{x\to a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \cdot \frac{g(x) - g(a)}{x - a}$$

which is by definition $f'(g(x))g'(x)$. However, if $g'(a) = 0$, then we would be dividing by zero; in the case $g'(a) = 0$, since $f'(g(a))$ necessarily exists we have $c, \gamma > 0$ such that any valid choice of y below satisfies

$$|g(a) - y| < \gamma \implies |f(y) - f(g(a))| \leq c\,|y - g(a)|\,;$$

by continuity of g, for any arbitrary $\varepsilon > 0$ there exists δ where $|x - a| < \delta$ implies $|g(x) - g(a)| < \varepsilon$. Combining the two definitions, we have

$$|f(g(x)) - f(g(a))| \leq c\,|g(x) - g(a)|\,,$$

where if we divide both sides by $x - a$ and take the limit $x \to a$, we have that the RHS is equal to zero (since $g'(a) = 0$) and the limit on the left is squeezed to zero as well. Thus, $(f \circ g)'(a) = 0$ in this case, and we are done. Alternatively, Almut's proof was to take a sequence $x_n \to a$ such that $g(x_n) = g(a)$, then showing $g'(a) = 0$, but I don't understand that... sorry :) [29]

[29] scuffed proof but i think it works. lmk if it doesn't.

§31 Day 25: Derivatives Pt. IV; Chain Rule, Second Derivative; Local Extrema Conditions (Nov. 13, 2023)

Roman numerals ftw! Today we will reprove the chain rule and talk about necessary and sufficient conditions to obtain points of local extrema; on Wednesday, we will be talking about graphing functions.

§31.1 Chain Rule, Revisited

Recall that the chain rule takes two differentiable functions f, g and states that $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$. This can be done more than two functions as well (i.e. applying the rule again, and well, more); ex,

$$(f \circ g \circ h)'(a) = f'(g(h(a))g'(h(a))h'(a).$$

To prove this, define $D(y)$ as follows,

$$D(y) = \begin{cases} \dfrac{f(y) - f(b)}{y - b}, & y \neq 0 \\ f'(b), & y = 0. \end{cases}$$

Note that here we are taking $b = g(a)$ and $y \in \operatorname{dom} f$. First, notice that $D(y)$ is continuous at y by definition of the derivative (i.e., $D(b) = \lim_{y \to b} D(y)$), and so we may write

$$\frac{d}{dx}(f \circ g)(x) = \lim_{x \to a} D(g(x)) \frac{g(x) - g(a)}{|x - a|} = D(b)g'(a),$$

where we may note D is continuous at b and g is differentiable at a. Thus, the above is equal to $f'(g(a))g'(a)$, and we are done.

§31.2 Higher Derivatives

The second derivative ("rate of change of the rate of change") describes the concavity of the graph. Let $f : [a, b] \to \mathbb{R}$ where $x \mapsto f'(x)$ be a differentiable function; then f is twice differentiable at $c \in (a, b)$ if f' is differentiable at c.

If we take any general twice-differentiable function f defined over an interval $[a, b]$, then if $f'(x) = 0$ and $f''(x) < 0$, then x is a local maxima; if $f''(x) > 0$, it is a local minima, and if $f''(x) = 0$, it is inconclusive and there might be a local minima, maxima, or an inflection point. [30]

In general, you can also use higher derivatives to approximate a given function; in fact, using the second derivative gives us the second order Taylor polynomial

$$y = f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2.$$

This was mentioned in class, but I don't think we went that in depth about it. That's it for today!

[30] we haven't gone over the proof so i won't be touching that for now, since I'd imagine we would be talking about concavity and convexity, and idk if i should assume that in the notes.

§32 Day 26: Derivatives Pt. V; Convexity/Concavity, Local Extrema Conditions (Nov. 15, 2023)

Today we proved the conditions for local extrema. Before we do that, we can first define convexity and concavity; we say a function $f : [a, b] \to \mathbb{R}$ is convex if its derivative is monotonically non-decreasing on that interval; if the function is differentiable and convex, then it is continuously differentiable on the interval it is defined on. Concavity works similarly, in the sense if f is monotonically non-increasing.

§32.1 Local Extrema

Let $f : [a, b] \to \mathbb{R}$ and fix $c \in (a, b)$. First, we prove the necessary condition. Assume f attains a local maximum at c. If f is differentiable at c, then $f'(c) = 0$ (i.e., c is a critical point); moreover, let f be differentiable on some small open interval containing c, and that $f''(c)$ exists; then $f''(c) \leq 0$. Start by letting $\delta > 0$ satisfy $f(x) \leq f(c)$ for all $|x - c| < \delta$. Then

$$f'(c) = \lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = \lim_{h\to\infty} \underbrace{\frac{f\left(c + \frac{1}{h}\right) - f(c)}{1/h}}_{\leq 0} \leq 0.$$

Construct a sequence $\{x_k\}_{k\in\mathbb{N}}$ with $\lim x_k = c$ and $x_k > c$ for all k such that $f'(x_k) \leq 0$ (for sufficiently large k). To do this, we may consider $f(c + \frac{1}{k}) - f(c) \leq 0$, where by MVT, we have

$$\exists x_k \in \left(c, c + \frac{1}{k}\right) \text{ such that } f'(x_k) = \frac{f\left(c + \frac{1}{k}\right) - f(c)}{1/k} \leq 0.$$

Thus, such a sequence necessarily exists, and we may compute

$$f''(c) = \lim_{h\to 0} \frac{f'(c+h) - f'(c)}{h} = \lim_{k\to\infty} \frac{f'(x_k) - f'(c)}{x_k - c} \leq 0,$$

which we may verify by checking signs; i.e., $f'(x_k) \leq 0$, $f'(c) = 0$, and $x_k - c > 0$.

Now, we prove the sufficient condition. Assume $f'(c) = 0$ and $f''(c) < 0$. Then we claim that there is a strict local maxima at c (proof for local minima is similar, leaving it as an exercise). To prove this, we simply observe that there must exist $\delta > 0$ such that $f'(x) < 0$ for $c < x < c + \delta$ and $f'(x) > 0$ for $c - \delta < x < c$; if not, then there exists $x \in (c, c + \delta)$ such that $f'(x) \geq 0$ (or $x \in (c - \delta, c)$ such that $f'(x) \leq 0$. Whichever you choose!). Construct a sequence $\{x_k\}_{k\in\mathbb{N}}$ converging to c with all $x_k > c$ such that $f'(x_k) \geq 0$. Directly compute

$$f''(c) = \lim_{k\to\infty} \frac{f'(x_k) - f'(c)}{x_k} \geq 0,$$

which we may see from evaluating the signs. This contradicts $f''(c) < 0$, and we are done.

§33 Day 27: Graphing (Nov. 17, 2023)

Today was kinda an overview on graphing functions! And uh, LaTeX doesn't play nice with graphing stuff ;w;... in general we just went over zeroes, local min/max, inflection points, asymptotes, fractional linear transformation, etc... review stuff!

§34 Day 28: Inverse Function Theorem (Nov. 20, 2023)

Assume that f is a continuously differentiable function defined on the real interval $[a, b]$. If f has a nonzero derivative at the point a, then f is injective in a neighborhood of a, f^{-1} is continuously differentiable near $b = f(a)$, and

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

To prove this, consider the neighborhoods (read: small open intervals) $U \subset [a, b]$ and $V \subset \mathbb{R}$ such that $f|_U : U \to V$ is bijective with inverse $f^{-1}|_V : V \to U$. By assumption, U and V are nonempty, and so we may take $\delta > 0$ such that $|x - h| < \delta$ and take $f(x) = y$, $f(x + h) = y + k$, and write the limit

$$(f^{-1})'(y) = \lim_{k \to 0} \frac{f^{-1}(y + k) - f^{-1}(y)}{k} = \lim_{h \to 0} \frac{h}{f(x + h) - f(x)} = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))},$$

where we may note $h = f^{-1}(y + k) - x$ for the exchange; we may do such an exchange of variable above since f^{-1} is continuous (check previous notes to see proof of continuity of f^{-1}).[31] We also went over some examples in class.

(a) Let $f(x) = x^n$ (and so $f^{-1}(x) = \sqrt[n]{x}$); pick whatever $a \in \mathbb{R}$, then $b = a^n$ (using our previous definition of $b = f(a)$). Using the power rule, we may evaluate

$$(f^{-1})'(b) = \frac{a^{1-n}}{n} = \frac{1}{f'(a)}.$$

(b) Let $f(x) = \tan(x)$ (and so $f^{-1}(x) = \arctan(x)$); pick whatever $a \in \mathbb{R}$, then $b = \tan(a)$. Note that

$$(\arctan x)' = \frac{1}{x^2 + 1}, \quad (\tan x)' = \sec^2(x),$$

and so we may evaluate

$$(f^{-1})'(b) = \frac{1}{\tan^2(a) + 1} = \frac{1}{\sec^2(a)} = \frac{1}{f'(a)},$$

which follows by evaluating $\sin^2(a) + \cos^2(a) = 1$.

(c) I forgor the third example :(someone lmk what it was. blahaj time

[31] i'm not entirely sure if this is right; if i'm wrong, ill put almut's proof in

§35 Day 29: Continuity of Inverse Functions (Nov. 22, 2023)

Today we went over problem 5 on assignment 7, as well as another proof of the Inverse Function Theorem. Let $I, J \subset \mathbb{R}$ be non-empty intervals, and let $f : I \to J$ be bijective, that is, for every $j \in J$ there exists exactly one $x \in I$ with $f(x) = y$.

(a) Suppose that f is continuous. Prove that f is either strictly increasing on I, or strictly decreasing on I.

Let $I = [a, b]$ so we may write $f : [a, b] \to J$; without loss of generality, let $f(a) < f(b)$ (the case $f(b) > f(a)$ follows a similar proof; $f(a) = f(b)$ cannot hold), and note that f is necessarily injective since it is bijective by assumption. By contradiction, suppose there exists $x > y$ such that $x \leq y$. Then we may find c in the open interval of $\operatorname{dom} f$ such that we may apply IVT to either get:

- If $c \in (a, x)$, then $f(c) = f(y)$;
- If $c \in (x, y)$, then $f(c) = f(a)$ or $f(c) = f(b)$;
- If $c \in (y, b)$, then $f(c) = f(x)$.

For intuition, check all the possibilities for x, y greater or less than $f(a), f(b)$. In all the above cases, we find a contradiction to the injectivity of f; thus f must be strictly increasing or decreasing.

(b) Conversely, suppose f is strictly increasing on I. Prove that f is continuous.

For any $c \in I$, pick $\varepsilon > 0$ small enough such that $[f(c) - \varepsilon, f(c) + \varepsilon] \subset J$. Since f is bijective, take the pre-images a, b such that

$$f(a) = f(c) - \varepsilon,$$
$$f(b) = f(c) + \varepsilon.$$

Note that since f is increasing, we have $a < c < b$. Thus, pick $\delta = \min\{c-a, b-c\} > 0$ so that for whatever $x \in \{c - \delta, c + \delta\} \subset I$, we may write

$$f(a) = f(c) - \varepsilon < f(x) < f(c) + \varepsilon = f(b) \implies |f(x) - f(c)| < \varepsilon,$$

which concludes f is continuous at c.

(c) Combine the previous two parts to conclude that if $f : I \to J$ is bijective and continuous, then its inverse $g := f^{-1}$ is continuous as well.

By part (a), we see f must be strictly increasing or decreasing; let us prove the former case, and the other will come in an analogous fashion. First, we show f^{-1} must be strictly increasing as well; if not, then there exists $a, b \in J$ with $a < b$ such that $x = f^{-1}(a)$, $y = f^{-1}(b)$ with $y \leq x$. Then $b = f(y) \leq f(x) = a$, which is contradictory. Thus f^{-1} is strictly increasing. Now, to show f^{-1} is continuous, let $c \in I$ and let $d = f(c)$. We will now show $\lim_{y \to d} f^{-1}(d) = c$. For any $\varepsilon > 0$ such that $(c - \varepsilon, c + \varepsilon) \subset I$, we have

$$f(c - \varepsilon) < f(c) < f(c + \varepsilon).$$

Now, take $\delta = \frac{1}{2}\min\{f(c) - f(c-\varepsilon), f(c+\varepsilon) - f(c)\}$, and notice $f(c-\varepsilon) < f(c) - \delta$ as well as $f(c) + \delta < f(c+\varepsilon)$. Thus, we see $|y - d| < \delta \implies \left|f^{-1}(y) - f^{-1}(d)\right| < \varepsilon$, and the limit is complete. Thus, f^{-1} is continuous. □

§36 Day 30: L'ĥopital's Rule and Cauchy's Mean Value Theorem (Nov. 24, 2023)

Today, we go over the L'ĥopital rule and Cauchy's Mean Value Theorem[32]. When computing limits of the type 0/0 (indeterminate form), such as

$$\lim_{x\to 0}\frac{\sin x}{x}, \qquad \lim_{x\to a}\frac{f(x)}{g(x)} \text{ with } f(a)=g(a)=0,$$

and like a bunch more inconveniently annoying examples, we want a better rule that gets to the limit in a more efficient manner. Anyways, I present a proof of L'hôpital's rule, one with Cauchy MVT, and the other direct. To start, Cauchy's mean value theorem is stated as follows; given $f, g : [a,b] \to \mathbb{R}$ continuous functions that are differentiable on the open intervals (a,b), there exists an element $c \in (a,b)$ such that

$$f'(c)(g(b)-g(a)) = (f(b)-f(a))g'(c).$$

In particular, if $g'(c) \neq 0$, then

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

To prove this, consider $h(x) = f(x)(g(b)-g(a)) - ((f(b)-f(a))g(x)$, of which is continuous and differentiable on its given closed and open intervals respectively. Computing $h(a)$, we have

$$h(a) = f(a)(g(b)-g(a)) - (f(b)-f(a))g(a) = f(a)g(b) - f(b)g(a) = h(b)$$

by Rolle's theorem, where we may pick $c \in (a,b)$ such that $h'(c) = 0$. We may now apply this to finish the L'hôpital rule, where we may note

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f(x)-f(a)}{g(x)-g(a)} \implies \frac{f'(c_x)}{g'(c_x)} \text{ with } c_x \in (x,a).$$

With this,

$$\lim_{x\to a}\frac{f'(c)}{g'(c)} = L,$$

which is our original desired limit. □

[32] i swear to god that hat on the o

§37 Day 31: L'hôpital's Rule, Revisited (Nov. 27, 2023)

Let $f, g : [a, b] \to \mathbb{R}$ be a continuous function that is differentiable on the open interval (a, b), except perhaps at $c \in (a, b)$, where we may assume $f(c) = g(c) = 0$. Then

$$\ell = \lim_{x \to c} \frac{f'(x)}{g'(x)} \text{ exists, and } \lim_{x \to c} \frac{f(x)}{g(x)} = \ell.$$

We consider a second case: suppose f, g diverge to $+\infty$, i.e.

$$\lim_{x \to c} f(x) = \lim_{x \to c} g(x) = +\infty.$$

Then consider the following; first, we may write

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

by Cauchy's Mean Value Theorem for some choice $c \in (x, y)$. If x, y is close to a, then c is naturally close to a as well, hence

$$\frac{f(x) - f(y)}{x - y} = \frac{f'(c)}{g'(c)} \text{ is close to } \ell.$$

Now, to see that the below holds for any y, we just write

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f(x) - f(y)}{g(x) - g(y)} = \lim_{x \to \infty} \frac{f(x)}{g(x)} \left(\frac{1 - \frac{f(y)}{f(x)}}{1 - \frac{g(y)}{g(x)}} \right),$$

of which the bracketed portion approaches 1 if we take $x \to a$; we may now estimate

$$\lim_{x \to a} \left| \frac{f(x)}{g(x)} - \ell \right| = \lim_{x \to a} \left| \frac{f(x) - f(y)}{g(x) - g(y)} - \ell \right| \leq \lim_{x \to a} \sup_{[c,a]} \left| \frac{f'(c)}{g'(c)} - \ell \right|.$$

If we pick a δ small enough such that $|x - a| < \delta \implies \left| \frac{f'(x)}{g'(x)} - \ell \right| < \varepsilon$, then notice $|c - a| \leq |y - a| < \delta$, and so

$$\left| \frac{f'(c)}{g'(c)} - \ell \right| < \varepsilon \implies \lim_{x \to a} \left| \frac{f(x)}{g(x)} - \ell \right| = 0. \quad \square$$

As an exercise, show

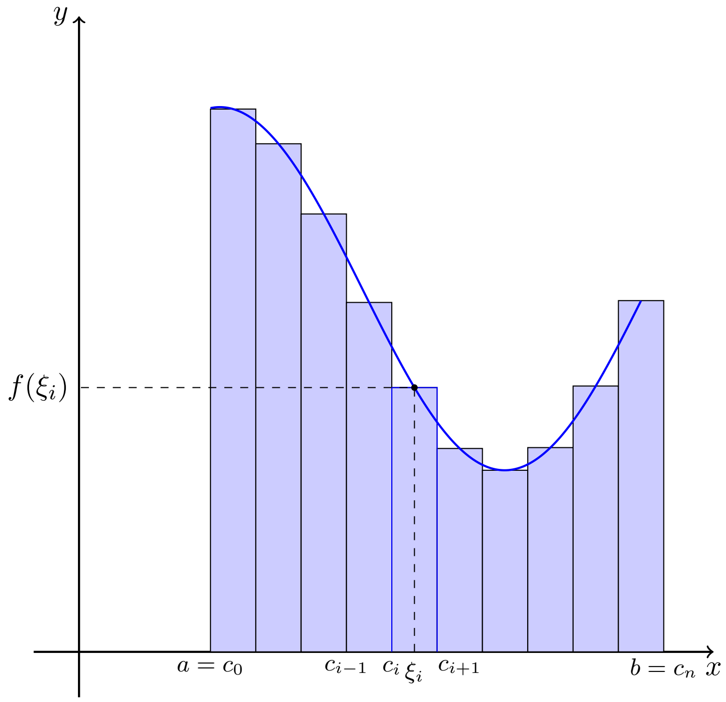
$$\lim_{x \to 0} \frac{\sin x - x}{x^3} = -\frac{1}{6}$$

through repeated applications of the L'Hôpital rule. We also went over some integration today, but I'll write it up in next time's notes.

Note: turns out my proof is scuffed so i will redo awhen i have time sorry!

§38 Day 32: Integration (Nov. 29, 2023)

Today we construct the integral. Given an integratable (more on this later) function f, construct $R(f,a,b)$ as the function describing the "area under the graph of f from a to b." As a basic idea, consider the following diagram, where we subdivide the graph of f into n rectangles, as seen below:



Consider $a = c_0$ and $b = c_n$, with c_i for all $0 \leq i \leq n$ being evenly distributed across where we are integrating (i.e. $[a,b]$). Note that ξ_i just denotes the center of the ith sub-interval[33] after splitting $[a,b]$. Intuitively, the larger n gets, the thinner the columns become, and the "finer" the approximation of $R(f,a,b)$. How do we exactly pick these rectangles though? Reasonably, we could look at a lower sum and an upper sum. Formally, let's write

$$P = \{c_0, \ldots, c_n\} \in [a,b] \text{ for some choice of } n \in \mathbb{N}.$$

Then define m_i, M_i as follows,

$$m_i = \inf_{[c_{i-1},c_i]} f(x),$$
$$M_i = \sup_{[c_{i-1},c_i]} f(x).$$

With this, we may bound $R(f,a,b)$ between the lower and upper sum as follows,

$$L(f,P) = \sum_{i=1}^{n} m_i(c_i - c_{i-1}) \leq R(f,a,b) \leq \sum_{i=1}^{n} M_i(c_i - c_{i-1}) = U(f,P),$$

It is much easier to compute these lower and upper sums to bound $R(f,a,b)$, since in the end they're just a bunch of rectangles anyways. Thus, let $U(f,P) - L(f,P)$ be an approximate error term. To refine this bound (i.e., minimize the error), suppose we have a different set, call it Q. Then Q is said to be a refinement of P if $Q \supset P$ (intuitively, Q

[33] god i don't know what numbering i'm following now

would split our interval $[a, b]$ into more sub-intervals than P does). As an exercise, check that

$$L(f, P) = \sum_{i=1}^{n} \inf_{[c_{i-1}, c_i]} f(x)(c_i - c_{i-1}) \leq \sum_{i=1}^{n} \underbrace{\inf_{[c_{i-1}, c_i]} f(x)(c_i - c_{i-1})}_{\text{index over elements of } Q} = L(f, Q),$$

and that a similar argument holds for $U(f, P) \geq U(f, Q)$.[34] From this, we see that if Q is a refinement of P, then clearly

$$U(f, P) - L(f, P) \geq U(f, Q) - L(f, Q),$$

which means it more closely approximates $R(f, a, b)$. Moreover, let

$$U \int_a^b f = \inf_P U(f, P),$$
$$L \int_a^b f = \sup_P L(f, P).$$

We say f is integratable over the interval $[a, b]$ if $U \int_a^b f = L \int_a^b f$ (for a suitably chosen P). The rest of class was spent talking about past problems on assignments.

[34] personally i'm leaving as exercise cuz my thoughts are scuffed on this one. tl;dr if an interval is unsplit, then it's the same in both summations. if it's split by a new term in Q, then it must be greater or equal. repeat and iterate... not sure how to formalize but that's the idea.

§39 Day 33: Integration II (Dec. 1, 2023)

Last time, we went over upper and lower sums of an interval of f, as well as talked about the refinements of a Riemann sum.

Anyways, right on- (okay how do you even spell integral-able in the correct english)

Theorem 39.1 (Continuity implies (Riemann) Integratibility)
Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is said to be Riemann integrable.

Recall that the definition of continuity is given by: if f is continuous at x, then for all $\varepsilon > 0$, there must exist a δ such that $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$. However, this is not enough; we need uniform continuity, defined as follows. A function $f : [a, b] \to \mathbb{R}$ is said to be uniformly continuous if, for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $x, y \in [a, b]$, we have

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Remark 39.2 (Continuity vs. Uniform Continuity)**.** The main difference here is that while ordinary continuity requires the existence of some δ that works at some point for whatever point we choose in $\operatorname{dom} f$, uniform continuity asks for a single δ that works for every possible $x \in \operatorname{dom} f$. In general, any given function that is continuous on a compact set is also uniformly continuous on said set: check here (Heine-Cantor Theorem).

To prove the above, let us first show that continuity implies uniform continuity.

Lemma 39.3 (Continuity $\implies$ Uniform Continuity (on $\mathbb{R}$))
Let $[a, b]$ be a closed interval, and let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous.

Proceed by contradiction. Suppose f is not uniformly continuous: then there must be some $\varepsilon > 0$ such that for all $\delta > 0$, we have $x, y \in [a, b]$ where $|x - y| < \delta$, but $|f(x) - f(y)| \geq \varepsilon$. Using this fact, construct sequences $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}}$ such that

$$|x_n - y_n| < \frac{1}{n}, \text{ but } |f(x_n) - f(y_n)| \geq \varepsilon.$$

Clearly, this sequence is bounded in $[a, b]$, so we may apply the Bolzano-Weierstrass theorem to get convergent subsequences $\{x_{n_k}\}_{k\in\mathbb{N}}$ and $\{y_{n_k}\}_{k\in\mathbb{N}}$ with limits x^* and y^* respectively. However, as we take $k \to \infty$, notice that

$$|x_{n_k} - y_{n_k}| < \frac{1}{n_k} \implies x^* = y^*.$$

This implies $f(x^*) = f(y^*)$, which is a contradiction (since $\varepsilon > 0$). Thus, f is clearly uniformly continuous. □

Now, we may prove that continuity implies Riemann integratibility[35]. Since $f : [a, b] \to \mathbb{R}$ is continuous, for any ε, there exists a δ (uniform continuity lemma above) such that $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$ for any choice of $x, y \in [a, b]$. Let P be a partition of $[a, b]$ into n equally spaced intervals with endpoints $c_0, \ldots, c_n$, with n large enough such that the size of each interval is less than δ. Now, directly compute the error term (as given by the difference between the lower and upper sums):

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_{i=1}^{n} (c_i - c_{i-1}) \underbrace{\left(\sup_{[c_{i-1}, c_i]} f(x) - \inf_{[c_{i-1}, c_i]} f(x) \right)}_{M_i - m_i} \\ &\leq n \left(\frac{b - a}{n} \cdot \varepsilon \right) \\ &= \varepsilon(b - a). \end{aligned}$$

For the second line above, note that $\sup f(x) - \inf f(x)$ on each of the sub-intervals is necessary less than ε because $c_i - c_{i-1} < \delta$. To conclude, for any $\varepsilon^* > 0$, if we pick $\varepsilon < \frac{\varepsilon^*}{b-a}$, then the error term $U(f, P) - L(f, P) < \varepsilon$, which concludes that f is indeed Riemann integrable on $[a, b]$. $\square$

[35]damn how do u spell this *cries*. i didn't catch almut's proof tho, so here's my own. sorry-

§40 Day 34: Integration III (Dec. 4, 2023)

Today, we proved that continuity implies Riemann integrable, basic properties of the integral, and the fundamental theorems of calculus[36]. Anyways; bits and pieces from class now!

(a) Let f, g be functions defined on the non-negative reals as given below,

$$f(x) = \begin{cases} \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}, \qquad g(x) = \begin{cases} \frac{1}{\sqrt{x}}, & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

Then f is not Riemann integrable over the interval $[0, c]$, while g is.[37]

(b) Let S be a finite set of points on the real line, $\{x_1, \ldots, x_k\}$. Then the indicator function I_S is Riemann integrable with

$$\int_a^b I_S = 0, \qquad \text{(Assuming } a \leq x_1, \ldots, x_k \leq b\text{)}$$

while the Dirichlet function $I_{\mathbb{Q}}$ is not. To check this, observe $L(f, P) = 0$ and $U(f, P) = 1$ for any set P by the density of rationals and irrationals in $\mathbb{R}$.

Lemma 40.1 ("Additivity of Domain")

Suppose we have f is given to be integrable on $[a, b]$ and $[b, c]$ with $a \leq b \leq c$. Then f is integrable on $[a, c]$, and satisfies

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

Given any $\varepsilon > 0$, observe that since f is integrable on $[a, b]$ and $[b, c]$ individually, let us have partitions Q and R of $[a, b]$ and $[b, c]$ respectively such that

$$U(f, Q) - L(f, Q) < \frac{\varepsilon}{2},$$
$$U(f, R) - L(f, R) < \frac{\varepsilon}{2}.$$

Then let $P = Q \cup R$, and observe $U(f, P) = U(f, Q) + U(f, R)$ as well as $L(f, P) = L(f, Q) + L(f, R)$. This automatically gives $U(f, P) - L(f, P) < \varepsilon$ as desired.

§40.1 Fundamental Theorems of Calculus

The fundamental theorem of calculus is given in two parts,

(a) Let f be a continuous, real-valued function on $[a, b]$. Then define F as follows,

$$F(x) = \int_a^x f.$$

Then F is differentiable on $[a, b]$ with $F' = f$.

[36] okay so I wrote Day 33 and Day 34 notes in 1 sitting and didn't realize we didn't cover the riemann integrable proof, so i ended up writing it in the previous day. my proof is very similar to almut's so we call it a day

[37] the proof is quick: take antiderivative and observe $\log 0$ dne.

(b) (Newton-Leibniz)[38] Assuming f is integrable on $[a, b]$ with $f = g'$ for some function g on $[a, b]$ (and g is differentiable on (a, b)). Then

$$\int_a^b f = g(b) - g(a).$$

Note from Almut: read f as g' in the integral above.

I think we're going to cover the proofs next time around? Blahaj time until Wednesday for now.

[38] damn this leibniz guy showin up everywhere my brother! think he gauss !! ??

§41 Day 35: Fundamental Theorem of Calculus (Dec. 6, 2023)

Today we proved the fundamental theorem of calculus.

(a) Let $f : [a,b] \to \mathbb{R}$ be a Riemann integrable function with an accompanying function F satisfying

$$F(x) := \int_a^x f = \int_a^x f(t)\,dt.$$

Then F is continuous (first fund. theorem). To prove this, start by observing that if f is integrable, then it must be bounded on the interval $[a,b]$. This means that there must exist some real $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in (a,b)$; we shall proceed to demonstrate continuity from the right. If f is integrable, then there must be a partition P of the interval $[a,x]$ such that

$$U(f,P) - L(f,P) < \frac{\varepsilon}{2}$$

for any $\varepsilon > 0$ of our choosing; moreover, for $h > 0$ (such that $x + h < b$), consider the partition Q given by

$$Q := P \cup \{x+h\} = (c_0, \ldots, c_n, c_{n+1}) = (c_0, \ldots, c_n, x+h).$$

Then we see

$$\begin{aligned} U(f,Q) &\leq U(f,P) + Mh, \\ L(f,Q) &\geq L(f,P) - Mh, \end{aligned}$$

which means $F(x+h) - F(x)$ is less than or equal to $U(f,Q) - L(f,P)$, but is greater or equal to $L(f,Q) - U(f,P)$, which we may refine on both ends to see

$$-\frac{\varepsilon}{2} - Mh \leq F(x+h) - F(x) \leq \frac{\varepsilon}{2} + Mh.$$

Therefore, we may pick $h = \frac{\varepsilon}{2M+1}$ (with the denominator as $2M+1$ as a "safety measure" instead of $2M$), we get

$$\begin{aligned} |F(x+h) - F(x)| &\leq \frac{\varepsilon}{2} + Mh \\ &\leq \varepsilon. \end{aligned}$$

The same argument holds for showing that F is continuous from the left as well.

(b) More specifically, let $f : [a,b] \to \mathbb{R}$ be integrable, and set $F(x) = \int_a^x f(t)\,dt$ as before; then given any $c \in [a,b]$, if f is continuous at c, then F must be differentiable at c with $F'(c) = f(c)$. If $c = a$ or $c = b$, then we obtain special cases of

$$\begin{aligned} F'(a^+) &= f(a), \\ F'(b^-) &= f(b). \end{aligned}$$

To prove this, take our c and let $h > 0$ be arbitrary; then by additivity of domain, we have

$$F(c+h) - F(c) = \int_a^{c+h} f - \int_a^c f = \int_c^{c+h} f.$$

Let P be a partition of $[c, c+h]$, given simply by $\{c, c+h\}$. Then we see

$$\frac{\inf_{c \leq x \leq c+h} f(x) \cdot h}{h} \leq \frac{F(c+h) - F(c)}{h} \leq \frac{\sup_{c \leq x \leq c+h} f(x) \cdot h}{h}$$

where the left and right hand sides both converge to $f(c)$ when we take $h \to 0$ and $x \to c$. This concludes the claim.

(c) Now for the second fundamental theorem of calculus! Let $f : [a, b] \to \mathbb{R}$ b e differentiable. Then $f = g'$ for some $g : [a, b] \to \mathbb{R}$ (that is continuous on $[a, b]$ and differentiable on (a, b)). Then

$$\int_a^b f = g(b) - g(a).$$

To prove this, let P be a partition of $[a, b]$. Directly compute for $g(b) - g(a)$:

$$g(b) - g(a) = g(c_n) - g(c_0) = \sum_{i=1}^{n} g(c_i) - g(c_{i-1}).$$

By MVT, let $t_i \in [c_{i-1}, c_i]$ such that

$$g'(t_i) = \frac{g(c_i) - g(c_{i-1})}{c_i - c_{i-1}}.$$

Since f is integrable, we see $|U(f, P) - L(f, P)| < \varepsilon$ for all $\varepsilon > 0$ (for a suitable P), which gives

$$g(b) - g(a) - \varepsilon \leq U(f, P) - \varepsilon \leq \int_a^b f \leq L(f, P) + \varepsilon = g(b) - g(a) + \varepsilon,$$

so

$$\left| \int_a^b f - (g(b) - g(a)) \right| < 2\varepsilon;$$

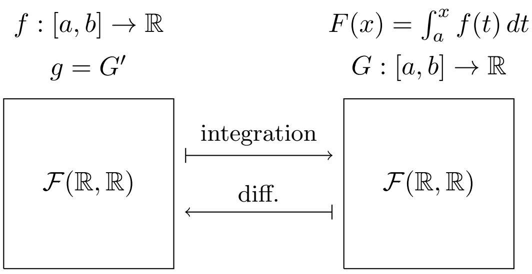
since ε is arbitrary, we are done. □

§42 Day 36: FTC; Relationship between Differentiation and Integration (Jan. 8, 2024)

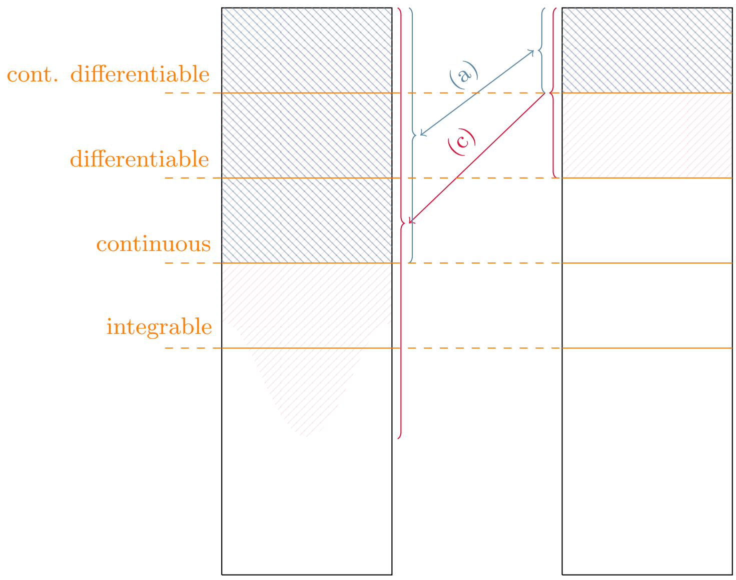
Today we go over the relationship between differentiating and integrating a function! Consider the following,

$$f : [a, b] \to \mathbb{R};\ F(x) = \int_a^x f(t)\, dt,$$
$$G : [a, b] \to \mathbb{R};\ g = G',$$

where F is the function obtained by integrating f (we call F to be the primitive of f, or the antiderivative; more on that in a sec), and g is the function obtained by differentiating G. Consider the following:



Observe that integrable functions are but a subset of $\mathcal{F}(\mathbb{R}, \mathbb{R})$, of which some are continuous; of the functions in $\mathcal{C}(\mathbb{R})$, some are differentiable, of which a smaller subset are continuously differentiable. Using this, we can graph the relationship between f, g, F, G as follows:[39]



(a) In the above north-east hatch (light blue), we see that all continuous functions g are the derivative of some continuously differentiable G; this comes straight from the fundamental theorem of calculus.

(b) In the opposite direction, while every continuously differentiable F is given to be the primitive (antiderivative) of some integrable f (again, by FTC), there exist differentiable functions that are not the primitive of some integrable f, as well as non-differentiable functions that are the primitive of some integrable f.

[39] i am losing my shit over this diagram... tikz is so finicky...

- For the former case, consider Volterra's Function, which is everywhere differentiable, but is in fact, not integrable.
- For the latter case, consider $f(x) = \text{sgn}(x)$ defined on the interval $I = [-1, 1]$. Then we see
$$F(x) = \int_{-1}^{x} f(t)\, dt\, |x| - 1,$$
which is clearly not differentiable at $x = 0$. Consider that such functions occur only if f is non-continuous.

(c) Given some differentiable G, we see $\mathcal{C}(\mathbb{R})$ is a subset of all possible $g = G'$; however, there are also examples of g being non-integrable, or being discontinuous and integrable while not being the derivative of some G.

- Recall that the Riemann integral requires the function to be bounded; if we take
$$G(x) = \begin{cases} x^2 \sin(1/x) & , x \neq 0 \\ 0 & , x = 0 \end{cases},$$
we see it is differentiable on $I = [-1, 1]$, but the derivative is unbounded, so $G' = g$ is non-integrable.
- I can't think of a function that is integrable but is not the derivative of something. sorry :3

§42.1 Notation for Indefinite Integrals

If $F' = f$, we say F is the primitive of f, as in $\int f = F$ (where $\int f$ is the indefinite integral) implies the existence of some a, c such that[40]

$$\int_a^x f(t)\, dt = F(x) + c.$$

Ideally, we have $F' = f \iff \int f = F$, which happens when we're working in the continuous differentiable case, as in (a) from above. Examples from class now!

(a) For $n \in \mathbb{N}$, we have

$$\frac{d}{dx} x^n = n x^{n-1} \implies \int x^n\, dx = \frac{x^{n+1}}{n+1} + C.$$

This also holds for n in the negative integers, with $\frac{d}{dx} x^{-n} = n \frac{n}{-x^{n+1}}$.

(b) In the $n = -1$ case from above, we define

$$\log x = \int_1^x \frac{1}{t}\, dt \implies \int \frac{1}{x} = \log x + c.$$

(c) In general, if we have $F = \int f$, we have

$$\int f(ax + b) = \frac{1}{a} F(ax + b).$$

Case in point, observe

$$\int \frac{1}{3x + 7}\, dx = \frac{1}{3} \log(3x + 7).$$

[40] someone check with me here

(d) Almut decided to do something unhinged for fun:

$$\int xe^{-x^2}\,dx = -\frac{1}{2}\int(-2x)e^{-x^2}\,dx = -\frac{1}{2}\int(f\circ g)'(x)\,dx = -\frac{1}{2}e^{-x^2} + C,$$

where $f(y) = e^y$ and $g(x) = -x^2$ being defined conveniently to apply the reverse chain rule.

Another small lemma at the end of the class: using the fact that $\log : (0,\infty) \to \mathbb{R}$ is bijective, we may define $\exp = \log^{-1}$; since $\exp(0) = 1$ and $\exp(a+b) = \exp(a) + \exp(b)$, observe

$$\exp'(y) = (\log^{-1})'(y) = \frac{1}{\frac{1}{\exp y}} = \exp y.$$

§43 Day 37: Linearity and Composition of Integral (Jan. 10, 2024)

As talked about on Monday, the "ideal" case for integrals is when we have f continuous, implying

$$F := \int_a^x f(t)\,dt$$

to be continuously differentiable (and so $F' = f$), which lets us apply the fundamental theorem of calculus to get $\int_a^b F'(x)\,dx = F(b) - F(a)$. In this case, our current long term goal is to figure out how to compute integrals explicitly, using an "integral table" of pre-established rules for simple functions. The table is not included in the notes because it's kinda elementary and I spent like 30 minutes trying to figure it out and I gave up sorry!!

Now, we prove the linearity of integrals; claim that $\int(\alpha f + g) = \alpha \int f + \int g$. To do this, we start by proving additivity: suppose $f, g : [a,b] \to \mathbb{R}$ to be integrable. Then let $P : [a,b]$ be a partition $P = \{t_0, \dots, t_n\}$. We may write,

$$\begin{aligned}
U(f+g,P) &= \sum_{j=1}^{n} \left(\sup_{x\in[t_{j-1},t_j]} (f(x)+g(x))\,(t_j - t_{j-1}) \right) \\
&\le \sum_{j=1}^{n} \left(\sup_{x\in[t_{j-1},t_j]} f(x) + \sup_{x\in[t_{j-1},t_j]} g(x) \right)(t_j - t_{j-1}) \\
&= U(f,P) + U(g,P).
\end{aligned}$$

By the same way, $L(f+g,P) \ge L(f,P) + L(g,P)$. Using these two we may obtain

$$\begin{aligned}
U\int_a^b (f+g) = \inf_P U(f+g,P) &\ge \inf_P (U(f,P) + U(g,P)) \\
&\ge \inf_P U(f,P) + \inf_P U(g,P) \\
&= \int f + \int g.
\end{aligned}$$

Once again, in a similar fashion, we may obtain $L\int_a^b (f+g) \le \int f + \int g$. By refining our partition P, we see $\int f + \int g$ is squeezed between two equal limits, and we are done. For scalar multiplication, simply consider that $U(\alpha f, P) = \alpha U(f,P)$ and $L(\alpha f, P) = \alpha L(f,P)$ for $\alpha \ge 0$, and $U(\alpha f, P) = -\alpha L(f,P)$ and $L(\alpha f, P) = -\alpha U(f,P)$[41] in the case $\alpha < 0$.

Now, for composition. Let $g : [a,b] \to [c,d]$ be continuously differentiable and bijective, and $f : [c,d] \to \mathbb{R}$ be integrable (so we may say $F = \int f$). Given that f is continuous, we have

$$\int (f \circ g) g' = F \circ g,$$

i.e. $\int_a^b f(g(x))g'(x)\,dx = \int_c^d f(y)\,dy$. To prove this, consider the chain rule on $(F \circ g)'(x)$. We get $F'(g(x))g'(x)$, which is simply $(f \circ g)g'$; using this, we may write

$$\int_a^b (f \circ g) g'(x)\,dx = \int_a^b (F \circ g)')(x)\,dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(x)\,dx.$$

[41] i think this is right, since we're considering the graph of f flipped

§44 Day 38: Integration by Parts (Jan. 12, 2024)

Typing this up like, one month late, so I don't have records of what happened in class, but I'll retrieve some stuff I had on my notes. The main idea is to reverse the product rule as follows; let f, g be integrable, real-valued functions. Then consider that

$$(fg)' = fg' + f'g,$$

and by careful manipulation of the terms, we obtain

$$\begin{aligned}(fg)' &= fg' + f'g \\ \int (fg)' &= \int fg' + \int f'g \\ \int fg' &= fg - \int f'g.\end{aligned}$$

Note that the bounds given for a definite integral is preserved in this operation. This is an elementary integration technique that ends up being very useful in a bunch of circumstances; for example, consider the below,

$$\int e^x \cos x \, dx.$$

Let us pick $f(x) = \cos x \implies f'(x) = -\sin x$, and $g(x) = e^x \implies g'(x) = e^x$. Then we have

$$\int e^x \cos x \, dx = e^x \cos x + \int e^x \sin x \, dx,$$

which upon applying integration by parts again, we get

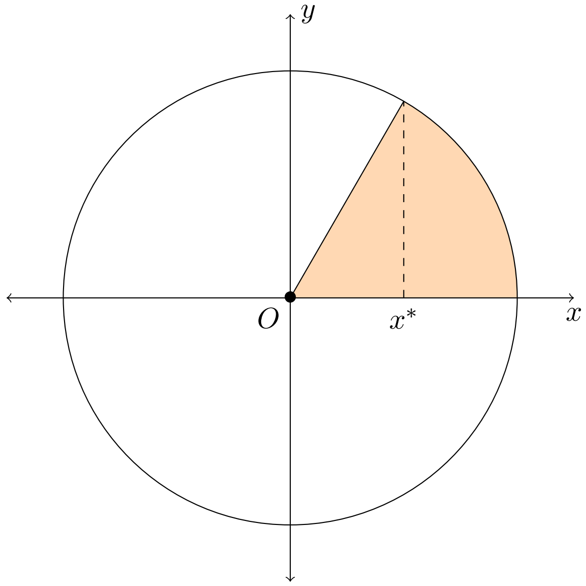
$$\begin{aligned}\int e^x \cos x \, dx &= e^x \cos x + e^x \sin x - \int e^x \cos x \, dx \\ 2 \int e^x \cos x \, dx &= e^x (\cos x + \sin x) + C \\ \int e^x \cos x \, dx &= \frac{1}{2} e^x (\cos x + \sin x) + C. \quad \square\end{aligned}$$

This concludes our example.

§45 Day 39: Derivation of Trigonometric Functions; "A Sanity Check" (Jan. 15, 2024)

Some administrative details first; there will be extra office hours TODAY and Wednesday at **1pm**, in **BA6234**. Diba (a TA) will also be hosting her office hours on Tuesday and Thursday in Gerstein; check her Quercus announcement for more information. The exam is from **7-9pm** in **EX100**.

Today, we will sanity check the trigonometric functions. Quick note that I got pretty lost during lecture today, so my notes are also pretty scuffed. First, though, a quick treatise on log; we assume log to be the natural log, i.e. ln or $\log_e$. Let $A(x)$ denote the area of a sector of the unit circle, where x is given as below (ok it is labeled as x^* to prevent confusion but you get me),



Then we may write

$$A(x) = \frac{1}{2}x\sqrt{1-x^2} + \int_x^1 \sqrt{1-t^2}\,dt,$$

where the former half describes the area of the triangle, and the latter half describes the area of the leftover sector. Differentiating this, we get $A'(x) = -\frac{1}{2}\arcsin x$.[42]

Now, we want to do an angle conversion to x; i.e., find which angle corresponds to any given choice of $x \in [-1, 1]$. Let $B(x) = cA(x)$. For our purposes, let $c = 2$ for now. Then

$$B'(x) = 1 - \frac{1}{\sqrt{1-x^2}}^{43},$$

with $B(1) = 0$ and $B(-1) = \pi$. Define $\cos = B^{-1}$. Then we may find x such that $B(x) = \theta$, since $\sin\theta = \sqrt{1-\cos^2\theta}$ for $0 \le \theta \le \pi$. We may also check

$$\cos'\theta = (B^{-1})'(\theta) = \frac{1}{B'(\cos\theta)} = -\sqrt{1-\cos^2\theta} = -\sin\theta.$$

[42] this was not what i had on paper and not what i saw on the board i think...
[43] ok where does this come from now

§46 Day 40: Digression on Properties of Trigonometric Functions (Jan. 19, 2024)

Not sure where the Jan. 17th lecture notes went, sorry. As previously established in class, sin and cos are trigonometric functions operating on $\mathbb{R}$ with an image of $[-1, 1]$, i.e.

$$\cos : \mathbb{R} \to [-1, 1],$$
$$\sin : \mathbb{R} \to [-1, 1].$$

Moreover, they are periodic with a period of 2π (more specifically, $\cos(\theta + \pi) = -\cos\theta$, with the same about sin), which lets us see

$$\cos^2\theta + \sin^2\theta = 1$$

by construction. Observing that

$$\cos' = -\sin,$$
$$\sin' = \cos,$$

we see that any function $y(\theta)$ that solves $y'' = y$ (on some interval (a, b)) is equivalent to the form $y(\theta) = A\cos\theta + B\sin\theta$ with some constant A, B.

§46.1 Proof of Angle Addition Formula

Fix $\beta \in \mathbb{R}$, and consider $y(\theta) = \sin(\theta + \beta)$. Note that $y''(\theta) = -\sin(\theta + \beta) = -y(\theta)$; by our lemma above, we see

$$y(\theta) = A\cos\theta + B\sin\theta \text{ for some } A, B \in \mathbb{R};$$

to find A, B, let us compute

$$y(0) = \sin\beta = A \cdot 1 + B \cdot 0,$$
$$y'(0) = \cos\beta = A \cdot 0 + B \cdot 1,$$

implying $A = \sin\beta$, $B = \cos\beta$, and we have $\sin(\theta+\beta) = \sin\beta\cos\theta + \cos\beta\sin\theta$ as desired. This derivation works similarly for cos as well.

§46.2 Other Trigonometric Derivatives

Let us compute $\tan'(\theta)$ as follows,

$$\tan'(\theta) = \frac{d}{dt}\left(\frac{\sin\theta}{\cos\theta}\right) = \frac{\cos^2\theta + \sin^2\theta}{\cos^2\theta} = \frac{1}{\cos^2\theta} = \sec^2\theta.$$

Alternatively, we also have

$$1 + \tan^2\theta = 1 + \frac{\sin^2\theta}{\cos^2\theta} = \frac{1}{\cos^2\theta} = \sec^2\theta.$$

This means $\tan\theta$ is a solution to $y' = 1 + y^2$. Let us consider $\arctan(x) := \tan^{-1}(x)$; clearly, $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ is a smooth bijection (read: has a continuous derivative[44]), and observe

$$\begin{aligned}
\arctan'(\theta) &= \frac{1}{\tan'(\arctan\theta)} \\
&= \frac{1}{\sec^2(\arctan\theta)} \\
&= \cos^2(\arctan x) \\
&= \frac{1}{x^2}, \\
\implies \int \frac{1}{1+x^2}\,dx &= \arctan(x) + C.
\end{aligned}$$

I'll write about partial fraction decomposition next class.

[44] i think

§47 Day 41: Partial Fraction Decomposition and Trig Substitution Example (Jan. 22, 2024)

We went over partial fraction decomposition today. Here are the examples from class.

(a) First example gone over in class as below,

$$\begin{aligned}\int \frac{1}{x^2-1}\,dx &= \int \frac{A}{x+1}+\frac{B}{x-1}\,dx && \text{(Partial Fraction Decomp.)}\\ &= \underbrace{\frac{1}{2}\int \frac{1}{x-1}-\frac{1}{x+1}\,dx}_{\text{Solve } 1=A(x-1)+B(x+1)}\\ &= \frac{1}{2}(\log(x-1)-\log(x+1)) = \log\sqrt{\frac{x-1}{x+1}}+C.\end{aligned}$$

(b) Second example; not exactly sure where this one went, so I'll rewrite it. Note that in class, we used $Cx+D$ instead of C for the decomposition to maintain a degree of 3 for x across all terms; I just don't find it completely necessary so I'm writing it here.

$$\begin{aligned}\int \frac{x^2}{1-x^4}\,dx &= -\underbrace{\int \frac{A}{1-x}+\frac{B}{1+x}+\frac{C}{1+x^2}\,dx}_{(1-x)(1+x)(1+x^2)=1-x^4}\\ &= -\int \frac{1}{4}\left(\frac{1}{x-1}\right)-\frac{1}{4}\left(\frac{1}{x+1}\right)+\frac{1}{2}\underbrace{\left(\frac{1}{x^2+1}\right)}_{\text{std. integral}}\,dx\\ &= \frac{\log(x+1)-\log(x-1)}{4}-\frac{\arctan(x)}{2}+C.\end{aligned}$$

(c) Trig sub example; I also lost track of this one in class, so I'm rewriting it here. Start by taking $x=\tan u \implies dx=\sec^2(u)\,du$.

$$\begin{aligned}\int \sqrt{1+x^2}\,dx &= \int \sec^2(u)\sqrt{\tan^2(u)+1}\,du\\ &= \int \sec^3(u)\,du. && \left(\text{Note: } \sqrt{\tan^2(u)+1}=\sec u\right)\end{aligned}$$

Now, proceed with IBP as follows,

$$\begin{aligned}\int \sec u\sec^2 u\,du &= \sec u\tan u - \int \sec u\tan^2 u\,du\\ &= \sec u\tan u - \int \sec^3 u\,du + \underbrace{\int \sec u\,du}_{\text{std. integral}}\end{aligned}$$

$$\begin{aligned}&\implies 2\int \sec^3 u\,du = \sec u\tan u + \log|\sec u+\tan u|\\ &\implies \int \sec^3 u = \frac{1}{2}\left(\sec u\tan u+\log|\sec u+\tan u|\right).\end{aligned}$$

Using the fact that $u=\arctan(x)$, we see $\tan(\arctan(x))=x$ and $\sec(\arctan(x))=\sqrt{x^2+1}$, which gives

$$\int \sqrt{1+x^2}\,dx = \frac{1}{2}\left(\log\left|\sqrt{x^2+1}+x\right|+x\sqrt{x^2+1}\right)+C.$$

§48 Day 42: More Integration Techniques; Secant Integral; Weierstrass Substitution (Jan. 24, 2024)

Elementary functions are "nice" functions are single-variable functions that can be finitely expressed as sums, products, compositions, solving equations, etc... of polynomial, trigonometric, exponential/logarithmic functions, and so on. Note that elementary functions need not have elementary integrals. Examples include

$$e^{-x^2},\ \frac{\sin x}{x},\ \sin(x^2),$$

and so on.

§48.1 Secant Integral

We will be going over two ways to derive $\int \sec\theta\, d\theta$. First, start by directly computing as follows,

$$\begin{aligned}
\int \sec\theta\, d\theta &= \int \frac{1}{\cos\theta}\, d\theta \\
&= \int \frac{\cos\theta}{1-\sin^2\theta}\, d\theta \\
&= \int \frac{1}{1-y^2}\, dy && (y=\sin\theta \implies dy=\cos\theta\, d\theta) \\
&= \frac{1}{2}\int \frac{1}{1-y}+\frac{1}{1+y}\, dy && \text{(Partial Fraction Decomp.)} \\
&= \frac{1}{2}\log\left(\frac{1+\sin\theta}{1-\sin\theta}\right)+C \\
&= \log\left(\sec\theta+\tan\theta\right)+C.
\end{aligned}$$

Alternatively, we could proceed with a different substitution; let $t=\tan\frac{\theta}{2}$, and observe that $\theta = 2\arctan t$, yielding

$$d\theta = \frac{2}{1+t^2}\, dt \implies \cos\theta = \cos^2\left(\frac{\theta}{2}\right)-\sin^2\left(\frac{\theta}{2}\right) = \frac{1}{1+t^2}-\frac{t^2}{1+t^2}.$$

Using this substitution[45], we have

$$\begin{aligned}
\int \sec\theta\, d\theta &= \int \frac{1}{\frac{1-t^2}{1+t^2}}\frac{2}{1+t^2}\, dt && (\text{Substitute } t=\tan\tfrac{\theta}{2}) \\
&= \int \frac{2}{1-t^2}\, dt \\
&= \log\left(\frac{1+t}{1-t}\right)+C \\
&= \log\left(\frac{1+\tan\frac{\theta}{2}}{1-\tan\frac{\theta}{2}}\right)+C \\
&= \log\left(\sec\theta+\tan\theta\right)+C.
\end{aligned}$$

[45] written poorly sorry, idk how else to express it

§48.2 Weierstrass Substitution

I wasn't in class during this day, so I'm kinda just reading off of what someone told me had happened, and I'm just making my own judgment call to put this section in because I found it useful in Assignment 10 (albeit in retrospect...). When substituting $t = \tan \frac{\theta}{2}$ for other trigonometric expressions, we may substitute the following into the integrand as well:

$$\begin{aligned} \sin\theta &= \frac{2t}{1+t^2}, \\ \cos\theta &= \frac{1-t^2}{1+t^2}, \\ d\theta &= \frac{2}{1+t^2}\, dt. \end{aligned}$$

The above formulas can be quickly derived from double angle identities applied as follows,

$$\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2} = \frac{2\tan\frac{\theta}{2}}{1+\tan^2\frac{\theta}{2}}; \qquad (\sec^2\theta = 1 + \tan^2\theta)$$

the other identity is presented in the previous page.

§49 Day 43 and 44: Taylor Theorem (Jan. 26, 2024 & Jan. 29, 2024)

Combining Friday and Monday's stuff together! Let $f : I \to \mathbb{R}$ (where I is an open interval) be an n-times differentiable function at $a \in I$. In the case $n = 0$, we assume f is continuous at $x = a$. Define the Taylor polynomial (of order n at $x = a$) as

$$P_{n,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Then

$$\lim_{x\to a} \underbrace{\frac{f(x) - P_{n,a}(x)}{(x-a)^n}}_{h(x)} = 0.$$

Moreover, let us define the error to be $R_{n,a}(x) := f(x) - P_{n,a}(x)$. Then we claim the following,

(a) There exists some function h (as given above) such that $R_{n,a}(x) = h(x)(x-a)^n$ with $\lim_{x\to a} h(x) = 0$. Moreover, we have that h is continuous at $x = a$. This is called the Peano form of the remainder.[46]

If f is $n+1$ times differentiable with $f^{(n)}$ continuous on the closed interval between x and a[47], then

(b) There exists some constant t_c between x and a such that

$$R_{n,a}(x) = \frac{f^{(n+1)}(t_c)}{n!}(x-t_c)^n(x-a).$$

This is called the Cauchy remainder.

(c) There exists some constant t_l between x and a such that

$$R_{n,a}(x) = \frac{f^{(n+1)}(t_l)}{(n+1)!}(x-a)^{n+1}.$$

This is called the Lagrange remainder.

Now, for the proof. For (a), by construction, let us start by observing

$$\left(\frac{d}{dx}\right)^k P_{n,a}(x)\Big|_{x=a} = f^{(k)}(a) \text{ for } k = 0, 1, 2, \ldots, n.$$

To see this, observe that $P_{n,a}(x)$ is a polynomial where each term vanishes if it contains $(x-a)$ to some power. Then, similar to Assignment 10 Problem 4, observe that the fraction is of "type 0/0" (since $R_{n,a}(x) = h(x)(x-a)^n = 0$ at $x = a$), so we may apply L'Hôpital's n times to obtain

$$\begin{aligned}\lim_{x\to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} &= \frac{1}{n!} \cdot \frac{d}{dx}\left(f^{(n-1)}(x) - P_{n,a}^{(n-1)}(x)\right)\Big|_{x=a} \\ &= 0. \quad \square\end{aligned}$$

[46] i believe R refers to remainder here...

[47] not sure if this condition is needed...

For part (b), assuming f is n-times differentiable on I, and $f^{(n+1)}(a)$ exists, let us consider $S(a) = f(x) - P_{n,a}(x)$ as a function of a. Then we have,

$$S(t) = f(x) - P_{n,t}(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k,$$

which is differentiable w.r.t. t, as follows,

$$\frac{d}{dt}S(t) = 0 - \underbrace{\left(\sum_{k=0}^{n} f^{(k+1)}(t)\frac{(x-t)^k}{k!} - \sum_{k=1}^{n} f^{(k)}(t)\frac{(x-t)^{k-1}}{(k-1)!}\right)}_{\text{Product rule for } k\neq 0;\ k=0 \text{ case in first summation.}}.$$

Observe that the terms in both summations cancel each other out on the corresponding kth derivatives of f, which lets us finish with

$$S'(t) = \sum_{k=1}^{n} f^{(k)}(t)\frac{(x-t)^{k-1}}{(k-1)!} - \sum_{k=0}^{n} f^{(k+1)}(t)\frac{(x-t)^k}{k!} = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Now, using our formula above, we may apply the mean value theorem to see there exists some t_c between x and a such that

$$\begin{aligned} S'(t_c) &= \frac{S(x) - S(a)}{x-a} = -\frac{S(a)}{x-a} \\ \implies S(a) &= -(x-a)S'(t_c) = \frac{f^{(n+1)}(t_c)}{n!}(x-t_c)^n(x-a) \end{aligned}$$

Using the fact that $S(a) = f(x) - P_{n,a}(x) = R_{n,a}(x)$ as per definition, we see this completes the derivation of the Cauchy remainder.

For part (c), I don't know how to prove it with Cauchy MVT. Will try to figure it out later.

ruchir wrote this i will refine later...

$$\begin{aligned} R_{n,a} &:= f(x) - P_{n,a}(x) \\ P_{n,a}(x) &:= \sum_{k=0}^{n} \frac{f^k(x)}{k!}(x-a)^k \\ f(x) &= P_{n,a}(x) + \int_a^x \frac{f^{n+1}(t)}{n!}(x-t)^n dt \\ f(x) - P_{n,a}(x) &= \int_a^x \frac{f^{n+1}(t)}{n!}(x-t)^n dt \end{aligned}$$

Let

$$\begin{aligned} m &= \inf_{[x,a]} f^{n+1}(x), \\ M &= \sup_{[x,a]} f^{n+1}(x) \end{aligned}$$

$$m(x-t)^n \leq f^{n+1}(x)(x-t)^n \leq M(x-t)^n$$

$$\int_a^x m(x-t)^n dt \leq \int_a^x f^{n+1}(x)(x-t)^n dt \leq \int_a^x M(x-t)^n dt$$

$$m(x-t)^n \leq \int_a^x f^{n+1}(x)(x-t)^n dt \leq M(x-t)^n$$

$$m\frac{(x-t)^{n+1}}{n+1} \leq \int_a^x f^{n+1}(x)(x-t)^n dt \leq M\frac{(x-t)^{n+1}}{n+1}$$

$$m\frac{(x-t)^{n+1}}{(n+1)!} \leq \frac{1}{n!}\int_a^x f^{n+1}(x)(x-t)^n dt \leq M\frac{(x-t)^{n+1}}{n+1}$$

$$m \leq \frac{\frac{1}{n!}\int_a^x f^{n+1}(x)(x-t)^n dt}{\frac{(x-t)^{n+1}}{(n+1)!}} \leq M$$

$$m \leq f^{n+1}(x) \leq M \qquad \text{(by IVT)}$$

$$f^{n+1}(x) = \frac{\frac{1}{n!}\int_a^x f^{n+1}(x)(x-t)^n dt}{\frac{(x-t)^{n+1}}{(n+1)!}}$$

$$f^{n+1}(x)\frac{(x-t)^{n+1}}{(n+1)!} = \frac{1}{n!}\int_a^x f^{n+1}(x)(x-t)^n dt$$

$$R_{n,a} = f^{n+1}(x)\frac{(x-t)^{n+1}}{(n+1)!}$$

§50 Day 45: Examples and Concepts of Taylor Polynomials / Theorem (Jan. 31, 2024)

Today we go over examples of Taylor polynomials. Consider $f(x) = \cos x$, expanded around $a = 0$. Last class, it was shown that (I haven't written it down yet) that

$$\cos^{(k)}(0) = \begin{cases} 0, & n \text{ odd} \\ (-1)^n, & n = \frac{k}{2} \end{cases}.$$

To compute the $2k$'th Taylor polynomial, we may write

$$P_{2k,0}(x) = \sum_{j=0}^{k} \frac{(-1)^j}{(2j)!} x^{2j},$$

where we may note every other term vanishes, since for those terms, evaluating sin at 0 causes said term to vanish. Thus, we ask the following question: Does $P_{n,0}(x)$ converge to $\cos x$? Does it for all x, and how fast does it converge? Start by letting $n = 2k$ be even, and considering the integral form of the remainder, given by

$$R_n(x) = \cos x - P_{n,0}(x) = \int_0^x \frac{\cos^{(n+1)}(t)}{n!} x^n \, dt,$$

which we may note exists, since cos is infinitely differentiable. Furthermore, we may rewrite the integral as follows (using $n = 2k$),[48]

$$\int_0^x \frac{\cos^{(n+1)}(t)}{n!} x^n \, dt = \frac{(-1)^k}{(2k)!} \int_0^x x^{2k} \sin t \, dt = \frac{(-1)^k x^{2k}}{(2k)!} \sin x.$$

Ignoring the $\sin x$ term for now, we will prove that for all $x \in \mathbb{R}$ we have

$$\lim_{n\to\infty} \frac{x^n}{n!} = 0. \qquad \text{(Special Consequence of Stirling)}$$

We will prove this in two ways.

(a) Consider Stirling's Formula, given by

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n, \lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = 1.$$

Plugging into our original limit, we have

$$\begin{aligned} \lim_{n\to\infty} \frac{x^n}{n!} &= \lim_{n\to\infty} \frac{x^n}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} \\ &= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi n}} \cdot \frac{x^n}{\left(\frac{n}{e}\right)^n} \\ &= \left(\lim_{n\to\infty} \frac{1}{\sqrt{2\pi n}}\right)\left(\lim_{n\to\infty} \left(\frac{ex}{n}\right)^n\right) && \text{(Both Limits} \to 0) \\ &= 0. \quad \square \end{aligned}$$

[48] at this point i've lost the plot so i'm just winging

(b) Alternatively, consider the sequence $\{a_n\}$ defined by $a_n = \frac{x^n}{n!}$; if $x = 0$, it is trivial to see $a_n \to 0$; otherwise, if $x > 0$, consider that each term a_n is positive (since numerator and denominator are both positive), and that for $n \geq x$, we have

$$a_{n+1} = \frac{x^{n+1}}{(n+1)!} = \frac{x}{n+1} a_n < a_n;$$

by the monotone convergence theorem, we may conclude that $a_n \to 0$ as $n \to \infty$. For the case $x < 0$, simply append $(-1)^n$ to the term definition.

Since the limit of $\frac{x^n}{n!}$ goes to 0 and $\sin x$ is bounded, we may conclude that $R_n(x) = 0$ (with $n \to \infty$), meaning $\cos x$ coincides with $P_{n,0}(x)$ for large enough n. □

§51 Day 46: Interval of Convergence (Feb. 2, 2024)

Recall that we proved the special case of Taylor's Theorem (expansion around $x = 0$) from Assignment **10**, Problem **4**, as demonstrated below

$$\lim_{x\to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} = 0.$$

Today, let us fix x, and let $n \to \infty$. When does the limit still hold? Let's start with an example: let

$$P_{n,0,f}(x) = \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1-x}.$$

Then as $n \to \infty$, we have

$$\lim_{n\to\infty} P_{n,0,f}(x) = \begin{cases} \frac{1}{1-x} = f(x), & \text{if } |x| < 1 \\ +\infty, & \text{if } x > 1 \\ \text{divergence otherwise} \end{cases}$$

This means that for $-1 < x < 1$, we have $f(x) - P_{n,0,f}(x) = 0$, implying

$$\lim_{x\to 0} \frac{f(x) - P_{n,0,f}(x)}{(x-0)^n} = 0$$

as desired. Moreover, note that as $n \to \infty$, the remainder term $R_{n,a}(x) := f(x) - P_{n,a}(x)$ approaches 0 (which only occurs if $P_{n,0,f}(x)$ coincides with $f(x)$; in our case, if $-1 < x < 1$), which means we may write

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)(x-a)^k}{k!}.$$

Let us observe some examples.

- Let $g(x) = \frac{1}{1+x^2} = f(-x^2)$. Then we have the Taylor polynomial

$$P_{2n,0,g}(x) = P_{n,0,f}(x) = \frac{1 - x^{2n}}{1 + x^2}.$$

- Let us have an infinitely differentiable function on $\mathbb{R}$ that satisfies $f^{(n)}(0) = 0$ for all n. We claim that such an f is given by

$$f(x) = \begin{cases} 0, & x \le 0 \\ e^{-\frac{1}{x}}, & x > 0 \end{cases}$$

  I leave it as an exercise to check that this satisfies the conditions above; moreover, its Taylor polynomial vanishes as a result.

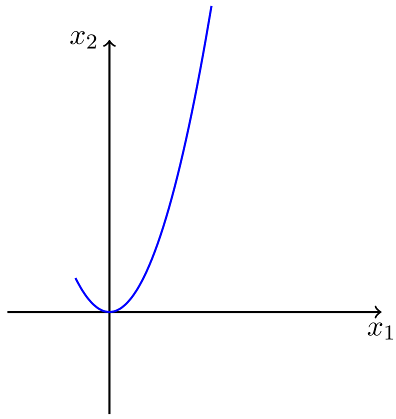
§52 Day 47: Curves (Feb. 5, 2024)

A (continuous) curve in $\mathbb{R}^n$ is a function given by

$$t \mapsto (x_1(t), \dots, x_n(t))$$

where each $x_1, \dots, x_n$ are individually continuous, real-valued functions on an interval I (given by $[a, b]$, or (a, b), or (a, ∞), etc...). For example, we have

- For $n = 2$ (i.e. over $\mathbb{R}^2$), then we have the curve $\gamma(t) = (x_1(t), x_2(t))$; we can also write it as two functions u, v where $\gamma(t) = (u(t), v(t))$ going over each component. For a visual example, suppose $u(t) = x$ and $v(t) = x^2$, and we are working over the interval $I = (-1, 5)$.[49]



- For $n = 3$, we may write $\gamma(t) = (u(t), v(t), w(t))$, and so on.

We can construct more examples; suppose we wish to parameterize a circle of radius 5, around an arbitrary center (h, k). This is given by the formula

$$(x_1 - h)^2 + (x_2 - k)^2 = 5^2,$$

which we may parameterize as

$$\begin{aligned} u(t) &= 5\cos t + h, \\ v(t) &= 5\sin t + k. \end{aligned}$$

Considering $\gamma(t) = (u(t), v(t))$, then we see γ is periodic with period 2π (verify this with the period of sin and cos individually). For another example, consider the Archimedean spiral, given by $\gamma(t) = (t\cos t, t\sin t)$ over the interval $I = [0, \infty)$; (sorry no graph for now, I can't figure out how to make LaTeX cooperate).

Another example we had in class was $\gamma(t) = (\cos t, \sin t, \frac{t}{10})$, which takes the form of a coil (I'll graph this in mathematica and upload here, maybe. That seems more viable as a long-term solution).

[49] i'll be real this is not from -1 to 5 if it was the graph would explode off the page. it's a *visual* not the damn real thing

§52.1 Polar Coordinate Transformation

Let's start by constructing a surjection between coordinates of the Cartesian form $(x, y) \in \mathbb{R}^2$ and coordinates of the polar form (r, θ):

$$x = r\cos\theta, y = r\sin\theta$$

$$r = \sqrt{x^2 + y^2}, \theta = \begin{cases} \arctan\frac{y}{x} & \text{if } x > 0 \\ \arctan\frac{y}{x} + \pi & \text{if } x < 0, y \geq 0 \\ \arctan\frac{y}{x} - \pi & \text{if } x, y < 0 \\ \frac{\pi}{2} & \text{if } x = 0, y > 0 \\ -\frac{\pi}{2} & \text{if } x = 0, y < 0. \end{cases}$$

Note that the origin is preserved between transformations (though θ would be undefined in this case, though I don't know if this is our convention in class). We constructed the polar equation of an ellipse given foci F_1, F_2 with the distance from its surface to the foci given by $a > |F_1 - F_2|$ in class today, but I didn't catch it so I'll leave it to next lecture for now! :¿

§53 Day 48: General Geometric Description of an Ellipse (Feb. 12, 2024)

An ellipse is determined by two points $A, B \in \mathbb{R}^2$ (called focus/foci), where

$$A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \text{ and } B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Let $a > |B - A|$ be the distance from these two foci. In this manner, we have

$$\{P \in \mathbb{R}^2 : |P - A| + |P - B| = 2a\},$$

where $|P - A|$ and $|P - B|$ represent the usual Euclidean metric. In cartesian coordinates, for the sake of simplicity, let us fix $a > c > 0$, where A is on $(c, 0)$ and B is on $(-c, 0)$. Then

$$P = \begin{pmatrix} x \\ y \end{pmatrix} \implies 2a = \sqrt{(x-c)^2 + y^2} + \sqrt{(x+c)^2 + y^2},$$

which we simplify as follows,

$$\begin{aligned} 2a &= \sqrt{(x-c)^2 + y^2} + \sqrt{(x+c)^2 + y^2} \\ \left(2a - \sqrt{(x-c)^2 + y^2}\right)^2 &= (x+c)^2 + y^2 \\ 2cx &= 4a^2 - 4a\sqrt{(x-c)^2 + y^2} - 2cx \\ (cx - a)^2 &= -a\left((x-c)^2 + y^2\right) \\ x^2(a^2 - c^2) + a^2 y^2 &= a^2(a^2 - c^2) \\ \frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} &= 1, \end{aligned}$$

as desired. To parameterize the curve, let us substitute $x = \gamma \cos\theta$ and $y = \gamma \sin\theta$ into our formula and simplify as follows,

$$\begin{aligned} \frac{(\gamma\cos\theta)^2}{a^2} + \frac{(\gamma\sin\theta)^2}{a^2 - c^2} &= 1 \\ (a^2 - c^2)(\gamma\cos\theta)^2 + a^2(\gamma\sin\theta)^2 &= a^2(a^2 - c^2) \\ \gamma &= \frac{2a^2(a^2 - c^2)}{-2a^2 + c^2\cos(2\theta) + c^2}. \end{aligned}$$

This was not what we got in class, I think, but I don't see how to get to the expression done in class (it doesn't exactly make sense to me?). Sorry about that!

§54 Day 49: Differentiating Curves (Feb. 9, 2024)

Consider the continuous curve $\gamma : I \to \mathbb{R}^n$ (where I is some interval $[a, b]$) given by

$$\gamma(t) := \begin{pmatrix} a(t) \\ b(t) \\ \vdots \end{pmatrix},$$

where $a(t), b(t), \ldots$ are continuous. If γ is differentiable at a point $t_0 \in I$ and $\gamma'(t_0) \neq 0$, then the line $\gamma(t_0) + \gamma'(t_0)(t - t_0)$ is tangent to the curve at $\gamma(t_0)$, where the derivative of γ is given by

$$\gamma'(t) = \begin{pmatrix} a'(t) \\ b'(t) \\ \vdots \end{pmatrix}.$$

Note that γ itself is differentiable only if its components are differentiable on their own.

We may refer to $\gamma(t_0)$ for some given t_0 as a "position," and $\gamma'(t_0)$ as the velocity (w.r.t. time) in $\mathbb{R}^n$; $|\gamma'(t_0)|$ yields the speed (this can be thought of as a magnitude). There are a handful of things to note regarding the length of a vector. Let $v \in \mathbb{R}^n$, and consider the Euclidean abs as follows,

$$|v| = \sqrt{\sum_{i=1}^{n} |v_i|^2}.$$

Then we have the following properties,

- "Positivity": $|v| \geq 0$. If $|v| = 0$, then the vector itself is the zero vector.
- "Homogeneity": $|cv| = |c|\,|v|$ where c is some real scalar, and $v \in \mathbb{R}^n$ as given above.
- "Triangle Inequality": $|u + v| \leq |u| + |v|$. As an exercise, prove this with the Cauchy-Schwarz inequality.

Using the above, let us define the length of γ. Consider a partition $P = \{t_0, \ldots, t_n\}$ of I, and define $\ell(\gamma, P)$ to be the length of the curve with respect to the partition P, as given below,

$$\ell(\gamma, P) = \sum_{i=1}^{n} |\gamma(t_i) - \gamma(t_{i-1})|.$$

Then we define the length of γ to be $\sup_P \ell(\gamma, P)$.

Proposition 54.1 (Cont. Differentiable Curves have Finite Length)

Let $\gamma : [a, b] \to \mathbb{R}^n$ be a continuously differentiable curve. Then γ has finite length, and it is given by

$$\int_a^b |\gamma'(t)|\, dt.$$

Almut's proof took the Taylor series of γ; I didn't really follow, so here's my attempt. Let $L(\gamma)$ be the length defined by $\sup_P \ell(\gamma, P)$: by continuously refining P, we may write

$$L(\gamma) = \lim_{n\to\infty} \underbrace{\sum_{i=1}^{\infty} |\gamma(t_i) - \gamma(t_{i-1})|}_{\ell(\gamma,P)} = \lim_{n\to\infty} \sum_{i=1}^{\infty} \frac{|\gamma(t_i) - \gamma(t_{i-1})|}{t_i - t_{i-1}}(t_i - t_{i-1}) = \int_a^b |\gamma'(t)|\, dt$$

by appealing to Riemann integration.

We must note that any curve may be expressed/parameterized in as many ways as we wish; consider a continuously differentiable bijection $\varphi : [a,b] \to [c,d]$. Then $\hat{\gamma} = \gamma \circ \varphi^{-1} : [c,d] \to \mathbb{R}^n$ would parameterize the same curve parameterized by γ while also being continuously differentiable at the same time. Thus, we may write

$$\begin{aligned} L(\gamma) &= \int_a^b \left|\gamma'(t)\right| \, dt \\ &= \int_a^b \left|\hat{\gamma}'(\varphi(t))\varphi'(t)\right| \, dt \\ &= \int_a^b \left|\hat{\gamma}'(\varphi(t))\right| \varphi'(t) \, dt && \text{(Homogeneity)} \\ &= \int_c^d \left|\hat{\gamma}'(u)\right| \, du && (u = \varphi(t)) \\ &= L(\hat{\gamma}), \end{aligned}$$

by definition. This implies that the length of the curve is invariant under however we choose to parameterize the curve.

§55 Day 50: Transcendence of e and π (Feb. 12, 2024)

We start with the definition of algebraic numbers and transcendetal numbers; if there exists a nonzero polynomial P such that $p(\alpha) = 0$, then we say α is *algebraic*. Otherwise, it is *transcendental*. Let's immediately start proving that e is transcendental.

Theorem 55.1 (Transcendence, Part 1)
e is transcendental.

Let us start by defining the following three expressions,

$$M := \frac{1}{(p-1)!}\int_0^\infty x^{p-1}((x-1)\dots(x-n))^p e^{-x}\,dx,$$
$$M_k := \frac{e^k}{(p-1)!}\int_k^\infty x^{p-1}((x-1)\dots(x-n))^p e^{-x}\,dx,$$
$$\varepsilon_k := \frac{e^k}{(p-1)!}\int_0^k x^{p-1}((x-1)\dots(x-n))^p e^{-x}\,dx.$$

Observe that $M_k + \varepsilon_k = e^k M \implies e^k = \frac{M_k+\varepsilon_k}{M}$. Suppose that e is not transcendental, i.e. there exists some polynomial P such that

$$P(e) = a_n e^n + a_{n-1}e^{n-1} + \dots + a_1 e + a_0 = 0,$$

where $a_j \in \mathbb{Z}$ and $a_0 \neq 0$. Now, let us take p prime where $p > \max\{n, |a_0|\}$ (so $p \nmid n, p \nmid a_0$), and write

$$((x-1)\dots(x-n))^p = (x^n \pm \dots \pm n!)^p = \sum_{j=0}^{np} c_j x^j. \qquad (c_j \in \mathbb{Z}, c_{np} = 1, c_0 = (\pm n!)^p)$$

Note that $p \nmid (\pm n!)^p$, since it is evident that $n!$ does not contain any factors of p ($p > n$). Rewriting M, we get

$$\begin{aligned} M &= \frac{1}{(p-1)!}\sum_{j=0}^{np} c_j \int_0^\infty x^{j+p-1}e^{-x}\,dx \\ &= \frac{1}{(p-1)!}\sum_{j=0}^{np} c_j (j+p-1)! && \text{(Gamma Function)} \\ &= (\pm n!)^p + \sum_{j=1}^{np} c_j (j+p-1)! && (*) \\ &\implies p \nmid M, \end{aligned}$$

since $j + p - 1 \geq p$ for $j \geq 1$, meaning p divides into the summation on $(*)$, but it does not divide into $(\pm n!)^p$. Using u-sub, let $u = x - k$, which yields

$$M_k = \frac{1}{(p-1)!}\int_0^\infty (u+k)^{p-1}(\underbrace{(u+k-1)}_{\geq 0}\dots\underbrace{(u+k-n)}_{\leq 0})^p e^{-x}\,dx.$$

Clearly, $(u+k-1)\dots(u+k-n)$ contains a term of u, meaning M_k is of the form

$$M_k = \sum_{j=1}^{np} d_j \frac{(p-1+j)!}{(p-1)!}; \qquad \text{(for coefficients } d_j)$$

this concludes that $p \mid M_k$ for all k. Considering the polynomial equation for e as described earlier, write

$$a_n\left(\frac{M_n+\varepsilon_n}{M}\right)+\cdots+a_1\left(\frac{M_n+\varepsilon_n}{M}\right)+a_0=0 \implies Ma_0+\sum_{j=1}^{n}a_j(M_j+\varepsilon_j)=0.$$

Observe that $Ma_0+\sum_{j=1}^{n}a_jM_j$ is not divisible by p, but it is an integer, meaning it is nonzero. Moreover, we may minimize $\sum_{j=1}^{n}a_j\varepsilon_j$ as small as we want by taking large p. To see this, directly compute

$$\begin{aligned}|\varepsilon_k| &\le \frac{e^k}{(p-1)!}\int_0^n |x|^{p-1}\,|(x-1)\dots(x-n)|^p\,e^{-x}\,dx\\ &\le \frac{e^k}{(p-1)!}\int_0^n \underbrace{|(x-1)\dots(x-n)|^p}_{\text{bounded above by }A}e^{-x}\,dx\\ &\le \frac{e^k}{(p-1)!}n^pA^p\int_0^\infty e^{-x}\,dx\\ &\le e^k\frac{(nA)^p}{(p-1)!},\end{aligned}$$

which goes to 0 as $p\to\infty$. However, this would imply that our polynomial $P(e)\neq 0$ (since the M terms do not vanish). This concludes that e is transcendental. □

Theorem 55.2 (Irrationality of π)
π is irrational.

Suppose $\pi=\frac{a}{b}$. Consider the following function,

$$A_n(b):=b^n\int_0^\pi \frac{x^n(\pi-x)^n}{n!}\sin x\,dx>0.$$

Also consider that

$$|A_n(b)|\le b^n\int_0^\pi \left|\frac{(\frac{\pi}{2})^{2n}}{n!}\right|dx=\frac{\pi}{n!}\left(\frac{b\pi^2}{4}\right)^n<1. \qquad \text{(For large enough } n\text{)}$$

Now, let us define f as below,

$$f(x)=b^n\frac{x^n(\pi-x)^n}{n!}=\frac{x^n(a-bx)^n}{n!}=\sum_{j=0}^{2n}\frac{c_j}{n!}x^j$$

for coefficients c_j. Observe that $f^{(k)}(0)\in\mathbb{Z}$ and $\frac{k!}{n!}c_k\in\mathbb{Z}$; this means $f^{(k)}(\pi)\in\mathbb{Z}$ as $f(x)=f(\pi-x)$; however,

$$A_n(b)=\int_0^\pi f(x)\sin x\,dx=\chi+\underbrace{\int_0^\pi f^{(2n+1)}(x)\cos x\,dx}_{=0}$$

for some nonzero χ after repeated applications of IBP. This is a contradiction, which lets us conclude that $\pi\notin\mathbb{Q}$.

§56 Day 53: Review of Riemann Sums (Feb. 26, 2024)

Let $f : [a, b] \to \mathbb{R}$. A Riemann sum is an expansion of the form

$$S = \sum_{j=1}^{n} f(x_j)(t_j - t_{j-1})$$

where $n \in \mathbb{N}$, $P = \{a = t_0, \dots, t_n = b\}$ is a partition of $[a, b]$, and $x_j \in [t_{j-1}, t_j]$. How do we make sure S is a good approximation for $\int_a^b f$? We claim that for any integrable f over $[a, b]$, for all $\varepsilon > 0$, there exists $\delta > 0$ such that every partition $P = \{t_0, \dots, t_n\}$ of $[a, b]$ has $\max_{j=1,\dots,n}(t_j - t_{j-1}) < \delta$. And for every choice $x_j \in [t_{j-1}, t_j]$, we have

$$\left| \sum_{j=1}^{n} f(x_j)(t_j - t_{j-1}) - \int_a^b f \right| < \varepsilon.$$

Typically, an application of this would mean that even partitions given by

$$P = \left\{ a + \frac{j(b-a)}{n} \mid j \in [n] \right\}$$

satisfies

$$S_n = \sum_{j=1}^{n} f(x_j)(t_j - t_{j-1}) \xrightarrow[n\to\infty]{} \int_a^b f.$$

To prove our claim, consider the special case where f is given to be continuous (uniformly continuous, even, since we're working in $\mathbb{R}$). In the general case, we simply approximate a uniformly continuous function to collapse it to the special case. We proceed as follows.

- For any $\varepsilon > 0$, start by picking $\delta > 0$ small enough such that $|x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{b-a}$. Given a partition $P = \{t_0, \dots, t_n\}$ of $[a, b]$, estimate

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_{i=1}^{n} \left| \max_{x\in[t_i,t_{i-1}]} f(x) - \min_{x\in[t_i,t_{i-1}]} f(x) \right| (t_i - t_{i-1}) \\ &= \sum_{i=1}^{n} |f(x_i) - f(y_i)| (t_i - t_{i-1}) \quad \text{(Let } x_i, y_i \text{ be max and min)} \\ &\le \frac{\varepsilon}{b-a} \sum_{i=1}^{n} (t_i - t_{i-1}) \\ &= \varepsilon. \end{aligned}$$

Since

$$L(f, P) \le \sum_{j=1}^{n} f(x_j)(t_j - t_{j-1}) \le U(f, P),$$

$$L(f, P) \le \int_a^b f \le U(f, P),$$

we easily claim $\left| S_n(f) - \int_a^b f \right| < \varepsilon$. This concludes the uniformly continuous case.

- For the general case, by our previous assignment question (**A13**, **Q4**), let us have continuous functions g, h such that $g \leq f \leq h$ and $\int_a^b h - g < \frac{\varepsilon}{3}$. Choose δ small enough such that $P = \{t_0, \ldots, t_n\}$ with $\max_j |t_j - t_{j-1}| < \delta$, we have

$$0 \leq U(h,P) - \int_a^b h < \frac{\varepsilon}{3}, \quad 0 \leq U(g,P) - \int_a^b g < \frac{\varepsilon}{3}$$

$$0 \leq \int_a^b h - L(h,P) < \frac{\varepsilon}{3}, \quad 0 \leq \int_a^b g - L(g,P) < \frac{\varepsilon}{3}$$

Let $S := \sum_{j=1}^n f(x_j)(t_j - t_{j-1})$ be a Riemann sum for P. Clearly, $L(g,P) \leq S \leq U(h,P)$ from $g \leq f \leq h$. Then

$$U(h,P) - L(g,P) \leq \left(\int_a^b h + \frac{\varepsilon}{3}\right) - \left(-\int_a^b g + \frac{\varepsilon}{3}\right) = \int_a^b h - g + \frac{2\varepsilon}{3} < \varepsilon. \quad \square$$

§57 Day 54: Curves and Parameterizations (Feb. 28, 2024)

For any given curve γ, there are many properties that may be discussed; for intrinsic quantities, we may discuss arc length, tangents, and curvature, and for dynamic quantities, we may discuss position, velocity, and acceleration. Consider a smooth curve $\gamma : [a, b] \to \mathbb{R}$ that is at least twice continuously differentiable, where

$$\gamma'(t) \neq 0 \quad \text{for all } t \in [a, b];$$

except perhaps at the endpoints, where we may consider $\gamma'(a^+)$ and $\gamma'(b^-)$. One such example is given by the logarithmic curve

$$\gamma(t) = e^{at} \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, t \in \mathbb{R};$$

Recall that the dynamic quantities position, velocity, and acceleration are given as $\gamma(t), \gamma'(t)$, and $\gamma''(t)$ respectively, and are given below

$$\begin{aligned} \gamma(t) &= e^{at} \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \\ \gamma'(t) &= e^{at} \begin{pmatrix} a\cos t - \sin t \\ a \sin t + \cos t \end{pmatrix}, \\ \gamma''(t) &= e^{at} \begin{pmatrix} (a^2 - 1)\cos t - 2a \sin t \\ (a^2 - 1)\sin t + 2a\cos t \end{pmatrix}. \end{aligned}$$

Some application was talked in class to Newton's law $F = ma$, but I didn't quite catch that...

§58 Day 55: Reparameterization; Conclusion of Curves (Mar. 4, 2024)

Let $\gamma : [a,b] \to \mathbb{R}^2$ be a smooth planar curve. We say $\tilde{\gamma} : [c,d] \to \mathbb{R}^2$ is a *reparameterization* of γ if $\tilde{\gamma}(\tilde{t}) = (\gamma \circ \varphi)(t)$ for some continuous invertible function $\varphi : [a,b] \to [c,d]$. We say that φ is orientation-preserving if φ is increasing; if not, we call it orientation-reversing.

Remark 58.1. From the fact that if $\psi : S \to T$ is a distance-preserving bijection (with S, T surfaces), intrinsic functions f satisfy $f_S(p) = (f_T \circ \psi)(p)$, I'm going to say that arc length, tangents are preserved under φ, while position, velocity, and acceleration are not. (I might be wrong.)

Usually, we operate under the convention that $\varphi'(t) \neq 0$ for all $t \in [a,b]$. Since every differentiable curve can be parameterized w.r.t. arc length, let us have

$$\varphi(t) = \int_a^b |\gamma'(t)| \, dt,$$

and let us compute the following (where length is defined as the Euclidean abs),

$$\begin{aligned}\frac{d}{dt}|\gamma'(t)| &= \frac{d}{dt}\sqrt{x'^2 + y'^2} \\ &= \frac{1}{2\sqrt{x'^2+y'^2}}\left(\frac{d}{dt}(x'^2 + y'^2)\right) \\ &= \frac{x'x'' + y'y''}{|\gamma'|} \\ &= \begin{pmatrix} x'' \\ y'' \end{pmatrix} \cdot \frac{1}{|\gamma'|}\begin{pmatrix} x' \\ y' \end{pmatrix}.\end{aligned}$$

In particular, we refer to $\frac{1}{|\gamma'|}\begin{pmatrix} x' \\ y' \end{pmatrix}$ as the unit tangent to the curve γ at t; it can be interpreted as the component of γ'' in the tangential direction. Assume for now that γ is parameterized over arc length with $|\gamma'(t)| = 1$ for all t; i.e., γ is a curve with constant speed. In this case, the double derivative of γ would vanish, and we see

$$0 = \frac{d}{dt}|\gamma'(t)| = \begin{pmatrix} x'' \\ y'' \end{pmatrix} \cdot \frac{1}{|\gamma'|}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x'' \\ y'' \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

This means $(x'', y'')^\top$ and $(x', y')^\top$ are orthogonal (read: velocity and acceleration are orthogonal); i.e., $\gamma''(t) = k(t)\text{normal}(t)$ for some normal vector and scalar $k(t)$. We call $k(t)$ to be the curvatureof γ at the point t. In particular, the formula is given by

$$k(t) := \frac{x'(t)y''(t) - y'(t)x''(t)}{|\gamma'(t)|^3};$$

and the proof will be given later on. Consider the following example where $R > 0$ (i.e. radius of curvature),

$$\gamma(t) = \begin{pmatrix} R\cos\frac{t}{R} \\ R\sin\frac{t}{R} \end{pmatrix}.$$

Then we have

$$\gamma'(t) = \begin{pmatrix} -\sin\frac{t}{R} \\ \cos\frac{t}{R} \end{pmatrix};$$

by the Pythagorean property, we necessarily have $|\gamma'(t)| = 1$. Thus, the speed is identically one.

§59 Day 56: Convexity (Mar. 6, 2024)

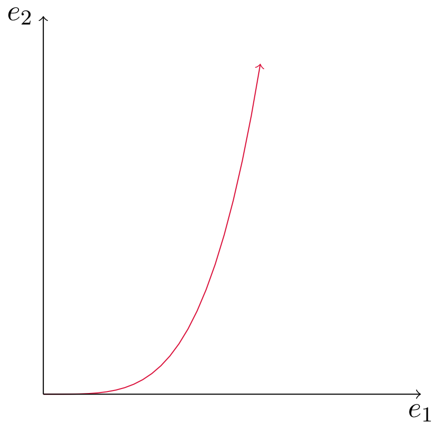
Consider the curve $\gamma : [a,b] \to \mathbb{R}^2$ given by

$$\gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \text{ with speed } \left|\gamma'(t)\right|.$$

Let the signed curvature (at point $\gamma(t)$) be given by

$$k(t) = \frac{x'y'' - x''y'}{|\gamma'|^3} = \frac{1}{|\gamma'|^3} \det \begin{pmatrix} x' & x'' \\ y' & y'' \end{pmatrix} = \frac{\gamma' \times \gamma''}{|\gamma'|^3}.$$

In particular, if we consider $\beta = \{e_1, e_2\}$ to be the standard basis of $\mathbb{R}^2$, then $k(t) > 0$ implies γ bends towards e_2, and $k(t) < 0$ implies γ bends towards e_1. This looks like Let



$I \subset \mathbb{R}$ be a non-empty interval, and define $f : I \to \mathbb{R}$ to be a real-valued function. We say that f is strictly convex on I if, for any subinterval $[a,b] \subset I$, the line segment joining $(a, f(a))$ and $(b, f(b))$ lies strictly above the graph of f. If f is convex (not necessarily strict), it is allowed for f to touch any aforementioned line segments. From this, we have the following inequality,

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a}, \tag{$*$}$$

with a strict inequality observed over a strictly convex interval of f. A more convenient way to write $(*)$, though, would be as follows; let $(a,b) \subset I$ be any subinterval of I, and let us have a parameterization index $t \in (0,1)$. Then

$$f(at + (1-t)b) \leq tf(a) + (1-t)f(b)$$

is an equivalent statement (with strict inequality once again observed for strict convexity). To prove this, simply observe that all points on the line segment between $(a, f(a))$ and $(b, f(b))$ are given by

$$L_{a,b} = \{(a + \lambda(b-a), f(a) + \lambda(f(b) - f(a))) \mid 0 < \lambda < 1\}.$$

This transforms into the mean value theorem inequality given in $(*)$ immediately.

§60 Day 57: More on Convexity (Mar. 8, 2024)

Let us have $f : I \to \mathbb{R}$. Then

- If f is convex, it is equivalent to say f' is non-decreasing on I; alternatively, $f'' \geq 0$.
- If f is strictly convex, then f' is strictly increasing on I, and $f'' > 0$.
- If we flip the definitions around, then we have concave and strictly concave.

Below are some examples to check.

(a) Let $f(x) = e^{\alpha x}$. Then $f''(x) = \alpha^2 e^{\alpha x} > 0$ (for $\alpha \neq 0$), and so it is strictly convex on $\mathbb{R}$.

(b) Let $f(x) = \log |x|$. Then $f''(x) = -\frac{1}{x^2} \leq 0$, which means f is concave on $\mathbb{R}$.

(c) Let $f(x) = x^\alpha$. Then $f''(x) = \alpha(\alpha - 1)x^{\alpha-2}$. f is strictly convex if $\alpha > 1$ and $\alpha < 0$, and strictly concave if $0 < \alpha < 1$.

If f is strictly convex, we know f' is strictly increasing; alternatively, if f is convex (but not strictly convex), that implies the existence of some $a, b \in I$ where $f'(a) = f'(b)$.

We now prove the first bulleted claim. Clearly, the rest follow analogously. Supposing f is convex on I, let us have $a < c < b$ on I and consider

$$\frac{f(c) - f(a)}{c - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(c)}{b - c}.$$

Let us have

$$f'(a) = \lim_{c \to a^+} \frac{f(c) - f(a)}{c - a} \leq \lim_{c \to a^+} \frac{f(b) - f(c)}{b - c} = \frac{f(b) - f(a)}{b - a}$$

and

$$\frac{f(c) - f(a)}{c - a} = \lim_{c \to b^-} \frac{f(c) - f(a)}{c - a} \leq \lim_{c \to b^-} \frac{f(b) - f(c)}{b - c} = f'(b).$$

This yields $f'(b) \geq f'(a)$ whenever $b > a$, implying f' is monotonically increasing. This argument goes both directions, and we are done.

§61 Day 58: Applications of Differentiation (Mar. 11, 2024)

This week, we will be going over applications of differentiation in geometric problems and converting multivariable problems into single variable ones. Next week will be integration applications.

We start by considering the cylinder. Its surface area and volume can be expressed by its height h and radius r as follows,

$$S = 2\pi r^2 + 2\pi rh, \qquad \text{(Can also be expressed as } A\text{)}$$
$$V = h\pi r^2.$$

In this way, using the surface area, we may write h as a function of r as follows,

$$h = \frac{A - 2\pi r^2}{2\pi r} = \frac{A}{2\pi r} - r.$$

Substituting into the volume formula, we get

$$V = \left(\frac{A}{2\pi r} - r\right)\pi r^2.$$

Sometimes, when minimizing area or maximizing volume w.r.t. a given area, we may differentiate V w.r.t. as follows,

$$V'(r) = \frac{A}{2\pi} - 3r^2 = 0 \implies R = \sqrt{\frac{A}{6\pi}}.$$

Today was mostly an introduction to the given relationship above, with an example problem on maximizing the area of a figure with respect to boundary length. Read: Dido's Problem (sorry, I understand it but I don't feel comfortable going in detail since I don't have everything from class down).

§62 Day 59: Surface Area and Volume of Rotational Solids (Mar. 13, 2024)

Today we went over the derivation and approximation of the surface area and volume of rotational solids. Consider a continuous function $f : I \to \mathbb{R}$; naturally, the area under the curve (in the two dimensional sense) is given by

$$\int_a^b f(x)\,dx.$$

Let us consider $P = \{a = t_0, \dots, t_n = b\}$ to be a partition of $I = [a, b]$, and start by considering the rotational solid formed by rotating f around the x-axis. Supposing P is an evenly distributed partition, we have that $t_i - t_{i-1} = \frac{b-a}{n}$, and we start by approximating the volume of said solid through the summation

$$\sum_{i=1}^{n} \pi(f(c_i)^2)(t_i - t_{i-1}) \xrightarrow[n\to\infty]{} \int_a^b \pi f(x)^2\,dx$$

where $c_i \in [t_{i-1}, t_i]$ (so that we may choose lower or upper sums however we desire; this is just an approximation). Think of the above summation as approximating our rotational solid with increasingly "thin slices" of cylinders; with $n \to \infty$ simply leading to the integral that yields the desired volume.

More specifically, for intuition, observe that for any $\varepsilon > 0$, by uniform continuity of f, there exists $\delta > 0$ where $x - y < \delta \implies |f(x) - f(y)| < \varepsilon$. Taking n large enough such that

$$\frac{b-a}{n} < \delta,$$

we see that creating our partition $P = \{a = t_0, \dots, t_n = b\}$ tells us that each term

$$\pi(f(c_i)^2)(t_i - t_{i-1})$$

is at most ε off in terms of "radius"; formally, we may write

$$\pi\left(\max_{x\in[t_{i-1},t_i]} f(x)\right)^2 (t_i - t_{i-1}) - \pi\left(\min_{x\in[t_{i-1},t_i]} f(x)\right)^2 (t_i - t_{i-1}) = O(1)(t_i - t_{i-1}),$$

which allows us to neglect the error with $n \to \infty$.[50]

Similarly, the surface area of a rotational solid is given by (intuition: $2\pi r$ on a slant)

$$\sum_{i=1}^{n} 2\pi f(c_i)\sqrt{1 + f'(c_i)^2}(t_i - t_{i-1}) \xrightarrow[n\to\infty]{} \int_a^b 2\pi f(x)\sqrt{1 + f'(x)^2}\,dx.$$

I leave it as an exercise to derive the surface area and volume of a rotational solid that rotates around the y-axis or z-axis. Now, we proceed with examples from class.

- Let us compute the volume of the unit ball. By revolving $y = \sqrt{1 - x^2}$ about the x-axis, we get the volume of a sphere when we "evaluate for each disk" over $x \in [-1, 1]$. Proceed as follows,

  $$\int_{-1}^{1} \pi\sqrt{1 - x^2}^2\,dx = \pi\int_{-1}^{1} 1 - x^2\,dx = \frac{4\pi}{3}.$$

  By considering $y = \sqrt{r^2 - x^2}$, we may derive the general volume of a sphere of radius r.

[50] I didn't exactly follow the board here, so I'm just writing my own personal intuition. correct me if i'm wrong

- Gabriel's Horn / Torricelli's Trumpet: Consider $f(x) = \frac{1}{1+x}$. This figure, when rotated around the x-axis, has infinite surface area but finite volume. Observe below,

$$V = \int_0^L \pi \left(\frac{1}{1+x} \right)^2 dx = -\frac{\pi}{1+x}\bigg|_0^L = -\frac{\pi}{1+L} + \pi.$$

Taking $L \to \infty$, we see $V = \pi$. However,

$$S = \int_0^L 2\pi \frac{1}{1+x} \underbrace{\sqrt{1 + \frac{1}{(1+x)^2}}}_{\geq 1} dx \geq \int_0^L \frac{2\pi}{1+x}\, dx,$$

which clearly diverges to infinity as $L \to \infty$.

It should be noted that in general, if the surface area is finite, then so is the volume (informal statement; just take some continuously differentiable $f : [1, \infty) \to [0, \infty)$, which should suffice).

§63 Day 60: More on Solids of Revolution and Shell Method (Mar. 15, 2024)

Given an integrable function $f : [a, b] \to \mathbb{R}$, let us rotate $y = f(x)$ around the x-axis. Using $P = \{a = t_0, \ldots, t_n = b\}$ as a partition of $[a, b]$ with $x_i \in (t_{i-1}, t_i)$, the resulting solid would have surface area and volume given as follows;

$$V = \sum_{i=1}^{n} \pi f(x_i)^2 (t_i - t_{i-1}) \xrightarrow[n\to\infty]{} \int_a^b \pi f(x)^2 \, dx,$$

$$S = \sum_{i=1}^{n} 2\pi f(x_i)\sqrt{1 + f'(x_i)^2}(t_i - t_{i-1}) \xrightarrow[n\to\infty]{} \int_a^b 2\pi f(x)\sqrt{1 + f'(x)^2} \, dx,$$

of which the integral formula for the surface area is valid only if f is continuously differentiable. Now, consider rotating $y = f(x)$ around the y-axis instead. We use the shell method,[51] as below; again, for a partition P as given above, we have

$$V = \sum_{i=1}^{n} \pi f(x_i)(t_i^2 - t_{i-1}^2) \xrightarrow[n\to\infty]{} 2\pi \int_a^b x f(x) \, dx.$$

Intuitively, take a rectangular box approximation of the integral of $f(x)$, then rotate each of the boxes individually; it's the same as splitting the solid of revolution into hollow cylinders with height $f(x_i)$ and outer radius t_i, inner radius t_{i-1} (the hollow part). An example from class is given by the cone; suppose we consider it as a solid of revolution around the y-axis, with function

$$f(x) = h\left(1 - \frac{x}{R}\right),$$

where h is the height and R is the radius of the cone. Then we may directly integrate as follows,

$$2\pi \int_0^R xh\left(1 - \frac{x}{R}\right) dx = 2\pi h \int_0^R x - \frac{x^2}{R} \, dx = (2\pi h)\left(\frac{1}{2}x^2 - \frac{1}{3}\frac{x^3}{R}\right)\Bigg|_{x=0}^{x=R} = \frac{\pi r^2 h}{3},$$

which makes sense.

Note: i'm not sure if we covered washer method; will write in here if we did. most likely we did tho.

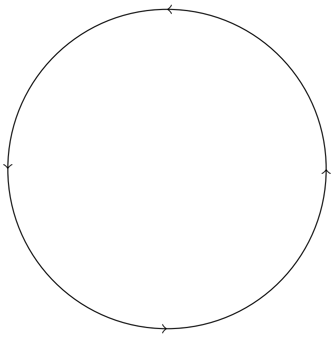
[51] tl;dr washer method is perpendicular to axis of revolution, shell is parallel "circular slice"

§64 Day 61: Enclosed Area of a Curve (Mar. 18, 2024)

Let $\gamma : [a, b] \to \mathbb{R}^2$ be a continuously differentiable closed curve with $\gamma'(t) \neq 0$ for all $t \in [a, b]$. Suppose γ is "simple"; i.e.,

$$\gamma(s) = \gamma(t) \iff \begin{cases} s = t \\ s = a, t = b \end{cases} .$$

The last case can be considered as the endpoints of γ; in order for it to be a closed curve, the endpoints must coincide. In this manner, let us consider A as the region enclosed by γ.[52]



In the case that γ surrounds A in a counter-clockwise fashion as shown above, the area of A is given by

$$\int_a^b (xy')(t)\,dt = -\int_a^b (yx')(t)\,dt = \frac{1}{2}\int_a^b (xy' - yx')(t)\,dt = \int_a^b \det\left(\gamma \mid \gamma' \right).$$

If γ surrounds A clockwise, then we simply negate the above formulas. Note that this is a special case of Green's Theorem / Stokes' Theorem. Here are a few sanity checks;

- Notice that $\int_a^b xy'\,dt$ scales with the size of γ; if we have $x \mapsto 2x$ and $y \mapsto 2y$, the area of A increases twofold twice as expected.
- Using the form $\frac{1}{2}\int_a^b xy' - yx'\,dt$, we see that the area is rotation-invariant.
- The area is also invariant under reparameterization. In particular, let us consider η as a reparameterization of γ, where $t = \phi(s)$ with

  $$\phi : [c, d] \to [a, b],$$
  $$\eta = \gamma \circ \phi.$$

  Assuming ϕ is orientation preserving, let us directly compute the area induced by η as follows,

  $$\begin{aligned} &\frac{1}{2}\int_c^d ((x \circ \phi)(y \circ \phi)' - (y \circ \phi)(x \circ \phi)')(t)\,dt \\ &= \frac{1}{2}\int_a^b ((x \circ \phi)(y \circ \phi)' - (y \circ \phi)(x \circ \phi)'(\phi^{-1}(t))\,dt \qquad \text{(Substitute } t \text{ for } \phi^{-1}(t)) \\ &= \frac{1}{2}\int_a^b (xy' - yx')(t). \qquad \text{(Chain Rule cancels out } (\gamma \circ \phi)' \text{ components)} \end{aligned}$$

[52] the curve below is supposed to represent an example of counter-clockwise parameterization gamma, but idk, make it more wonky n stuff for maximum visualization effect. i ain't doin all that with latex i've had enough trouble putting the DAMN ARROW POINTERS ON THE CIRCLE

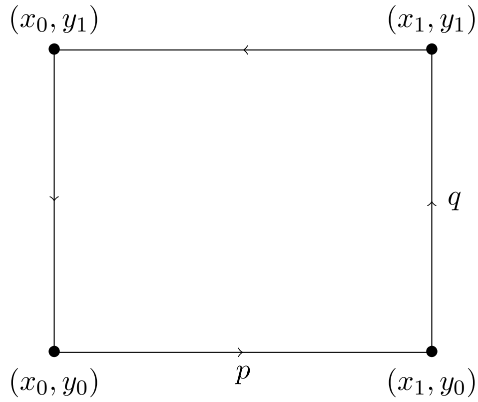
We now do a few examples; suppose $\gamma : [0, 2\pi] \to \mathbb{R}^2$ is the curve representing a circle, given by

$$\gamma(t) = \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix}.$$

Naturally, it is clear that the area is given by πr^2; we now see that the integral also gives the same output, as follows,

$$\int_a^b (xy')(t)\,dt = \int_0^{2\pi} r^2 \cos t(\sin t)'\,dt = \frac{r^2}{2}\left(t + \sin t\cos t\right)\Bigg|_{t=0}^{t=2\pi} = \pi r^2.$$

Now, let's consider a rectangle parameterized counter-clockwise by γ, with side lengths p, q as below,[53]



I leave it as an exercise to parameterize this and verify the formula; note that in class, this was done by splitting the rectangle into two right triangles.

[53] i am going to scream the p and q look off-center even though they ARE centered...

§65 Day 62: Improper Integrals (Mar. 20, 2024)

Suppose we have an integral on the interval $[a, b]$, with $a \to -\infty$ or $b \to \infty$ (colloquially speaking) with perhaps unbounded f. There are three scenarios to consider; either $f \geq 0$, $\int |f|$ exists, or $\int |f|$ does not exist, while $\int f$ can be defined. We proceed to cover these three.

Suppose f is integrable on $[a, T]$ for every $T > a$. Let us define

$$\int_a^\infty f(x)\,dx = \lim_{T\to\infty} \int_a^T f(x)\,dx,$$

assuming the limit exists;

- If $f \geq 0$, then $F(x) = \int_a^x f(t)\,dt$ is non-decreasing; we see

$$\lim_{T\to\infty} \int_a^T f(x)\,dx = \lim_{T\to\infty} F(T) = \sup F,$$

  which exists if $\sup F$ is finite; if not, then our limit diverges. For example, consider $f(x) = x^n e^{-x}$ with

$$\int_0^\infty x^n e^{-x}\,dx = \lim_{T\to\infty} \int_0^T x^n e^{-x}\,dx.$$

  To see that this converges, observe that

$$g(x) = x^n e^{-\frac{x}{2}} \implies g'(x) = (nx^{n-1} \cdot -\frac{1}{2}x^n)e^{-x},$$

  where $g(0) = 0$ and g is maximized at $2n$; this means

$$\lim_{x\to\infty} g(x) = 0 \implies \int_0^\infty g(x)\,dx \leq M$$

  for a large enough choice of M. Estimating our original integral, we have

$$\int_0^\infty f(x)\,dx \leq M \int_0^\infty e^{-\frac{x}{2}}\,dx = M \lim_{T\to\infty} \int_0^T e^{-\frac{x}{2}}\,dx = \lim_{T\to\infty} 2M\left(1 - e^{-\frac{T}{2}}\right) < \infty.$$

  We conclude that such an improper integral takes on a finite value, and so the integral exists. As another example, let us consider the following functional equation (given by the Gamma function specifically),

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t}\,dt = \left(-t^x e^{-t}\right)\Big|_{t=0}^{t=\infty} + \int_0^\infty x t^{x-1} e^{-t}\,dt = x\Gamma(x),$$

  which gives $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$.

- If f changes sign with $\int_a^\infty |f(x)|\,dx < \infty$, we have that $F(T) = \int_a^T f(x)\,dx$ is bounded, since

$$|F(T)| = \left|\int_a^T f(x)\,dx\right| \leq \int_a^T |f(x)|\,dx \leq \int_a^\infty |f(x)|\,dx < \infty.$$

By considering the increasing sequence $\{n_j\}_{j\geq 1}$ with $L = \lim_{j\to\infty} F(n_j)$ and using the fact that F is bounded, we may apply the Bolzano-Weierstrass theorem to obtain a convergent subsequence, showing

$$L = \int_a^\infty f(x)\, dx.$$

For now, let T be large enough such that

$$\int_0^\infty |f(x)|\ dx - \int_0^T |f(x)|\ dx < \varepsilon \implies \int_T^\infty |f(x)|\ dx < \varepsilon.$$

Let k be a large enough index such that $n_j > T$ for all $j \geq k$; then for all $t \geq T$, we have

$$\begin{aligned} |F(t) - L| = \left|\int_a^t f(x)\, dx - L\right| &\leq \left|\int_a^{n_j} f(x)\, dx - L\right| + \left|\int_{n_j}^T f(x)\, dx\right| \\ &\leq \int_T^\infty |f(x)|\ dx \leq 2\varepsilon. \end{aligned}$$

§66 Day 63: Improper Integrals (Mar. 22, 2024)

Let $f : [a, \infty) \to \mathbb{R}$ be such that f is integrable over $[a, b]$ for every $b > 0$. If $\int_0^\infty |f|$ exists, we claim that $\int_0^\infty f$ exists as well, and that

$$\left| \int_0^\infty f(x)\, dx \right| \leq \int_0^\infty |f(x)|\, dx.$$

Assume $|f|$ is integrable on $[a, \infty)$; that is,

$$\int_0^\infty |f(x)|\, dx = \sup_{T > a} \int_0^T |f| < \infty.$$

Then set

$$f_+(x) = \max\{f(x), 0\},$$
$$f_-(x) = \min\{f(x), 0\};$$

by definition, $|f(x)| = f_+(x) + f_-(x)$, and $f(x) = f_+(x) - f_-(x)$. Clearly, we may write

$$\int_a^\infty f_+(x)\, dx \leq \sup_{T>a} \int_a^\infty f_+(x)\, dx \leq \sup_{T>a} \int_a^T |f(x)|\, dx < \infty,$$

$$\int_a^\infty f_-(x)\, dx \leq \sup_{T>a} \int_a^\infty f_-(x)\, dx \leq \sup_{T>a} \int_a^T |f(x)|\, dx < \infty.$$

In particular, this means $\lim_{T\to\infty} \int_0^T f_\pm(x)\, dx$ must exist. From this, we may conclude that

$$\int_0^\infty f(x)\, dx = \lim_{T\to\infty} \int_0^T f(x)\, dx = \lim_{T\to\infty} \int_0^T f_+(x)\, dx - \lim_{T\to\infty} \int_0^T f_-(x)\, dx < \infty,$$

as both limits are finite as demonstrated above.

- For example, consider a bounded function $f : \mathbb{R} \to \mathbb{R}$; we claim that

  $$\int_{-\infty}^\infty f(x) e^{-\frac{x^2}{2}}\, dx$$

  exists. To see this, consider

  $$\int_{-\infty}^0 f(x) e^{-\frac{x^2}{2}}\, dx, \qquad \int_0^\infty f(x) e^{-\frac{x^2}{2}}\, dx$$

  separately. It suffices to show that they are absolutely integrable, i.e.

  $$\begin{aligned} \int_0^T \left| f(x) e^{-\frac{x^2}{2}} \right| dx &\leq M \int_0^T e^{-\frac{x^2}{2}}\, dx \\ &\leq M \left(1 + \int_0^T x e^{-\frac{x^2}{2}}\, dx \right) \\ &= M \left(1 - e^{-\frac{x^2}{2}} \right) \Bigg|_{x=0}^{x=\infty} \\ &\leq 2M. \end{aligned}$$

- Consider the following interesting example,

$$\int_0^\infty \frac{\sin x}{x}\,dx.$$

For our first try: is such a function absolutely integrable? Observe that

$$|f(x)| = \left|\frac{\sin x}{x}\right| \leq \min\{1, \frac{1}{|x|}\}.$$

Let's integrate as follows;

$$\int_0^1 \frac{\sin x}{x}\,dx + \int_1^\infty \frac{\sin x}{x}\,dx = \int_0^\infty \frac{\sin x}{x}$$
$$\int_1^\infty \frac{1}{x}\,dx = \log x\Big|_1^\infty = \infty,$$

which is incoherent. This method won't work; let us try integration by parts, on an interval

$$\begin{aligned}\int_{2\pi k-\frac{\pi}{2}}^{2\pi k+\frac{\pi}{2}} \frac{\sin x}{x}\,dx &\geq \left(\frac{1}{2\pi k+\frac{\pi}{2}}\right)\int_{2\pi k-\frac{\pi}{2}}^{2\pi k+\frac{\pi}{2}} \sin x\,dx \\ &= \frac{1}{2\pi k+\frac{\pi}{2}} \\ &\geq \frac{1}{\pi}\int_{2\pi k+\frac{\pi}{2}}^{2\pi k+\frac{3\pi}{2}} \frac{1}{x}\,dx.\end{aligned}$$

With this, we conclude

$$\int_{2\pi k-\frac{\pi}{2}}^\infty \left|\frac{\sin x}{x}\right|\,dx \geq \frac{1}{\pi}\int_{2\pi k+\frac{\pi}{2}}^\infty \frac{1}{x}\,dx = \infty.$$

This concludes that the function is not absolutely integrable.

§67 Day 64: (Mar. 25, 2024)