21-269 Recitation A

Thomas Lam

2022 Spring

Contents

1 Real Axioms, Natural Numbers, Induction, and AM-GM

1.1 Warm-Up

I have two ropes. The first burns up in one minute when lit from one end, and the second similarly burns up in two minutes. Burning rate is not necessarily uniform over the lengths of these ropes.

Determine a method to measure 75 seconds.

Hint: Gel yvtugvat obgu fvqrf bs n ebcr ng bapr.

1.2 Messing Around With Axioms

We defined the real numbers, $\mathbb{R}$, as an ordered field satisfying the *supremum property*. We'll tackle supremums later, but here are the takeaways from "ordered field":

- That $\mathbb{R}$ is a field just means that it's a nice place where we can add, subtract, multiply, and even divide (but not by zero). Multiplication and addition are related via the distributive property.
- It's an *ordered* field, meaning that in addition to this nice stuff, there is an ordering relation called $\leq$. The following are the axioms that construct such a relation.
- (Anything can be compared) For any $x, y \in \mathbb{R}$, we have either $x \leq y$ or $y \leq x$.
- (Transitivity) If $x \leq y \leq z$ then $x \leq z$.
- (Anti-Symmetry) If $x \leq y$ and $y \leq x$ then $x = y$.
- (Reflexivity) $x \leq x$ always.
- (Additive Preservation) If $x \leq y$ then $x + z \leq y + z$.
- (Multiplicative Preservation) If $x \geq 0$ and $y \geq 0$, then $xy \geq 0$.
- If $x \geq y$ but $x \neq y$, we say that $x > y$.

Let's show that $1 > 0$. First we need a lemma:

Lemma 1.1

$(-1)(-1) = 1$

Proof. We have $(1+(-1)) = 0$. So $(1+(-1))(1+(-1)) = (0)(0) = 0$. Expanding, we get:

$$1 + (-1) + (-1) + (-1)(-1) = 0$$

Where we have used $1 \cdot x = x$. Now, using again $1 - 1 = 0$, we deduce that $(-1)(-1) = 1$. □

Exercise 1.1: Prove that $1 > 0$.

Proof. Well, $1 \neq 0$. So either $1 > 0$ or $1 < 0$.

Case 1: $1 > 0$

Then we are done lol.

Case 2: $1 < 0$

Then $1+(-1) < 0+(-1)$, hence $0 < -1$. Also, $0 < -1$. Since both $0 < -1$ and $0 < -1$, we get that $0 < (-1)(-1)$. By the previous lemma, this is equivalent to $0 < 1$.

In both cases we got $1 > 0$ so we are done. □

1.3 The Natural Numbers

What are natural numbers?

- Are they $\mathbb{N} := \{1, 2, 3, 4, \cdots\}$? Well no, what the heck is 3?
- Are they $\mathbb{N} := \{1, 1+1, 1+1+1, \cdots\}$? Well no, the $\cdots$ is an implicit use of induction. Induction requires natural numbers in the first place! nooooo

Well, naturals numbers should be like, "if x is in there, then $x+1$, and also 1 is in there, and nothing else should be there". We can formalize the first two conditions we want like this:

Definition 1.1 (Inductive Set)

A subset $E \subseteq \mathbb{R}$ is *inductive* if:

- $1 \in E$
- $n \in E \implies n+1 \in E$

Also yes, 0 is not natural. Fight me.

Example 1.2: Verify that the following sets are inductive:

- $E = \mathbb{R}$
- $E = \mathbb{Q}$ (Pretend you know what $\mathbb{Q}$ is lol)
- $E =$ The Half-Integers (Pretend you know what integers are lol)
- $E = \{1, 2\} \cup [3, \infty)$

Ok, now how do we formalize that "nothing else should be in there"? The way you do it is say that $\mathbb{N}$ is just the "smallest set" satisfying these conditions. You can do this by just intersecting all such sets.

Definition 1.2 (Natural Numbers)

$$\mathbb{N} := \bigcap_{E \subseteq \mathbb{R}, E \text{ is inductive}} E$$

Example 1.3: Is $3 \in \mathbb{N}$?

Solution. We want to show that 3 is in every inductive set. So let E be inductive. Then:

- $1 \in E$
- $1 \in E \implies 1 + 1 = 2 \in E$
- $2 \in E \implies 2 + 1 = 3 \in E$

So 3 is in every inductive set, hence it is certainly in the intersection of all inductive sets, which is $\mathbb{N}$. ■

Example 1.4: Is $3.14 \in \mathbb{N}$?

Solution. If it were, then it is in every inductive set. But $\{1, 2, 3\} \cup [4, \infty)$ is inductive, and 3.14 isn't in it. ■

1.4 Induction

I love induction! You love induction! Everyone loves induction! Let's get induction.

Theorem 1.1

Let $\{p_n : n \in \mathbb{N}\}$ be a family a propositions such that:

- p_1 is true
- If p_n is true then p_{n+1} is true, for all $n \in \mathbb{N}$

Then p_n is true for all $n \in \mathbb{N}$.

Proof. Define $E = \{n \in \mathbb{N} : p_n \text{ is true}\}$. We WTS $E = \mathbb{N}$.

($\subseteq$) Obviously $E \subseteq \mathbb{N}$.

($\supseteq$) Note that E is an inductive set! Since $\mathbb{N}$ is the intersection of all inductive sets, we must have $\mathbb{N} \subseteq E$. □

Here are some examples on induction.

Example 1.5: Let $x > -1$, $n \in \mathbb{N}$. Prove that $(1+x)^n \geq 1 + nx$.

Proof. **Base Case:** $(1+x)^1 \geq 1 + 1 \cdot x$. Yay.

Hypothesis: Assume that $(1+x)^n \geq 1 + nx$.

Want To Show: $(1+x)^{n+1} \geq 1 + (n+1)x$.

We know $(1+x)^n \geq 1 + nx$. Since $x + 1 > 0$, we may multiply both sides by it to get: $(1+x)^{n+1} \geq (1+nx)(1+x) = 1 + (n+1)x + nx^2 \geq 1 + (n+1)x$ Yay! □

Example 1.6 (Binomial Theorem): Prove that for all $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$, we have:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Proof. It is clear for $n = 1$. Now assume that $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$. We have:

$$(x+y)^{n+1} = (x+y)(x+y)^n = (x+y) \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Distribute:

$$= x \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} + y \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Shove x and y into the sums:

$$= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k}$$

Shift the index on the left:

$$= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k}$$

Get the sum indices to match again:

$$= x^{n+1} + y^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=1}^{n} \binom{n}{k} x^k y^{n+1-k}$$

Combine sums:

$$= x^{n+1} + y^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] x^k y^{n+1-k}$$

Use Pascal's Identity:

$$= x^{n+1} + y^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} x^k y^{n+1-k}$$

Let the first and last terms back in:

$$= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}$$

□

Remark: Technically we need to assume that $0^0 = 1$. I am very adamant this is true. If you disagree then fight me c:

1.5 AM-GM

This might be important so I'm covering it. Assume that we've defined nth root.

Theorem 1.2

Let $a_1, \cdots, a_n \geq 0$. Then:

$$\frac{a_1 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \dots a_n}$$

With equality if and only if $a_1 = a_2 = \cdots = a_n$.

Proof. We use "backward induction". As an exercise, see if you can show that proving these clauses is sufficient:

- It is true for $n = 1$.

- If it is true for some n, then it is true for $2n$.
- If it is true for some n, then it is true for $n-1$.

$n=1$ is obvious. That's the base case. Also for $n=2$, you can get it by expanding $(a_1-a_2)^2 \geq 0$ to get $a_1^2+2a_1a_2+a_2^2 \geq 4a_1a_2$. Then square root. We need $n=2$ because we'll use it later.

Now let's do the "inductive steps".

($n \implies 2n$) Consider $2n$ variables $a_1, \cdots, a_{2n}$. Apply AM-GM for n variables on each "half":

$$\frac{a_1+\cdots+a_n}{n} \geq \sqrt[n]{a_1\dots a_n}$$

$$\frac{a_{n+1}+\cdots+a_{2n}}{n} \geq \sqrt[n]{a_{n+1}\dots a_{2n}}$$

Now average these equations:

$$\frac{a_1+\cdots+a_{2n}}{2n} \geq \frac{\sqrt[n]{a_1\dots a_n}+\sqrt[n]{a_{n+1}\dots a_{2n}}}{2}$$

Now use AM-GM for 2 variables on the right!

$$\geq \sqrt[2n]{a_1\dots a_{2n}}$$

This concludes this step.

($n \implies n-1$) The key is that intuitively, AM-GM for n variables should be somehow "stronger" than $n-1$ because it uses more variables, hence it handles more "information". We somehow need to "waste information".

Suppose we have $a_1, \cdots, a_{n-1}$ and we want to plug this into the AM-GM for n variables. What should the nth variable be? A "wasteful" idea is to let the nth variable be the arithmetic mean μ of $a_1, \cdots, a_{n-1}$. This gives:

$$\frac{a_1+\cdots+a_{n-1}+\mu}{n} \geq \sqrt[n]{a_1\dots a_{n-1}\mu}$$

Notice that the left side is just μ.

$$\mu \geq \sqrt[n]{a_1\dots a_{n-1}\mu}$$

$$\mu^n \geq a_1\dots a_{n-1}\mu$$

$$\mu^{n-1} \geq a_1\dots a_{n-1}$$

$$\mu \geq \sqrt[n-1]{a_1\dots a_{n-1}}$$

Hence we have proven the AM-GM inequality.

Showing the claim about when equality occurs is left as an exercise. □

Ok I didn't talk about this in recitation but I want to write about it anyway because I feel like it and it gets me paid. So here we go, examples!

Example 1.7: What is the largest possible area of a rectangle with perimeter 4?

Solution. If sides are x, y then $x + y = 2$. By AM-GM:

$$\frac{x+y}{2} \geq \sqrt{xy}$$

This gives $xy \leq 1$. Hence the maximum area is 1...

...no! This is a RUSHED CONCLUSION. I have only showed that 1 is an upper bound. I need to show that it can be obtained. To do this, study the equality case. If $xy = 1$ then we must have had $\frac{x+y}{2} \geq \sqrt{xy}$, meaning $x = y$. Combine this with $x + y = 2$ to deduce that $x = 1, y = 1$ should obtain the equality case. This is enough to verify that 2 is obtainable, but if you really care to do so you could always just plug in $x = 1, y = 1$ to verify it, or even "guess" that is works and verify it. ■

Example 1.8: Minimize $x + 7/x$ over $x > 0$.

Solution. By AM-GM:

$$x + \frac{7}{x} \geq 2\sqrt{x \cdot \frac{7}{x}} = 2\sqrt{7}$$

So theoretical minimum is $2\sqrt{7}$. Can we get it? Yes, with $x = \sqrt{7}$. Done. ■

I have more examples, will update this maybe.

2 Supremums

2.1 Warm-Up

Find a way to hang a painting around two nails such that the painting will fall when *either* nail is taken away.

Hint: Nofgenpgyl, erzbir gur jnyy sebz gur cvpgher naq guvax bs rnpu anvy nf na vasvavgryl ybat cbyr va 3Q fcnpr.

Hint: Ebgngr bar bs gur cbyrf 90 qrterrf. Guvf qbrf abg punatr gur ceboyrz, naq guvf znl fhttrfg n irel, irel pbby jnl gb gvr n ybbc bs ebcr nebhaq gurz.

2.2 One More Induction Example

> **Exercise 2.1:** Suppose $x \in \mathbb{R}$ such that $x + \frac{1}{x} = 42$. Prove that $x^n + \frac{1}{x^n}$ is an integer, for all $n \in \mathbb{N}$.

Proof. This is an example of induction where we want two base cases (or, alternatively, strong induction). Here specifically, we need two base cases because the inductive step will need to the two previous assumptions. That is, we're going to need "$n - 1$ and n imply $n + 1$" instead of just "n implies $n + 1$".

Base Case 0: It is obvious for $n = 0$. I know n isn't natural but it's totally chill to take it as a base case.

Base Case 1: It is obvious for $n = 1$ because 42 is an integer.

Inductive Assumption: Let us assume that $x^n + \frac{1}{x^n}$ and $x^{n-1} + \frac{1}{x^{n-1}}$ are integers. Let's give these integers names:

$$\text{OwO} := x^n + \frac{1}{x^n}$$

$$\text{uwu} := x^{n-1} + \frac{1}{x^{n-1}}$$

Inductive Step: We want to show that $x^{n+1} + \frac{1}{x^{n+1}}$ is an integer. One natural way to go about this is to expand $(x^n + 1/x^n)(x + 1/x)$. So:

$$(x^n + 1/x^n)(x + 1/x) = 42 \cdot \text{OwO}$$

$$x^{n+1} + \frac{1}{x^{n+1}} + x^{n-1} + \frac{1}{x^{n-1}} = 42 \cdot \text{OwO}$$

$$x^{n+1} + \frac{1}{x^{n+1}} + \text{uwu} = 42 \cdot \text{OwO}$$
$$x^{n+1} + \frac{1}{x^{n+1}} = 42 \cdot \text{OwO} - \text{uwu}$$

Since $42, \text{OwO}$, and uwu are integers, we have that $42 \cdot \text{OwO} - \text{uwu}$ is an integer, hence $x^{n+1} + \frac{1}{x^{n+1}}$ is an integer as needed. □

2.3 Supremums

I like to think of sups as "kinda like maximums". It's like what you'd like the maximum to be, if there was no maximum.

Example 2.2: Let $S = \{1, 1 + 1/2, 1 + 1/2 + 1/4, \cdots\}$. Find $\sup S$.

Solution. Because of Zeno, it's intuitively 2. Let's prove it.

Step 1 (Upper Bound): We need to show that 2 is $\geq$ everyone in S.

So pick an arbitrary element $2 - \frac{1}{2^n}$ in S (I leave it as an exercise to show that all numbers in S take this form).

We need to show that $2 \geq 2 - \frac{1}{2^n}$. Fortunately this is very true, so indeed 2 is an upper bound.

Step 2 (Least such upper bound): Suppose M were another upper bound. That is:

$$M \geq 2 - \frac{1}{2^n} \qquad \forall n \in \mathbb{N}$$

We need to show that $2 \leq M$.

Assume for contradiction that actually $M < 2$. By the Archimedean Property, let us find some $n \in \mathbb{N}$ so large that:

$$n(2 - M) > 1$$

Since $2^n > n$ (exercise), we have in fact that:

$$2^n(2 - M) > 1$$

This rearranges to $M < 2 - \frac{1}{2^n}$. This contradicts the assumption that $M \geq 2 - \frac{1}{2^n}$ for all n.
■

2.4 Another Way To Think About Supremums

In lecture, you learned that $M = \sup S$ iff:

1. M is an upper bound for S.
2. If M' is also an upper bound, then $M \leq M'$.

Here, I provide you an alternate way to think about this. I claim that $M = \sup S$ iff:

1. M is an upper bound for S.
2. S contains an element of $(M - \varepsilon, M]$ for all $\varepsilon > 0$.

Let's prove this.

Theorem 2.1

Suppose M is an upper bound for S. Then $M \leq M'$ for all upper bounds M' if and only if $S \cap (M - \varepsilon, M] \neq \emptyset$ for all $\varepsilon > 0$.

Proof. ($\implies$) Suppose $M \leq M'$ for every upper bound M' for S. Fix $\varepsilon > 0$. I claim that S contains an element of $(M - \varepsilon, M]$.

If not, then let $M' = M - \varepsilon$. Notice that M' is an upper bound! In fact, it is an upper bound strictly smaller than M. Contradiction.

($\impliedby$) Suppose that S contains someone in $(M - \varepsilon, M]$ for all $\varepsilon > 0$.

Let M' be an upper bound for S. We need to show that $M \leq M'$. To do this, we suppose otherwise. Then $M > M'$, and so $M - M' > 0$. Taking $\varepsilon = M - M'$, we must have that S contains someone in $(M - \varepsilon, M]$, or $(M', M]$. But this is really bad! This means that there is some $x \in S$ satisfying $M' < x \leq M$. Since M' is not $\geq x$, and $x \in S$, we see that actually M' was not an upper bound, contradiction. $\square$

How should we think about this new condition? I think of it like this: "If M is the sup, then M is $\geq$ everyone in S, AND MOREOVER you can get **as close to M as you want**, while staying in S".

Let's use this intuition to prove an important theorem.

Definition 2.1

For $S, T \subseteq \mathbb{R}$ we define $S + T := \{x + y : x \in S, y \in T\}$.

Theorem 2.2

$\sup(S+T) = \sup S + \sup T$

Proof. Let $M = \sup S + \sup T$.

Step 1 (Upper Bound): We need to show that M is $\geq$ everyone in $S+T$. So let's pick an arbitrary element $x+y \in S+T$, where $x \in S$ and $y \in T$. Is $M = \sup S + \sup T \geq x+y$? Well I'd sure hope so! Since $x \leq \sup S$ and $y \leq \sup T$, we win here.

Step 2 (Can Get Very Close To M): Fix an arbitrary $\varepsilon > 0$. We want to show that some element $x+y$ of $S+T$ is less than ε away from M. That is, we need to find x and y such that:

$$M - \varepsilon < x + y \leq M$$

Can we do it? The trick is to choose x and y really close to $\sup S$ and $\sup T$. In theory this should work!

Since $\sup S$ is the supremum of S, we must be able to choose an element $x \in S$ such that:

$$\sup S - \frac{\varepsilon}{2} < x \leq \sup S$$

Since $\sup T$ is the supremum of T, we must be able to choose an element $y \in T$ such that:

$$\sup T - \frac{\varepsilon}{2} < y \leq \sup T$$

Now let's add these equations together for fun.

$$\sup S + \sup T - \varepsilon < x + y \leq \sup S + \sup T$$

$$M - \varepsilon < x + y \leq M$$

This is exactly what we wanted! Since we found an $x+y \in S+T$ satisfying the above, for any choice of $\varepsilon > 0$, we have proven the second condition. Thus indeed $M = \sup(S+T)$. $\square$

3 Normed Spaces and Inner Product Spaces

3.1 Warm-up

There are 5 apples for sale with 5 different sizes and 5 different positive integer prices from \$1 to \$5. In dollars, what is the price of the apple that's bigger than the apple that costs more than the apple that's smaller than the apple that's cheaper than the apple that's green, given that it is red?

3.2 Review of Space!

Definition 3.1 (Vector Space)

A **vector space** is a nice place where you can add things and the things can stretch!
A **vector space** is a module over a field!
A **vector space** (over $\mathbb{R}$) is a set endowed with a commutative, associative vector addition and scalar multiplication, closed under these operations, such that blah blah blah help

Definition 3.2 (Normed Space)

A **normed space** is a vector space where you can judge the size of things.
A **normed space** is a vector space endowed with a *norm* $\|\cdot\|$ satisfying the following:

1. $\|x\| \geq 0$ always, and equality occurs iff $x = 0$
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$
3. $\|x\| + \|y\| \geq \|x + y\|$

Definition 3.3 (Inner Product Space)

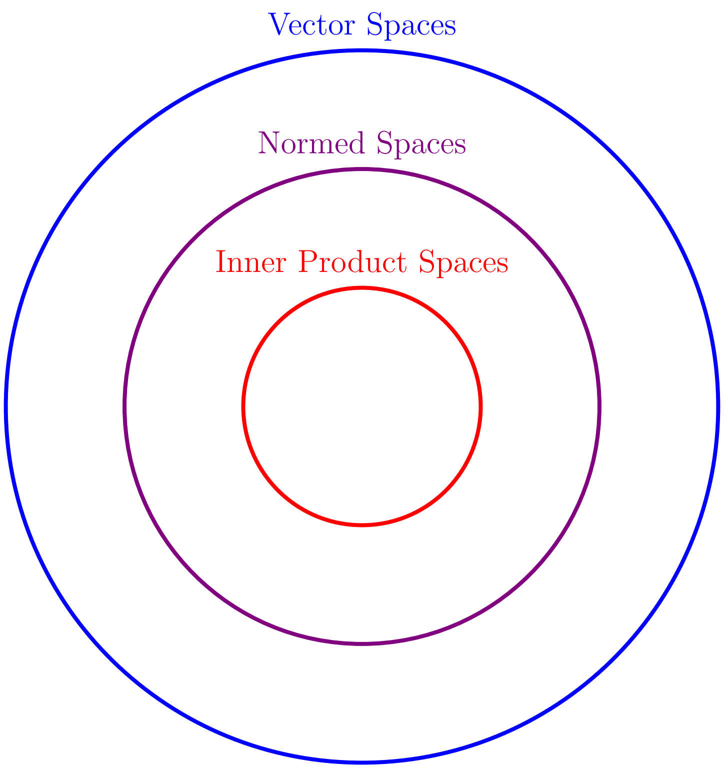
An **inner product space** is a vector space endowed with a binary operation $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$, sometimes denoted $\cdot$, satisfying the following:

1. $\langle x, y \rangle = \langle y, x \rangle$
2. $\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle$
3. $\langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle$ ("Bilinearity")
4. $\langle x, x \rangle \geq 0$ with equality exactly when $x = 0$

Theorem 3.1

All inner product spaces are normed spaces by taking $\|x\| := \sqrt{\langle x, x\rangle}$.

Proof. Exercise. □



3.3 In $\mathbb{R}^N$

Some norms for $\mathbb{R}^N$:

- Standard (Euclidean): $\|x\| = \sqrt{\sum x_i^2}$
- Taxicab: $\|x\|_1 = \sum |x_i|$
- L^p norm: $\|x\|_p = (\sum |x_i|^p)^{1/p}$ (One-liner proof: "Google Minkowski")
- Max norm: $\|x\|_\infty = \max |x_i|$
- Weird norm I just made up: $\|x\| := |x_1| + \max(|x_2|, |x_3|) + \sqrt{x_4^2 + x_5^2}$
- **Non-Example:** $\sum x_i^2$ (Issue: Scaling)

Some inner products for $\mathbb{R}^N$:

- Standard one: $\langle \vec{x}, \vec{y} \rangle := \sum_{i=1}^N x_i y_i$
- Some others: $\langle \vec{x}, \vec{y} \rangle := x^T A y$ for symmetric, positive-definite $N \times N$ matrix A
- NOTHING ELSE!!! (Proof: Sacrifice your soul to the eternal Cummings)
- **Non-Example:** $(a, b) \cdot (c, d) := \begin{vmatrix} a & c \\ b & d \end{vmatrix}$ (Issue: $(1, 1) \cdot (1, 1) = 0$ but $(1, 1) \neq \vec{0}_X$

3.4 More Exotic Stuff

Definition 3.4

For $1 \leq p < \infty$, l^p is the set of all infinite sequences $\{x_i\}$ for which $\sum_{i=1}^\infty |x_i|^p < \infty$. We call this the *sequence space*.

Definition 3.5

Let I be an interval. Then $C_b(I)$ is the set of all continuous and bounded functions on I.

Examples of exotic normed spaces:

- l^1 with $\|\{x_i\}\|_1 := \sum_{i=1}^\infty |x_i|$
- l^2 with $\|\{x_i\}\|_2 := \sqrt{\sum |x_i|^2}$
- $C_b[0, 1]$ with $\|f\|_\infty := \sup_{[0,1]} |f|$ (What is the dimension of this space?)
- $C_b[0, 1]$ with $\|f\|_{L^p} = \left(\int_0^1 |f|^p \right)^{1/p}$
- $\mathbb{R}^{N \times N}$ with $\|A\|_{X^*} := \sup\{\|Av\| : \|v\| = 1\}$
- **Non-Example:** Let X be the space of sequences of real numbers, and endow it with the "sup norm" $\|\{x_i\}\|_\infty := \sup_{1 \leq i \leq \infty} |x_i|$. (Issue: Not actually a norm because norms must always be finite!)
- **Non-Example:** Space of "integrable" functions on $[0, 1]$ with $\|f\| = \int_0^1 |f| \, dx$. (Issue: There are many functions that have "norm" 0, which is bad. Only the zero element is allowed to have zero norm, hence this integral thing is NOT a norm on this space.)

Examples of exotic inner product spaces:

- l^2 with $\sum |x_i y_i|$
- $C_b[0,1]$ with $\langle f, g\rangle := \int_0^1 f(x)g(x)\,dx$

3.5 Ok but why do I care???

Lots of analysis is done using just the standard stuff on $\mathbb{R}^N$. But if you open your mind, you'll see that a lot of the stuff we will do (topologies, continuity, directional derivatives, minimization, Lagrange multipliers, etc.) will extend to a bunch of really exotic spaces, and this lets us do some insanely powerful stuff.

Example:

- We all know that in the normed space $\mathbb{R}$, we can minimize a differentiable function $f : \mathbb{R} \to \mathbb{R}$ by setting its derivative to 0 (and arguing that it works via e.g. convexity).
- Some of you know that in the normed space $\mathbb{R}^N$, we can minimize a differentiable function $f : \mathbb{R}^N \to \mathbb{R}$ by setting its gradient to $\vec{0}$ (and arguing that it works via e.g. the Hessian matrix).
- But these methods actually extend to infinite dimensions, in a way! For example, $F(f) := \int_0^1 f'(x)^2 + (f(x) - x)^2\,dx$ is a function $F : C^1(0,1) \to \mathbb{R}$. That is, it's a function that takes in continuously differentiable functions and spits out a real number. The vector space here is an infinite-dimensional *function space*, and it turns out that you can minimize $F(f)$ because we can take directional derivatives in infinite dimensions and set them to 0. Google "Calculus of Variations" if you're interested.

4 Metric Spaces and Open Sets

4.1 Warm-up

Consider the **empty topological space**, which is the empty set endowed with the unique topology on the empty set.

How many open covers are there of $\{\}$ in this space?

Answer & Reason: Gur nafjre vfa'g mreb be bar, vg vf npghnyyl gjb. Guvf vf orpnhfr hfvat ab bcra frgf ng nyy jvyy pbire gur rzcgl frg, naq gnxvat gur fvatyr bcra frg - gur rzcgl frg - jvyy nyfb pbire gur rzcgl frg.

4.2 Definition Recall

I forgot to say this in recitation but in analysis-y math stuff, the two most important types of sets/domains are the **open sets** and the **compact sets**. Open sets give you really good wiggle room which let you figure out changes near things without abrupt trouble. Compact sets are the opposite - they are *really* abrupt, and prevent things from changing too much.

Today we focus more on the open sets and related things.

Definition 4.1 (Open)

A set is open if it never ends abruptly.
A set is open if there's always wiggle room inside.
Let (X, d) be a metric space. A set $E \subseteq X$ is **open** if for all $x \in E$, there exists $r > 0$ such that $B(x, r) \subseteq E$.

Definition 4.2 (Closed)

$E \subseteq X$ is closed if $X \setminus E$ is open.

Definition 4.3 (Interior)

$x \in E$ is an **interior point** if it can freely wiggle around. That is, you can find a ball $B(x, r) \subseteq E$.
E° is the set of all interior points of E.

Exercise 4.1: Do we have $E^\circ \cup F^\circ = (E \cup F)^\circ$?

Solution. No. Take $E = \mathbb{Q}$ and $F = \mathbb{R} \setminus \mathbb{Q}$. ■

Exercise 4.2: Do we have $E^\circ \cap F^\circ = (E \cap F)^\circ$?

Solution. Yes. Suppose $x \in E^\circ \cap F^\circ$. Then $x \in E^\circ$ so there is $B(x, r_1) \subseteq E$, and $x \in F^\circ$ so there is $B(x, r_2) \subseteq F$. Taking $r = \min(r_1, r_2)$ we get $B(x, r) \subseteq E$ and $B(x, r) \subseteq F$ so $B(x, r) \subseteq E \cap F$, so $x \in (E \cap F)^\circ$.

For the other direction, note that if $x \in (E \cap F)^\circ$ then we can find $B(x, r) \subseteq E \cap F$. Particularly $B(x, r) \subseteq E$ and $B(x, r) \subseteq F$ so both $x \in E^\circ$ and $x \in F^\circ$. ■

Definition 4.4 (Accumulation)

A point $x \in X$ is an **accumulation point** of a set E if you can get really close to x while staying in E (without ever reaching x itself). That is, for all $r > 0$ there exists $y \in E$ with $y \neq x$ such that $d(y, x) < r$. Or, alternatively, $E \cap (B(x, r) \setminus \{x\}) \neq \emptyset$ for all $r > 0$.

Finding the closure $\overline{E}$ can be very annoying. But finding $\operatorname{acc} E$ might be easier on the mind, if you're anything like me. The next theorem helps with determining $\overline{E}$.

Theorem 4.1

For (X, d) and $E \subseteq X$, we have:

$$\overline{E} = E \cup \operatorname{acc} E$$

Moreover, we may deduce that $C \subseteq X$ is closed iff C contains all its accumulation points.

Proof. (Draw a picture to follow along!)

($\subseteq$) Take $x \in \overline{E}$. Then $x \in C$ for every closed $C \supseteq E$. Equivalently, $x \notin U$ for every open $U \subseteq X \setminus E$.

If $x \in E$ then this direction is done, so assume that $x \notin E$. We want to prove that $x \in \operatorname{acc} E$. To wit, fix $r > 0$. We must prove that there exists $y \in E$ with $0 < d(x, y) < r$, i.e. $y \in E \cap B(x, r) \setminus \{x\}$.

Suppose not. Then no such y exists, which means $E \cap (B(x, r) \setminus \{x\}) = \emptyset$. In fact, since $x \notin E$, we can get more strongly that $E \cap B(x, r) = \emptyset$. This means that $B(x, r)$ is an open set satisfying $B(x, r) \subseteq X \setminus E$, so from the first paragraph we obtain $x \notin B(x, r)$, a clear contradiction.

($\supseteq$) For the other direction, suppose $x \in E \cup \operatorname{acc} E$. There are two cases. If $x \in E$, then clearly x is in every closed set that contains E, so $x \in \overline{E}$. Else, if $x \in \operatorname{acc} E$, then it suffices to show that if $C \supseteq X$ is closed then $x \in C$.

Assume otherwise. Then $x \in X \setminus C$, and this is open. So we may find $B(x, r) \subseteq X \setminus C$. But now $B(x, r) \setminus \{x\}$ contains no points in E, contradicting $x \in \operatorname{acc} E$. Hence $x \in C$, and the theorem is proven by double containment.

To show the last remark, note that if C is closed then $C = \overline{C} = C \cup \operatorname{acc} C$, or $\operatorname{acc} C \subseteq C$. Conversely if $\operatorname{acc} C \subseteq C$ then $C \subseteq \overline{C} = C \cup \operatorname{acc} C \subseteq C$, so $C = \overline{C}$ and C is closed. $\square$

4.3 Some Examples

Example 4.3: Let $E = \mathbb{R} \setminus \{1/n : n \in \mathbb{N}\} \subseteq \mathbb{R}^2$. Take the standard metric.

- Is E open? Is E closed?
- What is E°?
- What is $\overline{E}$?
- What is $\operatorname{acc} E$?
- What is ∂E?

Solution.

- E is NOT open. (Look at $0 \in E$) E is not closed either. (Look at $1 \in \mathbb{R} \setminus E$)
- $E^\circ = \mathbb{R} \setminus (\{0\} \cup \{1/n : n \in \mathbb{N}\})$ (To prove: First argue that all points inside this proposed sets are indeed interior points. For instance, if $1/3 < x < 1/2$ you can pick $r = \min(1/2 - x, x - 1/3)$, so that $B(x, r) \subseteq (1/3, 1/2) \subseteq E$. Then, argue that all other points of E cannot be interior points... and, well, the only other point is 0.)
- $\overline{E} = \mathbb{R}$ (Use $\overline{E} = E \cup \operatorname{acc} E$ and use the next line)
- $\operatorname{acc} E = \mathbb{R}$ (All interior points are accumulation points, so just need to check 0 and $1/n \, \forall n \in \mathbb{N}$.)
- $\partial E = \{0\} \cup \{1/n : n \in \mathbb{N}\}$ (Just compute $\overline{E} \setminus E^\circ$)

■

Example 4.4: Let $E = (1, 2) \times (1, 2) \subseteq \mathbb{R}^2$. Take the standard metric.

- Is E open?
- What is E°?
- What is $\overline{E}$?
- What is $\operatorname{acc} E$?
- What is ∂E?

Solution.

- E is open. (For $(x, y) \in E$ pick $r = \min(x - 1, 2 - x, y - 1, 2 - y)$ and argue that this works)
- $E^\circ = E$ (The interior of any open set is itself.)
- $\overline{E} = [1, 2] \times [1, 2]$ (Union E with its accumulation, which we find in the next line.)
- $\operatorname{acc} E = [1, 2] \times [1, 2]$ (First use casework to argue that if (x, y) is outside of $[1, 2] \times [1, 2]$, then you can draw a small ball around (x, y) that does not intersect E, which implies that $(x, y) \notin \operatorname{acc} E$. Next, every $(x, y) \in E$ is clearly in $\operatorname{acc} E$ because E is open. It remains to handle the "border points", and this is handled by casework. For example, to show $(1, 1) \in \operatorname{acc} E$, we need to find some point in E that is less than r away from $(1, 1)$, no matter how small r is. Fortunately we can pick something like $(1 + r/50, 1 + r/50)$ and you can argue that this must be close enough.)
- $\partial E = \{1, 2\} \times [1, 2] \cup [1, 2] \times \{1, 2\}$ (Just compute $\overline{E} \setminus E^\circ$.)

■

Definition 4.5

The **French Railway metric** is a metric d on $\mathbb{R}^2$ defined as follows:

$$d(x,y) := \begin{cases} \|x-y\|, & x,y,0 \text{ are collinear} \\ \|x\| + \|y\|, & \text{Otherwise} \end{cases}$$

Example 4.5: Let $E = (1,2) \times (1,2) \subseteq \mathbb{R}^2$. Take the French Railway metric.

- Is E open?
- What is E°?
- What is $\overline{E}$?
- What is $\operatorname{acc} E$?
- What is ∂E?

Solution. Yeah I don't want to prove any of these formally, this example is best for capturing intuition for reasoning about weird metrics.

- E is open. (There is wiggle room along every ray/railroad intersecting E)
- $E^\circ = E$ (Again, the interior of any open set is itself)
- $\overline{E} = [1,2] \times [1,2] \setminus \{(1,2),(2,1)\}$ (See next line)
- $\operatorname{acc} E = [1,2] \times [1,2] \setminus \{(1,2),(2,1)\}$ (You can approach every point on the boundary along the railways... EXCEPT those two darn corners!)
- $\partial E = \{1,2\} \times [1,2] \cup [1,2] \times \{1,2\} \setminus \{(1,2),(2,1)\}$

■ k

Example 4.6: Consider the normed space $(C[0,1], \|\cdot\|_\infty)$ and take

$$E = \{f \in C[0,1] : \|f\|_\infty < 1\}$$

- Is E open?
- What is E°?
- What is $\overline{E}$?
- What is $\operatorname{acc} E$?
- What is ∂E?

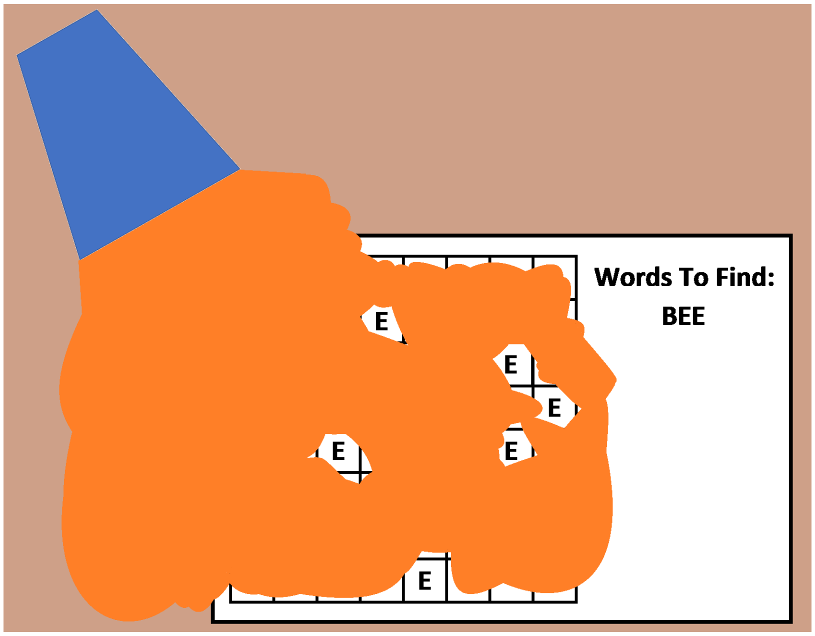
Solution. Exercise. (Not as hard as it looks!) ■

5 Limits

This recitation will be a bit more practical rather than rigorous at some points.

5.1 Warm-up

I made a really special wordsearch made up of only two letters! But I spilled my orange juice all over it. Too bad. Solve it.



1. Words are found in any of the 8 standard compass directions (in particular, a word *may* appear backwards).
2. The word you need to find appears once and exactly once.
3. Again, the wordsearch contains only two distinct letters.

Hint: Bapr n orr unf orra sbezrq, ab zber orrf pna or sbezrq, naq guvf vf n cbjreshy pbaqvgvba. Bar jnl gb fgneg vf nf sbyybjf: Hfvat "purffobneq abgngvba" jvgu obggbz-yrsg orvat n1, pbafvqre jung jbhyq unccra vs s7 jrer n o. Jung yrggref ner sbeprq gb or r? Nf lbh xrrc tbvat, jvyy lbh eha vagb n pbagenqvpgvba va gung zhygvcyr orrf nccrne? Vs fb, gura s7 vf abg n o. Guvf zvtug gnxr n juvyr gubhtu, fb frr vs lbh pna fcbg gur tenaq fpurzr bs gur chmmyr.

5.2 Some Starting Limits

Example 5.1:

$$\lim_{x\to 3} x^3$$

Solution. **Key idea:** Can always assume $\varepsilon < 1$.

We claim the limit is 27.

Fix $\varepsilon > 0$, assume for ease that $\varepsilon < 1$. We choose $\delta = \varepsilon/1000$. To see that this works, note first that if $0 < |x-3| < \delta$ then in particular, $x < 3 + \delta < 3.001 < 4$. This helps us get the following inequality for all x with $0 < |x-3| < \delta$:

$$|x^3 - 27| = |x-3| \cdot |x^2 + 3x + 9| \leq \frac{\varepsilon}{1000} \cdot (|x^2| + |3x| + 9)$$

$$\leq \frac{\varepsilon}{1000} \cdot (16 + 12 + 9) < \varepsilon$$

■

Example 5.2:

$$\lim_{x\to 42} 1_{\mathbb{Q}}(x)$$

Solution. We claim that the limit does not exist.

Suppose the limit were L. Take $\varepsilon = 1/4$ or something. We want to show that for any $\delta > 0$, there is some x with $0 < |x - 9001| < \delta$ such that $|1_{\mathbb{Q}}(x) - L| \geq \varepsilon$.

To wit, take any $\delta > 0$. There are two cases.

- If $L \geq 1/2$, we use density of irrationals to pick $x \in (9001, 9001 + \delta)$ irrational. This gives $1_{\mathbb{Q}}(x) = 0$ so $|1_{\mathbb{Q}}(x) - L| \geq 1/2 \geq \varepsilon$.
- If $L < 1/2$, use density of rationals instead.

So the limit does not exist. ■

Example 5.3: $\lim_{x \to 1} f(f(x))$, where

$$f(x) := \begin{cases} -(x+1)^2, & x < 0 \\ 0, & x = 0 \\ (x-1)^2, & x > 0 \end{cases}$$

Solution. We claim the limit exists (!), and is equal to 1.

As usual, fix $\varepsilon > 0$. We want to find δ such that

$$|f(f(x)) - 1| < \varepsilon$$

for all x with $0 < |x - 1| < \delta$.

At this point you should tell the reader your choice of δ and then show that it works. I will not do that in order to make the choice of δ more intuitive.

Let us plan to take a δ that is less than 1, in order to guarantee that $x > 0$ (why does this work?). Then for all x within δ of 1, we have $f(f(x)) = f((x-1)^2)$.

But $|x - 1| > 0$ so particularly $(x-1)^2 > 0$ (strictly!), so we can further write

$$f((x-1)^2) = ((x-1)^2 - 1)^2.$$

(Remember, this is for all x with $0 < |x - 1| < \delta$.)

At this point you can probably just "plug 1 in", but if you want to be really super duper rigorous, we can continue as follows: Write, for all $0 < |x - 1| < \delta$:

$$|f(f(x)) - 1| = |((x-1)^2 - 1)^2 - 1| = |(x-1)^2| \cdot |(x-1)^2 - 2|$$

$$= (x-1)^2 \cdot |x^2 - 2x - 2|$$

Let's use our forcing of $\delta < 1$ to deduce $x < 2$, giving:

$$\leq \delta^2(|x^2| + |-2x| + |-2|) \leq \delta^2(4 + 4 + 2) = 10\delta^2$$

Hence we see that we can choose $\delta = \min(\sqrt{\varepsilon/10}, 1)$. So we're done.

(If we impose the assumption $\varepsilon < 1$ like before, we don't need the min; we can just do $\delta = \sqrt{\varepsilon/10}$ and this already guarantees $\delta < 1$ from $\varepsilon < 1$.)

■

5.3 To Infinity and Beyond

1. What is $\lim_{x \to +\infty} \dfrac{1}{x}$???

2. What is $\lim_{x \to 0} \dfrac{1}{x^2}$???

We need new definitions.

Definition 5.1 (Infinity Limits)

We say that $\lim_{x\to\infty} f(x) = L$ if for all $\varepsilon > 0$ there exists N_ε such that $|f(x) - L| < \varepsilon$ for all $x \geq N_\varepsilon$.
We say that $\lim_{x\to x_0} f(x) = \infty$ if for all $N > 0$ there exists $\delta > 0$ such that $f(x) \geq N$ for all x with $0 < |x - x_0| < \delta$.
Analogously you can define limits such as $\lim_{x\to-\infty} f(x) = L$, $\lim_{x\to\infty} f(x) = \infty$, etc.

Limits to infinity can also be defined using this sort of "extended real" metric, i.e. assigning this really niche metric on $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, but I don't think it's really that useful.

5.4 Beware of L'Hopital

Theorem 5.1

Suppose $f, g : \mathbb{R} \to \mathbb{R}$, $x_0 \in \overline{\mathbb{R}}$. Assume that f, g are differentiable in $B(x_0, r) \setminus \{x_0\}$ for some r, and the limits $\lim\limits_{x\to x_0} f(x)$ and $\lim\limits_{x\to x_0} g(x)$ exist and are equal to 0. Then, if the limit

$$L := \lim_{x\to x_0} \frac{f'(x)}{g'(x)}$$

exists, then $\lim\limits_{x\to x_0} \dfrac{f(x)}{g(x)}$ exists and is equal to L.

Example 5.4 (Stolen from Gan): Compute:

$$\lim_{x\to+\infty} \frac{x}{2x + \sin x}$$

An incorrect solution would take the derivative of numerator and denominator to conclude that the limit is the limit of $\frac{1}{2+\cos x}$, which does not exist.

This does NOT mean that the original limit does not exist! This is because L'Hopital is *inconclusive* if the limit of the expression you get after differentiation does not exist. Read the statement of the theorem carefully!

Solution. We claim that the limit is $\frac{1}{2}$. To see this, fix $\varepsilon > 0$. We choose $N_\varepsilon = \frac{1}{2} + \frac{1}{4\varepsilon}$. Then for all $x \geq N_\varepsilon$ we have:

$$\left|\frac{x}{2x+\sin x} - \frac{1}{2}\right| = \left|\frac{\sin x}{4x + 2\sin x}\right|$$
$$\leq \frac{1}{|4x+2\sin x|} \leq \frac{1}{|4x| - |2\sin x|} \leq \frac{1}{4x-2} < \varepsilon$$

(Do you see where each of these inequalities come from?) ∎

5.5 Using Standard Limits

You are basically allowed to assume all of these guys unless otherwise specified:

- $\lim_{x\to 0} \frac{\sin x}{x} = 1$
- $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}$
- $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$
- $\lim_{x\to 0} \frac{e^x - 1}{x} = 1$
- $\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e$

Example 5.5:

$$\lim_{x\to 0} \frac{\sin(1-\cos\sqrt{|x|})}{|\log(1+x)|}$$

Solution. For all $x \neq 0$, note that we have the following:

$$\frac{\sin(1-\cos\sqrt{|x|})}{|\log(1+x)|} = \frac{\sin(1-\cos\sqrt{|x|})}{1-\cos\sqrt{|x|}} \cdot \frac{1-\cos\sqrt{|x|}}{|x|} \cdot \frac{|x|}{|\log(1+x)|}$$

Now look at the right side as we send $x \to 0$.

- As $x \to 0$ we have $1 - \cos\sqrt{|x|} \to 0$, so we may apply the standard $\frac{\sin y}{y} \to 1$ limit to deduce that the first term goes to 1.
- As $x \to 0$ we have that $\sqrt{|x|} \to 0$, so we may apply the standard $\frac{1-\cos y}{y^2} \to \frac{1}{2}$ limit to deduce that the second limit goes to $\frac{1}{2}$.

- Directly, the third limit goes to 1 by a standard limit.
- Since each of these three limits exist, we have that the limit of their product exists, and is equal to the product of the respective limits.
- Therefore the limit is $1 \cdot \frac{1}{2} \cdot 1 = \boxed{\frac{1}{2}}$.

■

5.6 Limits in more dimensions!

Question: If limit exists and $= L$ along every line $y = mx$, must the limit be L?

Answer: No, consider $f(x,y) = \begin{cases} x^2/y, & y \neq 0 \\ 0, & y = 0 \end{cases}$, and consider the limit $\lim_{(x,y)\to(0,0)} f(x,y)$.

Along $y = mx$ the limit as we approach $(0,0)$ is $\lim_{x\to 0} \frac{x^2}{mx} = 0$. But the original 2D limit does not exist! Consider approaching $(0,0)$ along the curve $y = x^2$. Then $f(x,y)$ will approach 1...

Now let's find other limits. Solutions at the end.

Example 5.6:

$$\lim_{(x,y)\to 0} \frac{xy}{x^2 + y^2}$$

Example 5.7:

$$\lim_{(x,y)\to 0} \frac{\sqrt{|x|} \cdot y}{x^2 + y^2}$$

Example 5.8:

$$\lim_{(x,y)\to 0} \frac{x^2 y^3}{|x|^3 + y^4}$$

Example 5.9 (HARD):

$$\lim_{(x,y)\to 0} \frac{x^2 y^2}{|x|^3 + y^4}$$

Example 5.10:

$$\lim_{(x,y)\to(0,0)} \frac{\log(1 + \sin(x^2 + y^4))}{x^2 + y^2}$$

Solution. (5.6) The limit does not exist. Approach $(0,0)$ along $y = mx$. The value approaches along this path is

$$\lim_{x\to 0} \frac{x \cdot mx}{x^2 + m^2x^2} = \lim_{x\to 0} \frac{m}{1+m^2}$$

Choosing like $m = 1$ and $m = 2$ we see that we can get tons of different limits depending on m, so the limit does not exist. ∎

Solution. (5.7) The limit does not exist.

First approach along $y = 0$. Then the limit we get is... 0. So if the limit exists then it must be 0.

On the other hand, we can approach $(0,0)$ along the curve $y = x^{3/2}$ (specifically, along the positive part of it). Then the value approached is

$$\lim_{x\to 0^+} \frac{\sqrt{|x|} \cdot x^{3/2}}{x^2 + x^3} = \lim_{x\to 0^+} \frac{x^2}{x^2 + x^3} = \lim_{x\to 0^+} \frac{1}{1+x} = 1.$$

Since we got two different limits, the limit does not exist.

(I just used 0^+ to be careful regarding signs.) ∎

Solution. (5.8) The limit DOES exist, and we claim that the limit is 0. The key observations to deduce this are as follows:

- For all (x,y) sufficiently close to $(0,0)$, we have that $|x|^3 \geq x^4$. Rigorously, we can ensure this inequality occurs (which will help us compute the limit) by forcing our $\delta > 0$ to be smaller than 1.
- By the AM-GM inequality, $x^4 + y^4 \geq 2x^2y^2$.

Thus for all (x,y) close enough to $(0,0)$, we are able to write:

$$\begin{aligned} 0 \leq \left| \frac{x^2y^3}{|x|^3 + y^4} \right| &= |y| \cdot \frac{x^2y^2}{|x|^3 + y^4} \\ &\leq |y| \cdot \frac{x^2y^2}{x^4 + y^4} \\ &\leq |y| \cdot \frac{x^2y^2}{x^2y^2} = |y| \end{aligned}$$

Since $|y| \to 0$ as $(x,y) \to (0,0)$, we may conclude by the squeeze rule.

Note: It is intuitive to guess that the limit should be 0 because the degree of the numerator is larger than the degree of the denominator, so the numerator should tend more rapidly to 0. This is no proof, though... you must show it! ∎

Solution. (5.9 - Thomas Lam) With apologies to everyone that reads this.

Fix $\varepsilon > 0$. We choose

$$\delta = \min\left(\left(\frac{7}{4}\right)^2 \varepsilon^2, \left(\frac{7}{3}\right)^{3/2} \varepsilon^{3/2}\right)$$

We show that this works. First observe via AM-GM that:

$$\frac{|x|^{7/2} + |x|^{7/2} + |x|^{7/2} + |x|^{7/2} + |y|^{14/3} + |y|^{14/3} + |y|^{14/3}}{7} \geq x^2y^2$$

From which it follows that:

$$x^2y^2 \leq \tfrac{4}{7}|x|^{7/2} + \tfrac{3}{7}|y|^{14/3} = \tfrac{4}{7}|x|^3|x|^{1/2} + \tfrac{3}{7}|y|^4|y|^{2/3} \qquad (*)$$

But for all (x, y) with $\|(x, y)\| = \sqrt{x^2 + y^2} < \delta$, we have that:

- $|x| \leq \sqrt{x^2 + y^2} < \delta \leq \left(\frac{7}{4}\right)^2 \varepsilon^2$, so $\frac{4}{7}|x|^{1/2} < \varepsilon$.
- $|y| \leq \sqrt{x^2 + y^2} < \delta \leq \left(\frac{7}{3}\right)^{3/2} \varepsilon^{3/2}$, so $\frac{3}{7}|y|^{2/3} < \varepsilon$.

Applying these inequality to $(*)$ gets us:

$$x^2y^2 \leq \varepsilon|x|^3 + \varepsilon|y|^4$$

Or:

$$\frac{x^2y^2}{|x|^3 + y^4} < \varepsilon$$

And this holds for all $0 < \|(x, y)\| < \delta$, so the limit is $\boxed{0}$. ■

Solution. (5.9 - David Altizio) By AM-GM:

$$\frac{1}{|x|}\left(\frac{x}{y}\right)^2 + \left(\frac{y}{x}\right)^2 \geq \frac{2}{\sqrt{|x|}}$$

Thus

$$0 \leq \frac{x^2y^2}{|x|^3 + y^4} = \frac{1}{\frac{1}{|x|}(x/y)^2 + (y/x)^2} \leq \frac{1}{2/\sqrt{|x|}} = \frac{\sqrt{|x|}}{2}$$

Now apply the squeeze rule. (Exercise for you: Is $x = 0$ a problem? How can you reason with it?) ■

Solution. (5.9 - Edward Hou) Substitute $(|x|^{3/2}, y^2) = (r\cos\theta, r\sin\theta)$ so that $|x|^3 = r^2\cos^2\theta$ and $y^4 = r^2\sin^2\theta$. Now we need to prove that:

$$\frac{r^{4/3}|\cos\theta|^{4/3} r|\sin\theta|}{r^2} \to 0$$

"As $r \to 0^+$ no matter how θ misbehaves". That is, for all $\varepsilon > 0$ there is $\delta > 0$ such that the above expression is $< \varepsilon$ for all (r, θ) with $0 < r < \delta$.

But, this is clear, since we can upper bound via $r^{1/3}$ which tends to zero with no dependence on θ. Tada.

...

If you have doubts about the polar substitution, we can pretend to do the substitution without actually doing it: Write

$$\frac{x^2 y^2}{|x|^3 + y^4} = \frac{\sqrt{|x|^3 + y^4}^{4/3} \cdot \frac{x^2}{\sqrt{|x|^3+y^4}^{4/3}} \cdot \sqrt{|x|^3 + y^4} \cdot \frac{y^2}{\sqrt{|x|^3+y^4}}}{\sqrt{|x|^3 + y^4}^2}$$

and go up via $\sqrt{|x|^3 + y^4}^{4/3} \geq \sqrt{|x|^3}^{4/3} = x^2$ and $\sqrt{|x|^3 + y^4} \geq \sqrt{y^4} = y^2$:

$$\leq \frac{\sqrt{|x|^3 + y^4}^{4/3} \cdot \sqrt{|x|^3 + y^4}}{\sqrt{|x|^3 + y^4}^2}$$

$$= \sqrt{|x| + y^4}^{1/3}$$

And this $\to 0$ as $(x, y) \to (0, 0)$. ■

Solution. (5.10) (I'm writing this proof "properly" but remember that this is motivated by trying to use the "unfolding" trick from Example 5.5, and then seeing that it doesn't work.)

We claim that the limit does NOT exist.

Assume for contradiction that it did exist. Then the limit $\lim_{(x,y)\to(0,0)} \frac{x^2 + y^4}{x^2 + y^2}$ must exist, because we may write

$$\frac{x^2 + y^4}{x^2 + y^2} = \frac{x^2 + y^4}{\sin(x^2 + y^4)} \cdot \frac{\sin(x^2 + y^4)}{\log(1 + \sin(x^2 + y^4))} \cdot \frac{\log(1 + \sin(x^2 + y^4))}{x^2 + y^2}$$

and each of the three terms on the right side have limits as $(x, y) \to (0, 0)$ (the first two by "Standard Limits", the third by assumption).

But this is a contradiction: $\lim_{(x,y)\to(0,0)} \frac{x^2 + y^4}{x^2 + y^2}$ does *not* exist, and this can be seen by taking the path $y = 0$ (approaches 1) and taking the path $x = 0$ (approaches 0). Thus the original limit does not exist. ■

5.7 Some More Practice If You Want Idk Help Aaa

Exercise 5.11: Compute $\lim_{x\to 0} \dfrac{\sin\sin\sin\sin\sin\sin\sin x}{x}$.

Exercise 5.12: Find all $\alpha > 0$ for which

$$\lim_{(x,y)\to(0,0)} \frac{(x^2+y^2)^\alpha}{(|x|+|y|)^3}$$

exists.

6 Midterm Review!

Do **not** do all the problems! Put more focus on any areas you want to review more.

6.1 Warm-up

Find the last digit of $11^{10} - 10^{11}$.

Hint: vg'f artngvir yzsnb

6.2 True or False

1. The exam is tomorrow at 5 PM.
2. The exam will be held in Baker Hall A36.
3. The exam will have four problems.
4. I am going to do very well on the exam because I've been studying.
5. 0 is natural.
6. S has an upper bound, so $\sup S$ exists.
7. If $M := \sup S$ exists for some $S \subseteq \mathbb{R}$, then $M \in S$.
8. No open subset of $\mathbb{R}$ contains its supremum.
9. Every closed subset of $\mathbb{R}$ contains its supremum.
10. If there exist $\sup S$ and $\inf T$, then $\sup S - \inf T = \sup\{x - y : x \in S, y \in T\}$
11. If U is open and C is closed then $U \setminus C$ is open.
12. In a normed space, any closed and bounded set is compact.
13. In $\mathbb{R}^N$, the finite union of compact sets is compact.
14. The finite union of compact sets is compact.
15. Fix $R > 0$. In a metric space, any $E \subseteq B(0, R)$ with an infinite number of points has non-empty accumulation.
16. $\operatorname{acc} E \subseteq E$
17. $\overline{E} = \operatorname{acc} E \cup E$

18. $\partial E = \overline{E} \setminus E^\circ$
19. $\partial E = \overline{E^c} \cap \overline{E}$
20. ∂E must be closed.
21. (X, d_X), (Y, d_Y) metric spaces, $E \subset X$, and $f : E \to Y$. If $x_0 \in X \setminus E$ then $\lim_{x \to x_0} f(x)$ can *never* be defined.
22. $E = (-42, \pi) \cup \{1337\}$, $f : E \to \mathbb{R}$ with $f(x) = \begin{cases} 1, & -42 < x < \pi \\ 2, & x = 1337 \end{cases}$. Then
$$\lim_{x \to 1337} f(x) = 2.$$
23. (X, d_X), (Y, d_Y) metric spaces, $E \subset X$, $x_0 \in \operatorname{acc} E$, and $f : E \to Y$. If there exists $\lim_{x \to x_0} f(x)$, then this limit is *unique*.
24. $f, g, h : E \to \mathbb{R}$, $f(x) = g(x)h(x)$ for all $x \in E$, and $x_0 \in \operatorname{acc} E$. Then
$$\lim_{x \to x_0} f(x) = \left(\lim_{x \to x_0} g(x)\right)\left(\lim_{x \to x_0} h(x)\right).$$
25. I will get plenty of sleep the night before the exam, because it is a very good idea and Thomas said so.

6.3 Geometry

Consider the metric space $\mathbb{R}^2$ under the Euclidean metric. Let $E = B((0,0),1)$ and $F = B((4,0),2)$. Define:
$$S := \{d(x,y) : x \in E, y \in F\}$$

1. Does there exist $\sup S$ and $\inf S$?
2. Compute, with proof, $\inf S$ and $\sup S$.
3. Are either the inf or sup obtained?

6.4 Topologist's Sine Curve

Let:

$$E := \left\{(x,y) \in \mathbb{R}^2 : x \neq 0, y = \sin\left(\frac{1}{x}\right)\right\} \cup \{(0,0)\}$$

1. Is E open?
2. Is E closed?
3. Compute E°.
4. Compute $\operatorname{acc} E$.
5. Deduce what $\overline{E}$ and ∂E are.

6.5 Random Limits

1. Compute, with proof, $\displaystyle\lim_{x\to 0} \frac{1x + 2\sin x}{3x + 4\sin x}$.
2. Compute, with proof, $\displaystyle\lim_{x\to +\infty} \frac{1x + 2\sin x}{3x + 4\sin x}$.
3. Compute $\displaystyle\lim_{x\to 0} \frac{\sin\sin\sin\sin\sin\sin\sin x}{x}$.
4. Find all $\alpha > 0$ for which $\displaystyle\lim_{x\to 0} \frac{1-\cos(\sin x^2)}{x^\alpha}$ exists in $\mathbb{R}$. Is there an α for which the limit is non-zero?
5. Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{|x| + |y|}$.
6. Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^2 + y^8}$.
7. Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x\log(1+x)(1-\cos y)}{x^4 + y^4}$
8. Find all $\alpha > 0$ for which $\displaystyle\lim_{(x,y)\to(0,0)} \frac{(x^2+y^2)^\alpha}{(|x|+|y|)^3}$ exists in $\mathbb{R}$.
9. Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x+y}{e^x - e^{-y}}$.
10. Solve 2021 Putnam A2 if you haven't already.

6.6 Theorem Roulette

I play this game with my classmates before every exam. There are several ways to formulate it. To start, let X be the set of all possible theorems and lemmas that Leoni could test for Problem 1.

- (Measure Theory / Probability Method) *Bet a probability measure* μ on the space X. Your score is $\mu(\{\text{Theorem Tested}\})$.
- (The Eternal Cummings Method) Bet a *formal linear combination* $\sum_{i=1}^{n} \lambda_i x_i \in \mathbb{R}X$ with non-negative coefficients summing to 1. That is, $x_i \in X$ and $\lambda_i \in \mathbb{R}$, $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$. Here, the "multiplication" $\lambda_i x_i$ doesn't mean anything, it's basically just attaching a coefficient onto an object. That's why it's called a "formal linear combination". Anyways, your score is the coefficient of the x_i that appears on the exam, or if you *really* want, "$(\sum_i \lambda_i x_i)$(Theorem Tested) where we "view" $\sum_i \lambda_i x_i$ as a map $X \to \mathbb{R}$ in the "natural way"", if you're some kind of sucker for abstraction. I know this sounds like complete abstract nonsense but it does have the practical benefit of being able to communicate bets so much more easily. For instance, I bet .3(Radon Nikodym)+.3(Riesz Representation)+.2(Extension of measures)+.2(Fubini) for the 21-720 final last semester, and it's crystal clear what this bet represents in relatively few extra characters.
- (The Boring Sane Method) Bet $f \in \mathbb{R}^X$ satisfying $f \geq 0$ and $\sum_{x \in X} f(x) = 1$, your score being f(Theorem Tested).

Anyways, your exercise is to play some Theorem Roulette.

JUSTIFY ALL YOUR ANSWERS

6.7 (Basically) All Warm-up Solutions

Burning Ropes: Yvtug obgu raqf bs gur bar zvahgr ebcr, NAQ bar raq bs gur gjb zvahgr ebcr, nyy ng gur fnzr gvzr. Gur bar zvahgr ebcr jvyy arprffnevyl ohea hc va guvegl frpbaqf. Ol guvf gvzr, gur gjb zvahgr ebcr unf avargl frpbaqf erznvavat. Ol yvtugvat gur frpbaq raq bs guvf ebcr, vg jvyy ohea hc va sbegl svir frpbaqf. Gbgny gvzr ryncfrq vf guvegl cyhf sbegl svir.

Hanging a Painting: Ynl qbja pvepyr bs ebcr ba gur sybbe. Gnxr n ybat zrgny cbyr naq ynl vg ba gbc, ubevmbagnyyl. Gnxr n frpbaq zrgny cbyr naq cynpr vg iregvpnyyl ba gbc bs gur svefg zrgny cbyr gb znxr n cyhf fvta, naq jrnir vg haqre gur pvepyr bs ebcr. Fb guvf cbyr fubhyq tb haqre gur ebcr, bire gur svefg cbyr, naq haqre gur ebcr ntnva. Gurfr cbyrf ner lbhe cvaf, naq gur ebcr vf lbhe cnvagvat. Gel vg lbhefrys: Pebff lbhe nezf gb znxr na K funcr, naq jrnir n chefr guebhtu lbhe nezf va gur jnl V unir qrfpevorq nobir. Bapr qbar, hapebff lbhe nezf.

Apples for Sale: Ernfba onpxjneqf. Gur terra nccyr zhfg or gjb qbyynef, fvapr vg vf vzcyvrq gung gurer rkvfgf n havdhr nccyr purncre guna vg. Guvf vf gur bar qbyyne nccyr, naq vg zhfg or frpbaq fznyyrfg orpnhfr gurer vf n havdhr nccyr fznyyre guna vg. Gur fznyyrfg nccyr, va ghea, zhfg or gur frpbaq zbfg rkcrafvir, naq gur zbfg rkcrafvir nccyr zhfg or gur frpbaq ynetrfg. Gur nccyr jr jnag, juvpu vf erq, vf gur ynetrfg nccyr. Ol rkunhfgvba bs nyy cbffvoyr cevprf, gur nafjre vf rvgure gjb be guerr qbyynef, naq vg pnaabg or gjb qbyynef orpnhfr gung vf gur cevpr bs gur terra nccyr! Gur nafjre vf guerr.

One Word Wordsearch: See https://puzzling.stackexchange.com/questions/50832/mystery-of-the-one-word-wordsearch.

6.8 True or False Solutions

1. True
2. True
3. True
4. This better be true >:c
5. False, $\mathbb{N}$ starts at 1.
6. False, consider $S = \{\}$. EDIT: Apparently you guys did not define "upper bound" for empty sets? If we're going with that definition then this is true, but this distinction isn't important. Point is, don't take the sup of sets that might be empty!
7. False, consider $S = [0, 1)$

8. True. Suppose $M = \sup U \in U$ for U open. Then since U open, there exists $r > 0$ such that $B(M, r) \subseteq U$. But now $M < M + r/2 \in U$, contradicting that M is an upper bound.

9. False. Consider $\mathbb{R}$. (What if we changed it to compact?) EDIT: Some people have a good point, it seems implied that a supremum exists in the first place. If the supremum exists in $\mathbb{R}$ then the answer would be true.

10. True. Use $\sup(-S) = -\inf S$ and apply $\sup S + \sup T = \sup(S + T)$.

11. True. $U \setminus C = U \cap (C^c)$, an intersection of open sets, which is open.

12. False. Heine-Borel holds specifically in $\mathbb{R}^N$ (or *finite* dimensional vector spaces). This statement happens to always break in infinite dimensions. I won't prove it.

13. True. Use Heine-Borel for an easy argument: Finite union of closed sets is closed, finite union of bounded sets is easily bounded.

14. True, still. Suppose $K_1, \cdots, K_n$ compact and let $\bigcup_{\alpha\in\Lambda} U_\alpha \supseteq \bigcup_{i=1}^n K_i$. Then $\bigcup_{\alpha\in\Lambda} U_\alpha \supseteq K_i$ for each i. Thus for each i we may find a finite subcollection of indices $A_i \subseteq \Lambda$ for which $\bigcup_{\alpha\in A_i} U_\alpha \supseteq K_i$. Now $B = \bigcup_{i=1}^n A_i$ is also a finite collection of indices with $\bigcup_{\alpha\in B} U_\alpha \subseteq \bigcup_{i=1}^n K_i$.

15. False. Bolzano-Weierstrass holds specifically in $\mathbb{R}^N$ endowed with the Euclidean metric (or anything "equivalent"). Some counterexamples: $\mathbb{R}$ endowed with the *discrete* metric and taking $\mathbb{N} \subseteq B(0, 2)$. The vector space of infinite sequences of real numbers that are eventually constant at 0 with the metric $d(x, y) = \sum |x_i - y_i|$ and taking $(1, 0, 0, \cdots), (0, 1, 0, \cdots), (0, 0, 1, \cdots) \in B(0, 2)$.

16. False. Accumulations points of a set do not need to be in the set.

17. True

18. True

19. True

20. True (use any of the two previous statements!)

21. False. Limits can be defined at accumulation points, even if the accumulation points is not in the domain.

22. False. The limit is not defined because 1337 is not an accumulation point of E.

23. True (why?)

24. False. Consider $x_0 = 1$, $f(x) = x^2$, $g(x) = \begin{cases} \pi, & x \in \mathbb{Q} \\ 42, & x \notin \mathbb{Q} \end{cases}$, $h(x) = x^2/g(x)$. The issue is that for this to be true, you need to assert that the limits $\lim_{x \to x_0} g(x)$ and $\lim_{x \to x_0} h(x)$ exist in the first place.

25. Go to sleep.

6.9 Geometry Solutions

1. Yes. S is non-empty, and a sort of triangle inequality, like I will soon show, gives an upper and lower bound.

2. We claim that $\sup S = 7$ and $\inf S = 1$.

   For ease, let $P = (0, 0)$ and $Q = (4, 0)$. To see that 7 is an upper bound, write:

$$d(x, y) \leq d(x, P) + d(P, Q) + d(Q, y) \leq 1 + 4 + 2 = 7$$

   To see that it is the least upper bound, fix $\varepsilon > 0$. We may assume that $\varepsilon < 1$. Choosing $x = (-1 + \varepsilon/4, 0)$ and $y = (6 - \varepsilon/4, 0)$, we see that $d(x, y) = 7 - \varepsilon/2$, so:

$$7 - \varepsilon < d(x, y) \leq 7$$

   Thus $(7 - \varepsilon, 7] \cap S$ is non-empty for all $\varepsilon > 0$ (or equivalently, all $0 < \varepsilon < 1$, which was what we did), so $\sup S = 7$.

   To see that 1 is a lower bound, write instead:

$$d(P, Q) \leq d(P, x) + d(x, y) + d(y, Q)$$

   And I'll let you work out the details. It's basically isomorphic to the above work.

3. No. You can show this by strengthening some of the above inequalities to be strict by definition of a ball.

6.10 Topologist's Sine Curve Solutions

1. E is not open. Take pretty much any point (x, y) on the curve. Then for contradiction, if $B((x, y), r) \subseteq E$ then surely $(x, y + r/2) \in E$, which is impossible since E is the graph of a *function*.

2. E is not closed. Take $(0, 1) \notin E$. Suppose $B((0, 1), r) \subseteq E^c$. Note that $(x, 1) \in E$ for all x with $x = (2\pi + k\pi/2)^{-1}$, $k \in \mathbb{N}$. Using the Archimedean Property you can choose k large enough so that $x < r$, so that $(x, 1) \in E$ and $(x, 1) \in B((0, 1), r)$, contradiction.

3. $E^\circ = \emptyset$. This follows from the logic in (1), which implies that no point is an interior point.

4. $\operatorname{acc} E = E \cup (\{0\} \times [-1,1])$. Use some kind of continuity argument to argue that every point of E is an accumulation point. Then everything in $(\{0\} \times [-1,1])$ is an accumulation point by an argument akin to that used in (2). Nothing else is an accumulation point, and you can argue this by casing on whether you're on the y-axis or not.

5. Use $\overline{E} = E \cup \operatorname{acc} E$ to deduce that $\overline{E} = E \cup (\{0\} \times [-1,1])$. Use $\partial E = \overline{E} \setminus E^\circ$ to deduce that $\partial E = E \cup (\{0\} \times [-1,1])$.

6.11 Random Limits Solutions

1. L'Hopital is chill to use here, and this gives 1/3. Alternatively, trying dividing numerator and denominator by x.

2. L'Hopital is NOT chill to use here! Easiest way is to divide numerator and denominator by x. Then numerator tends to 1 and denominator tends to 3, giving 1/3.

3. Let's just pretend there are three sines for sake of my fingers. Then you can write
$$\frac{\sin\sin\sin x}{x} = \frac{\sin\sin\sin x}{\sin\sin x} \cdot \frac{\sin\sin x}{\sin x} \cdot \frac{\sin x}{x}$$
and apply the standard limits on each term to get 1.

4. Claim $\alpha \leq 4$. Write:
$$\frac{1-\cos(\sin x^2)}{x^\alpha} = \frac{1-\cos(\sin x^2)}{(\sin x^2)^2} \cdot \frac{(\sin x^2)^2}{(x^2)^2} \cdot \frac{x^4}{x^\alpha}$$
By standard limits it suffices to ensure limit of $\frac{x^4}{x^\alpha}$ exists.
Limit $= \frac{1}{2} > 0$ when $\alpha = 4$.

5. Write:
$$0 \leq \frac{|x^2-y^2|}{|x|+|y|} = |x-y| \cdot \frac{|x+y|}{|x|+|y|} \leq |x-y| \cdot \frac{|x|+|y|}{|x|+|y|} = |x-y| \leq |x|+|y| \to 0$$
So limit is 0.

6. Doesn't exist. Go along $x = y^4$.

7. Write:
$$\frac{x\log(1+x)(1-\cos y)}{x^4+y^4} = \frac{\log(1+x)(1-\cos y)}{x \cdot y^2} \cdot \frac{x^2 \cdot y^2}{x^4+y^4}$$
From here you can conclude that the limit does not exist by showing that the limit of $\frac{x^2 \cdot y^2}{x^4+y^4}$ does not exist.

8. I claim $\alpha > 3/2$. You won't get anything this hard on the exam, probably. I mean if you do then everyone fails which is ok. [TBD]

9. uh its 3 am and i need to do my probability homework and grade your psets

10. apply the Google Theorem

7 Series

7.1 Warm-up

(Stanford Math Tournament) Compute:

$$\sum_{n=0}^{\infty} \frac{\left(\frac{-2}{5}\right)^{\lfloor \sqrt{n} \rfloor}}{\sqrt{n}+\sqrt{n+1}}$$

Hint 1: Guvf vf n ceboyrz gung V guvax qrzbafgengrf jryy gur neg bs zngurzngvpny ceboyrz fbyivat. Vg ybbxf fpnel, ohg vg pna or fbyirq vs lbh gnpxyr rnpu fpnel cneg vaqvivqhnyyl. Fgneg ol erzbivat gur qrabzvangbe ol zhygvcylvat ol n pregnva enqvpny pbawhtngr.

Hint 2: Arkg, fgneg erneenatvat gur grezf n ovg. Gur tbny abj vf gb hfr gur njshy rkcbarag gb lbhe nqinagntr. Pregnva grezf fubhyq pnapry bhg.

7.2 Bunch of Convergence Tests

Theorem 7.1 (Stupid Test)

If $\lim_{n\to\infty} |a_n| \neq 0$ then $\sum_{n=1}^{\infty} a_n$ does not converge.

Theorem 7.2 (p-Test et. al.)

$\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges for $p > 1$ and diverges for $p \leq 1$.

$\sum_{n=1}^{\infty} a^n$ converges for $|a| < 1$ and diverges for $|a| \geq 1$.

Theorem 7.3 (Direct Comparison Test)

Let $a_n \geq 0$ be a sequence. If you can find a b_n that *eventually dominates* a_n (i.e. there is N such that $b_n \geq a_n$ for all $n \geq N$), such that $\sum_{n=1}^{\infty} b_n < \infty$, then $\sum_{n=1}^{\infty} a_n < \infty$ i.e. converges.
Similarly, if instead you found a b_n for which eventually $0 \leq b_n \leq a_n$ forever, with $\sum_{n=1}^{\infty} b_n = +\infty$, then $\sum_{n=1}^{\infty} a_n = +\infty$ i.e. diverges.

We're going to be invoking this implicitly quite a lot.

Also this is a thing:

Theorem 7.4 (Limit Comparison Test)

If two sequences are close together, they behave the same way. That is, if $a_n \geq 0$ and $b > 0$ are two sequences for which there exists the limit $\lim_{n\to\infty} \frac{a_n}{b_n} = L \in (0, \infty)$, then $\sum_{n=1}^{\infty} a_n$ converges iff $\sum_{n=1}^{\infty} b_n$ converges (so if either converges then the other converges, and if either diverges then the other diverges!).

Next are some tests inspired by "geometric series".

Theorem 7.5 (Ratio Test)

Consider the limit of the ratio between successive terms, $L = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$, if it exists.

- If $L < 1$, then $\sum_{n=1}^{\infty} a_n$ converges.
- If $L > 1$, then $\sum_{n=1}^{\infty} a_n$ diverges.
- If $L = 1$, you know nothing.

Remark: The hyper-analysis-savvy reader would be delighted to know that there is a version of this theorem that can be used if the limit does not exist, by using instead the $\limsup$ and $\liminf$.

Theorem 7.6 (Root Test)

Consider $L = \lim_{n\to\infty} \sqrt[n]{|a_n|}$, if it exists.

- If $L < 1$, then $\sum_{n=1}^{\infty} a_n$ converges.
- If $L > 1$, then $\sum_{n=1}^{\infty} a_n$ diverges.
- If $L = 1$, you know nothing. *

Remark 1: (*) I'm lying slightly for the case $L = 1$ for simplicity. If you know that $\sqrt[n]{a_n}$ converges to L *strictly from above*, then you can conclude divergence.

Remark 2: The hyper-analysis-savvy reader would be delighted to know that you can replace lim with lim sup.

Remark 3: You can use the root test to prove the ratio test. I leave this as an exercise if you haven't done it.

For dealing with series that have positive and negative terms:

Theorem 7.7 (Alternating Series Test)

If $a_n \geq 0$ is **monotone decreasing** and **tends to** 0, then $\sum_{n=1}^{\infty}(-1)^n a_n$ converges.

Theorem 7.8 (Absolute Convergence)

If $\sum_{n=1}^{\infty} |a_n| < \infty$ then $\sum_{n=1}^{\infty} a_n$ converges.

7.3 Examples Using Tests

Example 7.1: $\sum_{n=1}^{\infty} \log\left(1 + \frac{1}{n^3}\right)$

Solution. We use the fancy inequality $\log(1+x) \le x$ (comes from $1 + x \le e^x$) to get that this converges. ∎

Example 7.2: $\sum_{n=1}^{\infty} \frac{1}{(n!)^{1/n}}$

Solution. We use the very fancy inequality (see appendix) $n! \le (n+1)^{n+1}e^{-n}$ to go down:

$$\sum_{n=1}^{\infty} \frac{1}{(n!)^{1/n}} \ge \sum_{n=1}^{\infty} \frac{1}{(n+1)^{\frac{n+1}{n}} e}$$

We claim this new series diverges. We do this by limit comparison with the series $\sum_{n=1}^{\infty} \frac{1}{n+1}$, which diverges. For the comparison to work, it suffices to prove that:

$$\lim_{n\to\infty} \frac{n+1}{(n+1)^{\frac{n+1}{n}}} = L \in (0, \infty)$$

A sketch: Show that $(n+1)^{1/n}$ is decreasing and ≥ 1, so the limit of this exists. ∎

Example 7.3: $\sum_{n=1}^{\infty} \frac{1}{2^{\sqrt{n}}}$

Solution. We claim convergence. It suffices to show that $\frac{1}{2^{\sqrt{n}}} \le \frac{1}{n^2}$ for all large enough n. This rearranges to $n^2 \le 2^{\sqrt{n}}$ or $2\log n \le \sqrt{n}\log 2$. Since $\sqrt{n}$ grows faster than $\log n$ (why?), the claim is true. ∎

Solution. (Alternate) You can try writing $\sum_{n=1}^{\infty} \frac{1}{2^{\sqrt{n}}} \le \sum_{n=1}^{\infty} \frac{1}{2^{\lfloor \sqrt{n} \rfloor}}$. Then this new series is exactly equal to like $\sum_{k=1}^{\infty} \frac{2k+1}{2^k}$, and this converges by your favorite test. ∎

Example 7.4: Find all x for which $\sum_{n=1}^{\infty} \frac{x^n}{2+x^n}$ converges.

Solution. Resolve $|x| \ge 1$ using the stupid test. For $|x| < 1$, note that:

$$\sum_{n=1}^{\infty} \frac{|x^n|}{|2+x^n|} \le \sum_{n=1}^{\infty} |x|^n < \infty$$

■

Example 7.5: Find all x for which $\sum_{n=1}^{\infty} \frac{n! x^n}{n^n}$ converges.

Solution. For $|x| < e$, apply ratio test:

$$\frac{(n+1)!|x|^{n+1}/(n+1)^{n+1}}{n!|x|^n/n^n} = \frac{(n+1)|x|n^n}{(n+1)^{n+1}} = |x|\left(\frac{n}{n+1}\right)^n = \frac{|x|}{(1+1/n)^n} \to \frac{|x|}{e} < 1$$

So we have convergence for all such x. For $|x| \geq e$, note that:

$$\frac{(n+1)!|x|^{n+1}/(n+1)^{n+1}}{n!|x|^n/n^n} = \frac{|x|}{(1+1/n)^n} \geq 1$$

So the sequence is actually non-decreasing (in fact it is strictly increasing)! Since the first term is > 0, the terms don't go to 0, so we apply the stupid test. ■

Example 7.6: Find all x for which $\sum_{n=1}^{\infty} \frac{\log(1+2^n)}{n^2 + x^{2n}}$ converges.

Solution. Claim $|x| < 1$. Essentially use the bounds $n \log 2 \leq \log(1+2^n) \leq 1 + n\log 2$. ■

7.4 Bonus Round

We didn't get to this in recitation.

Theorem 7.9 (Integral Test)

In the series $\sum_{n=1}^{\infty} f(n)$, You can replace the sum with an integral and nothing changes convergence-wise (as long as f is non-negative and monotone decreasing).

Theorem 7.10 (Cauchy Condensation)

For a decreasing and non-negative sequence a_n, we have that $\sum_{n=1}^{\infty} a_n$ converges iff $\sum_{n=1}^{\infty} 2^n a_{2^n}$ converges!

Example 7.7: $\sum_{n=1}^{\infty} \frac{1}{n \log^2 n}$

Solution. (Integral Test) u-sub $u = \log x$. ■

Solution. (Condensation) $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{2^n \log^2(2^n)} = \sum_{n=1}^{\infty} \frac{1}{n^2 \log^2(2)} < \infty$ ■

7.5 Appendix: Hella Lit Inequalities To Keep In Mind

Lemma 7.1

$1 + x \leq e^x$ for ALL $x \in \mathbb{R}$.

Proof. This is essentially a statement of how the tangent line to e^x at $x = 0$ lies under the curve forever. So, show that e^x is strictly convex and that should do the trick.

Alternatively, show that the minimum value of $e^x - x$ is 1 by using some calculus. □

Lemma 7.2

$\log(1 + x) \leq x$ for $x > -1$.

Proof. log both sides of the previous inequality :p □

Who knows if the next one is useful :shrug: I'll include it in case Leoni puts a factorial somewhere, and the odds of that are kinda low but eh whatever it's interesting math why not include it here yay learning

Lemma 7.3

$$n^n e^{-n} \leq n! \leq (n+1)^{n+1} e^{-n}$$

Proof. So uh you do this by integration. First note that:

$$\int_{k-1}^{k} \log x \, dx \leq \log k \leq \int_{k}^{k+1} \log x \, dx$$

We essentially got that by using the fact that log is increasing (so like, $\int_a^b \log a \, dx \leq \int_a^b \log x \, dx \leq \int_a^b \log b \, dx$). Now sum each side from $k = 2$ to n.

$$\int_{1}^{n} \log x \, dx \leq \log(n!) \leq \int_{2}^{n+1} \log x \, dx$$

In fact I'm just going to increase the bounds on the right side to make the integral cleaner:

$$\int_{1}^{n} \log x \, dx \leq \log(n!) \leq \int_{1}^{n+1} \log x \, dx$$

Now evaluate!

$$(x \log x - x)\big|_{x=1}^{n} \leq \log(n!) \leq (x \log x - x)\big|_{x=1}^{n+1}$$

$$n \log n - n + 1 \leq \log(n!) \leq (n+1)\log(n+1) - n$$

Yeet the log:

$$e^{n \log n - n + 1} \leq n! \leq e^{(n+1)\log(n+1) - n}$$

$$n^n e^{-n+1} \leq n! \leq (n+1)^{n+1} e^{-n}$$

This proves the lemma, and in fact we have an extra factor of e if you want it. $\square$

8 Pointwise & Uniform Convergence, Liminf & Limsup

8.1 Warm-up

Must every totally-ordered subset of $P(\mathbb{N})$ be countable?

$P(\mathbb{N})$ is the power set of $\mathbb{N}$. A totally-ordered set of sets $\mathcal{F}$ is such that for any $S, T \in \mathcal{F}$, we either have $S \subseteq T$ or $S \supseteq T$. For example, $\{\{1\}, \{1, 2\}, \{1, 2, 3, 4\}, \{1, 2, \cdots, 7, 8\}, \cdots\}$ is totally-ordered.

(Hint 1: Nafjre vf ab.)

(Hint 2: Vafgrnq bs angheny ahzoref, pbafvqre fbzr bgure pbhagnoyr frg gb znxr guvatf rnfvre.)

(Hint 3: Pbafgehpgvba bs gur ernyf.)

(Hint 4: Tbbtyr "Qrqrxvaq Phgf".)

8.2 Pointwise Convergence

A sequence of functions f_n tends to f *pointwise* if, point by point, we have $f_n(x) \to f(x)$.

Definition 8.1 (Pointwise Convergence)

Let $f_n, f : E \to \mathbb{R}$. Then $f_n \to f$ *pointwise* if for every $x \in E$ we have:

$$\lim_{n\to\infty} f_n(x) = f(x)$$

Example 8.1: The sequence of functions $f_n(x) := \sin(x) + \frac{e^{-x^2}}{n}$ converges pointwise to $\sin(x)$. This is because if we *fix* any $x \in \mathbb{R}$, then $\frac{e^{-x^2}}{n} \to 0$.

Example 8.2: The sequence of functions $f_n(x) := x^n$ on $[0, 1]$ converges pointwise to:

$$f(x) := \begin{cases} 0, & 0 \le x < 1 \\ 1, & x = 1 \end{cases}$$

This is because $1^n = 1$ for all n, whereas if $0 \le x < 1$ then x^n exponentially decays toward 0 as n grows big.

8.3 Uniform Convergence

A sequence of functions f_n tends to f *uniformly* if, roughly speaking, the "entirety of f_n" tends to f, all at once, at the same time.

Definition 8.2 (Uniform Convergence)

$f_n \to f$ *uniformly* if:

$$\lim_{n\to\infty} \sup_E |f_n - f| = 0$$

An alternate way to view it that demonstrates the power of uniform convergence is this: $f_n \to f$ *uniformly* if for every $\varepsilon > 0$ you can find N_ε such that $|f_n(x) - f(x)| < \varepsilon$ for all $n \geq N_\varepsilon$. Oh whoops I forgot to specify for which x that holds. But that's the thing! I don't need to; **when $n \geq N_\varepsilon$ we get $|f_n(x) - f(x)|$ for ALL x.** For big enough n, EVERYONE is close to f. ALL POINTS are getting close to f at the *same time*.

Example 8.3: The sequence of functions $f_n(x) := \sin(x) + \frac{e^{-x^2}}{n}$ converges uniformly to $f(x) = \sin(x)$. This is because:

$$|f_n(x) - f(x)| = \frac{e^{-x^2}}{n} \leq \frac{1}{n} \to 0$$

Notice how I didn't care about what x is. Heuristically, I'm not allowed to, I have to prove that this difference $|f_n(x) - f(x)|$ is upper-bounded by something in terms of ONLY n, such that this upper bound goes to 0.

Explicitly, I've managed to argue that, since $|f_n(x) - f(x)| \leq \frac{1}{n}$ for ALL x, we have that $\sup_E |f_n - f| \leq \frac{1}{n}$.

Example 8.4: The sequence of functions $f_n(x) := x^n$ on $[0,1]$ converges pointwise to:

$$f(x) := \begin{cases} 0, & 0 \leq x < 1 \\ 1, & x = 1 \end{cases}$$

As we saw. **But it does NOT converge uniformly!** This is because $\sup_{[0,1]} |f_n - f| = 1$ always, no matter how big n is. There's always some stubborn point that refuses to get close to f.

Explicitly, one can e.g. choose $x = \frac{1}{\sqrt[n]{2}}$. Then $|f_n(x) - f(x)| = 1/2$, so $\sup_{[0,1]} |f_n - f| \geq \frac{1}{2}$ for all n. This shows that $\sup_{[0,1]} |f_n - f|$ can never go to 0.

See https://www.desmos.com/calculator/uxrkf0zmbj for a visualization of these examples.

8.4 Determining Pointwise and Uniform Convergence: Examples

Example 8.5: Consider $f_n : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ defined by $f_n(x) = \frac{\sin(nx)}{nx}$. Prove that f_n converges pointwise to a function f. Does $f_n \to f$ uniformly? Does $f_n \to f$ uniformly over $E = (-\infty, -0.1) \cup (0.1, \infty)$?

Solution. For any $x \neq 0$ we have that:

$$\frac{|\sin(nx)|}{|nx|} \leq \frac{1}{n|x|}$$

And this $\to 0$ as $n \to +\infty$. Thus the pointwise limit is $f(x) = 0$.

We do not have uniform convergence. This is because for any n, we can find x near 0 such that $\sin(nx)/(nx)$ is close to 1 (and therefore, above e.g. $1/2$, and thus not close to 0). Such an x can be found by using the standard limit $\lim_{x\to 0} \frac{\sin x}{x} = 1$.

However we do have uniform convergence over E, because for all $x \in E$:

$$\frac{|\sin(nx)|}{|nx|} \leq \frac{1}{n|x|} \leq \frac{1}{n \cdot 0.1}$$

This holds for ALL x, so in fact:

$$\sup_{x\in E} \frac{|\sin(nx)|}{|nx|} \leq \frac{1}{n \cdot 0.1}$$

And the RHS goes to 0. ■

Example 8.6: Consider $\sum_{k=0}^{\infty} x^2(1-x)^k$ over $E = [0,1]$. What is the pointwise limit? Do we have uniform convergence to the pointwise limit?

Solution. Remember that when we're discussing pointwise/uniform convergence of a *series* of functions, we're just talking about the pointwise/uniform convergence of the *partial sums*.

I'll leave it to you to check that the pointwise limit is x. For uniform convergence, note that the nth partial sum is given by:

$$\frac{x^2(1-(1-x)^{n+1})}{1-(1-x)} = x(1-(1-x)^{n-1})$$

And so the difference between this partial sum and the pointwise limit is:

$$|x - x(1-(1-x)^{n-1})| = x(1-x)^{n-1}$$

One can now compute the maximum of this over $[0,1]$ using Calculus and show that this maximum tends to 0 to conclude. Alternatively, use AM-GM! ■

8.5 liminf and limsup: WHY???

I'll tell you why we like them!

- Limits don't always exist. **limsup and liminf ALWAYS exist.** So you can "take limits of things" even if you don't know that a limit exists a priori.

> **Example 8.7:** $\lim_{x\to\infty} \sin(x)$ does not exist. However, $\limsup_{x\to+\infty} \sin(x) = 1$ and $\liminf_{x\to+\infty} \sin(x) = -1$.

- ...but how does that help find limits? Well basically **if the liminf equals the limsup, then the limit exists and is equal to both of them!** (Likewise if the limit exists then it is equal to the limsup and liminf)
- limsup and liminf have a bunch of cool properties that can help you.

8.6 liminf and limsup: WHEN???

I'll tell you when we use them!

- Many of the convergence tests we talked about can be improved (i.e. work for more cases) by sliding in a limsup or liminf. Or both.
- Complexity theory! We write $f(n) = O(g(n))$ iff $\limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty$.
- Measure theory! The liminf in particular will have its 15 minutes of fame while we're on the road towards the Lebesgue Dominated Convergence theorem.
- Calculus of Variations! Something something Gamma convergence! Something something research with Leoni!
- To reiterate, we can spam limsups and liminfs when we want to talk about a limit that we're not quite sure exists yet.

8.7 liminf and limsup: WHO???

I will now tell you *who* these creatures are. I will define them five different times. These definitions are all equivalent!

(I define lim sup but the definition for lim inf is analogous.)

Definition 8.3 (Very Layman)

The lim sup is a lim that's very biased to look at big values of f.

Definition 8.4 (Layman)

If you drape a curtain over the top of the graph of f, and the curtain gets closer and closer to L, then $\limsup_{x \to +\infty} f(x) = L$.

The next definition is the first "real" definition, and it's also the one we'll be using most often. Er, in fact, basically always. All the other definitions are there to just help your intuition.

Definition 8.5 (Decreasing Sups)

$$\limsup_{x \to +\infty} f(x) := \inf_{a \in \mathbb{R}} \sup_{x > a} f(x)$$

Note that the limsup is (at the top level) an *infimum*, not a supremum!

See https://www.desmos.com/calculator/tiwrgpoa0x for a visualization.

What the a controls in the definition is "how we we can look at" for the sup. As a gets bigger, we are allowed to look at less and less of the domain, and hence the sup gets smaller and smaller, approaching the limsup.

Compare with normal limits: lim really cares that *every* $f(x)$ *with* $x > a$ is controlled, both from above and from below. lim sup, on the other hand, only seems to care about some control from above.

Definition 8.6 (Best Bound on the Tail)

Let U be the set of all "eventual upper bounds" on f. That is, if $y \in U$, then that means that *eventually* we'll get $f(x) < y$ for all x large enough (rigorously, there exists A_y such that $f(x) < y$ for all $x \geq A_y$).
Think of these as "upper bounds" on the "tail" or "end behavior" of f.
The $\limsup_{x \to +\infty} f(x)$ is the "best" such upper bound, i.e. $\inf U$. (But note that $\inf U$ might not actually itself be an eventual upper bound!)

Compare with normal limits: The tail *is* the limit. If the tail behaves badly, then there is no limit. But surely you can upper bound the bad behavior, and that's where lim sup is born.

Definition 8.7 (Best Subsequence)

$$\limsup_{x\to+\infty} f(x) = \sup\left\{\lim_{n\to\infty} f(x_n) : x_n \to +\infty \text{ and } \lim_{n\to\infty} f(x_n) \text{ exists}\right\}$$

Intuitively, we're examining the "end behavior" of f by being picky and only choosing points x_n where it's big, and taking the limit.

Compare with normal limits: If the normal limit exists, then actually *all* subsequences $x_n \to +\infty$ will have $f(x_n)$ tending to this limit.

Exercise 8.8: Prove that all these definitions are equivalent.

Exercise 8.9: Prove that in the "Best Subsequence" definition, the sup is actually a max.

8.8 Definitions in other contexts

You should be able to guess these definitions, but here they are just in case.

Definition 8.8 (Limsup to a point)

$$\limsup_{x\to x_0} f(x) = \inf_{r>0} \sup_{0<|x-x_0|<r} f(x)$$

See https://www.desmos.com/calculator/va3cdionyv for a visualization.

As an exercise, reformulate the "alternative definitions" into this context.

Definition 8.9 (Limsup of a sequence)

$$\limsup_{n\to+\infty} a_n = \inf_{N\in\mathbb{N}} \sup_{n\geq N} a_n$$

Viewing sequences as functions $\mathbb{N} \to \mathbb{R}$, this isn't very surprising.

Also, try reformulating the "Best Subsequence" alternate definition into this context. Then you'll see why I'm calling it "Best *Subsequence*".

8.9 liminf and limsup: WHAT???

Now I will tell you the properties that make these nice. I'll be using the standard definition for $x \to x_0$ but they work for the other contexts.

- The lim sup is indeed the limit of a sup. That is, you can write:

$$\limsup_{x\to x_0} f(x) = \lim_{r\to 0^+} \sup_{0<|x-x_0|<r} f(x)$$

  This is because the function $r \mapsto \sup_{0<|x-x_0|<r} f(x)$ is *decreasing as* r *decreases*, so its infimum coincides with its limit.

- $\limsup_{x\to x_0} f(x) = \liminf_{x\to x_0} f(x)$ if and only if $\lim_{x\to x_0} f(x)$ exists. And in this case all of the "three limits" are equal.

- We always have $\liminf_{x\to x_0} f(x) \le \limsup_{x\to x_0} f(x)$

- If $f(x) \le g(x)$ then $\limsup_{x\to x_0} f(x) \le \limsup_{x\to x_0} g(x)$ and $\liminf_{x\to x_0} f(x) \le \liminf_{x\to x_0} g(x)$

- We have $\limsup_{x\to x_0} -f(x) = -\liminf_{x\to x_0} f(x)$. That is, if you pull a negative out then you must flip from sup to inf and vice versa.

- (Subadditivity) We have:

$$\limsup_{x\to x_0}(f(x)+g(x)) \le \limsup_{x\to x_0} f(x) + \limsup_{x\to x_0} g(x)$$

  And:

$$\liminf_{x\to x_0}(f(x)+g(x)) \ge \liminf_{x\to x_0} f(x) + \liminf_{x\to x_0} g(x)$$

8.10 liminf and limsup: HOW???

Ok here's a stupid example to show how one can use limsup and liminf to be happy.

Example 8.10: Prove the squeeze rule lol. That is, if $g_1(x) \le f(x) \le g_2(x)$, and both $\lim_{x\to x_0} g_1(x)$ and $\lim_{x\to x_0} g_2(x)$ exist and are equal to L, then $\lim_{x\to x_0} f(x) = L$.

Proof. [WRONG PROOF] Just take $g_1(x) \le f(x) \le g_2(x)$ and take the limit of all three parts, to get $L \le \lim_{x\to x_0} f(x) \le L$, so $\lim_{x\to x_0} f(x) = L$. Tada? □

This is **very very very very very very very very wrong** because I don't actually know that $\lim_{x\to x_0} f(x)$ exists in the first place! So this is very bad and horrible and terrible.

...

But I *do* know that $\limsup_{x\to x_0} f(x)$ and $\liminf_{x\to x_0} f(x)$ exist. Because they *always* exist.

Proof. [Actual Proof] Taking limsup on both sides of the right inequality and liminf on both side of the left inequality, we get the following for free:

$$\liminf_{x\to x_0} g_1(x) \le \liminf_{x\to x_0} f(x) \le \limsup_{x\to x_0} f(x) \le \limsup_{x\to x_0} g_2(x)$$

Ok but, $\lim_{x\to x_0} g_1(x)$ exists, so $\liminf_{x\to x_0} g_1(x) = \lim_{x\to x_0} g_1(x) = L$... and similarly, we know that $\limsup_{x\to x_0} g_2(x) = \lim_{x\to x_0} g_2(x) = L$. So actually this is just saying that:

$$L \le \liminf_{x\to x_0} f(x) \le \limsup_{x\to x_0} f(x) \le L$$

So the liminf and limsup of f were equal, and in fact both are equal to L, so the limit exists and is L. Uwu. □

9 MVT and L'Hôpital

9.1 Warm-up

I have a wooden cube. Drill a hole through it such that a larger cube can be passed through the hole.

(Clarification because I know at least one person is gonna nitpick: "Drill a hole" means take the set difference between the cube and some infinitely long prism with some a base of some shape, such that the resulting set is non-empty and homeomorphic to a torus.)

(Hint 1: Jybt gur fvqr yratgu vf 1. Svefg pbafvqre ebgngvat gur phor fb gung vgf cebwrpgvba hagb gur kl cynar vf n erthyne urkntba (...bs jung fvqr yratgu?). Hfvat "jvfushy guvaxvat" jr ubcr gung gur ynetrfg fdhner vafpevonoyr vagb guvf urkntba vf ynetre guna 1, naq vs fb gura jr pna qevyy gung ubyr. Guvf vf abj n pynffvpny trbzrgel ceboyrz.)

(Hint 2: Gb svaq gur fvqr yratgu bs gur fdhner gung lbh ner vafpevovat, hfr fvzvyne gevnatyrf (cnegvphyneyl gur guvegl fvkgl avargl gevnatyrf) naq nytroen. Qb abg or nsenvq gb qenj fbzr nhkvyynel yvarf.)

9.2 Does differentiable imply that the derivative is continuous?

Nonononononononononononononononononoononononononononononononono!! It does not!!

I know it's hard to imagine how that's possible but here's an example.

Let:

$$f(x) = \begin{cases} x^2 \sin(1/x), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Graphing it in Desmos... it sure does look troublesome!

CLAIM: f is differentiable everywhere.

It is not hard to show that x^2 and $\sin(1/x)$ are differentiable at all $x \neq 0$, so the product is also differentiable at all such points by the product rule. So what happens at $x = 0$? We claim that in fact $f'(0)$ exists and $= 0$.

We just need to show that:

$$\lim_{x \to 0} \frac{x^2 \sin(1/x) - 0}{x - 0} = 0$$

But it is, because this is just $\lim_{x\to 0} x\sin(1/x)$ and this tends to 0 (because sin is bounded).

CLAIM: f' is NOT continuous.

By the product rule, and what we computed above, we have:

$$f'(x) = \begin{cases} 2x\sin(1/x) - \cos(1/x), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

So if f' were continuous, we would have $\lim_{x\to 0} 2x\sin(1/x) - \cos(1/x) \overset{?}{=} 0$. But this limit does not exist! (Reason: If it did exist, then since $2x\sin(1/x) \to 0$, we must have $\lim_{x\to 0}\cos(1/x)$ existing. But it doesn't.) So no continuity. Sad!

9.3 Rolle's Theorem and MVT

Theorem 9.1 (Rolle)

Suppose f is continuous on $[a, b]$ and differentiable on (a, b) with $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.

The reason why we don't just say "differentiable on $[a, b]$" is because this is a stronger condition than what we need. That is, the theorem still works for functions that go *nyoom* near the endpoints, like $\sqrt{1-x^2}$ over $[-1, 1]$.

Proof. The idea is simple: Just use the fact that $f'(c) = 0$ where c is where the maximum is obtained. We just have to make sure $c \neq a, b$.

So let's start like this: If f is a constant function then clearly any $c \in (a, b)$ will do. So we may assume it's not constant.

$[a, b]$ is compact and f is continuous (by virtue of being differentiable), so f must obtain a maximum at some $c_1 \in [a, b]$ and a minimum at some $c_2 \in [a, b]$.

Can c_1 and c_2 both be endpoints of the interval (either $= a$ or $= b$)? No! This is because if they were, then $f(c_1) = f(c_2)$ from the assumption that $f(a) = f(b)$. But $f(c_1)$ is the maximum and $f(c_2)$ is the minimum, and since we're assuming that f is NOT constant, they cannot be equal, contradiction!

So one of c_1 or c_2 is in the *open* interval (a, b), from which we conclude by a theorem that the derivative of f at such a point is 0. □

A direct application of Rolle's it the Lagrange Mean Value Theorem.

Theorem 9.2 (Lagrange / MVT)

Suppose f is continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that:

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

Proof. Since this is just a "rotated" version of Rolle's, surely it can be proven by rotating f and then applying Rolle's... so let's just follow our nose!

To "rotate" f, let's like add a line or something. Let $g : [a, b] \to \mathbb{R}$ with:

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

Since $x \mapsto \frac{f(b)-f(a)}{b-a} \cdot x$ is differentiable, we have that g is differentiable. Next, note that:

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

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

So $g(a) = g(b)$. Now we may apply Rolle's!

We find $c \in (a, b)$ such that $g'(c) = 0$. In fact:

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

Or, $f'(c) = \dfrac{f(b) - f(a)}{b - a}$. Tada! □

Remark: The calculations are easier if you instead subtract $\frac{f(b)-f(a)}{b-a} \cdot (x - a)$.

9.4 Cauchy Mean Value Theorem

There is, in fact, *another* MVT! Here it is:

Theorem 9.3 (Cauchy MVT)

Suppose f and g are continuous on $[a, b]$ and differentiable on (a, b), with $g'(x) \neq 0$ for all $x \in (a, b)$. Then there exists $c \in (a, b)$ such that:

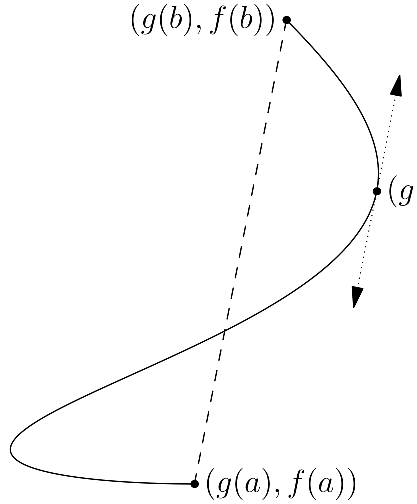
$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$$

Before we prove this, I need to describe what the hell this means.

Key Point 1: Cauchy MVT *generalizes* the Lagrange MVT.

To see this, take $g(x) := x$.

Key Point 2: Cauchy MVT is like a 2D version of MVT.



Essentially, if you plot the point $(g(t), f(t))$ as t runs from a to b, then this traces out a curve going from $(g(a), f(a))$ to $(g(b), f(b))$. The Cauchy MVT just states that at some point in time, the curve is running "parallel" to the line segment between those endpoints.

Now we turn to the proof. The motivation is just to work backwards. If we have $\dfrac{f(b) - f(a)}{g(b) - g(a)} = \dfrac{f'(c)}{g'(c)}$, then this rearranges to:

$$(f(b) - f(a))g'(c) - (g(b) - g(a))f'(c) = 0$$

So by "wishful thinking" we hope that we can apply Rolle's to $(f(b) - f(a))g(x) - (g(b) - g(a))f(x)$. Hm.

Proof. Define $h : [a, b] \to \mathbb{R}$ as:

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

Whoa, $h(a) = f(b)g(a) - g(b)f(a) = h(b)$. Moreover h is the linear combination of differentiable functions, so it is differentiable. Thus by Rolle's there is $c \in (a, b)$ such that:

$$(f(b) - f(a))g'(c) - (g(b) - g(a))f'(c) = 0$$

So:

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

It remains to prove that we can divide stuff.

Since $g'(c) \neq 0$ by assumption, we can divide each side by $g'(c)$. But can we divide each side by $g(b) - g(a)$?

Suppose for contradiction that $g(a) = g(b)$. Then by Rolle's again (!), there exists x such that $g'(x) = 0$, which contradicts the assumption. So in fact $g(a) \neq g(b)$, and we may indeed divide by $g(b) - g(a)$. In conclusion:

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

Voíla. ☐

Remark: Indeed, the condition $g'(x) \neq 0 \quad \forall x \in (a, b)$ is only needed at the very end. So if we were to state the Cauchy MVT as "there exists c such that $[f(b) - f(a)] \cdot g'(c) = [g(b) - g(a)] \cdot f'(c)$", then there is no need for that condition.

9.5 Le Théorème de L'Hôpital

The Cauchy MVT can prove L'Hopital's Rule!

Uh, I think I told several lies last time I stated L'Hopital, so we're going to do it correctly now.

Theorem 9.4 (L'Hopital)

Let $[a, b]$ be an interval, let $x_0 \in (a, b)$.
Suppose f and g are continuous in $[a, b]$ and differentiable in $(a, b) \setminus \{x_0\}$, with $g(x) \neq 0$ and $g'(x) \neq 0$ for all $x \in (a, b) \setminus \{x_0\}$. Assume moreover that $f(x_0) = g(x_0) = 0$. Then if the limit

$$L := \lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

exists in $\overline{\mathbb{R}}$ (we can have $L = \pm\infty$), then we may conclude that:

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = L$$

Proof. Consider some $x \in (a, b) \setminus \{x_0\}$. Apply the Cauchy MVT to the interval between x and x_0 to find c_x between x and x_0 such that:

$$\frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(c_x)}{g'(c_x)}$$

Wait but $f(x_0) = g(x_0) = 0$, so actually this just says:

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}$$

Now if we send $x \to x_0$, we have that c_x approaches x_0 (because it's in between x and x_0), so $f'(c_x)/g'(c_x)$ approaches the limit L. Thus $f(x)/g(x)$ approaches the limit L. Done. $\square$

Remark: The next example shows how you'd deal with functions that are "essentially continuous" but aren't technically defined at x_0, even though they *continuously extend* to x_0. In short, we can just continuously extend in that way.

Example 9.1: Compute:

$$\lim_{x\to 0}\left(\frac{\sin x}{x}\right)^{1/x}$$

Solution. Since log is continuous, let us instead find:

$$\lim_{x\to 0}\frac{\log(\sin(x)/x)}{x}$$

We *plan* to use L'Hopital. To wit, we study the limit:

$$\lim_{x\to 0}\frac{\frac{d}{dx}\log(\sin(x)/x)}{\frac{d}{dx}x} = \lim_{x\to 0}\frac{x}{\sin x}\left(\frac{\cos(x)x-\sin(x)}{x^2}\right) = \lim_{x\to 0}\frac{x\cos x-\sin x}{x\sin x}$$

This still is not good. Let us plan to use L'Hopital a second time. So we study the limit:

$$\lim_{x\to 0}\frac{\frac{d}{dx}x\cos x-\sin x}{\frac{d}{dx}x\sin x} = \lim_{x\to 0}\frac{\cos x - x\sin x-\cos x}{\sin x + x\cos x} = \lim_{x\to 0}\frac{-x\sin x}{\sin x + x\cos x} = \lim_{x\to 0}\frac{-\sin x}{\frac{\sin x}{x}+\cos x}$$

Aha! THIS limit exists and $= 0$. That's because the top goes to 0 whereas the bottom goes to 2 (use standard $\sin(x)/x \to 1$). Now, *rigorously*, we argue as follows:

- Since $\lim_{x\to 0}\frac{\frac{d}{dx}x\cos x-\sin x}{\frac{d}{dx}x\sin x} = 0$, $\frac{d}{dx}x\sin x \neq 0$ near (but not necessarily at) $x = 0$, and both the numerator and denominator of $\frac{x\cos x-\sin x}{x\sin x}$ are continuous and equal 0 when $x = 0$, we conclude by L'Hopital that:

$$\lim_{x\to 0}\frac{x\cos x-\sin x}{x\sin x} = 0$$

- It follows that $\lim_{x\to 0}\frac{\frac{d}{dx}\log(\sin(x)/x)}{\frac{d}{dx}x} = 0$. Since $\frac{d}{dx}x \neq 0$, and the numerator and denominator of $\frac{\log(\sin(x)/x)}{x}$ are "continuous" (can extend the numerator to be continuous at $x = 0$ since $\sin(x)/x \to 1$; see remark!) and equal 0 when $x = 0$, we conclude by L'Hopital that:

$$\lim_{x\to 0} \frac{\log(\sin(x)/x)}{x} = 0$$

- Since log is continuous, we conclude that:

$$\lim_{x\to 0} \left(\frac{\sin x}{x}\right)^{1/x} = e^0 = 1$$

■

10 Taylor and Various Modes of Differentiating

10.1 Warm-up

Baka the Bunny has a carrot. In one bite, Baka eats some random amount of it (from no carrot to entire carrot). Baka continues taking bites, with each bite eating the same amount of carrot as the first bite, until there is no longer enough carrot to take a full bite. What is the expected fraction of the carrot eaten by Baka?

Hint: Vs gur svefg ovgr vf bs fvmr k, gura gur ohaal jvyy gnxr sybbe bs bar bire k ovgrf. Fb gung vf k gvzrf sybbe bs bar bire k pneebg-senpgvba rngra.

Hint 2: Gurersber, lbh ner pbzchgvat gur vagrteny bs k gvzrf sybbe bs bar bire k bire gur vagreiny sebz mreb gb bar. Rinyhngr guvf ol fcyvggvat vg bire n ohapu bs vagreinyf, rnpu qrgrezvarq ol n inyhr gnxra ol sybbe bs bar bire k. Lbh jvyy or yrsg jvgu na vasvavgr fhz.

Hint 3: Lbhe vasvavgr fhz fubhyq or or gur fhz bs fbzrguvat zvahf n fbzrguvat. V jbhyq abg pbzovar gur fbzrguvatf. Ohg gel jevgvat bhg fbzr grezf bs guvf frevrf naq frr jung unccraf.

10.2 Remarks on Differentiability

In lecture, you learned about this notion of differentiability, called "Frechét Differentiability", which generalizes differentiability to arbitrary normed spaces (not just $\mathbb{R}^N$).

In these points, we assume $f : E \to Y$ is some function, and x_0 is an interior point of E, just so that things are nice (otherwise you have to account for weird things that I personally don't really care about :/).

Key Point #0: If f is differentiable at x_0, then all partial/directional derivatives exist at x_0.

This is because if there exists a linear $df(x_0)$ for which

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - df(x_0)(x - x_0)}{\|x - x_0\|_X} = 0$$

then surely

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0) - df(x_0)(x_0 + tv - x_0)}{\|x_0 + tv - x_0\|_X} = 0$$

or

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0) - t \cdot df(x_0)(v)}{|t|} = 0$$

and you can massage this (in one way or another) to:

$$\lim_{t\to 0}\frac{f(x_0+tv)-f(x_0)}{t}=df(x_0)(v)$$

Key Point #1: *Therefore,* $df(x_0)(v)=\frac{\partial f}{\partial v}(x_0)$.

This follows from what we just computed! The motto is "$df(x_0)$ stores information on all partial derivatives".

Key Point #2: *Therefore,* we can determine what $df(x_0)$ is by computing the directional derivative $\frac{\partial f}{\partial v}(x_0)$ for each direction v.

Of course, this doesn't mean that a function is differentiable if all the partial derivatives exist. I'm just saying that if you knew all the partial derivatives, then you know what the differential $df(x_0)$ would *have* to be, if it exists from differentiability. Trying to evaluating the actual "differentiability limit" for your proposed continuous and linear $df(x_0)$ will either confirm or reject the hypothesis that $df(x_0)$ exists, i.e. that f is differentiable at x_0... which is often very much of interest!!!

Key Point #3: *In fact,* when in $\mathbb{R}^N$, we have that $df(x_0)(v)=\nabla f(x_0)\cdot v$.

To reason this out, first let $v=(\lambda_1,\lambda_2\cdots,\lambda_N)\in\mathbb{R}^N$ be your favorite vector. Define $x_1,\cdots,x_N$ to be the basis vectors in $\mathbb{R}^N$ (as will be the usual notation) and note that:

$$\begin{aligned}df(x_0)(x_1)&=\frac{\partial f}{\partial x_1}(x_0)\\ df(x_0)(x_2)&=\frac{\partial f}{\partial x_2}(x_0)\\ &\vdots\\ df(x_0)(x_N)&=\frac{\partial f}{\partial x_N}(x_0)\end{aligned}$$

If we take a linear combination of these N equations, with coefficients $\lambda_1,\lambda_2,\cdots,\lambda_N$, then we get:

$$\sum_{i=1}^N\lambda_i df(x_0)(x_i)=\sum_{i=1}^N\lambda_i\frac{\partial f}{\partial x_i}(x_0)$$

But remember, $df(x_0)$ is a linear map! So we can shove stuff inside on the left:

$$df(x_0)\left(\sum_{i=1}^N\lambda_i x_i\right)=\sum_{i=1}^N\lambda_i\frac{\partial f}{\partial x_i}(x_0)$$

But hey, isn't $\sum_{i=1}^N\lambda_i x_i=v$? Also, isn't $\sum_{i=1}^N\lambda_i\frac{\partial f}{\partial x_i}(x_0)=v\cdot\nabla f(x_0)$? This means that $df(x_0)(v)=\nabla f(x_0)\cdot v$, as I claimed.

Remark: We tend to view/associate vectors as/with column matrices. That includes $\nabla f(x_0)$.

Extra Point: What happens for $f : \mathbb{R}^N \to \mathbb{R}^M$?

Then $df(x_0)(v) = J_f(x_0)v$ where $J_f(x_0)$ is the *Jacobian Matrix* of f evaluated at x_0. You will learn this soon.

10.3 Is $df(x_0)$ unique?

Recall: If there exists *a* linear and continuous L satisfying the differentiability limit thing, *then* we say $df(x_0) = L$. But is it the *only* linear and continuous L that works?

As usual suppose x_0 is an interior point of the domain, otherwise it gets stupid (and I believe you can construct a weird domain in which $df(x_0)$ is not unique). In this case, the answer is absolutely **yes**, $L = df(x_0)$ is unique.

I'm actually just going to assume further that $f : X \to Y$ (that is, the domain is everything) so that my head doesn't hurt.

Proof. Suppose there were two such linear and continuous maps L and K that both satisfy the differentiability definition. That is:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|_X} = 0$$

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - K(x - x_0)}{\|x - x_0\|_X} = 0$$

Subtract them to get:

$$\lim_{x \to x_0} \frac{(L - K)(x - x_0)}{\|x - x_0\|_X} = 0$$

We want to show that the linear map $L - K$ is the 0 map.

To show this, let $y \in X$. We want to show $(L - K)(y) = 0$. If $y = 0_X$ then there is nothing to show because $T(0) = 0$ for any linear T. So assume $y \neq 0$. Letting $x = x_0 + ty$, we have:

$$\lim_{t \to 0^+} \frac{(L - K)(ty)}{\|ty\|_X} = 0$$

$$\lim_{t \to 0^+} \frac{(L - K)(y)}{\|y\|_X} = 0$$

...but there's no t anymore, so actually:

$$\frac{(L - K)(y)}{\|y\|_X} = 0$$

From which it follows that $(L-K)(y) = 0$. Done. □

10.4 Differentiability Examples

Example 10.1: Suppose we define a function $f : \mathbb{R}^2 \to \mathbb{R}$ as:

$$f(x,y) := \begin{cases} 0, & x^2 \le y \le 2x^2 \\ \sin(y), & \text{Otherwise} \end{cases}$$

1. Where is f continuous?
2. Where do $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist?
3. Where is f differentiable?

Solution.

Continuity

Let $E = \{(x,y) : x^2 \le y \le 2x^2\}$. Also let $P = \partial E = \{(x,y) : y = x^2 \text{ or } y = 2x^2\}$We claim that f is continuous everywhere in this set:

$$(\mathbb{R}^2 \setminus P) \cup \bigcup_{n \in \mathbb{N}_0} \left\{ (\pm\sqrt{\pi n}, \pi n), \left(\pm\sqrt{\frac{\pi n}{2}}, \pi n \right) \right\}$$

To see this, note that if $z_0 = (x_0, y_0) \notin P$, then z_0 is either in the interior of E or its complement $F = \mathbb{R}^2 \setminus P$. If $z_0 \in E^0$, then we may draw a ball around z_0 for which $f(z) = 0$ for all z in the ball, implying continuity at z_0. Similarly, if $z_0 \in F^0$, then $f(x,y) = \sin y$ for all z in some ball around z_0, and we know that $\sin y$ is continuous so we have continuity at all such z_0.

As for the set $\bigcup_{n \in \mathbb{N}_0} \left\{ (\pm\sqrt{\pi n}, \pi n), \left(\pm\sqrt{\frac{\pi n}{2}}, \pi n \right) \right\}$, note that this is just all points (x,y) in P for which $\sin(y) = 0$. One way to see that we have continuity at these points is as follows: Let $z_0 = (x_0, y_0)$ be such a point. Then $f(z_0) = 0$. By continuity of $\sin(y)$, we have that for any $\varepsilon > 0$, there is $\delta > 0$ such that $|\sin(y)| < \varepsilon$ for all z with $0 < \|z - z_0\| < \delta$. But for the same δ, we have $|f(z)| < \varepsilon$ for all z with $0 < \|z - z_0\| < \delta$ (because either $f(z) = \sin y$ or $f(z) = 0$, and $|f(z)| < \varepsilon$ is true in both cases), so we have continuity of f at z_0.

Now we show that f is not continuous at the points z_0 in P for which $\sin(y_0) \ne 0$. In fact, we'll show that for any $z_0 \in P$, the limit $\lim_{z \to z_0} f(z)$ does not exist. But this is obvious: Consider an approach to z_0 along a horizontal restriction, i.e. consider $g(x) = f(x, y_0)$. Then if x is sufficiently close to x_0, x is in E from one direction, and x is in F from the other.

From the direction from E, the limit is just 0. But from the direction from F, the limit is $\sin(y_0) \neq 0$, so the limit at z_0 doesn't exist, so continuity certainly cannot exist.

Partial Derivatives

We claim that the $\frac{\partial f}{\partial x}$ exists exactly in E^0, F^0, and those points on P for which $\sin y = 0$ (essentially the same answer from the previous part). Whereas, $\frac{\partial f}{\partial y}$ exists only in E^0, F^0, and at $(0,0)$. For E^0 and F^0 it is obvious that those partials exist, so it remains to analyze the partials in P.

Note that $f(x,y) = \sin(y)$ everywhere along the restrictions $x = 0$ and $y = 0$, so we have that both partials exist at $(0,0)$ since $\sin(y)$ is differentiable everywhere. Now we handle the four "branches" of P. Let $z_0 = (x_0, x_0)^2 \in P$ with $z_0 > 0$. Define $g(t) = f(x_0, x_0^2 + t)$. Then since $x_0 \neq 0$, we have for sufficiently small $|t|$ that:

$$g(t) = \begin{cases} 0, & t \geq 0 \\ \sin(x_0^2 + t), & t < 0 \end{cases}$$

So:

$$\lim_{t \to 0^-} \frac{g(t) - g(0)}{t} = \lim_{t \to 0^-} \frac{\sin(x_0^2 + t)}{t}$$

This diverges unless x_0^2 is some integer multiple of π. If so, then it is well known that this left limit is either -1 or $+1$, depending on whether x_0^2 is an even or odd multiple. As for the right limit:

$$\lim_{t \to 0^+} \frac{g(t) - g(0)}{t} = \lim_{t \to 0^-} \frac{0}{t} = 0$$

Since these limits are not equal, the partial with respect to y does not exist anywhere on the positive branch of $y = x^2$. A similar calculation shows that this partial doesn't exist anywhere on the other three branches.

As for the partial with respect to x, consider instead $h(t) = f(x_0 + t, x_0^2)$. Then, for sufficiently small $|t|$, we have:

$$h(t) = \begin{cases} 0, & t \leq 0 \\ \sin(x_0^2), & t > 0 \end{cases}$$

For x_0^2 not a multiple of π, h is discontinuous at $t = 0$ so there is no limit with $t \to 0$ and hence no partial with respect to x. Otherwise if x_0^2 is a multiple of π (so this is a point on the boundary with $\sin y = 0$), then $h(t)$ is simple the zero function for sufficiently small t, in which case we clearly have differentiability of $h(t)$ at $t = 0$, and thus a partial derivative with respect to 0 at z_0. This logic follows for the other three "branches". Hence our claims are proven.

Differentiability

We claim that differentiability exists inside E^0, F^0, and at $(0,0)$. To see this, note that if we have differentiability, the partials must exist, so we only need to consider E^0, F^0, and $(0,0)$. At $(0,0)$, the gradient is $\begin{pmatrix} \frac{\partial \sin y}{\partial x}(0,0) \\ \frac{\partial \sin y}{\partial y}(0,0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ since we noted in the previous part that the partials at the origin are just those partials of $\sin(y)$ at the origin, so differentiability at $(0,0)$ holds iff:

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - 0 - (0x + 1y)}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{f(x,y) - y}{\sqrt{x^2+y^2}} = 0$$

To show this, let f_E be the restriction of f to E. Then:

$$\lim_{(x,y)\to(0,0)} \frac{f_E(x,y) - y}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{-y}{\sqrt{x^2+y^2}}$$

For all $(x,y) \in E$ with $(x,y) \neq (0,0)$, we have $y > 0$. So if we rewrite the expression as $\frac{-1}{\sqrt{\frac{x^2}{y^2}+1}}$, this decreases as y increases. But for $(x,y) \in E$, we have the bounds $x^2 \leq y \leq 2x^2$, hence:

$$\frac{-1}{\sqrt{\frac{1}{x^4}+1}} \leq \frac{-1}{\sqrt{\frac{x^2}{y^2}+1}} \leq \frac{-1}{\sqrt{\frac{1}{4x^4}+1}}$$

So as $(x,y) \in E$ tends to $(0,0)$, x tends to 0, and both the above upper bound and lower bound clearly tend to 0, so by Squeeze theorem:

$$\lim_{(x,y)\to(0,0)} \frac{f_E(x,y) - y}{\sqrt{x^2+y^2}} = 0$$

Similarly, let f_F be the restriction of f to F. Then:

$$\lim_{(x,y)\to(0,0)} \frac{f_F(x,y) - y}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{\sin(y) - y}{\sqrt{x^2+y^2}} = 0$$

This limit is zero as well because $\sin(y)$ is differentiable at $(0,0)$.

Now let $g(x,y) = \frac{f(x,y)-y}{\sqrt{x^2+y^2}}$. Let g_E and g_F be the restrictions of g over E and F, respectively. Now note that for all $\varepsilon > 0$, there exists $\delta_1, \delta_2 > 0$ such that:

$$0 < \|z\| < \delta_1 \implies |g_E(z)| < \varepsilon$$

$$0 < \|z\| < \delta_2 \implies |g_F(z)| < \varepsilon$$

Now choose $\delta = \min(\delta_1, \delta_2)$. Then for all $z \in \mathbb{R}^2$ satisfying $0 < \|z\| < \delta$, either $z \in E$ or $z \in F$. If $z \in E$, then $|g(z)| = |g_E(z)| < \varepsilon$, and similarly if $z \in F$ then $|g(z)| < \varepsilon$, so $|g(z)| < \varepsilon$ for all such z. This implies that $\lim_{z\to(0,0)} g(z) = 0$, so we have differentiability at $(0,0)$.

Now we just show that there is differentiability within E^0 and F^0. But the partials exist everywhere, **and are continuous everywhere**, inside these sets, and these sets are open, so by a theorem we have differentiability everywhere in both sets. ■

Example 10.2 (Evaluation Map): Let X be the normed space of all bounded functions on $\mathbb{R}$, endowed with the norm $\|f\|_\infty := \sup_{\mathbb{R}} |f|$. Define an operator $F : X \to \mathbb{R}$ via:

$$F(f) := f(0) \qquad \forall f \in X$$

Is F differentiable at f_0?

Solution. We need to find a linear L such that:

$$\lim_{f \to f_0} \frac{F(f) - F(f_0) - L(f - f_0)}{\|f - f_0\|_\infty} = 0$$

How in the world can we possibly go about finding L... or even guessing L?

Let's instead think about what it would *have* to be by considering some directional derivatives. This is because for a direction $v \in$ with $\|v\|_\infty = 1$, we *must* have $L(v) = \frac{\partial F}{\partial v}(f_0)$. (Why?)

So let's take a direction v and compute:

$$\begin{aligned}\frac{\partial F}{\partial v}(f_0) &= \lim_{t \to 0} \frac{F(f_0 + tv) - F(f_0)}{t} \\ &= \lim_{t \to 0} \frac{f_0(0) + tv(0) - f_0(0)}{t} \\ &= \lim_{t \to 0} v(0) = v(0)\end{aligned}$$

We conclude that for every direction v we must have $L(v) = \frac{\partial F}{\partial v}(f_0) = v(0)$. Hm, so what must L be?

If we now instead look at any $g \in X$, we can compute $L(g)$ using the linearity of L!:

$$L(g) = \|g\|_\infty L\left(\frac{g}{\|g\|_\infty}\right) = \|g\|_\infty \cdot \frac{g(0)}{\|g\|_\infty} = g(0)$$

So L is just the evaluation map $L : g \mapsto g(0)$. This indeed is a continuous and linear map! (Exercise: Why is it continuous?) Now we can verify the original limit (which ends up being really stupid):

$$\lim_{f \to f_0} \frac{F(f) - F(f_0) - L(f - f_0)}{\|f - f_0\|_\infty} = \lim_{f \to f_0} \frac{f(0) - f_0(0) - (f - f_0)(0)}{\|f - f_0\|_\infty} = \lim_{f \to f_0} 0 = 0$$

Thus f is differentiable at every $f_0 \in X$. ■

That was dumb. Let's try a different one:

Example 10.3 (Evaluation Map v2): Let X be the normed space of all bounded functions on $\mathbb{R}$, endowed with the norm $\|f\|_\infty := \sup_{\mathbb{R}} |f|$. Define an operator $F : X \to \mathbb{R}$ via:

$$F(f) := f(0)f(1) \qquad \forall f \in X$$

Is F differentiable at f_0?

Solution. Let's again begin by finding partial derivatives. For a direction v:

$$\begin{aligned}\frac{\partial F}{\partial v}(f_0) &= \lim_{t\to 0} \frac{F(f_0+tv) - F(f_0)}{t} \\ &= \lim_{t\to 0} \frac{(f_0(0)+tv(0))(f_0(1)+tv(1)) - f_0(0)f_0(1)}{t} \\ &= \lim_{t\to 0} f_0(1)v(0) + f_0(0)v(1) + tv(0)v(1) = f_0(1)v(0) + f_0(0)v(1)\end{aligned}$$

We conclude that for every direction v we must have $dF(v) = \frac{\partial F}{\partial v}(f_0) = f_0(1)v(0) + f_0(0)v(1)$ if dF were to exist, and we deduce that $dF(g) = f_0(1)g(0) + f_0(0)g(1)$, which is linear and continuous. We claim that this is the desired continuous and linear dF. Indeed:

$$\begin{aligned}&\lim_{f\to f_0} \frac{F(f) - F(f_0) - dF(f-f_0)}{\|f-f_0\|_\infty} \\ &= \lim_{f\to f_0} \frac{f(0)f(1) - f_0(0)f_0(1) - (f_0(1)(f-f_0)(0) + f_0(0)(f-f_0)(1))}{\|f-f_0\|_\infty} \\ &= \lim_{f\to f_0} \frac{f(0)f(1) - f_0(0)f_0(1) - (f_0(1)f(0) - f_0(1)f_0(0) + f_0(0)f(1) - f_0(0)f_0(1))}{\|f-f_0\|_\infty} \\ &= \lim_{f\to f_0} \frac{(f(0)-f_0(0))(f(1)-f_0(1))}{\|f-f_0\|_\infty} \\ &= 0\end{aligned}$$

Why is the last equality true? Two reasons:

1. $|f(0) - f_0(0)| \le \sup_{x\in\mathbb{R}} |f(x) - f_0(x)| = \|f - f_0\|_\infty$, thus $\frac{f(0)-f_0(0)}{\|f-f_0\|_\infty}$ is bounded.
2. $f(1) - f_0(1) \le \|f - f_0\|_\infty \to 0$ as $f \to f_0$. (Remember that the "$f \to f_0$" under the lim means, literally, that $f \to 0$ under $\|\cdot\|_\infty$, i.e. $\|f - f_0\|_\infty \to 0$.)

Thus f is differentiable at every $f_0 \in X$. ■

The next example is left as an exercise for the bold.

Example 10.4 (Higher Order Frechét Differentiation): Recall, in the definition of differentiability, that $df(x_0)$ is a continuous linear function from X to Y. That is, it is an element of $\mathcal{L}(X,Y)$, the space of all continuous linear maps from X to Y. In fact, if X and Y are normed spaces, then $\mathcal{L}(X,Y)$ is a normed space with the following norm:

$$\|T\|_{\mathcal{L}(X,Y)} := \sup\{\|Tx\|_Y : \|x\|_X = 1\}$$

Since $df(x_0) \in \mathcal{L}(X,Y)$ for each $x_0 \in X$, we can view df as a function that takes in values in X and spits out continuous linear maps in $\mathcal{L}(X,Y)$.
Your mission is this: Suppose we are to define $f : \mathbb{R}^N \to \mathbb{R}$ via $f(\vec{x}) := \|\vec{x}\|^2$. Observe that f is differentiable everywhere, so that $df(\vec{x})$ is well-defined for all $\vec{x} \in \mathbb{R}^N$. Is df differentiable at every $\vec{x_0} \in \mathbb{R}^N$? If so, compute $ddf(\vec{x})(\vec{y})(\vec{z})$ for all $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^N$.

10.5 Taylor's Theorem

We first introduce a new notion of "small asymptotics", called *little-o notation*.

Definition 10.1

Fix some function g. We say that a function f is $o(g(x))$, as $x \to x_0$, if:

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0$$

Essentially, we use $o(g(x))$ as a "placeholder" for some function/expression that vanishes when divided by $g(x)$ (and sending $x \to x_0$).

Examples (as $x \to 0$):

- x^2 is $o(x)$ as $x \to 0$. That's because $\frac{x^2}{x} \to 0$ as $x \to 0$.
- x is NOT $o(x)$.
- $\sin(x)$ is $o(1)$. That's because $\sin(x)/1 \to 0$.
- $\sin(x)$ is NOT $o(x)$. Recall that $\sin(x)/x \to 1$, not 0.
- We can write $\frac{o(x^5)}{x^2}$ as $o(x^3)$. This is because

$$\frac{o(x^5)/x^2}{x^3} = \frac{o(x^5)}{x^5} \to 0$$

by definition of $o(x^5)$.

- We can write $o(o(x))$ as $o(x)$. This is because

$$\frac{o(o(x))}{x} = \frac{o(o(x))}{o(x)} \cdot \frac{o(x)}{x}$$

There is heavy notation abuse here: Not all of these $o(x)$'s represent the same expression being replaced. But some do, which is important. Particularly I'm multiplying and dividing by the *inside* $o(x)$. So in some sense I'm doing this:

$$\frac{o_1(o_2(x))}{x} = \frac{o_1(o_2(x))}{o_2(x)} \cdot \frac{o_2(x)}{x}$$

Anyways, this tends to 0 because...

  - As $x \to 0$ we have $o_2(x)/x \to 0$ by definition of $o_2(x)$.
  - In particular, we must have that $o_2(x) \to 0$ as $x \to 0$.
  - Since $\frac{o_1(y)}{y} \to 0$ as $y \to 0$, and $o(x) \to 0$ as $x \to 0$, we deduce that $\frac{o_1(o_2(x))}{o_2(x)} \to 0$ as $x \to 0$ (replace y with $o_2(x)$...)

So we conclude that $\frac{o(o(x))}{x} \to 0$, hence $o(o(x))$ is $o(x)$ by definition of $o(x)$.

Theorem 10.1 (Taylor)

Let $f : [a, b] \to \mathbb{R}$ be n-times differentiable. Then:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2} + \ldots + f^{(n)}(x_0)\frac{(x - x_0)^n}{n!} + o((x - x_0)^n)$$

Remember, by definition of little-o, this $o((x - x_0)^n)$ thing is a placeholder for an expression that satisfies the key property $\lim_{x \to x_0} \frac{o((x - x_0)^n)}{(x - x_0)^n} = 0$.

Proof. We use induction. Shocking.

For $n = 1$ (...the avid philosopher can toy with trying to use $n = 0$ instead), we want to show if f is differentiable, then:

$$f(x) \stackrel{?}{=} f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$$

Moving things around using the power of algebra, what we really want to show is that:

$$\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \stackrel{?}{=} \frac{o(x - x_0)}{x - x_0}$$

By definition of $o(x - x_0)$, it remains to show that the left side tends to 0 as $x \to x_0$. And indeed it does because $\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$ by... uh, definition.

Now assume that whenever f is n-times differentiable, we have

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2} + \ldots + f^{(n)}(x_0)\frac{(x - x_0)^n}{n!} + o((x - x_0)^n)$$

For the inductive step, we want to show that, for f $(n + 1)$-times differentiable, we have:

$$f(x) \stackrel{?}{=} f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2} + \ldots + f^{(n+1)}(x_0)\frac{(x - x_0)^{n+1}}{(n + 1)!} + o((x - x_0)^{n+1})$$

Again, by the power of algebra, we must show that:

$$\frac{f(x) - f(x_0) - f'(x_0)(x - x_0) - f''(x_0)\frac{(x-x_0)^2}{2} - \ldots - f^{(n+1)}(x_0)\frac{(x-x_0)^{n+1}}{(n+1)!}}{(x - x_0)^{n+1}} \stackrel{?}{=} \frac{o((x - x_0)^{n+1})}{(x - x_0)^{n+1}}$$

In other words, by definition of little-o, we need to show that this left side goes to 0 as $x \to x_0$.

Apply L'Hopital's Rule (verify that the conditions for its use are met!), and so we want to prove that:

$$\lim_{x \to x_0} \frac{f'(x) - f'(x_0) - f''(x_0)(x - x_0) - \ldots - f^{(n+1)}(x_0)\frac{(x-x_0)^n}{n!}}{(n + 1)(x - x_0)^n} \stackrel{?}{=} 0$$

But f' is n-times differentiable, so **by the inductive hypothesis applied to** f', we can rewrite the numerator, so that we want to show:

$$\lim_{x \to x_0} \frac{o((x - x_0)^n)}{(n + 1)(x - x_0)^n} \stackrel{?}{=} 0$$

But this is true by definition of little-o. Thus Taylor has been proven. □

10.6 Examples of Taylor

First a simple example.

Example 10.5: Compute $\lim_{x\to 0} \frac{\sin x - x}{x^3}$.

Solution. Write $\sin x = x - x^3/6 + o(x^4)$. This is an EQUALITY. I can replace $\sin x$ with this "approximation" and you can't stop me.

$$\lim_{x\to 0} \frac{\sin x - x}{x^3} = \lim_{x\to 0} \frac{x^3/6 + o(x^4)}{x^3} = \lim_{x\to 0} \frac{1}{6} + \frac{o(x^4)}{x^3}$$

But $\frac{o(x^4)}{x^3} \to 0$ because $x \to 0$ and $\frac{o(x^4)}{x^4} \to 0$ so their product $\frac{o(x^4)}{x^4}$ goes to 0 too. So the limit is $\boxed{1/6}$. ■

Example 10.6: Compute $\lim_{x\to 0} \left(\frac{\sin x}{x}\right)^{1/x}$.

Solution. Write $\sin x = x + o(x^2)$ (that's all we need). Then:

$$\lim_{x\to 0} \left(\frac{\sin x}{x}\right)^{1/x} = \exp\left(\lim_{x\to 0} \frac{\log(\sin(x)/x)}{x}\right)$$

$$= \exp\left(\lim_{x\to 0} \frac{\log((x + o(x^2)/x)}{x}\right)$$

Using $o(x^2)/x = o(x)$:

$$= \exp\left(\lim_{x\to 0} \frac{\log(1 + o(x))}{x}\right)$$

Write $\log(1 + y) = y + o(y)$. Taking $y = o(x)$:

$$= \exp\left(\lim_{x\to 0} \frac{o(x) + o(o(x))}{x}\right)$$

But as we remarked a while back, $o(o(x)) = o(x)$. Moreover clearly we have $o(x) + o(x) = o(x)$. Thus this is:

$$= \exp\left(\lim_{x\to 0} \frac{o(x)}{x}\right)$$

Apply definition of little-o:

$$= \exp(0) = \boxed{1}$$

■

11 Midterm Review II : The Two Towers

11.1 Warm-up

Can the average of two consecutive prime numbers be prime?

Hint 1: Vs lbh'er ernqvat guvf uvag gura lbh ner bireguvaxvat gur ceboyrz...

Hint 2: Frevbhfyl lbh fubhyq ernq gur dhrfgvba irel pnershyyl

Hint 3: V nz fb, fb fbeel gung lbh unira'g fbyirq vg lrg

Hint 4: Lbh ner tbvat gb xvpx lbhefrys jura lbh svanyyl ernyvmr gur rkgerzryl nagvpyvzngvp fbyhgvba

Hint 5: Guvf unf abguvat gb qb jvgu cevzr ahzoref.

11.2 True or False

1. I'm going to do very well on the exam because I've been studying the proofs.
2. $E \subseteq \mathbb{R}^N$, $f : E \to \mathbb{R}$, $x_0 \in E^\circ$, $\frac{\partial f}{\partial x_i}(x_0$ exists for all $1 \leq i \leq N$. Then f is differentiable at x_0.
3. The following logic is correct: Suppose $0 \leq f(x) \leq g(x)$ and $\lim\limits_{x\to 0} g(x) = 0$. Let $L = \lim\limits_{x\to 0} f(x)$. We claim that $L = 0$. To see this, note that $0 \leq L = \lim\limits_{x\to 0} f(x) \leq \lim\limits_{x\to 0} g(x) = 0$. So $0 \leq L \leq 0$ and hence $L = 0$.
4. The negation of the proposition "$\lim\limits_{x\to x_0} f(x) = 42$" is "$\lim\limits_{x\to x_0} f(x) \neq 42$".
5. Suppose $\sum\limits_{n=1}^{\infty} f_n(x)$ converges pointwise for all $x \in \mathbb{R}$. Take $x_0 \in \mathbb{R}$ and assume that $\lim\limits_{x\to x_0} f_n(x)$ exists for each n. Then:
$$\lim_{x\to x_0} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \lim_{x\to x_0} f_n(x)$$
6. The terms of a series tend to 0 so it must converge.
7. The series $\sum\limits_{n=1}^{\infty} \frac{2n+1}{4+3n+2n^2+n^3}$ converges.

11.3 Differentiability and Stuff

1. Define a function $f : \mathbb{R}^2 \to \mathbb{R}$ as follows:

$$f(x,y) := \begin{cases} \frac{1-\cos(xy)}{y}, & y \neq 0 \\ 0, & y = 0 \end{cases}$$

   sTuDy ThE dIfFeReNtIaBiLiTy Of f.

2. Is the function $g(x,y) := \begin{cases} \frac{xy}{\sqrt{x^2+y^2}}, & (x,y) \neq 0 \\ 0, & (x,y) = 0 \end{cases}$ differentiable at $(0,0)$?

11.4 Uniform Convergence and Stuff

Consider the series $\displaystyle\sum_{n=1}^{\infty} \frac{x}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}}$.

1. Find all $x \in \mathbb{R}$ for which the series converges pointwise.
2. Sketch a graph of $\frac{x}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}}$ if you want.
3. Find all sets $E \subseteq \mathbb{R}$ for which the series converges uniformly.

11.5 Theorem Roulette

Have fun ;)

11.6 (Basically) All Warm-up Solutions

Last digit: Jr pynvz gung gur ynfg qvtvg vf avar. Gb frr guvf, abgr gung rivqragyl gur rkcerffvba zhfg or rdhvinyrag gb bar zbq gra. Vg erznvaf gb fubj gung npghnyyl gur rkcerffvba vf artngvir. Lbh pna cebonoyl trg guvf sebz gur zbabgbavpvgl bs ybt k bire k be fbzrguvat.

Horrific Stanford Sum: Yeah recitation is in half an hour and I don't have time to write this lol... will update later

Uncountable (?) totally-ordered subset of $P(\mathbb{N})$: Svefg ovwrpg gur anghenyf vagb gur engvbanyf, fb gung jr pna whfg jbex jvgu engvbanyf vafgrnq. Gura jr pbafvqre gur Qrqrxvaq phgf, juvpu ner va ovwrpgvba jvgu gur erny ahzoref (naq vaqrrq gurl freir nf n pbafgehpgvba bs gur ernyf). Fvapr gur ernyf ner gbgnyyl beqrerq, jr unir gung gur snzvyl bs Qrqrxvaq phgf vf gbgnyyl beqrerq haqre gur fhofrg cnegvny beqre, naq guvf snzvyl vf hapbhagnoyr orpnhfr gur ernyf ner hapbhagnoyr.

Fitting a larger cube into a smaller cube: Please enjoy the solution in picture form: https://i.imgur.com/b6KXbi9.png

Baka the Bunny:

$$\int_0^1 x \left\lfloor \frac{1}{x} \right\rfloor dx = \sum_{n=1}^{\infty} \int_{1/(n+1)}^{1/n} nx\, dx = \sum_{n=1}^{\infty} \frac{n}{2}\left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right)$$

$$= \frac{1}{2}\sum_{n=1}^{\infty} \frac{1}{n^2} + \frac{n-1}{n^2} - \frac{n}{(n+1)^2} = \frac{\pi^2}{12}$$

11.7 True or False Solutions

1. This better be true!
2. False. You need partial derivatives to exist and be continuous everywhere in a ball around x_0. (Though as in the relevant theorem, you can let one partial derivative be a little less nice.)
3. False. I assumed that L existed. A priori, I don't know that. So I can't play with writing $L = \lim_{x\to 0} f(x)$. To fix the logic, do not write this limit, and instead just directly apply the squeeze rule. Either that or use a limsup, but you won't be required to know that.
4. False. There is a third possibility: The limit doesn't exist. The correct negation is as follows: "There exists $\varepsilon > 0$ such that for all $\delta > 0$ you can find $x \in B(x_0, \delta)$ with $x \neq x_0$ such that $|f(x) - f(x_0)| \geq \varepsilon$."

5. False. You should always be suspicious about switching two limits! See this terrible graph (https://www.desmos.com/calculator/rjhokztgnp) for a dumb counterexample I came up with.

6. False. Consider the harmonic series.

7. True. Note that for all n large enough, we have:

$$\frac{2n+1}{4+3n+2n^2+n^3} \leq \frac{3n}{\frac{1}{2}n^3}$$

This is because $3n$ will "beat" $2n+1$ in the long run, and $4+3n+2n^2+n^3$ "beat" $\frac{1}{2}n^3$ in the long run. Thus this inequality must hold for all $n \geq N$ where N is some huge number that I couldn't care less about. It follows that:

$$\sum_{n=N}^{\infty} \frac{2n+1}{4+3n+2n^2+n^3} \leq \sum_{n=N}^{\infty} \frac{3n}{\frac{1}{2}n^3} = 6\sum_{n=N}^{\infty} \frac{1}{n^2} < \infty$$

Which is enough to conclude that $\sum_{n=N}^{\infty} \frac{2n+1}{4+3n+2n^2+n^3} < \infty$. (Why?)

11.8 Differentiability and Stuff Solution

Problem 1

For all (x_0, y_0) with $y_0 \neq 0$, it is clear that f is differentiable. Just verify that the partial derivatives are continuous everywhere in $\mathbb{R}^2 \setminus \{y = 0\}$.

Now consider a point $(x_0, 0)$. We claim that f is differentiable at $(x_0, 0)$ with $df(x_0, 0)(x, y) = \frac{1}{2}x_0^2 y$. Indeed, this $df(x_0, 0)$ is linear and continuous, and so it remains to verify the limit associated with differentiability. That is:

$$\lim_{(x,y)\to(x_0,0)} \frac{f(x,y) - f(x_0,0) - \frac{1}{2}x}{\sqrt{(x-x_0)^2+y^2}} \stackrel{?}{=} 0$$

First let's take the following upper bound, for all $y \neq 0$:

$$\left|\frac{f(x,y) - f(x_0,0) - \frac{1}{2}x_0^2 y}{\sqrt{(x-x_0)^2+y^2}}\right| = \left|\frac{\frac{1-\cos(xy)}{y} - \frac{1}{2}x_0^2 y}{\sqrt{(x-x_0)^2+y^2}}\right| \leq \left|\frac{\frac{1-\cos(xy)}{y} - \frac{1}{2}x_0^2 y}{y}\right|$$

$$= \left|\frac{1-\cos(xy)}{y^2} - \frac{1}{2}x_0^2\right|$$

$$= \left|\frac{1}{2}x^2 \cdot 2\frac{1-\cos(xy)}{x^2y^2} - \frac{1}{2}x_0^2\right|$$

As $(x,y) \to (x_0,0)$ we have $xy \to 0$ (why?), so that $2\frac{1-\cos(xy)}{x^2y^2} \to 1$. Moreover, $\frac{1}{2}x^2 \to \frac{1}{2}x_0^2$. Thus the upper bound we have obtained tends to 0.

If otherwise $y = 0$ then:

$$\left|\frac{f(x,y) - f(x_0,0) - \frac{1}{2}x_0^2 y}{\sqrt{(x-x_0)^2+y^2}}\right| = \left|\frac{0}{\sqrt{(x-x_0)^2}}\right| = 0$$

Which is stupid. In all, we conclude that for *all* $(x,y) \neq (x_0,0)$ we have:

$$\left|\frac{f(x,y) - f(x_0,0) - \frac{1}{2}x_0^2 y}{\sqrt{(x-x_0)^2+y^2}}\right| \le \max\left(\left|\frac{1}{2}x^2 \cdot 2\frac{1-\cos(xy)}{x^2y^2} - \frac{1}{2}x_0^2\right|, 0\right)$$

And we have shown that the RHS tends to 0 as $(x,y) \to (x_0,0)$ by the above casework. So indeed we have the desired limit by the squeeze rule. □

Remark: The claimed $df(x_0,0)$ was not pulled out of thin air! You can determine what it would *have* to be by examining the partial derivatives. This is a very important point!

Problem 2

We claim that g is not differentiable at $(0,0)$. As in the above remark, we assume for contradiction that g is differentiable. We then may compute what $dg(0,0)$ would *have* to be. Indeed, note that:

$$\frac{\partial g}{\partial x}(0,0) = \lim_{x\to 0}\frac{g(x,0)-g(0,0)}{x} = \lim_{x\to 0} 0 = 0$$

And similarly $\frac{\partial g}{\partial y}(0,0) = 0$. Thus we *must* have that $dg(0,0)(x,y) = 0$ for all x,y. That is, $dg(0,0)$ is the identically 0 linear map.

So if g were differentiable, we *must* have the following limit:

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - dg(0,0)(x,y)}{\sqrt{x^2+y^2}} = 0$$

Or:

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = 0$$

But this is bogus by approaching $(0,0)$ along the curve $y = x$, over which a limit of $1/2$ is obtained, and this is enough to obtain a contradiction. □

11.9 Uniform Convergence and Stuff Solution

Pointwise Convergence

Evidently:

$$\sum_{n=1}^{\infty} \frac{|x|}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}} \le \sum_{n=1}^{\infty} \frac{|x|}{n^2} < \infty$$

So we have absolute convergence (thus pointwise convergence) everywhere.

Uniform Convergence

I claim that we have uniform convergence on every set $E \subseteq \mathbb{R}$ that is bounded from above.

Sufficiency

Take such a set E. Since it is bounded from above, there is M such that $x \le M$ for all $x \in E$.

There are two cases: Either $x < -M$ or $|x| \le M$. Each of these cases will be "small" for a different reason.

If $x < -M$, then "the e term wins".

$$\left| \sum_{n=N}^{\infty} \frac{x}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}} \right| \le \sum_{n=N}^{\infty} \frac{|x|}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}} \le \sum_{n=N}^{\infty} |x| e^{nx}$$

This is just a geometric series:

$$= \frac{|x|e^{Nx}}{1-e^x}$$

One way or another, argue that replacing x with $-M$ will increase this expression. For example, you can differentiate $\frac{-xe^{Nx}}{1-e^x}$ and show that the derivative is positive over $(-\infty, 0)$, and thus is increasing. Whatever you do, write:

$$= \frac{Me^{-NM}}{1-e^{-M}}$$

This tends to 0 as $N \to +\infty$, and this is a bound that does not depend on x, so we have uniform convergence over all $x < -M$.

Now suppose that $|x| \le M$. Then "the $1/n^2$ term wins".

$$\left| \sum_{n=N}^{\infty} \frac{x}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}} \right| \le \sum_{n=N}^{\infty} \frac{|x|}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}} \le \sum_{n=N}^{\infty} \frac{M}{n^2}$$

Since $\displaystyle\sum_{n=1}^{\infty} \frac{M}{n^2} < \infty$, we have that $\displaystyle\sum_{n=N}^{\infty} \frac{M}{n^2} \to 0$ as $N \to +\infty$, and once again this is a bound that does not depend on x, hence we have uniform convergence in $[-M, M]$.

Therefore we have uniform convergence over $(-\infty, M]$ and thus uniform convergence over E. If you don't buy that we can just "combine" these uniform convergences, note that we can write, for all $x \leq M$:

$$\left|\sum_{n=N}^{\infty} \frac{x}{n^2} \cdot \frac{e^{nx}}{1+e^{nx}}\right| \leq \max\left(\frac{Me^{-NM}}{1-e^{-M}}, \sum_{n=N}^{\infty} \frac{M}{n^2}\right)$$

And this weird bound is an upper bound that does not depend on x, and vanishes when we send $N \to +\infty$, hence uniform convergence.

Necessity

Suppose E has no upper bound. Appealing to the negation of uniform convergence: Fix $\varepsilon = 1/4$ or something (that probably will work). Suppose my "enemy" picks some large $N \in \mathbb{N}$. My job is to find $n \geq N$ and some $x \in E$ such that:

$$\left|\sum_{k=n}^{\infty} \frac{x}{k^2} \cdot \frac{e^{kx}}{1+e^{kx}}\right| \geq \varepsilon$$

In fact, I'll do you one better: I'll find $x > 0$, $x \in E$ such that $\frac{x}{N^2} \cdot \frac{e^{Nx}}{1+e^{Nx}} \geq \varepsilon$. That will definitely be good enough!

Indeed, observe that $\frac{x}{N^2} \cdot \frac{e^{Nx}}{1+e^{Nx}} \to +\infty$ as $x \to +\infty$, so there is some very large $K > 0$ such that $\frac{x}{N^2} \cdot \frac{e^{Nx}}{1+e^{Nx}} \geq 1/2 = \varepsilon$ for all $x > K$. Since K is not an upper bound of E, I can find $x \in E$ with $x \in K$ and I choose that x (and, implicitly, I am choosing $n = N$). This contradicts uniform convergence. $\square$

12 Taylor Abuse and Lagrange Multipliers

12.1 Warm-up

Thomas is trying to have a completely normal conversation with Wanlin. Fill in the blank with the correct number.

Thomas: "Did you know that your favorite number is the sum of the ages of my stuffed animal turtles, and that my favorite number is their product?"

Wanlin: "I wouldn't know because I don't know your favorite number. If you tell me your favorite number and how many stuffed animal turtles you have, would I know the ages of your stuffed animal turtles?"

Thomas: "No."

Wanlin: "Oh, so your favorite number is ___________!"

"Hint": Guvf vf uneq naq n ovg jbex-vagrafvir. V qb abg xabj bs n dhvpx zrgubq. Vg vf bayl urer orpnhfr gur ceboyrz unf n pbby ahzrevpny nafjre, qrfcvgr univat ab ahzoref va gur fgngrzrag.

12.2 Taylor Abuse

Ok, please don't panic:

Example 12.1: Compute:

$$\lim_{(x,y)\to(0,0)} \frac{\cos(x+y)+\sin(xy)-e^{x^2+y^2}}{\cos(x^2+y^2)-\sqrt{1+x^2+y^2}}$$

Solution. Alright let's start:

- $\cos(x+y) = 1 - \frac{(x+y)^2}{2} + o((x+y)^2)$
- $\sin(xy) = xy + o(xy)$
- $e^{x^2+y^2} = 1 + x^2 + y^2 + o(x^2+y^2)$
- $\cos(x^2+y^2) = 1 - \frac{(x^2+y^2)^2}{2} + o((x^2+y^2)^2)$
- $\sqrt{1+x^2+y^2} = (1+x^2+y^2)^{1/2} = 1 + \frac{1}{2}(x^2+y^2) + o(x^2+y^2)$

Therefore the expression we're finding the limit of is exactly equal to:

$$\frac{1-\frac{(x+y)^2}{2}+o((x+y)^2)+xy+o(xy)-(1+x^2+y^2+o(x^2+y^2))}{1-\frac{(x^2+y^2)^2}{2}+o((x^2+y^2)^2)-\left(1+\frac{1}{2}(x^2+y^2)+o(x^2+y^2)\right)}$$

Let's do some algebra:

$$=\frac{-\frac{3}{2}(x^2+y^2)+o((x+y)^2)+o(xy)+o(x^2+y^2)}{-\frac{1}{2}(x^2+y^2)-\frac{1}{2}(x^2+y^2)^2+o((x^2+y^2)^2)+o(x^2+y^2)}$$

Uh, wow this looks horrific. There are a bunch of o's that are unwanted. In particular it'd be great if we can write all the o's in terms of some $o((x^2+y^2)^k)$.

- We claim $o((x+y)^2)=o(x^2+y^2)$. This is because

$$\frac{o((x+y)^2)}{x^2+y^2}=\frac{o((x+y)^2)}{(x+y)^2}\cdot\frac{(x+y)^2}{x^2+y^2}$$

  which vanishes as $(x,y)\to(0,0)$ because $\frac{o((x+y)^2)}{(x+y)^2}\to 0$ by definition and $\frac{(x+y)^2}{x^2+y^2}$ is bounded by $3/2$ or something.

- We claim $o(xy)=o(x^2+y^2)$. This is because

$$\frac{o(xy)}{x^2+y^2}=\frac{o(xy)}{xy}\cdot\frac{xy}{x^2+y^2}\to 0$$

- Note that $o((x^2+y^2)^2)+o(x^2+y^2)=o(x^2+y^2)$ because "a pair of people walking together will walk at the pace of the slower person".

Thus our expression has become:

$$=\frac{\frac{-3}{2}(x^2+y^2)+o(x^2+y^2)}{-\frac{1}{2}(x^2+y^2)-\frac{1}{2}(x^2+y^2)^2+o(x^2+y^2)}$$

We have $(x^2+y^2)^2=o(x^2+y^2)$, so we'll let that be absorbed:

$$=\frac{\frac{-3}{2}(x^2+y^2)+o(x^2+y^2)}{-\frac{1}{2}(x^2+y^2)+o(x^2+y^2)}$$

You might be wondering how in the world we can continue since the bottom is blegh. One way to see what's going on is to write it as this now:

$$=\frac{\frac{-3}{2}(x^2+y^2)+o(x^2+y^2)}{x^2+y^2}\div\frac{-\frac{1}{2}(x^2+y^2)+o(x^2+y^2)}{x^2+y^2}$$

$$=\left(-\frac{3}{2}+\frac{o(x^2+y^2)}{x^2+y^2}\right)\div\left(-\frac{1}{2}+\frac{o(x^2+y^2)}{x^2+y^2}\right)\to\boxed{3}$$

■

12.3 Lagrange Multipliers

Example 12.2: Find the maximum possible value of $x - y + 2z$, given that $x^2 + y^2 + 2z^2 \leq 2$.

Solution. To be precise, let $f(x, y, z) = x - y + 2z$. We wish to maximize f over the set $E = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + 2z^2 \leq 2\}$.

STEP 1: Maxima in the interior?

If a maximum occurs at some $(x_0, y_0, z_0) \in E^\circ$, then we know that $\nabla f(x_0, y_0, z_0) = \vec{0}$. But we don't necessarily have this if instead we had $(x_0, y_0, z_0) \in \partial E$. That's why there are two cases to consider.

So let's suppose that there was a maxima at some interior point $(x_0, y_0, z_0) \in E^\circ$. Then x_0, y_0, z_0 would *have* to satisfy the equation

$$\nabla f(x_0, y_0, z_0) = \vec{0}$$

or

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which... seems kinda hard to satisfy not gonna lie. We conclude that there are NO maxima inside the interior. In particular, if there is a maxima then it must occur on the boundary $\partial E = \{(x, y, z) \in \mathbb{R}^2; x^2 + y^2 + 2z^2 = 2\}$.

STEP 2: Maxima on the boundary?

Note that ∂E is closed and bounded (why?), so it is compact. Since f is continuous, we have that it must obtain a maximum over ∂E at some $(x_0, y_0, z_0) \in \partial E$. In fact, since we argued that the absolute maximum does not occur in E°, we have that this maximum at (x_0, y_0, z_0) must be the desired absolute maximum.

To find (x_0, y_0, z_0), we use Lagrange Multipliers with the constraint function $g(x, y, z) = x^2 + y^2 + 2z^2 - 2$, because $g(x, y, z) = 0$ iff $(x, y, z) \in \partial E$. To justify the use of the theorem, we need to verify that the set $\{\nabla g(x_0, y_0, z_0)\}$ is linearly independent. Is it? Well, as long as $\nabla g(x_0, y_0, z_0) \neq 0$, then yes, so that's what we have to check. Indeed:

$$\nabla g(x_0, y_0, z_0) = \begin{bmatrix} 2x_0 \\ 2y_0 \\ 4z_0 \end{bmatrix} \neq \vec{0}$$

Note that we know this is $\neq \vec{0}$ because otherwise $(x_0, y_0, z_0) = 0$, contradicting $(x_0, y_0, z_0) \in \partial E$.

Ok, so since f and g are C^1 blah blah blah, we have by the Lagrange Multipliers Theorem that for some $\lambda \in \mathbb{R}$:

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$$

Or:

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \lambda \begin{bmatrix} 2x_0 \\ 2y_0 \\ 4z_0 \end{bmatrix}$$

This gives a system of three equations and four variables. The constraint $g(x_0, y_0, z_0) = 0$ gives the fourth:

$$1 = \lambda 2x_0$$

$$-1 = \lambda 2y_0$$

$$2 = \lambda 4z_0$$

$$x_0^2 + y_0^2 + 2z_0^2 = 2$$

This is now just an algebra problem. Evidently we have $\lambda \neq 0$. So we have $x_0 = \frac{1}{2\lambda}$, $y_0 = \frac{-1}{2\lambda}$, and $z_0 = \frac{1}{2\lambda}$. Plugging this into the fourth equation, we can proceed to solve for λ:

$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} + \frac{1}{2\lambda^2} = 2$$

Therefore $\lambda \in \{1/\sqrt{2}, -1/\sqrt{2}\}$. This gives two possible locations for (x_0, y_0, z_0): It is either $(\sqrt{2}/2, -\sqrt{2}/2, \sqrt{2}/2)$ or $(-\sqrt{2}/2, \sqrt{2}/2, -\sqrt{2}/2)$.

By plugging these possibilities into f, we get the values $2\sqrt{2}$ and $-2\sqrt{2}$. We know that one of these guys is the absolute maximum. So, the absolute maximum would have to be $\boxed{2\sqrt{2}}$. (And in fact, the other is going to end up being the absolute minimum.) ■

Example 12.3: Consider the set:

$$E = \{(x, y, z) \in \mathbb{R}^3 : x^2 - xy + y^2 - z^2 = 1, x^2 + y^2 = 1\}$$

What point(s) of E have minimal distance to the origin?

Solution. This will not be a complete solution, we're just going to set it up (and honestly if you can do just this step then I'm satisfied).

First we set things up. We want to minimize $f(x, y, z) = x^2 + y^2 + z^2$ (this is the *distance squared*, but it is equivalent to minimize this, and it's nicer because there is no nasty square root). The constraint function has multiple components:

$$g(x, y) = (g_1(x, y), g_2(x, y)) = (x^2 - xy + y^2 - z^2 - 1, x^2 + y^2 - 1)$$

First we must justify that a minimum over E exists. Can we show that E is compact?

Here is one method: Show that the set $\{x^2 - xy + y^2 - z^2 = 1\}$ is closed (e.g. it is the union of continuous function graphs). Show that the set $\{x^2 + y^2 = 1\}$ is closed (e.g. via some manual casework to show that the complement is open). Conclude that the intersection of these sets, which is E, is closed. For boundedness, note that both x^2 and y^2 are bounded since $x^2 + y^2 = 1$, so see if you can get a bound on z^2 using the first equation. If we're in E, the first equation simplifies to $z^2 = -xy$, so in fact we just need to show that xy is bounded. This can be done with e.g. AM-GM.

From these arguments, we may deduce that E is compact and hence f obtains an absolute minimum somewhere in E by continuity. Specifically, we can say that it obtains this minimum at some $(x_0, y_0, z_0) \in E$. To find this minimum, we apply Lagrange Multipliers. But we first must justify its use by showing that $\{\nabla g_1(x_0, y_0, z_0), \nabla g_2(x_0, y_0, z_0)\}$ is linearly independent. Is it? Well, this set is just:

$$\left\{ \begin{bmatrix} 2x_0 - y_0 \\ -x_0 + 2y_0 \\ -2z_0 \end{bmatrix}, \begin{bmatrix} 2x_0 \\ 2y_0 \\ 0 \end{bmatrix} \right\}$$

To see that this is linearly independent, note first that since $x_0^2 + y_0^2 = 1$, we have that one of x_0, y_0 is non-zero. From here, if we have that $z_0 \neq 0$, then we have the linear independence (why?).

Otherwise, if $z_0 = 0$, then we must show that $\left\{ \begin{bmatrix} 2x_0 - y_0 \\ -x_0 + 2y_0 \end{bmatrix}, \begin{bmatrix} 2x_0 \\ 2y_0 \end{bmatrix} \right\}$ is linearly independent. But from $0 = z_0^2 = -x_0 y_0$, we have that one of x_0, y_0 is 0 and the other is non-zero. In both cases, you can verify that the above set will be independent!

By this independence (and by the fact that f, g are class C^1 yadda yadda) we may apply Lagrange Multipliers to deduce that there exist $\lambda_1, \lambda_2 \in \mathbb{R}$ such that:

$$\nabla f(x_0, y_0, z_0) = \lambda_1 \nabla g_1(x_0, y_0, z_0) + \lambda_2 \nabla g_2(x_0, y_0, z_0)$$

Or:

$$\begin{bmatrix} 2x_0 \\ 2y_0 \\ 2z_0 \end{bmatrix} = \lambda_1 \begin{bmatrix} 2x_0 - y_0 \\ -x_0 + 2y_0 \\ -2z_0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2x_0 \\ 2y_0 \\ 0 \end{bmatrix}$$

This is a system of three equations and five variables. The constraint equalities $g_1(x_0, y_0, z_0) = 0$ and $g_2(x_0, y_0, z_0) = 0$ give the other two we need. Now we solve. Which I won't do. Because why would I. This is awful. Blegh. Finish the solution at your own risk. ■

13 Implicit Function Theorem

13.1 Warm-up

Countably infinite prisoners $p_1, p_2, \cdots$ are standing in a line, facing in the same direction such that p_n can only see the heads of $p_{n+1}, p_{n+2}, \cdots$.

The warden puts a hat on every prisoner, each one with a real number written on it. p_n cannot see the number written on their hat, but can see the numbers written on the hat of p_m for all $m > n$.

Starting at p_1, the warden asks every prisoner for the number on their hat. If the prisoner guesses correctly, they live. Otherwise, they are shot.

The prisoners may devise a plan beforehand. Can they come up with a strategy such that they can guarantee that only finitely many prisoners die?

Oh also, the prisoners are all deaf, so they have no idea what happens behind them.

Hint 1: Gur cevfbaref jvyy arrq gur nkvbz bs pubvpr.

Hint 2: Pbafgehpg na rdhvinyrapr eryngvba ba gur frg bs nyy cbffvoyr frdhraprf bs ung ahzoref. Vg fubhyq or fhpu gung rirel cevfbare xabjf juvpu rdhvinyrapr pynff gurl ner va.

Hint 3: Hfr gur rdhvinyrapr eryngvba "...vf riraghnyyl gur fnzr nf...".

13.2 Implicit Function Theorem

Example 13.1: Let $f : \mathbb{R}^4 \to \mathbb{R}^2$ with $f(w, x, y, z) = (y^2 + z^2 - 2wx, y^3 + z^3 - w^3 + x^3)$. Let $\vec{x_0} = (1, 1, 1, -1)$.
Prove that near $\vec{x_0}$, the equation $f(w, x, y, z) = 0$ induces an implicit function in that (w, x) may be written as a function of (y, z). (Make this statement precise!)
Of what differentiability class is such a function? Compute the first order partial derivatives of the function at $(1, -1)$.

Solution. To be precise, we seek $g = (g_1, g_2)$, $g : B((1, -1), r_1) \to B((1, 1), r_2)$ such that:

- $y^2 + z^2 - 2g_1(y, z)g_2(y, z) = 0$
- $y^3 + z^3 - g_1(y, z) + g_2(y, z) = 0$

First we check that $f(\vec{x_0}) = 0$ otherwise we fail pretty badly. It is, so we're doing well.

Next, since (w, z) is the "output variable", we need to show that the "change in this variable isn't 0" or else we "fail the vertical line test". That is, to use the Implicit Function Theorem, we must verify:

$$\det \frac{\partial f}{\partial(w,z)}(\vec{x_0}) \neq 0$$

Indeed, we may compute:

$$\frac{\partial f}{\partial(w,z)}(w,x,y,z) = \begin{bmatrix} -2x & -2w \\ -3w^2 & 3x^2 \end{bmatrix}$$

So that:

$$\det \frac{\partial f}{\partial(w,z)}(1,1,1,-1) = \begin{vmatrix} -2 & -2 \\ -3 & 3 \end{vmatrix} = -12 \neq 0$$

So indeed we may apply the Implicit Function Theorem to conclude that the desired g exists.

Since f is of differentiability class C^∞, we are guaranteed that g is of class C^∞. In particular, we are justified in taking its first order partial derivatives, which we will now compute.

As $y^2 + z^2 - 2g_1(y,z)g_2(y,z) = 0$ for all (y, z) near $(1, -1)$, we can take the partial derivative with respect to y on both sides:

$$y - \frac{\partial g_1}{\partial y}(y,z)g_2(y,z) - g_1(y,z)\frac{\partial g_2}{\partial y}(y,z) = 0$$

Plugging in $(y, z) = (1, -1)$ and using $g(1, -1) = (1, 1)$, we get:

$$-1 = \frac{\partial g_1}{\partial y}(1,-1) + \frac{\partial g_2}{\partial y}(1,-1)$$

Doing the same thing with the equation $y^3 + z^3 - g_1(y, z) + g_2(y, z) = 0$, we obtain:

$$3 = \frac{\partial g_1}{\partial y}(1,-1) - \frac{\partial g_2}{\partial y}(1,-1)$$

Hence, we have a system of equations! Solving, we obtain the partials $\boxed{\frac{\partial g_1}{\partial y}(1,-1) = 1}$ and $\boxed{\frac{\partial g_2}{\partial y}(1,-1) = -2}$.

Observe that by symmetry, the partial derivatives with respect to z at $(1, -1)$ are exactly the same. ■

Example 13.2: Let $f : \mathbb{R}^3 \to \mathbb{R}$ with $f(x, y, z) = x^2 + 4y^2 - 2yz - z^2$. Let $\vec{x_0} = (2, 1, -4)$.
Prove that near $\vec{x_0}$, the equation $f(x, y, z) = 0$ induces an implicit function in that z may be written as a function of (x, y). If this function is g, compute $\frac{\partial^2 g}{\partial x \partial y}(2,1)$ and $\frac{\partial^2 g}{\partial y \partial x}(2,1)$.

Solution. Since $f(2,1,-4)=0$ and $\frac{\partial f}{\partial z}(2,1,-4)=-2(1)-2(-4)=6\neq 0$, we may apply the Implicit Function Theorem to obtain the existence of g.

To compute partials, we rely on the equality

$$x^2+4y^2-2yg(x,y)-g(x,y)^2=0 \qquad (*)$$

for all (x,y) near $(2,1)$.

Differentiating $(*)$ with respect to y:

$$8y-2g(x,y)-2y\frac{\partial g}{\partial y}(x,y)-2g(x,y)\frac{\partial g}{\partial y}(x,y)=0 \qquad (**)$$

Differentiating with respect to x:

$$-2\frac{\partial g}{\partial x}(x,y)-2y\frac{\partial^2 g}{\partial x\partial y}(x,y)-2g(x,y)\frac{\partial^2 g}{\partial x\partial y}(x,y)-2\frac{\partial g}{\partial x}(x,y)\frac{\partial g}{\partial y}(x,y)=0 \qquad (***)$$

It appears that some work needs to be done. First we need $\frac{\partial g}{\partial x}(2,1)$. Let differentiate $(*)$ with respect to x:

$$2x-2y\frac{\partial g}{\partial x}(x,y)-2g(x,y)\frac{\partial g}{\partial x}(x,y)=0$$

Plugging in $(2,1)$ we get:

$$4-2\frac{\partial g}{\partial x}(2,1)-2g(2,1)\frac{\partial g}{\partial x}(2,1)=0$$

$$4-2\frac{\partial g}{\partial x}(2,1)+8\frac{\partial g}{\partial x}(2,1)=0$$

$$\boxed{\frac{\partial g}{\partial x}(2,1)=-2/3}$$

Next, we need $\frac{\partial g}{\partial y}(2,1)$. Plugging in $(2,1)$ into $(**)$:

$$8-2g(2,1)-2\frac{\partial g}{\partial y}(2,1)-2g(2,1)\frac{\partial g}{\partial y}(2,1)=0$$

$$8+8-2\frac{\partial g}{\partial y}(2,1)+8\frac{\partial g}{\partial y}(2,1)=0$$

$$\boxed{\frac{\partial g}{\partial y}(2,1)=-8/3}$$

Finally we stuff all that garbage into $(***)$:

$$-2\frac{\partial g}{\partial x}(2,1)-2\frac{\partial^2 g}{\partial x\partial y}(2,1)-2g(2,1)\frac{\partial^2 g}{\partial x\partial y}(2,1)-2\frac{\partial g}{\partial x}(2,1)\frac{\partial g}{\partial y}(2,1)=0$$

$$-2(-2/3) - 2\frac{\partial^2 g}{\partial x \partial y}(2,1) - 2(-4)\frac{\partial^2 g}{\partial x \partial y}(2,1) - 2(-2/3)(-8/3) = 0$$

After three years of college I am literally incapable of doing this basic algebra, so by Mathematica this comes out to $\boxed{\frac{\partial^2 g}{\partial x \partial y}(2,1) = 10/27}$.

Note that f is of class C^∞, so in particular g is of class C^∞ and so its second derivatives exist and are continuous. So by Schwarz, we have $\boxed{\frac{\partial^2 g}{\partial y \partial x}(2,1) = 10/27}$ as well. ■

13.3 Chain Rule 101

We begin with a terrible example.

Example 13.3: Let $f(x, y, z) = x + 2y + 3z$.

1. What is $\frac{\partial}{\partial x} f(x, y, x)$?
2. What is $\frac{\partial f}{\partial x}(x, y, x)$?

Solution.

1. $f(x, y, x) = x + 2y + 3x = 4x + 2y$, so $\frac{\partial}{\partial x} f(x, y, x) = 4$.
2. But $\frac{\partial f}{\partial x}(x, y, z) = 1$, so in particular $\frac{\partial f}{\partial x}(x, y, x) = 1$.

■

What the hell is the difference?

KEY POINT:

- When we speak of $\frac{\partial f}{\partial x}$, we are speaking of the partial derivative of f with respect to *the compnonent of f whose name is x*.
- When we speak of $\frac{\partial}{\partial x} f(\cdots)$, we are viewing the expression $f(\cdots)$ as a *function*, one of whose variables is x, and taking the partial derivative with respect to x. f need not have a component named x for this to be well-defined.

In the previous example, when I wrote $f(x, y, z) = x + 2y + 3z$, I am implicitly giving names to the first, second, and third components of f as x, y, z. So when I write $\frac{\partial f}{\partial x}$, this means the derivative with respect to the first component. Then $\frac{\partial f}{\partial x}(x, y, x)$ means, I'm taking $\frac{\partial f}{\partial x}$ and plugging in (x, y, x) for some variables x, y, x.

But, when I write $\frac{\partial}{\partial x} f(x, y, x)$, I interpret $f(x, y, x)$ as a function that sends (x, y) to $f(x, y, x)$. **The "x" here has absolutely nothing to do with the first component of** f, which by coincidence has the same name.

If you find this awful and confusing, it may be helpful to absolutely never reuse component names as variables, e.g. you could rewrite the example as finding $\frac{\partial}{\partial s} f(s, t, s)$ and $\frac{\partial f}{\partial x}(s, t, s)$. Trying different notations such as f_x or $\left.\frac{\partial f}{\partial x}\right|_{(x,y,z)}$ etc. could also help.

Whatever path you choose, keep this awkward difference in mind as we work through the following example.

Example 13.4: Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is C^1 or something. Suppose that (x_0, y_0) is a point for which $f(x_0, y_0) = 0$ and $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$.
By the Implicit Function Theorem, there is $g(x)$ such that $f(x, g(x)) = 0$ for all x near x_0. Find a formula for $g'(x_0)$ in terms of the partials of f at (x_0, y_0).

Solution. We have $f(x, g(x)) = 0$ for all x near x_0. To avoid confusion, let's change the name of the variable. Ok, so $f(t, g(t)) = 0$ for all t near x_0, and so we are permitted to take the derivative with respect to t:

$$\frac{d}{dt} f(t, g(t)) = 0$$

Now we just apply the chain rule! Easy right?

...

If you're having trouble seeing it, don't worry. I like to view this as defining an "argument unpacking" function to help out.

Let $h(t) = (t, g(t))$. Then we're trying to compute $\frac{d}{dt} f(h(t))$. *Now* this is direct by the chain rule:

$$0 = \frac{d}{dt} f(h(t)) = \nabla f(h(t)) \cdot \frac{dh}{dt}(t)$$

Now let's compute each of these things:

- ∇f is just the matrix $\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$. We're just plugging in $h(t)$ into it, so:

$$\nabla f(h(t)) = \begin{bmatrix} \frac{\partial f}{\partial x}(h(t)) \\ \frac{\partial f}{\partial y}(h(t)) \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x}(t, g(t)) \\ \frac{\partial f}{\partial y}(t, g(t)) \end{bmatrix}$$

- If h is a function with multiple components, then $\frac{dh}{dt}$ is just the derivative of each of its components, so:

$$\frac{dh}{dt}(t) = \begin{bmatrix} \frac{d}{dt}t \\ \frac{d}{dt}g(t) \end{bmatrix} = \begin{bmatrix} 1 \\ g'(t) \end{bmatrix}$$

Taking the dot product of these two things, we finally get:

$$0 = \nabla f(h(t)) \cdot \frac{dh}{dt}(t) = \frac{\partial f}{\partial x}(t, g(t)) + g'(t)\frac{\partial f}{\partial y}(t, g(t))$$

Plug in $t = x_0$:

$$0 = \frac{\partial f}{\partial x}(x_0, y_0) + g'(x_0)\frac{\partial f}{\partial y}(x_0, y_0)$$

At last, since $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ (wow!), we can divide to obtain:

$$g'(x_0) = \boxed{-\frac{\frac{\partial f}{\partial y}(x_0, y_0)}{\frac{\partial f}{\partial x}(x_0, y_0)}}$$

■

13.4 Rest of the Warm-Up Soutions

Average of two consecutive primes: Vs gjb cevzrf ner pbafrphgvir gura gurer ner ab cevzrf orgjrra gurz. Va cnegvphyne, gurve nirentr pnaabg or cevzr.

(Ab)Normal Conversation: https://puzzling.stackexchange.com/questions/84401/a-mathematical-discussion-fill-in-the-blank

Infinite Deaf Prisoners and Real-Numbered Hats: Qrsvar na rdhvinyrapr eryngvba E ba gur frg bs nyy cbffvoyr frdhraprf bs ernyf nf sbyybjf: N E O vss gur frdhraprf N naq O ner riraghnyyl gur fnzr. Gung vf, gurl bayl qvssre va svavgryl znal cynprf. Ol Pbaprcgf, guvf rdhvinyrag eryngvba cnegvgvbaf gur frg bs nyy erny frdhraprf vagb rdhvinyrapr pynffrf.

Gur cevfbaref' fgengrtl vf gb ybbx ng gurfr rdhvinyrapr pynffrf naq nccyl gur Nkvbz bs Pubvpr gb cvpx n ercerfragngvir frdhrapr sebz rnpu pynff. Gurl zrzbevmr gurfr ercerfragngvirf. Bapr gurl ner yvarq hc naq gur jneqra nffvtaf erny ahzoref gb nyy gur ungf, rirel cevfbare xabjf juvpu rdhvinyrapr pynff gur vaqhprq frdhrapr bs erny ahzoref vf va orpnhfr rirel cevfbare pna frr nyy gur ahzoref va sebag bs gurz (be, nygreangviryl, gurl pna frr nyy ohg svavgryl znal ryrzragf bs gur frdhrapr, juvpu vfa'g vzcbegnag sbe qrgrezvavat lbhe rdhvinyrapr pynff). Rnpu cevfbare thrffrf gur erny ahzore gung gurve ung jbhyq or va va gur zrzbevmrq ercerfragngvir frdhrapr gung gur cevfbaref nterrq ba.

Fvapr gur ercerfragngvir frdhrapr bayl qvssref sebz gur npghny frdhrapr va svavgryl znal cynprf, bayl svavgryl znal cevfbaref qvr.

A while back, I received some feedback from a student concerning my recitation notes. I quote, my "recitation notes read like a Dr Seuss book", but in a "positive way".

This honestly confused me a bit, and I wasn't sure how to respond. I mulled it over for a day or two before drafting my reply.

Here is that reply.

Dank Memes and Spam by Thomas Lam

I am Thomas Lam
Lam I am

That Lam-I-am! That Lam-I-am!
I do not like that Lam-I-am!

Do you like analysis?

I do not like it, Lam-I-am.
I do not like analysis.

Would you like it here or there?

I would not like it here or there.
I would not like it anywhere.
I do not like analysis.
I do not like it, Lam-I-am.

Would you like it in Wean?
Would you like it with caffeine?

I do not like it in Wean.
I do not like it with caffeine.
I do not like it here or there.
I do not like it anywhere.
I do not like analysis.
I do not like it, Lam-I-am.

*

Would you learn it in Doherty?
Would you learn it with Leoni?

Not in Doherty. Not with Leoni.
Not in Wean. Not with caffeine.
I would not learn it here or there.
I would not learn it anywhere.
I would not learn analysis.
I do not like it, Lam-I-am.

Would you? Could you? In Qatar?
Learn it! Learn it! Here are the axioms.

I would not, could not, in Qatar.

You may like it. You will see.
You may like it after department tea!

I would not, could not after deparent tea.
Not in Qatar! You let me be.
I do not like it in Doherty. I do not like it with Leoni.
I do not like it in Wean. I do not like it with caffeine.
I do not like it here or there. I do not like it anywhere.
I do not like analysis. I do not like it, Lam-I-am.

Infimums, supremums,
fixed point theorems!
Don't you want to learn the theorem of Brouwer's?

Not Brouwer's! Not after tea!
Not in Qatar! Lam! Let me be!

I would not, could not, in Doherty.
I could not, would not, with Leoni.
I will not learn it in Wean.
I will not learn it with caffeine.

I will not learn it here or there.
I will not learn it anywhere.
I do not like analysis.
I do not like it, Lam-I-am.

Say! At a math talk?
Here at this talk!
Would you, could you,
at a talk?

I would not, could not, at a talk.

Would you, could you, at office hours?

I would not, could not, at office hours.
Not at a talk. Not with the theorem of Brouwer's.
Not in Qatar. Not after tea
I do not like it, Lam, you see.
Not in Wean. Not in Doherty.
Not with caffeine. Not with Leoni.
I will not learn it here or there.
I do not like it anywhere!

You do not like analysis?

I do not like it, Lam-I-am.

Could you, would you, learn the lemma of Fatou?

I would not, could not, learn Fatou!

Would you, could you, at an REU?

I could not, would not, at an REU.
I will not, will not, learn Fatou.
I will not learn it at office hours.
I will not learn the theorem of Brouwer's.

Not at a talk! Not after tea!
Not in Qatar! You let me be!
I do not like it in Doherty.
I do not like it with Leoni.
I will not learn it in Wean.
I do not like it with caffeine.
I do not like it here or there.
I do not like it anywhere!
I do not like analysis!
I do not like it, Lam-I-am.

You do not like it. So you say.
Try it! Try it! And you may.
Try it and you may, I say.

Lam! If you will let me be, I will try it. You will see.

...

Say! I like analysis! I do!
I like it, Lam-I-am!
And I would learn Fatou.
And I would learn it at an REU...
And I will learn it at office hours
And at a talk. And the theorem of Brouwer's.
And in Qatar. And after tea.
It is so fun, so fun, you see!
So I will learn in Doherty.
And I will learn it with Leoni.
And I will learn it in Wean.
And I will learn it with caffeine.
And I will learn it here and there.
Say! I will learn it anywhere!
I do so like analysis!
Thank you! Thank you, Lam-I-am!