

UChicago Honors Analysis Notes: 20700

Notes by Agustín Esteva, Lectures by Luis Silvestre, Books by Charles Pugh and Hubbard & Hubbard

Academic Year 2024-2025

Contents

1	Lectures	2
1.1	Monday, Sept 30: The sup Property of \mathbb{R} .	2
1.2	Wed, Oct 2: Continuity of Real-Valued Functions	5
1.3	Fri, Oct 4: Metric Spaces and Topology	7
1.4	Mon, Oct 7: Sequential Compactness	9
1.5	Wed Oct 9: Uniformly continuity and Connectedness	11
1.6	Fri, Oct 11: Covering Compactness and the Cantor Set	13
1.7	Mon, Oct 14: Differentiability	15
1.8	Wed, Oct 16: Taylor Series and Integration	17
1.9	Fri, Oct 18: Midtem I	19
1.10	Mon, Oct 21: Series	20
1.11	Wed, Oct 23: Convergence of Functions	22
1.12	Fri, Oct 25: Arzela-Ascoli	25
1.13	Mon, Oct 28: Weierstrass Approximation Theorem	27
1.14	Wed, Oct 30: Banach Contraction Principle and Picard's Theorem for ODEs	28
1.15	Fri, Nov 1: Baire Category Theorem and Consequences	30
1.16	Mon, Nov 4: The Derivative in Higher Dimensions	31
1.17	Wed, Nov 6: C^r Space	34
1.18	Fri, Nov 8: The Inverse and Implicit Function Theorem	35
1.19	Mon, Nov 11: Lagrange Multipliers	37
1.20	Wed, Nov 13: Multivariate Integration	38
1.21	Fri, Nov 15: Line Integrals and Green's Theorem	40
1.22	Fri, Nov 17: Midterm II	41
1.23	Mon, Nov 20: Surface Integrals and the Divergence Theorem	43
1.24	Fri, Nov 22: Differential Forms	44

1.25 Wed, Dec 4: Pullbacks and Stokes' Theorem	52
1.26 Fri, Dec 6: Translating Stokes, Hairy Ball Theorem, and Brower's Theorem	53
1.27 Thu, Dec 13: Final	53

1 Lectures

1.1 Monday, Sept 30: The sup Property of \mathbb{R} .

Definition 1. A **Dedekind cut** $A|B$ are two sets in \mathbb{R} such that $A, B \subset \mathbb{Q}$ and:

- $A \cap B = \emptyset$;
- $A \cup B = \mathbb{Q}$;
- $A, B \neq \emptyset$;
- If $a \in A$ and $b \in B$, then $a < b$.
- A contains no largest element.

\mathbb{R} is made up of all the Dedekind cuts. Dedekind cuts are terrible and unwieldy (see PSET 1), but at least the following theorem, a very important one, is easy to prove with them.

Theorem 1. (sup property of \mathbb{R}) Suppose that $X \subset \mathbb{R}$ is nonempty and bounded above, then $c = \sup X$ exists.

Proof. Since $X \subset \mathbb{R}$, then we let

$$X = \{A_\alpha|B_\alpha\},$$

where $A_\alpha|B_\alpha$ is the collection of cuts in X . Let $\mathcal{A} = \bigcup_\alpha A_\alpha$ and $\mathcal{B} = \bigcap_\alpha B_\alpha$. From here, it is easy to show that $\mathcal{A}|\mathcal{B} = \sup X$. \square

This is how most of the proofs go in this class, we state the main idea, and leave the verification of the idea to the reader. The ease of the above proof is, from what I can tell, why we care about Dedekind cuts. We go on to construct \mathbb{R} with them, which is a pain. Alternatively, we could have constructed \mathbb{R} with the completion of the Cauchy sequences in \mathbb{Q} , but then Theorem 1 becomes a pain to prove. In my experience, this second method is much more natural, and it should be provided as an exercise to anyone:

We now formalize this with the completeness of \mathbb{R} .

Definition 2. A sequence (a_n) is **Cauchy** if for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that if $n, m \geq N$, we have that

$$|a_n - a_m| < \epsilon.$$

Theorem 2. \mathbb{R} is **complete** in the sense that every Cauchy sequence converges to some limit in \mathbb{R} .

This was not proved in class, but it follows pretty immediately from Theorem 1:

Proof. Let $(a_n) \in \mathbb{R}$ be Cauchy and let $\epsilon > 0$. Let X be its image. Since (a_n) is Cauchy, its image must be bounded (exercise). Thus, X is bounded and nonempty. Create a new set:

$$S = \{s \in X : a_n < s \text{ finitely often}\}.$$

Let $b = \sup S$, indeed, this is by definition its limsup (see Def 3) We claim that $a_n \rightarrow b$. We have that by the sup property, there exists some $N_1 \in \mathbb{N}$ such that if $n \geq N_1$, then

$$|a_n - b| < \frac{\epsilon}{2}.$$

Let $N_2 \in \mathbb{N}$ be the natural from (a_n) 's Cauchy. Then take $N = \max\{N_1, N_2\}$ and we get that if $n \geq N$:

$$|a_n - b| \leq |a_n - a_N| + |a_N - b| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

□

In the proof, we showed that if (a_n) is Cauchy, then it must converge to its lim sup. Later we will show that if (a_n) has a convergent subsequence and it is Cauchy, then it must converge. To see a metric space that is not complete, consider \mathbb{Q} and consider the Cauchy Sequence:

$$1, 1.4, 1.41, 1.414, 1.4142, \dots$$

Which would converge to $\sqrt{2}$, but $\sqrt{2} \notin \mathbb{Q}$!!! Or maybe consider $\mathbb{R} \setminus \{0\}$ and the sequence $\frac{1}{n}$. Cauchy, but does not converge to a limit in $\mathbb{R} \setminus \{0\}$ since $0 \notin \mathbb{R} \setminus \{0\}$!!!

Example 1.1. (From Rudin). *This exercise is not that easy.*

Let X be a metric space.

- (a) Call two Cauchy sequences $\{p_n\}, \{q_n\}$ in X *equivalent* if

$$\lim_{n \rightarrow \infty} d(p_n, q_n) = 0.$$

Prove that this is an equivalence relation.

- (b) Let X^* be the set of all equivalence classes so obtained. If $P \in X^*, Q \in X^*, \{p_n\} \in P, \{q_n\} \in Q$, define

$$\Delta(P, Q) = \lim_{n \rightarrow \infty} d(p_n, q_n);$$

by Exercise 23, this limit exists. Show that the number $\Delta(P, Q)$ is unchanged if $\{p_n\}$ and $\{q_n\}$ are replaced by equivalent sequences, and hence that Δ is a distance function in X^* .

- (c) Prove that the resulting metric space X^* is complete.
- (d) For each $p \in X$, there is a Cauchy sequence all of whose terms are p ; let P_p be the element of X^* which contains this sequence. Prove that

$$\Delta(P_p, P_q) = d(p, q)$$

for all $p, q \in X$. In other words, the mapping φ defined by $\varphi(p) = P_p$ is an isometry (i.e., a distance-preserving mapping) of X into X^* .

- (e) Prove that $\varphi(X)$ is dense in X^* , and that $\varphi(X) = X^*$ if X is complete. By (d), we may identify X and $\varphi(X)$ and thus regard X as embedded in the complete metric space X^* . We call X^* the *completion* of X .

Definition 3. Let (x_n) be a sequence, then we say that

$$\limsup_{n \rightarrow \infty} x_n = \inf_{m \rightarrow \infty} \{\sup\{x_m ; m \geq n\}\} = \lim_{n \rightarrow \infty} \sup_{n \geq m} \{x_m\}$$

It immediately follows that

$$\lim_{n \rightarrow \infty} a_n = a \quad \Longleftrightarrow \quad \limsup_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} a_n = a.$$

To prove this, consider that subsequences converge to their mother sequence and for the back way, take $\frac{\epsilon}{2}$.

1.2 Wed, Oct 2: Continuity of Real-Valued Functions

Definition 4. Let X be a metric space. The following are equivalent:

- We say $f : X \rightarrow \mathbb{R}$ with $X \subset \mathbb{R}$ is **continuous** at some $x \in X$ if for all $\epsilon > 0$, there exists a $\delta > 0$ such that if $y \in X$ with $|y - x| < \delta$, we have that $|f(x) - f(y)| < \epsilon$.
- We say $f : X \rightarrow \mathbb{R}$ with $X \subset \mathbb{R}$ is continuous at some $x \in X$ if when $x_n \rightarrow x$, we have that $f(x_n) \rightarrow f(x)$.

Proof. To show these are equivalent, consider first (1) \rightarrow (2). Suppose $x_n \rightarrow x$. By the $\epsilon - \delta$ definition, we have that for any $\epsilon > 0$, there exists some $\delta > 0$ such that if $y \in X$ with $|y - x| < \delta$, then $|f(x) - f(y)| < \epsilon$. Since $x_n \rightarrow x$, we have that for large n , $|x_n - x| < \delta$, and thus $|f(x_n) - f(x)| < \epsilon$. Because this is true for any ϵ , then $f(x_n) \rightarrow f(x)$.

To show (2) \rightarrow (1), we suppose it does not fulfill the $\epsilon - \delta$ criterion. That is, for some $\epsilon > 0$, we have that for all $\delta > 0$, if $y \in X$ with $|y - x| < \delta$, we have that $|f(x) - f(y)| \geq \epsilon$. Let $\delta_n = \frac{1}{n}$, then we can build a sequence x_n such that $|x_n - x| < \frac{1}{n}$. Thus, we have that $x_n \rightarrow x$ but $f(x_n) \not\rightarrow f(x)$, a contradiction! \square

After this, we discussed the EVT and IVT. I hate these proofs, so I will not write them up. They really just amount to using Theorem 1 and the definition of continuity, but we will end up proving them later in much cooler ways. It is still nice to see that you can prove them using continuity alone without the fancy machinery.

Definition 5. Let $x, y \in \mathbb{R}^n$. The **dot product** of x and y is a scalar such that

$$(x, y) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Remark 1. The dot product is bilinear, symmetric, and positive definite. The dot product is a special case of an inner product. There exists inner products in infinite-dimensional vector spaces, such as continuous functions:

$$(f, g) = \int_a^b f(x)g(x)dx$$

Remark 2. You might see some people use $\langle \cdot, \cdot \rangle$ instead of (\cdot, \cdot) for dot products. This might seem like a good notation until you get to next quarter when it becomes terrible notation.

Definition 6. A **norm** on a vector space V is any function $\| \cdot \| : V \rightarrow \mathbb{R}$ such that for any $v, w \in V$ and $\lambda \in \mathbb{R}$, we have that

- $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$;
- $\|\lambda v\| = |\lambda| \|v\|$;
- $\|v + w\| \leq \|v\| + \|w\|$.

Remark 3. Every inner product defines a norm with $\|v\| = \sqrt{(v, v)}$, but the converse fails. For a partial converse, it might interest the reader (and it might be a good exam question (!)) to read up on how to induce the norm with the parallelogram law satisfied by any dot product: <https://math.stackexchange.com/questions/21792/norms-induced-by-inner-products-and-the-parallelogram-law>

We now come to the most important inequality of all time ever (until you get to Hölder's Inequality):

Theorem 3. (Cauchy-Schwarz Inequality) For all $x, y \in \mathbb{R}^m$, we have that

$$(x, y) \leq |x||y|.$$

Proof. Let $z = x + ty$, where $t \in \mathbb{R}$. Then we have that by bilinearity:

$$(z, z) = (x, x) + 2t(x, y) + t^2(y, y) = c + bt + at^2.$$

Since the expression must be nonnegative, then we must have a nonpositive discriminant, i.e, $b^2 - 4ac \leq 0$. Thus, we have that

$$4(x, y)^2 - 4(x, x)(y, y) \leq 0,$$

and so

$$(x, y) \leq \sqrt{(x, x)}\sqrt{(y, y)} = |x||y|.$$

□

This is a special case of the Hölder inequality, which becomes very important next quarter!

1.3 Fri, Oct 4: Metric Spaces and Topology

We stop talking about \mathbb{R} and start talking about more interesting, more general metric spaces.

Definition 7. A metric space (M, d) is a set M equipped with a metric d that satisfies the same requirements as Definition 6.

Example 1.2. • The usual Euclidean Metric on \mathbb{R}^n :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- The discrete metric on any set M :

$$d(x, y) = 1 \quad \forall x, y \in M.$$

Except when $x = y$ in which case $d(x, y) = 0$. This metric is extremely useful for coming up with counterexamples. As you will see later, every set is both open and closed here, so any function from this space is continuous.

- Let $M = \{(x_n) ; (x_i) \text{ is a bounded sequence in } \mathbb{R}\}$, then

$$d((x_n), (y_n)) = \sup |x_i - y_i|.$$

- Let $M = C^0([a, b], \mathbb{R})$ be the space of real-valued continuous functions defined on $[a, b]$. Then we define the sup-norm to be

$$d(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)|.$$

Now we start talking about topological properties of abstract metrics. For the following, let (M, d) be a metric space and suppose $X \subset M$

Definition 8. A **limit point** of X is some x such that there exists a sequence $x_n \in X$ such that $x_n \rightarrow x$.

Definition 9. We say that x is an **interior point** of X if there exists some $r > 0$ such that

$$B_r(x) \subset M,$$

where $B_r(x) := \{q \in M : d(q, x) < r\}$ is the **ball of radius r** around x .

Definition 10. We say that X is **open** if for all $x \in X$, x is an interior point of X .

Definition 11. We say that X is **closed** if X^c (the complement of X) is open. This is equivalent (prove) to saying that if x is a limit point of X , then $x \in X$.

The topological criterion for continuity is very easy to state:

Proposition 1. We say that $f : X \rightarrow Y$ is continuous if and only if for all closed $F \subset Y$, $f^{-1}(F)$ is closed in X .

I leave the simple proof as an enjoyable exercise. Hint: Use the sequential definition of continuity instead of $\epsilon - \delta$.

Definition 12. We say that $f : X \rightarrow Y$ is a **homeomorphism** if f is continuous and if $f^{-1} : Y \rightarrow X$ is continuous.

Definition 13. An **exterior point** y of X is such that there exists some $r > 0$, such that $B_r(y) \subset X^c$.

Definition 14. A **boundary point** x of X is a point such that for all $r > 0$, we have that

$$B_r(x) \not\subset X \quad B_r(x) \not\subset X^c.$$

What is the boundary of \mathbb{Q} ?

Proposition 2. • The arbitrary union of open sets is open, the arbitrary intersection of closed sets is closed.

- The finite intersection of closed sets is open, the finite union of closed sets is closed.

Proof. • Let G_α be open for all $\alpha \in \mathcal{A}$, we claim that

$$\mathcal{G} = \bigcup_{\alpha \in \mathcal{A}} G_\alpha$$

is open. Let $x \in \mathcal{G}$, then $x \in G_\alpha$ for some α , and thus there exists some $r_\alpha > 0$ such that

$$B_{r_\alpha}(x) \subset G_\alpha \implies B_{r_\alpha}(x) \subset \mathcal{G}.$$

Thus, \mathcal{G} is open. Using DeMorgan's law, let F_α be closed for all $\alpha \in \mathcal{A}$, then if

$$\mathcal{F} = \bigcap_{\alpha \in \mathcal{A}} F_\alpha \implies \mathcal{F}^c = \left(\bigcap_{\alpha \in \mathcal{A}} F_\alpha \right)^c = \bigcup_{\alpha \in \mathcal{A}} F_\alpha^c,$$

which is open by the above. Thus, \mathcal{F} is open.

- Let G_i be open for all $i \in [n]$. We claim that

$$\mathcal{G} = \bigcap_{i=1}^n G_i$$

is open. Let $x \in \mathcal{G}$, then we have that $x \in G_i$ for all $i \in [n]$, and thus for each i , there exists some $r_i > 0$ such that $B_{r_i}(x) \subset G_i$. Take the minimum of such r_i , and we obtain the result.

□

Remark 4. Consider

$$\bigcap_{i=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right).$$

We have that each set is open, but the intersection is just $\{0\}$, which is a singleton set and thus closed.

For the other counterexample, consider

$$\bigcup_{i=1}^n \left[\frac{1}{n}, 1 - \frac{1}{n}\right] = (0, 1).$$

1.4 Mon, Oct 7: Sequential Compactness

Recall that $X \subset M$ is **bounded** if there exists some $R \geq 0$ such that $S \subset B_R(x)$ for any $x \in S$. We now talk about what Charles Pugh calls “the single most important concept in Real Analysis.” He obviously has not heard about Fat Cantor sets.

Definition 15. We say that $X \subset M$ is **compact** if for any $(x_n) \in X$, there exists a subsequence (x_{n_k}) that converges to some $x \in X$.

We can go from the infinite to the finite with compactness. Its effects will echo across every page of these notes. Worship it.

Proposition 3. Compact sets are closed and bounded.

Proof. Suppose $X \subset M$ is compact. Let x be a limit point of X , then there exists some $(x_n) \in X$ with $x_n \rightarrow x$. By compactness, there exists some $(x_{n_k}) \rightarrow x$ and $x \in X$, but since subsequences converge to the same limit as their mothers, we have that $x_n \rightarrow x$. To prove bounded, we suppose it's not. Thus, let $x_0 \in X$, then create a sequence such that $d(x_n, x_0) \geq n$. But then by compactness we have that there exists some subsequence such that $d(x_{n_k}, x_0) \rightarrow d(x_n, x_0) \not\rightarrow \infty$. \square

We try (and succeed) to prove the Heine-Borel theorem in \mathbb{R} . We make heavy use of the sup property of the reals.

Lemma 1. $[a, b] \subset \mathbb{R}$ is compact.

Proof. We use Theorem 1. Let $(x_n) \in X$, and define

$$S := \{s \in [a, b] : x_n < s \text{ finitely often}\}.$$

By Theorem 1, we have that there exists some $s = \sup S$. Take $\epsilon = \frac{1}{k}$, then by the sup property, there must exist infinitely many x_n such that

$$x_n \in (s - \epsilon, s + \epsilon),$$

and thus just choose (x_{n_k}) to be so. Then we have that $d(x_{n_k}, s) < \epsilon$. \square

Lemma 2. Suppose A, B are compact, then $A \times B$ is compact.

Proof. Let $(x_n) \in A \times B$. Thus, we have that $x_n = (a_n, b_n)$ for any $n \in \mathbb{N}$. Since A is compact, we have that there exists some $(a_{n_k}) \rightarrow a$. Since B is compact, we have that there exists some $(b_{n_j}) \rightarrow b$. Sample the a_{n_k} from the n_j and sample the b_{n_j} from the n_k , then we have subsubsequences of both which converge at the same rate to (a, b) . Thus, we have created a subsequence $(x_{n_\ell}) \rightarrow (a, b) \in A \times B$. \square

The moral of the proof is that we can make convergent sequences converge at the same rate by sampling from each other.

Proposition 4. Suppose $X \subset M$ is closed and M is compact, then X is compact.

Proof. Let (x_n) be a sequence in X . Thus, $(x_n) \in M$ and since M compact, we have that $x_{n_k} \rightarrow x$ with $x \in M$. Thus, x is a limit point of X , but since X is closed, it must be that $x \in X$. \square

We now arrive to a very useful theorem, Heine(!)-Borel.

Theorem 4. (Heine-Borel). A set $X \subset M$ is compact if and only if it is closed and bounded.

Proof. We proved the forward direction in Proposition 3. For the backwards direction, we note that any bounded set fits inside of an $[a, b]$ box. Thus, since X is closed, we use proposition 4 and claim X is compact. \square

Remark 5. (Generalize Heine-Borel) A compact set is complete and **totally bounded**, i.e, for any $\epsilon > 0$, there exists a finite covering of the set by ϵ balls.

Now we get to the cool proof of the extreme value theorem, which states the continuous functions achieve their extrema.

Theorem 5. (EVT) Suppose X is compact and $f : X \rightarrow Y$ is continuous, then we have that $f(X)$ is compact.

Proof. Let $(y_n) \in f(X)$. Thus, there exists $(x_n) \in X$ such that $f(x_n) = y_n$. By compactness of X , we have that $(x_{n_k}) \rightarrow x$ with $x \in X$. Continuity carries limits, so we have that $y_n = f(x_n) \rightarrow f(x)$. Thus, $f(X)$ is compact. \square

Since $f(X)$ is compact, then it is closed and bounded. By bounded it must have a sup and and inf, and by closed those sups and infs must be in $f(X)$.

1.5 Wed Oct 9: Uniformly continuity and Connectedness

We arrive at a stronger notion of continuity.

Definition 16. We say that $f : M \rightarrow N$ is **uniformly continuous** if for any $\epsilon > 0$, there exist a $\delta > 0$ such that if $x, y \in M$ with $d(x, y) < \delta$, we have that $d(f(x), f(y)) < \epsilon$.

In words, the δ depends not on our choice of x . Uniform continuity obviously implies continuity.

Theorem 6. Suppose X is compact and $f : X \rightarrow Y$ is continuous, then f is uniformly continuous.

Proof. Suppose not, then we have that there exist some $\epsilon > 0$ such that if $\delta > 0$ and $x, y \in X$ with $d(x, y) < \delta$, then $d(f(x), f(y)) \geq \epsilon$. Let $\epsilon = \frac{1}{n}$ and let $(x_n), (y_n)$ such that $d(x_n, y_n) < \frac{1}{n}$. Then $d(f(x_n), f(y_n)) \geq \frac{1}{n}$. There exists (x_{n_k}) and (y_{n_k}) such that $d(x_{n_k}) \rightarrow x \in X$. Thus, we have that $d(y_{n_k}) \rightarrow x$. Thus, by continuity, we have that $f(x_{n_k}), f(y_{n_k}) \rightarrow f(x)$ and thus $d(f(x_n), f(y_n)) < \epsilon$. \square

Here we implicitly used another fact without realizing it was a fact. f is uniformly continuous if it sends Cauchy sequences to Cauchy sequences.

Corollary 1. Suppose $f : M \rightarrow N$ is bijective and continuous, then if M is compact, we have that f is homeomorphism.

We leave this simple proof as an exercise of applying previous results. Hint: Use the topological description of continuity. This corollary hints that compactness is a topological property. Thus, (a, b) is not homeomorphic to $[a, b]$ since one is not compact and the other is.

Definition 17. A metric space M is **disconnected** if $M = A \sqcup B$ with A, B clopen and nonempty.

Remark 6. You will see in some textbooks alternative definitions of disconnected sets. The most common one I have seen is by splitting M with A, B open. However, since A is open, then $A^c = B$ is closed, and thus B is both open and closed. Similar for A . Another common one is by just saying that there exists some nonempty $A \subset M$ clopen.

A metric space M is connected if it isn't disconnected.

Proposition 5. The interval $[a, b] \subset \mathbb{R}$ is connected.

Proof. As usual, we use the sup property of the reals. First, suppose $[a, b]$ is disconnected, then let $[a, b] = A \sqcup B$ and define

$$S := \{x : (a, x) \subset A\}.$$

Let $c = \sup S$. It takes little work to show that c is a limit point of both A and B , and thus by closedness of both, $c \in A, B$, a contradiction to their disjointness. \square

Time for the IVT!

Proposition 6. Suppose $f : M \rightarrow N$ is continuous with M connected and $N = A \sqcup B$, then either $f(N) \subset A$ or $f(N) \subset B$.

Proof. Suppose not, then take $f^{-1}(A)$ and $f^{-1}(B)$. Continuity leads to clopen, and we have nonempty by assumption. Thus, $M = f^{-1}(A) \sqcup f^{-1}(B)$. A contradiction. \square

Corollary 2. (IVT) Suppose $f : M \rightarrow N$ is continuous with M connected, then $f(M)$ is connected.

Remark 7. Any path connected set is connected, and in \mathbb{R}^n , any connected set is path connected. The only connected sets in \mathbb{R} are the intervals and \mathbb{R} .

Example 1.3. From Pugh, the topologists sine curve is a compact connected set that is not path connected. Let $M = G \sqcup Y$, with

$$G = \{(x, y) : y = \sin\left(\frac{1}{x}\right), \quad x \in (0, \frac{1}{\pi})\}$$

$$Y = \{(0, y) : y \in [-1, 1]\}$$

1.6 Fri, Oct 11: Covering Compactness and the Cantor Set

We see now a more unwieldy yet weirdly useful definition of compactness. It is more clear now how to go to the finite, but less clear to prove the previous results.

Definition 18. A set X is **covering compact** if for all open covers \mathcal{G} of X , there exists a finite open subcover.

Theorem 7. Let M be a metric space. M is sequentially compact if and only if it is covering compact.

Proof. One implication is much easier than the other.

- (\Leftarrow) Let $(x_n) \in M$ be some sequence, and assume it does not have a convergent subsequence. Thus, for all $x \in M$, there exist some $\epsilon > 0$ such that $x_n \in B_{\epsilon_x}(x)$ only finitely often. Cover M with these $B_{\epsilon_x}(x)$ balls, then by covering compactness, there exists some finite covering of M by $B_{\{\epsilon_i\}}(x_i)$, but then we still have that $x_n \in B_{\{\epsilon_i\}}(x_i)$ only finitely often. How is it that infinitely many things fit into finitely many things? The pigeonhole principle states that x_n must be in some ball infinitely often, a contradiction!
- (\Rightarrow) Let $M \subset \bigcup_{\alpha \in \mathcal{A}} G_\alpha$ be an open cover of M . Let $\rho : M \rightarrow \mathbb{R}$ such that

$$\rho(x) := \sup\{r : B_r(x) \subset G_\alpha\}.$$

We claim that ρ is continuous. To do this, convince yourself that

$$\rho(y) \geq \rho(x) - d(x, y), \quad \rho(x) \geq \rho(y) - d(x, y) \implies |d(\rho(y), \rho(x))| \leq d(x, y).$$

Theorem 5, we have that ρ achieves its minimum, and thus there exist some $x_0 \in M$ such that

$$\min \rho = \rho(x_0) = \rho_0 > 0.$$

Thus, for all $x \in M$ we have that $B_{\rho_0}(x) \subset G_\alpha$. This is what some books call a Lebesgue number. Let $p_1 \in M$, then we have that there exist some α_1 such that $B_{\rho_0}(p_1) \subset G_{\alpha_1}$. Let $p_2 \in M \setminus B_{\rho_0}(p_1)$, then there exists some α_2 such that $B_{\rho_0}(p_2) \subset G_{\alpha_2}$. If this process continues infinitely often, then for any $n, m \in \mathbb{N}$, we have that $d(p_n, p_m) \geq \rho_0$, and thus (p_n) does not have a convergent subsequence, a contradiction. Thus, we can cover M by finitely many G_{α_i} .

□

Luis gave a half hearted discussion of Cantor sets, so I will do the same. For an in depth discussion, Pugh glazes Cantor for the last part of Chapter 2 in his book, so go there.

Definition 19. We construct the middle-thirds **cantor set** by taking out intervals. Start with $C^0 = [0, 1]$, then $C^1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, i.e. $C^0 \setminus (\frac{1}{3}, \frac{2}{3})$. Continue this way by taking out middle-thirds. Then the Cantor set is

$$C = \bigcap C^n$$

We give a few extremely useful properties with zero proof for them! Do them as exercises or something.

Proposition 7. The Cantor set C has the following properties:

- Compact
- No interior (C is the boundary)

- Nowhere Dense in the sense that it contains no interval.
- Uncountable
- The Cantor set is **measure zero** (for all $\epsilon > 0$, there is a countable covering of C by open intervals (a_i, b_i) such that

$$\sum_{i=1}^{\infty} (b_i - a_i) < \epsilon$$

- C is totally disconnected in the sense that any point $c \in C$ has arbitrary small clopen neighborhoods.

1.7 Mon, Oct 14: Differentiability

We now talk about the infinitesimal!

Definition 20. Suppose $f : (a, b) \rightarrow \mathbb{R}$. We say f is **differentiable** at some $x \in (a, b)$ if

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

exists. If it exists, then we say that $f'(x)$ is the **derivative** of f at x . We say f is differentiable on (a, b) if it is differentiable for all $x \in (a, b)$.

Lemma 3. Suppose f has a max/min at some $x \in (a, b)$, then $f'(x) = 0$.

Proof. WLOG, suppose f has a local max at x . Then we have that $f(y) - f(x) \leq 0$, and thus

$$\frac{f(y) - f(x)}{y - x} \leq 0 \leq \frac{f(y) - f(x)}{y - x},$$

where $y \rightarrow x$ (hand-wavy, but that's Luis!). □

Now we get to Rolle's theorem. Important!

Theorem 8. (Rolle) Suppose $f : [a, b] \rightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) . Suppose $f(a) = f(b)$, then there exists some $\lambda \in (a, b)$ such that $f'(\lambda) = 0$.

Proof. The proof amounts to understanding that either f is constant on $[a, b]$, or else by continuity there must be a peak and/or crest somewhere on it. Since $[a, b]$ is compact, $f([a, b])$ is compact. Let $\lambda = \max f([a, b])$. If $\lambda \in (a, b)$, then by Lemma 3, we have that $f'(\lambda) = 0$. If it is not, then consider $\rho = \min f([a, b])$ and run it back. □

A very important theorem. The function achieves its derivative in some form.

Theorem 9. Suppose $f : [a, b] \rightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) . Then there exists some $\lambda \in (a, b)$ such that

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

Proof. We 'normalize' the function in some way by taking the ends and putting them on the same line. I.e., let $g : [a, b] \rightarrow \mathbb{R}$ such that

$$g(x) = f(x) - \left[f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \right].$$

Now we have that $g(a) = g(b) = 0$, and so we apply Rolle's theorem and find some $\lambda \in (a, b)$ such that

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

□

We now arrive at one of my favorite theorems for differentiability, the idea of Darboux continuity. That is, the derivative of a function never has a jump discontinuity. That is sick.

Theorem 10. Suppose $f : [a, b] \rightarrow \mathbb{R}$ is differentiable on (a, b) and continuous on $[a, b]$. If $f'(a) \leq D \leq f'(b)$, then there exists some $\gamma \in (a, b)$ such that $f'(\gamma) = D$.

Proof. Unfortunately, the proof shown in class is not too informative. The proof in Pugh chapter 3 is much better. Let $g(x) = f(x) - Dx$. Then $g'(a) = f'(a) - D \leq 0$ and $g'(b) = f'(b) - D \geq 0$, and thus since g is continuous, we must have that it achieves its minimum and thus for some $\gamma \in (a, b)$,

$$g'(\gamma) = 0 \implies f'(\gamma) = D.$$

□

We arrive at the one-dimension Inverse-Function Theorem. First, we say that $f \in C([a, b], \mathbb{R})$ if $f : [a, b] \rightarrow \mathbb{R}$ has a continuous derivative on $[a, b]$.

Theorem 11. (Inverse-Function) Suppose $f \in C([a, b], \mathbb{R})$ Suppose that for some $x \in (a, b)$, $f'(x) \neq 0$, then for some small neighborhood of x , we have that f restricted to this neighborhood is bijective and $f^{-1}(x)$ is differentiable and

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

Proof. Suppose $f'(x) \neq 0$. WLOG, suppose $f'(x) > 0$. By continuity of f' , there exists some $\epsilon > 0$ such that for any $p \in B_\epsilon(x)$, we have that $f'(p) \geq 0$. Thus, f is strictly increasing on $B_\epsilon(x)$ (prove using MVT) and is thus injective. Thus, we have found a bijection, and thus we have that $f^{-1}|_{f(B_\epsilon(x))}$ exists. Now we use the chain rule:

$$f^{-1}(f(x)) = x \implies (f^{-1})'(f(x))f'(x) = 1.$$

□

1.8 Wed, Oct 16: Taylor Series and Integration

Definition 21. We say f is **analytic** on (a, b) if for any $x \in (a, b)$, we can express $f(x + h)$ as a convergent series, where $h > 0$:

$$f(x + h) = \sum_{r=1}^{\infty} a_r h^r.$$

Definition 22. Suppose $f \in C^\infty$ (all derivatives of f are continuous), then the **Taylor Series** of f at x is

$$f(x + h) = \sum_{r=0}^{\infty} \frac{f^{(r)}(x)}{r!} h^r$$

Definition 23. Suppose $f \in C^r$. The **Taylor Polynomial** of f at x is

$$P(h) = \sum_{i=0}^r \frac{f^{(i)}(x)}{i!} h^i.$$

Theorem 12. (Taylor Remainder Theorem) Suppose $f \in C^r$, then

$$\lim_{h \rightarrow 0} \frac{f(x + h) - P_r(h)}{h^r} \rightarrow 0.$$

Suppose $f \in C^{r+1}$, then $f(x + h) - P_r(h) = \frac{f^{(r+1)}(\theta)}{(r+1)!} h^{r+1}$ for some $\theta \in (0, h)$.

Proof. We present the proof for the second statement first, making heavy use of the mean value theorem. Define a function $g : [a, b] \rightarrow \mathbb{R}$ such that

$$g(t) = f(x + t) - P(t) - R(h) \frac{t^{r+1}}{(r+1)!}.$$

Then $g(0) = 0$ (show this) and $g(h) = 0$ (convince yourself of this!). Thus, we can use MVT: there exists some $\theta_1 \in (0, h)$ such that

$$g(0) - g(h) = g'(\theta_1)(h) \implies g'(\theta) = 0.$$

Consider that $g'(h) = 0$, so use MVT to find some $\theta_2 \in (0, \theta_1)$ such that

$$g(\theta_2) = 0.$$

We can repeat this process $r + 1$ times. Then $g(\theta_{r+1}) = f^{(r+1)}(x + \theta_{r+1}) - R(h) = 0$. This proves the second statement. To prove the first statement, consider that if $f \in C^r$, then by what we showed above, we have that $f(x + h) = P_{r-1}(h) + \frac{f^{(r)}(\theta)h^r}{r!}$. Moreover, we have that $f^{(r)}$ is continuous, and thus as $\theta \rightarrow x$, we have that $f^{(r)}(\theta) - f^{(r)}(x) \rightarrow 0$.

$$\lim_{h \rightarrow 0} \frac{f(x + h) - P_r(h)}{h^r} = \lim_{h \rightarrow 0} \frac{P_{r-1}(h) + f^{(r)}(\theta) \frac{h^r}{r!} - (P_{r-1}(h) + f^{(r)}(x) \frac{h^r}{r!})}{h^r} \rightarrow 0.$$

□

I don't find these proofs entertaining. I don't find the result amusing. I don't like Taylor series if I am being honest with y'all. We introduce integrals!

Definition 24. (Riemann-Integrability) Let P, T be partition of $[a, b]$ such that if $P = \{x_i\}$ and $T = \{t_i\}$ then

$$a = x_0 \leq t_1 \leq x_1 \leq \dots x_{n-1} \leq t_n \leq x_n = b.$$

The **Riemann Sum**, is

$$R(f, P, T) = \sum_{i=1}^n f(t_i)[x_i - x_{i-1}].$$

We say a function is **Riemann-Integrable** if there exists some $I \in \mathbb{R}$ such that for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$||P|| < \delta \implies |R(f, P, T) - I| < \epsilon.$$

Here $||P||$ denotes the **mesh size** of P , i.e, the largest length between any x_i and x_{i-1} .

Definition 25. (Darboux-Integrability) Let f be bounded. Let P be a partition of $[a, b]$. The **lower sum** of f is

$$L(f, P) = \sum_{i=1}^n \inf_{[x_{i-1}, x_i]} f(t)(x_i - x_{i-1})$$

and the **upper sum** of f is

$$U(f, P) = \sum_{i=1}^n \sup_{[x_{i-1}, x_i]} f(t)(x_i - x_{i-1}).$$

We say that f is **Darboux-integrable** if for any $\epsilon > 0$, there exists a partition P such that

$$|U(f, P) - L(f, P)| < \epsilon.$$

Remark 8. Darboux-integrability and Riemann-integrability are equivalent.

This is pretty much all Luis had to say about integrals. There is one extremely useful result, which he stated without proof.

Theorem 13. (Riemann-Lebesgue) We can say that f is integrable if and only if it is bounded and if the set of discontinuities has measure zero (or equivalently, is a zero set).

1.9 Fri, Oct 18: Midtem I

We had an awesome midterm! Here it is!

Write explicit examples of the following, whenever possible:

- (a) A subset of \mathbb{R} that has no interior and no exterior points.
- (b) A subset of \mathbb{R}^2 that is homeomorphic to the open square but is not homeomorphic to the open circle.
- (c) A continuous bijective function $f : M \rightarrow N$ so that N is connected but M is not.
- (d) A continuous function $f : (0, 1) \rightarrow \mathbb{R}$ whose image is compact.
- (e) A metric space M with infinitely many points that contains exactly four subsets with empty boundary.
- (f) Two disjoint nonempty closed sets A and B in \mathbb{R} such that

$$\inf\{|x - y| : x \in A, y \in B\} = 0.$$

Proof. As a merciful overlord, we provide answers:

- (a) \mathbb{Q} .
- (b) Impossible.
- (c) $f : [0, 1) \cup \{2\} \rightarrow [0, 1]$ such that

$$f(x) = \begin{cases} x, & x \in [0, 1) \\ 1, & x = 2 \end{cases}.$$

- (d) $f(x) = \{420\}$.
- (e) $\mathbb{R} \setminus \{0\}$.
- (f) $A = \mathbb{N}$, $B = \{n - \frac{1}{n} : n \in \mathbb{N}\}$.

□

Theorem 14. Given two nonempty sets A and B and a metric space (M, d) , we define the distances between A and B as

$$d(A, B) = \inf\{d(x, y) : x \in A, y \in B\}.$$

Prove that if A and B are compact and disjoint, then $d(A, B) > 0$.

Proof. Suppose not. Thus, there exist sequences $(a_n) \in A$ and $(b_n) \in B$ such that $d(a_n, b_n) \rightarrow 0$. From compactness, we have that there exists some convergent subsequence $(a_{n_k}) \rightarrow a$ with $a \in A$. Thus, we have that $d(a_{n_k}, a) \rightarrow 0$, which implies that if we sample (b_n) with n_k , we get that $d(b_{n_k}, a) \rightarrow 0$, implying that a is a limit point of B , and thus since B is closed (Heine-Borel), we get that $a \in B$, a contradiction to the fact that the sets are disjoint. □

1.10 Mon, Oct 21: Series

We say that a sequence converges if its partial sums converge.

Proposition 8. (Comparison test) Suppose $|a_k| < b_k$ for any k . If $\sum b_k$ converges, then $\sum |a_k|$ converges. We say that $|a_k|$ **converges absolutely** here.

Proof. Using Cauchy's convergence criterion, we have that for any $\epsilon > 0$:

$$\left| \sum_{k=m}^n |a_k| \right| = \sum_{k=m}^n |a_k| \leq \sum_{k=m}^n b_k \leq \epsilon.$$

□

We come now to the integral test.

Theorem 15. (Integral test) Give $0 \leq a_k$, if $f(x) \geq a_k$ for $x \in [k-1, k]$ and $\int_0^\infty f(x)dx < \infty$, then the series converges. Suppose $g(x) \leq a_k$ for $x \in [k-1, k]$ and $\int_0^\infty g(x)dx$ diverges, then the series diverges.

Proof. This proof is not bad and amounts to showing that the partial sums are increasing and bounded. A picture is much more useful here. Look at Wikipedia or at next quarter's PSETs! □

Example 1.4. We prove that $\sum \frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \leq 1$.

Proof. We compare it to the integral

$$\int \frac{1}{x^p} = \begin{cases} \frac{x^{1-p}}{1-p}, & p \neq 1 \\ \ln(x), & p = 1 \end{cases}.$$

Thus, if $p < 1$, we have that the integral diverges. Conversely, if $p > 1$, the integral converges. □

Remark 9. As an important example, we have that for power series,

$$\sum_{k=0}^{\infty} \lambda^k = \begin{cases} \frac{1}{1-\lambda}, & |\lambda| < 1 \\ \infty, & |\lambda| > 1 \end{cases}$$

Theorem 16. (Root test) Let $\sum a_k$ be a series. The series converges absolutely if $\lambda = \limsup |a_k^{\frac{1}{k}}| < 1$. It diverges if $\lambda > 1$.

Proof. Since $\rho < 1$, we have that for k large, we have that there exists some $\beta \in \mathbb{R}$

$$\sqrt[k]{|a_k|} \leq \beta < 1 \implies |a_k| \leq \beta^k < 1.$$

$\sum \beta^k$ converges by a simple geometric series. Thus, we have that the series converges by the comparison test. It diverges if $\lambda > 1$ since the terms would not go to 0. □

We never use the ratio test since the root test is strictly stronger than it, but here it is regardless.

Theorem 17. Let $\lambda = \limsup \left| \frac{a_k}{a_{k+1}} \right|$ and $\rho = \liminf \left| \frac{a_{k+1}}{a_k} \right|$. The series converges if $\lambda < 1$ and diverges if $\rho > 1$.

Proof. We will prove for when it converges. We have that there exists some K large such that

$$\left| \frac{a_{K+1}}{a_K} \right| < 1 \implies |a_{K+1}| \leq |a_K| < 1 \implies |a_k| \leq |a_K|^{K-k},$$

the latter of which series converges by geometric series, and thus the series converges by comparison. \square

The alternating series test sucks but is lowkey important.

Theorem 18. (AST) Suppose $|a_n| \rightarrow 0$ monotonically and a_n alternates, then $\sum a_n$ converges.

Example 1.5. $\sum_{k=0}^{\infty} \frac{(-1)^k}{k}$ converges.

Definition 26. A **power series** is defined as the series

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

Theorem 19. A power series converges for all $|x| < R$, where R is the **radius of convergence** such that

$$R = \frac{1}{\limsup \sqrt[k]{|a_k|}}$$

Proof. This is just a straight up application of the ratio test. If $|x| < R$, then $|x| < \frac{1}{\limsup \sqrt[k]{|a_k|}}$. Thus, we have that applying the root test to $f(x)$, we have that

$$\limsup \sqrt[k]{|a_k x|} = |x| \limsup \sqrt[k]{|a_k|} < 1.$$

\square

1.11 Wed, Oct 23: Convergence of Functions

This is my favorite chapter in 20700, and it all has to do with spaces of functions.

Definition 27. Suppose that $f_n : X \rightarrow Y$ for any $n \in \mathbb{N}$. We say that $f_n \rightarrow f$ **uniformly** if for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that if $n \geq N$ and $x \in X$, we have that

$$d(f_n(x), f(x)) < \epsilon.$$

Remark 10. This is equivalent to convergence in the sup metric, i.e, $d(f_n, f) \rightarrow 0$ as $n \rightarrow \infty$.

We can imagine an ϵ tube around f . Does there exist some N such that for large n , we have that every f_n fits in this tube:

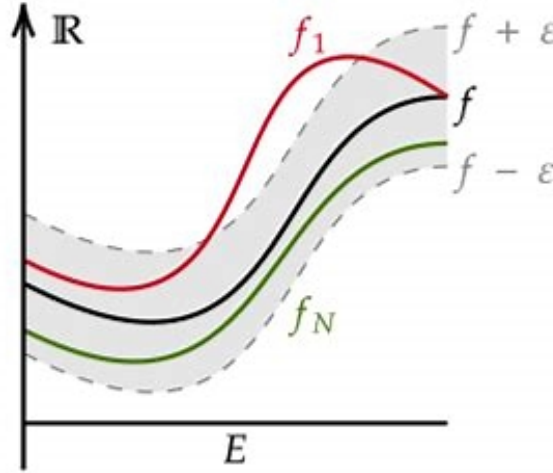


Figure 1: Epsilon-Tube

Definition 28. We say that $f_n \rightarrow f$ **pointwise** if for all $\epsilon > 0$, for each $x \in X$, there exists some $N_x \in \mathbb{N}$ such that if $n \geq N_x$, then we have that

$$d(f_n(x), f(x)) < \epsilon.$$

You can see now that uniform convergence is strictly stronger than pointwise convergence. The classic theorem to begin with is when do we have continuity of f ?

Theorem 20. Suppose f_n is continuous for all n . If $f_n \rightarrow f$ unif. then f is continuous.

Proof. Let $\epsilon > 0$ and $x \in X$. Since $f_n \rightarrow f$ uniformly, we have that there exists some $N \in \mathbb{N}$ such that if $n \geq N$, we have that

$$d(f_N(x), f(x)) < \frac{\epsilon}{3} \quad \forall x \in X.$$

Since f_N is continuous, then there exists some $\delta > 0$ such that if $y \in X$ with $d(x, y) < \delta$, then

$$d(f_N(x), f_N(y)) < \frac{\epsilon}{3}.$$

Thus,

$$d(f(x), f(y)) \leq d(f(x), f_N(x)) + d(f_N(x), f_N(y)) + d(f_N(y), f(y)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}.$$

□

Pointwise convergence is not enough to satisfy continuity as per the following example.

Example 1.6. Let $f_n : [0, 1] \rightarrow \mathbb{R}$ with $f_n = x^n$. To show that

$$\lim_{n \rightarrow \infty} f_n = \begin{cases} 0, & x \in [0, 1) \\ 1, & x = 1 \end{cases},$$

consider that if $x \in [0, 1)$, then $\lim_{n \rightarrow \infty} x^n = 0$ and if $x = 1$, then the limit is simply 1. To show that the convergence is not uniform, we must show that the N is dependent on both ϵ and x . Thus, let $\epsilon = \frac{1}{2}$, then assume there exists some N for uniform convergence. This N will then work for any x , say $x = (\frac{2}{3})^{\frac{1}{N}}$, but then

$$d(f_N(x), f(x)) = |x^N - 0| = x^N = (\frac{2}{3})^{\frac{1}{N} \cdot N} = \frac{2}{3} > \epsilon.$$

Definition 29. We say the **space of bounded functions** is C_b and the **space of continuous functions** is $C^0 = C^0([a, b], \mathbb{R})$.

Remark 11. Note that by the EVT, $C^0 \subset C_b$. Also note that by Theorem 20, if f is a limit point of C^0 , then we have a sequence of continuous functions converging (with respect to the sup metric, i.e., uniform convergence) to f implying that f is in fact continuous. Thus, C^0 is closed.

Theorem 21. C_b is complete.

Proof. Let (f_n) be Cauchy in C_b . For any x in the domain, we have that $(f_n(x))$ is a bounded Cauchy sequence in \mathbb{R} , and thus must converge. For each x , we let $\lim_{n \rightarrow \infty} f_n(x) = f(x)$. We claim that the convergence is also uniform! Let $\epsilon > 0$. By Cauchyness, we have that there exists some N such that if $n, m \geq N$, we have

$$d(f_n, f_m) < \frac{\epsilon}{2}.$$

By pointwise convergence, exists some $m \geq N$ dependent on x such that

$$d(f_{m(x)}, f(x)) < \frac{\epsilon}{2}.$$

Thus, we use the N from Cauchy and we know that if $n \geq N$, we have that:

$$d(f_n(x), f(x)) \leq d(f_n(x), f_{m(x)}(x)) + d(f_{m(x)}, f(x)) < \epsilon.$$

□

Note that we are using a different m for each x in this proof, but it doesn't matter at the end of the day because we know that given some N , then for any x , we can find this m by the pointwise convergence.

Corollary 3. C^0 is complete.

For this one-liner, use Remark 10 and Theorem 21.

What else does uniform convergence imply?

Theorem 22. Suppose $f_n \rightarrow f$ uniformly with each f_n integrable. Then f is integrable. Moreover,

$$\int f_n \rightarrow \int f, \quad \text{uniformly}$$

The spirit of the proof is that we can bound the upper and lower sums of f by upper and lower sums of f_n . Pugh presents a cleaner proof using the Riemann-Lebesgue Integrable Criterion:

Proof. Since each f_n is integrable, then each f_n has countably many discontinuities. Call this set D_n . Then each f_n is continuous for any $x \in [a, b] \setminus D_n$, and thus Theorem 20 states that f must be continuous for any $x \in [a, b] \setminus \bigcup D_n$. We have that f is bounded by Theorem 21. Thus by the Riemann-Lebesgue Theorem, f is integrable.

Let $\epsilon > 0$. From uniform convergence, there exists some N such that if $n \geq N$, we have that $d(f_n, f) < \frac{\epsilon}{b-a}$

Get the N from the uniform convergence of f_n , then

$$\left| \int_a^x f_n(t) dt - \int_a^x f(t) dt \right| = \left| \int_a^x (f_n(t) - f(t)) dt \right| \leq \int_a^x |f_n(t) - f(t)| dt \leq d(f_n, f)(b-a) < \epsilon.$$

□

Luis forgot (didn't care) to mention the Weierstrass-M test, but here it is. It is a useful criteria for deciding when a power series converges.

Theorem 23. (Weierstrass-M test) Let $(f_n) \in C_b$. Suppose that for each n , $\|f_n\| \leq M_n$, where M_n forms a convergent series. Then $\sum |f_n|$ converges uniformly.

Proof. We use our most powerful theorem, C_b is complete. We must show that the series is Cauchy, or better, that the partial sums, $F_n = \sum_{k=0}^n f_k$ are Cauchy. Suppose $m > n$, then we have that

$$d(F_n, F_m) \leq d(F_m, F_{m-1}) + \cdots + d(F_{n+1}, F_n) = \sum_{k=n+1}^m \|f_k\| \leq \sum_{k=n+1}^m M_k.$$

□

Looking back on this with the foresight of functional analysis, the M-test reminds me a ton of the Banach-Steinhaus Uniform Boundedness Principle (foreshadowing).

1.12 Fri, Oct 25: Arzela-Ascoli

The second half of the lecture fucks. Persevere!

Theorem 24. Suppose f_n is differentiable for all n . Suppose $f_n \rightarrow f$ uniformly and f'_n converges uniformly. Then f is differentiable and moreover $f'_n \rightarrow f'$ uniformly.

Proof. We weaken the statement and assume $f_n \in C^1$ just to make the proof enjoyable. We apply FTC and Theorem 22. Note that all the following convergences are uniform.

$$f_n(x) = f_n(a) + \int_a^x f'_n(t)dt \rightarrow f(a) + \int_a^x g(t)dt \implies f'(x) = g(x).$$

We leave the proof of the actual statement to the gods. □

We can now use the Weierstrass-M test!

Proposition 9. A power series converges uniformly on $[-r, r]$, where $|r| < R$ is the radius of convergence.

Proof. Let $\sum c_k x^k$ be a power series. We have that for large k , $\frac{1}{\sqrt[k]{c_k}} < \beta < R$, and thus we have that for large k :

$$x^k c_k \leq \left(\frac{x}{\beta}\right)^k,$$

where the right side forms a geometric power series. Thus, we have our result by Weierstrass-M. □

Now we get to the good stuff. First, a definition.

Definition 30. Let \mathcal{E} be a family of functions. We say \mathcal{E} is equicontinuous if for any $\epsilon > 0$, there exists some $\delta > 0$ such that if $d(x, y) < \delta$ and $f \in \mathcal{E}$, then

$$d(f(x), f(y)) < \epsilon.$$

To make sense of this definition, think of it as uniform continuity of a bunch of functions. There is some δ that determines the continuity of every function in the family.

Example 1.7. Suppose $f_\alpha : [a, b] \rightarrow \mathbb{R}$ and $\|f_\alpha\| < A$ for any α . Then we have by the MVT that if $\alpha \in \mathcal{A}$ and $x, y \in [a, b]$ then

$$|f_\alpha(x) - f_\alpha(y)| \leq |f'(\lambda)| |x - y| \leq A |x - y|.$$

Thus, the family is equicontinuous since they are governed by a **modulus of continuity**.

Here is a very very important theorem with a very very cool name.

Theorem 25. (Arzela-Ascoli) Any bounded equicontinuous sequence of functions in $C^0([a, b], \mathbb{R})$ has a uniformly convergent subsequence.

Proof. Since $[a, b]$ is compact, there exists some countably dense subset $D \subset [a, b]$ (exercise). Let $D = \{d_1, d_2, \dots\}$. Consider that $(f_n(d_1))$ forms a bounded sequence of reals, and thus by Bolzano-Weierstrass must have a convergent subsequence $(f_{n,1}(x))$ which converges at d_1 . Consider that $(f_{n,1}(d_2))$ forms a bounded sequence of reals, and must have a convergent subsequence $(f_{n,2}(x))$ which converges at d_2 . However, since

this subsequence is a subsequence of a sequence which converges at d_1 , then it also converges at d_1 . Keep going in this way, then consider the sequence $(f_{i,i})(x)$. That is, we are consider the following diagonal sequence:

$$\begin{array}{ccccccc} f_{1,1}(x) & f_{2,1}(x) & f_{3,1}(x) & f_{4,1}(x) & f_{5,1}(x) & \cdots & \\ & f_{2,2}(x) & f_{3,2}(x) & f_{4,2}(x) & f_{5,2}(x) & \cdots & \\ & & f_{3,3}(x) & f_{4,3}(x) & f_{5,3}(x) & \cdots & \\ & & & f_{4,4}(x) & f_{5,4}(x) & \cdots & \\ & & & & f_{5,5}(x) & \cdots & \\ & & & & & \ddots & \end{array}$$

For any $d_k \in D$, $f_{i,i}(x)$ converges at d_k . Thus, we have that for any d_k , there exists some large N such that if $i, j \geq N$, we have that

$$d(f_{i,i}(d_k), f_{j,j}(d_k)) < \frac{\epsilon}{3}.$$

We have that each $f_{i,i}$ is continuous on a compact set (and thus uniformly continuous), and so there exists some $\delta > 0$ such that if

$$d(x, y) < \delta \implies d(f_{i,i}(x), f_{i,i}(y)) < \frac{\epsilon}{3}.$$

For any $x \in [a, b]$, there exists some $d_k \in D$ such that $d(x, d_k) < \delta$. Thus, for any $\epsilon > 0$, there exists this δ such that if $x, y \in [a, b]$, then

$$d(f_{i,i}(x), f_{j,j}(x)) \leq d(f_{i,i}(x), f_{i,i}(d_k)) + d(f_{i,i}(d_k), f_{j,j}(d_k)) + d(f_{j,j}(d_k), f_{j,j}(y)) \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}.$$

Again, the first and last distances in the inequality above are bounded due to uniform continuity of the individual f_n , and the middle one is bounded from equicontinuity.

We have that $(f_{i,i})$ is Cauchy, and thus converges since C^0 is complete. □

This theorem has countless consequences, but the most important is the following.

Theorem 26. (Heine-Borel) $\mathcal{E} \subset C^0$ is compact if and only if it is equicontinuous, closed, and bounded.

Proof. We prove both directions.

- (\implies) Suppose \mathcal{E} is compact. Closed and bounded are immediate. We can find a finite $\epsilon - 3$ dense subset of \mathcal{E} by compactness (exercise). Let $\{f_{k_{\frac{\epsilon}{3}}} : k \in [n]\}$ be such a subset. Thus, if $f \in \mathcal{E}$, then there exists some f_k such that $d(f, f_k) < \frac{\epsilon}{3}$. Suppose $x, y \in [a, b]$ and $d(x, y) < \delta$, where δ is the one from uniform continuity, then

$$d(f(x), f(y)) \leq d(f(x), f_k(x)) + d(f_k(x), f_k(y)) + d(f_k(y), f(y)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}.$$

Again, the middle distance is bounded by uniform continuity, and outer ones by the dense subset.

- (\impliedby) Let $(f_n) \in \mathcal{E}$. By equicontinuity and boundedness, Arzela-Ascoli tells us that we have a convergent subsequence (f_{n_k}) and closedness tells us the limit lies within \mathcal{E} . □

Remark 12. f_n converges uniformly if and only if f_n are equicontinuous and converge pointwise.

The proof of the remark is an $\epsilon/3$ mess, but try it out.

1.13 Mon, Oct 28: Weierstrass Approximation Theorem

We discuss convolutions and the Weierstrass Approximation Theorem. It is a cool result with an unfortunately not cool (in my opinion) proof.

Definition 31. A **convolution** is defined by

$$f * g = \int_{\mathbb{R}} f(x-y)g(y)dy = \int_{\mathbb{R}} g(x-y)f(y)dy.$$

Remark 13. Suppose $f \in C^1$ and g have support (set where g outputs nonzero values) on $[a, b]$. Then

$$(f * g)'(x) = \lim_{h \rightarrow 0} \frac{1}{h} \int f(x+h-y)g(y)dy - \int f(x-y)g(y)dy = \int f'(x-y)g(y)dy = (f' * g)(x)$$

Proposition 10. The convolution of a not smooth function can be made smooth.

I apologize for this, Luis was pretty bad with it and I don't care enough about convolutions to fix his lecture.

Proof. Let $f : [a, b] \rightarrow \mathbb{R}$ be smooth with $f(a) = f(b) = 0$ with support on (a, b) . Let (g_n) such that g_n has support on $[-\frac{1}{n}, \frac{1}{n}]$ and $\int_{\mathbb{R}} g_n(y)dy = 1$. Then

$$\begin{aligned} |(f * g_n)(x) - f(x)| &= \int_{\mathbb{R}} |f(x-y) - f(x)| g_n(y)dy \\ &= \int_{-\delta}^{\delta} |f(x-y) - f(x)| g_n(y)dy + \int_{\mathbb{R} \setminus [-\delta, \delta]} |f(x-y) - f(x)| g_n(y)dy \\ &\leq \epsilon + C \int_{\mathbb{R} \setminus [-\delta, \delta]} g_n dy \rightarrow 0. \end{aligned}$$

□

Theorem 27. (Weierstrass Approximation Theorem) Polynomials are dense in $C^0([a, b], \mathbb{R})$.

The proof follows from the above and also the fact that if P is a polynomial, then $f * P$ is a polynomial.

1.14 Wed, Oct 30: Banach Contraction Principle and Picard's Theorem for ODEs

This lecture is awesome sauce, so strap in.

Definition 32. We say $f : M \rightarrow M$ is a **contraction** if there exists some $k < 1$ such that for any $x, y \in M$:

$$d(f(x), f(y)) \leq kd(x, y).$$

Definition 33. We say f has a **fixed point** at some x if $f(x) = x$.

Theorem 28. (Brouwer) Suppose $f : B_1(0) \rightarrow B_1(0)$ is continuous, then f has a fixed point.

The proof for this will hopefully be on the last page of this document. We turn to another fixed-point theorem, which has an awesome proof.

Theorem 29. (Banach) Suppose that M is complete and $f : M \rightarrow M$ is a contraction. Then f has a fixed point.

The idea for this proof is that eventually enough contraction results in a single point.

Proof. Let $x_0 \in M$. Let $x_1 = f(x_0)$, $x_2 = f(f(x_0))$, and $x_n = f^{(n)}(x_0)$. We wish to show that (x_n) is Cauchy. Consider that for any n :

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq kd(x_{n-1}, x_n) \leq k^n d(x_0, x_1).$$

Thus, if $n > m$, we have that

$$d(x_n, x_m) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{m-1}, x_m) \leq k^n d(x_0, x_1) + k^{n+1} d(x_0, x_1) + \dots + k^{m-1} d(x_0, x_1) = d(x_0, x_1) \left(\frac{k^n}{1-k} \right).$$

Since $k < 1$, we have that as $n \rightarrow \infty$, $d(x_n, x_m) \rightarrow 0$, and thus (x_n) is Cauchy, and by the completeness of M , we have $(x_n) \rightarrow x_\infty \in M$. We claim that $f(x_\infty) = x_\infty$. That is, x_∞ is our fixed point. To show this, consider that

$$d(f(x_\infty), x_\infty) \leq d(f(x_\infty), f(x_n)) + d(f(x_n), x_n) + d(x_n, x_\infty) \leq kd(x_\infty, x_n) + kd(x_\infty, x_n) + k^{n-1} d(x_0, x_1) + d(x_\infty, x_n) \rightarrow 0.$$

□

How can we guarantee the existence and uniqueness for a solution of an ODE?

Theorem 30. (Picard) Suppose $F : \Omega \rightarrow \mathbb{R}^d$ with Ω open is Lipschitz. Suppose $x_0 \in \Omega$, then there exists some unique trajectory $\gamma : (a, b) \rightarrow \Omega$ that if $\gamma(0) = x_0$, then $\gamma'(t) = F(\gamma(t))$.

Proof. Solving the ODE in the problem statement is equivalent to finding the solution to

$$\gamma(t) = x_0 + \int_0^t F(\gamma(s)) ds.$$

We can only find this solution locally, so restrict F to some compact $\overline{B_r(x_0)}$. By continuity of F , we have that F is locally bounded here by some M . Choose τ small such that

$$\tau M \leq r, \quad \tau L < 1.$$

Define $\mathcal{M} : \mathbb{R} \rightarrow \mathbb{R}$, where

$$\mathcal{M} = \{f : [-\tau, \tau] \rightarrow \Omega : f \text{ is continuous}\}.$$

With respect to d_{sup} , \mathcal{M} is complete. Define m by

$$m(\gamma(t)) = p + \int_0^t F(\gamma(s))ds.$$

By Banach's Contraction principle, it suffices to show that m is a contraction and it maps functions to \mathcal{M} . Let $\sigma, \gamma \in \mathcal{M}$, then we have that

$$d(m(\sigma), m(\gamma)) = \sup_t \left| \int_0^t F(\gamma(s)) - F(\sigma(s))ds \right| \leq \tau \sup_t |F(\gamma(t)) - F(\sigma(t))| \leq \tau L d(\gamma(t), d(\sigma(t))).$$

Since $\tau L < 1$, we have that m is a contraction. Suppose $\gamma \in \mathcal{M}$. Then we have that

$$m(\gamma(t)) = x_0 + \int_0^t F(\gamma(s))ds \leq x_0 + \tau M \implies d(x_0, m(\gamma(t))) \leq r \implies m(\gamma) \in \overline{B_r(x_0)}.$$

Thus, we have the existence of some γ that satisfies the conditions. □

Suppose we have a river (Ω) and some rock thrown at some point in the river ($x_0 \in \Omega$). Picard guarantees us that if the velocity/flow of the river is Lipschitz, then we can put some boat where the rock landed who's path ($\gamma(t)$) will move in the flow of the river in some time interval $t \in [-\tau, \tau]$.

Example 1.8. Suppose $f(x) = x^2$. Then f is not global Lipschitz, and so we have the consequence that the solution blows up in local time. I.e, we can find a solution in local but not global time:

$$\gamma'(t) = (\gamma(t))^2 \implies \gamma(t) = \frac{1}{2-t},$$

which blows up near $x = 2$.

Example 1.9. Suppose $f(x) = \sqrt{x}$, then $\gamma(t) = 0$ or $\gamma(t) = \frac{1}{4}t^2$. However, since we do not have Lipschitz, we also have that

$$\gamma(t) = \begin{cases} \frac{(x-a)^2}{4}, & x > a \\ 0, & x < 0 \end{cases}$$

and so our solution is not unique!

1.15 Fri, Nov 1: Baire Category Theorem and Consequences

The best theorem of all time. If I were a lemma I would go out with Baire Category Theorem, even though I hardly know him!

Theorem 31. Let (G_n) be a countable collection of open dense sets in M , a complete metric space. Then

$$\bigcap G_n$$

is dense in M .

Proof. Let $p \in M$ and $\epsilon > 0$. Since $B_\epsilon(p_0)$ is open, then there exists some $r_0 > 0$ and $p_1 \in M$ such that

$$\overline{B_{r_0}(p_1)} \subset B_\epsilon(p_0).$$

Since G_1 is open, there exists some $r_1 > 0$ and $p_2 \in M$ such that

$$\overline{B_{r_1}(p_2)} \subset B_{\frac{r_0}{2}} \cap G_1,$$

where the intersection is nonempty by the density of G_1 . Make the sequence (p_n) such that

$$\overline{B_{r_{n-1}}(p_n)} \subset B_{\frac{r_{n-2}}{2}} \cap G_1 \cap G_2 \cdots \cap G_{n-1}$$

We have that $r_n \rightarrow 0$ since $r_{n+1} \leq \frac{r_n}{2}$ and so

$$\lim_{n \rightarrow \infty} \text{diam} \overline{B_{\frac{r_n}{2}}(p_n)} \rightarrow 0,$$

implying that (p_n) is Cauchy. Moreover, we have that each $(p_n) \in B_\epsilon(p_0)$. Since M is complete, we have that $(p_n) \rightarrow p_\infty \in M$, and we must also have that $p_\infty \in B_\epsilon(p_0)$ and also that for each n , $p_\infty \in G_n$. Thus, for any $p_0 \in M$, and $\epsilon > 0$, there exists some $p_\infty \in M$ such that

$$p_\infty \in B_\epsilon(p_0) \cap \bigcap G_n,$$

implying that the intersection is dense in M . □

Corollary 4. Suppose M is complete and

$$M = \bigcap F_n$$

where each F_n is closed. Then at least one F_i must have some nonempty interior.

Proof. Suppose not. That is, suppose we can write

$$M = \bigcup F_n$$

where F_n closed has an empty interior for any n . Thus, we have that

$$\emptyset = M^c = (\bigcup F_n)^c = \bigcap F_n^c$$

is an intersection of open dense sets (the compliment of empty interior is a dense set). Thus, we have by Baire's theorem that \emptyset must be dense in M . Ridiculous! □

Baire Category Theorem allows us to prove some pretty insane stuff. Luis of course didn't prove them, but they are still cool!

Theorem 32. The set of continuous and nowhere-differentiable functions is general in C^0 . That is, this set is dense in C^0 .

Sketch: define

$$G_n := \{f \in C^0([a, b], \mathbb{R}) : \forall x \in [a, b - \frac{1}{n}], \exists y > x \text{ s.t. } |f(y) - f(x)| \geq n|x - y|\}.$$

We claim that G_n is open and that G_n is dense in C^0 .

1.16 Mon, Nov 4: The Derivative in Higher Dimensions

So get ready for the worst thing in analysis. In 1-D, we defined the derivative of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ to be the limit

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = f'(x) \iff f(x+h) = f'(x)h + f(x),$$

where we take the right hand in the limit as $h \rightarrow 0$.

Definition 34. The **total derivative** of $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ at $x \in \mathbb{R}^n$ if it exists, is the limit linear transform Df_x such that

$$f(x+h) = f(x) + Df_x h + R(h), \quad \lim_{|h| \rightarrow 0} \frac{|R(h)|}{|h|} \rightarrow 0,$$

where $||$ is the vector norm and $R(h)$ is an error term.

Since Df_x is a linear operator, it will be useful to talk about some linear algebra. In particular, let's define the stretch a linear transform T applies to vectors.

Definition 35. Suppose $T : V \rightarrow W$ is a linear transformation, we define the **operator norm** of T to be

$$||T|| = \sup_{v \in V} \frac{|T(v)|_W}{|v|_V},$$

where $|\cdot|_W$ and $|\cdot|_V$ are the norms of vectors in V and W .

You might have seen it with different definitions, we remark that this is equivalent to saying that

$$||T|| = \sup_{v \in V} |T(v)|.$$

In particular, we have the following theorem:

Theorem 33. The following are equivalent:

- (a) $||T||$ is finite.
- (b) T is uniformly continuous.
- (c) T is continuous at the origin (or any point in V for that matter).

Proof. Let $\epsilon > 0$. There exists a $\delta = \frac{\epsilon}{||T||}$ such that if $v, v' \in V$ with $|v - v'| < \delta$, then

$$|T(v) - T(v')| = |T(v - v')| \leq ||T|| |v - v'|$$

Two obviously proves three. Suppose T is continuous at 0. Let $v \in V$. Then let $\epsilon = 1$, and so we have that that for any $\delta > 0$, if $|u| < \delta$, then

$$|T(u)| \leq 1.$$

Thus, if we let $u = \frac{\delta v}{2|v|}$, then

$$\frac{|T(v)|}{|v|} = \frac{T(u)}{|u|} \leq \frac{2}{\delta} < \infty.$$

□

Moreover, we have the following theorem:

Theorem 34. Suppose W is a finite dimensional vector space, then any linear transform $T : \mathbb{R}^n \rightarrow W$ is continuous. In fact, if it is an isomorphism then it is a homeomorphism.

Definition 36. We define the **partial derivative** of $f_j : \mathbb{R}^n \rightarrow \mathbb{R}$ at x to be

$$(Df_j)_x(e_i) = D_i f_j(x) = \partial_i f_j(x) = \frac{\partial f_j}{\partial x_i}(x) = \lim_{|h| \rightarrow 0} \frac{f_j(x_1, x_2, \dots, x_i + h, \dots, x_n) - f_j(x)}{|h|} = \lim_{|h| \rightarrow 0} \frac{f_j(p + te_i) - f_j(p)}{|h|}.$$

Remark 14. In high school calculus textbooks, the total derivative of f at x would be denoted by the $m \times n$ matrix of partials:

$$Df_x = \begin{bmatrix} D_1 f_1 & D_2 f_1 & \cdots & D_n f_1 \\ D_1 f_2 & D_2 f_2 & \cdots & D_n f_2 \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f_m & D_2 f_m & \cdots & D_n f_m \end{bmatrix}.$$

Theorem 35. If every $\partial_i f_j(x)$ exists and is continuous for all $i \in [n]$ and $j \in [m]$, then f is differentiable.

This proof is severely unenjoyable and I refuse to type it up for now.

Proposition 11. The total derivative satisfies the following properties:

- (a) $D(f + g) = Df + Dg$
- (b) $D(f \circ g) = Df \circ Dg$
- (c) $D(f \cdot g) = Df \cdot g + f \cdot Dg$ (where \cdot is the dot product).

We give the best possible version of the Mean Value Theorem we can.

Theorem 36. Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable on U and $[p, q] \subset U$, then

$$|f(q) - f(p)| \leq \sup_{x \in U} Df_x |q - p|.$$

Proof. Let $u \in \mathbb{R}^n$ be the unit vector, then if $t \in [0, 1]$, we get that

$$g(t) = \langle u, f(p + t(q - p)) \rangle$$

is differentiable and real valued, and thus we use the 1-D MVT to get some $\theta \in [0, 1]$ such that

$$g(1) - g(0) = \langle u, f(q) - f(p) \rangle = g'(\theta)$$

and using Proposition 11, we get that

$$g'(\theta) = \langle u, Df_{p+\theta(q-p)}(q - p) \rangle \leq \sup_{x \in U} Df_x |q - p|.$$

Thus, we have that it has norm less than the desired quantity. □

We have that Df is a linear transform $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$. The second derivative $D^2 f$ is a function $D^2 f : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R}^m)$, where L is the space of linear transforms like T . Thus, v, w can take in two vectors $D^2(v)(w)$ and output a vector in \mathbb{R}^m .

Theorem 37. If $f : U \rightarrow \mathbb{R}^m$ is C^1 and if the segment $[p, q] \subset U$, then

$$f(q) - f(p) = T(q - p),$$

where T is the average derivative

$$T = \int_0^1 (Df)_{p+t(q-p)} dt.$$

Definition 37. Assume $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is second differentiable, then we say that D^2f_x is the **Hessian** of f . That is, the matrix of second partial derivatives.

Theorem 38. The Hessian is symmetric. That is, for any $v, w \in \mathbb{R}^n$ we have that

$$(D^2f)_x(v)(w) = (D^2f)_p(w)(v).$$

Remark 15. This does imply that partial double derivative are symmetric since

$$(D^2f_x)(e_i)(e_j) = \frac{\partial f_k}{\partial x_i \partial x_j} = \frac{\partial f_k}{\partial x_j \partial x_i} = (D^2)_x(e_j)(e_i)$$

Have fun proving this one.

1.17 Wed, Nov 6: C^r Space

Definition 38. We say $f \in C^r(U, \mathbb{R}^m)$ if the r th derivative of f exists and is continuous.

Definition 39. The C^r norm of a function $f \in C^r$ is defined by

$$\|f\|_{C^r} = \max\{\sup_{x \in U} |f(x)|, \dots, \sup_{x \in U} |D^{(r)}(x)|\}.$$

Definition 40. We say that a series f_k is C^r **uniformly convergent** to some f ($f_k \rightarrow f$ C^r unif) if

$$f_k \rightarrow f \quad \text{unif.}$$

$$(Df_k) \rightarrow Df \quad \text{unif}$$

$$\vdots$$

$$(D^{(r)}f_k) \rightarrow D^{(r)}f \quad \text{unif}$$

Definition 41. We say that a series f_k is C^r **Cauchy** if for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that if $k, \ell > N$, then

$$\|f_k - f_\ell\| < \epsilon$$

$$\|(Df_k) - (Df_\ell)\| < \epsilon$$

$$\vdots$$

$$\|(D^{(r)}f_k) - (D^{(r)}f_\ell)\| < \epsilon$$

Theorem 39. C^r is a **Banach Space**, i.e, a complete normed vector space.

Proof. Let f_k be C^r Cauchy. Since C^0 is complete, we have that $f_k \rightarrow f$ uniformly. We also know that $(Df_k) \rightarrow G$ uniformly. We claim that $(Df_k) \rightarrow (Df)$ uniformly (i.e, $(Df_k) = G$). Let $p \in U$ and consider $q \in U$ such that $[p, q] \subset U$. By Theorem 37, we have that

$$f_k(q) - f_k(p) = \int_0^1 (Df_k)_{p+t(q-p)} dt (q-p) \rightarrow f(q) - f(p) = \int_0^1 G(p+t(q-p)) dt (q-p).$$

Proceed by induction. □

1.18 Fri, Nov 8: The Inverse and Implicit Function Theorem

Both Pugh and Luis have godawful explanations for this. I defer to the Youtube Channel “The Bright Side of Mathematics.”

Theorem 40. Let $U, V \subset \mathbb{R}^n$ and $f \in C^1(U, V)$ and $x_0 \in U$. If $(Df)_{x_0}$ is invertible, then f is a local C^1 diffeomorphism at x_0 .

Proof. (STEP 1) If $y_0 = f(x_0)$, then define $\hat{f} = f - y_0$ and define $g(x) = \hat{f}(x + x_0)$. Thus, we have that $\hat{f}(0) = 0$ and $(Df)_0$ is invertible. Define $h(x) = (Df)_0^{-1}g(x)$, then $(Dh)_0 = \mathbf{I}$.

(STEP 2) Now that we have normalized the problem, we can begin. Let $z(x) = h(x) - x$, then $(Dz)_0 = \mathbf{0}$. Since $z \in C^1$, then by continuity we choose $\epsilon > 0$ such that

$$\left| \frac{\partial f_i(x)}{\partial x_j} \right| \leq \frac{1}{2} \frac{1}{n^2}, \quad \forall i, j \in [n], \quad \forall x \in \overline{B_\epsilon(0)}$$

and $(Dh)(x)$ is invertible for all $x \in B_\epsilon(0)$.

(STEP 3) Now that we have restricted our function to a convex ball, we can use the Mean Value Theorem when we define $a : [0, 1] \rightarrow \mathbb{R}$ by

$$a(t) = z_i(p + t(q - p)),$$

where $p, q \in B_\epsilon(0)$. Thus, we have that if

$$z(q) - z(p) = a(1) - a(0) = a'(\theta) = \langle \nabla z_i(p + \theta(q - p)), p - q \rangle.$$

Thus, we have by Cauchy-Schwartz that

$$|z(q) - z(p)| \leq \|\nabla z_i(p + \theta(q - p))\| \|p - q\| = \sum_j \left| \frac{\partial z_i}{\partial x_j} \right| \|q - p\| \leq \frac{1}{4} \frac{n}{n^4} \|q - p\|$$

Thus, we have that

$$\|z(q) - z(p)\|^2 = \sum_i |z_i(q) - z_i(p)|^2 \leq \frac{1}{16} \frac{1}{n^2} \|p - q\|^2,$$

thus, z is a contraction since

$$\|z(q) - z(p)\| \leq \frac{1}{2} \|q - p\|.$$

(STEP 4) Fix an element $y \in \overline{B_\epsilon(0)}$ and define $z_y : \overline{B_\epsilon(0)} \rightarrow \overline{B_\epsilon(0)}$ by

$$z_y(x) = y - z(x).$$

Thus, we have that

$$\|z_y(q) - z_y(p)\| \leq \frac{1}{2} \|q - p\|,$$

and by Banach's fixed point theorem, there exists some unique fixed point x such that

$$z_y(x) = x \iff y - z(x) = x \iff y - (h(x) - x) = x \iff y = h(x).$$

Thus, we have found a bijection in this closed ϵ ball.

The rest of the proof is left up to the reader. □

Remark 16. We have a nice formula for the inverse, that being

$$(Dh^{-1})_y = ((Dh)_{h^{-1}(y)})^{-1}$$

Theorem 41. Suppose $U \subset \mathbb{R}^n \times \mathbb{R}^m$ is open, and $F : U \rightarrow \mathbb{R}^m$ is C^1 . Let $z_0 = (x_0, y_0) \in U$ with $F(z_0) = 0$ and

$$DF_{z_0} = \begin{bmatrix} \frac{\partial F_1(z_0)}{\partial x_1} & \frac{\partial F_1(z_0)}{\partial x_2} & \dots & \frac{\partial F_1(z_0)}{\partial x_n} & \frac{\partial F_1(z_0)}{\partial x_{n+1}} & \dots & \frac{\partial F_1(z_0)}{\partial x_m} \\ \frac{\partial F_2(z_0)}{\partial x_1} & \frac{\partial F_2(z_0)}{\partial x_2} & \dots & \frac{\partial F_2(z_0)}{\partial x_n} & \frac{\partial F_2(z_0)}{\partial x_{n+1}} & \dots & \frac{\partial F_2(z_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m(z_0)}{\partial x_1} & \frac{\partial F_m(z_0)}{\partial x_2} & \dots & \frac{\partial F_m(z_0)}{\partial x_n} & \frac{\partial F_m(z_0)}{\partial x_{n+1}} & \dots & \frac{\partial F_m(z_0)}{\partial x_m} \end{bmatrix} = \begin{bmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{bmatrix}.$$

If $\frac{\partial F}{\partial y} \neq \mathbf{0}$, then there exists open $V_1 \subset \mathbb{R}^n$ and $V_2 \subset \mathbb{R}^m$ such that $z_0 \in V_1 \times V_2$ and a map $g \in C^1(V_1, V_2)$ such that

$$g(x, g(x)) = 0.$$

Proof. Define $f : U \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ by $f(x, y) = \begin{pmatrix} x \\ f(x, y) \end{pmatrix}$. Thus $f \in C^1(U, \mathbb{R}^n \times \mathbb{R}^m)$ and

$$Df_{z_0} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \frac{\partial F(z_0)}{\partial x} & \frac{\partial F(z_0)}{\partial y} \end{bmatrix}$$

Thus, we have that $\det Df_{z_0} = 1 \cdot \det DF_{z_0} \neq 0$. We can apply the inverse function theorem to f . We get open set $U'; V' \subset \mathbb{R}^n$ and $f^{-1} : V' \rightarrow U'$ is C^1 and $f(z_0) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix} \in R = V'_1 \times V'_2 \subset V'$. Let $S = f^{-1}(R)$, we know it is open and $z_0 \in S = V_1 \times V_2$. Thus, $f^{-1} : V_1 \times V_2 \rightarrow V_1 \times V_2$ with

$$\begin{pmatrix} x \\ y' \end{pmatrix} \rightarrow \begin{pmatrix} x \\ h(x, y') \end{pmatrix}$$

, where

$$h : V_1 \times V'_2 \rightarrow V_2$$

is C^1 . Thus, define $g : V_1 \rightarrow V_2$ such that $g(x) = h(x, 0) \in C^1$. □

1.19 Mon, Nov 11: Lagrange Multipliers

Theorem 42. Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ is C^1 for $j \in [m]$ (i.e, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$). Suppose f has a local extremum at x_0 with the constraint $g(x) = 0$ and $\text{rank}(Dg_{x_0}) = m$. Then there are real numbers, called **Lagrange Multipliers** $\lambda_j \in \mathbb{R}$ such that

$$\nabla f(x_0) = \sum_{j=1}^m \lambda_j \nabla g_j(x_0)$$

We will prove this for the $m = 1, n = 2$ case, which states that if $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ is C^1 , then if f has a local extremum at x_0 with the constraint of $g(x) = 0$, and $\nabla g(x_0) \neq 0$, then there exists some $\lambda \in \mathbb{R}$ such that

$$\nabla f(x_0) = \lambda \nabla g(x_0).$$

Proof. By the implicit function theorem, we have that $g(x_1, x_2) = 0$ can locally be written as either

$$g(x_1, \gamma(x_2)) = 0, \quad \text{or} \quad g(\beta(x_2), x_2) = 0.$$

Let's concentrate on the first. Then for all $x_1 \in U$, we write $g(x_1, \gamma(x_2)) = 0$. By assumption, we have that that by the chain rule

$$0 = \frac{d}{dx_1} g(x_1, \gamma(x_1)) = \langle \nabla g(x_1, \gamma(x_1)), \begin{pmatrix} 1 \\ \gamma'(x) \end{pmatrix} \rangle. \quad (1)$$

We have that the function $\hat{f} : U \rightarrow \mathbb{R}$ with $\hat{f}(x_1) = f(x_1, \gamma(x_1))$ has a local extremum at x_{0_1} . Thus, we have that

$$0 = \hat{f}'(x_{0_1}) = \langle \nabla f(x_{0_1}, \gamma(x_{0_1})), \begin{pmatrix} 1 \\ \gamma'(x_{0_1}) \end{pmatrix} \rangle$$

Because (1) holds for any $x \in U$, then we have the ∇f and ∇g are orthogonal to each other, and so there must exist some real λ such that

$$\nabla f = \lambda \nabla g.$$

□

You can see that the method of Lagrange Multipliers is an application of the implicit function theorem.

1.20 Wed, Nov 13: Multivariate Integration

We mostly talked about integration from $\mathbb{R}^2 \rightarrow \mathbb{R}$.

Definition 42. Suppose $R \subset \mathbb{R} \times \mathbb{R}$ is a rectangular region. Then we partition R into $R_{i,j} = [x_{i-1}, x_i) \times [y_{j-1}, y_j)$ such that $R = \bigcup_j \bigcup_i R_{i,j}$. Then we take a sampling of points S inside each $R_{i,j}$ and we call the Riemann Sum to be

$$R(f, P, S) = \sum f(s_{i,j})|R_{i,j}| = \sum f(s_{i,j})(x_i - x_{i-1})(y_i - y_{i-1}).$$

We say f is **Riemann Integrable** if there exists some $I \in \mathbb{R}$ such that for all $\epsilon > 0$, there exists some partition P such that

$$|P - I| < \epsilon.$$

Definition 43. We define the **upper sum** and the **lower sum** to be

$$U(f, P) = \sum \max f|R_{i,j}|, \quad L(f, P) = \sum \min f|R_{i,j}|.$$

We say that f is **Darboux integrable** if for any $\epsilon > 0$, there exists partitions P such that

$$|U(f, P) - L(f, P)| < \epsilon \iff \inf U = \sup L.$$

Remark 17. Notationally, we say that

$$\int f dx dy = \sup L, \quad \overline{\int} f dx dy = \inf U$$

Sliced Bread.

Theorem 43. (Fubini's Theorem) If f is integrable over $R = [a, b] \times [c, d]$, then

$$\int_R f dx dy = \int_c^d \left[\int_a^b f dx \right] dy.$$

Proof.

$$U = \inf \left\{ \int_R g dx dy; g \geq f, \quad g \text{ constant on } R_{i,j} \right\}$$

$$L = \sup \left\{ \int_R g dx dy; g \leq f, \quad g \text{ constant on } R_{i,j} \right\}$$

Thus,

$$\int_c^d \left[\int_a^b g dx \right] dy = \sum g(s_{i,j})|R_{i,j}|.$$

For any g with $g \leq f$ we have

$$\int_c^d \left[\int_a^b f dx \right] dy \geq \int_c^d \int_a^b g dx dy = \int_R g dx dy \leq \int_R f dx dy$$

and thus

□

We now pivot to a very important theorem in multidimensional calculus:

Theorem 44. (Inverse Function Theorem) Suppose $\varphi : U \rightarrow W$ is a C^1 diffeomorphism with U, W open. Suppose further that f is integrable and that $R \subset U$. Then we have that

$$\int_{\varphi(R)} f = \int_R f(\varphi) |\det(D\varphi)|$$

Proof. No.

□

1.21 Fri, Nov 15: Line Integrals and Green's Theorem

We leave the rigour for a few lectures to define some important topics in vector calculus. Suppose $\gamma : [a, b] \rightarrow \mathbb{R}^2$ and let f be an integrable function. If τ is the unit tangent, then we define the **line integral** of F , vector field, along the curve C parameterized by γ to be

$$\int_C F ds = \int_C F \cdot \tau ds = \int_{\gamma([a,b])} F \cdot \frac{\gamma'(t)}{|\gamma'(t)|} dt = \int_a^b F(\gamma(t)) \cdot |\gamma'(t)| dt \quad (2)$$

Theorem 45. The line integral is independent of parametrization.

Proof. Suppose $\alpha : [a, b] \rightarrow C$ and $\beta : [c, d] \rightarrow C$ are two parametrization of a curve C . Define $\gamma : \beta^{-1} \circ \alpha$. Thus, we have that $\gamma : [a, b] \rightarrow [c, d]$. Thus, we compute using the change of variables formula:

$$\begin{aligned} \int_C F ds &= \int_c^d f(\beta(t)) |\beta'(t)| dt \\ &= \int_a^b f(\beta(\gamma(s))) |\beta'(\gamma(s))| |\gamma'(s)| ds \\ &= \int_a^b f(\alpha(s)) |\alpha'(s)| ds \end{aligned}$$

□

Using this theorem, we can provide an alternate parametrization to (2). Let $\gamma(t) = (t, g(t))$ parameterize C . Then we have

$$\begin{aligned} \int_C f dx &= \int f(x, g(x)) dx = \int_C (f, 0) \cdot \tau ds \\ \int_C f dy &= \int_C (0, f) \cdot \tau ds \end{aligned}$$

Now we get that

$$\int_C F \cdot \tau ds = \int_C F_1 dx + F_2 dy.$$

Theorem 46. (Gradient Integral Theorem or the Fundamental Theorem for Line Integrals) Suppose $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and we have a curve C from $p \in \mathbb{R}^d$ to $q \in \mathbb{R}^d$, then

$$\int_C \nabla f \cdot \tau ds = \int_C \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n = f(q) - f(p)$$

Proof. Clear by FTC: if γ is a parameterization of C , then

$$\int_C \nabla f \cdot \tau ds = \int_a^b \nabla f(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \frac{d}{dt} f(\gamma(t)) dt = f(\gamma(b)) - f(\gamma(a))$$

□

Theorem 47. Suppose $F : \Omega \rightarrow \mathbb{R}^d$ for $\Omega \subset \mathbb{R}^d$ open, then the following are equivalent:

- (a) $\int_C F \cdot \tau ds$ depends only on the endpoints.
- (b)

$$\oint_C F \cdot \tau ds = 0$$

if C is a loop.

(c) There exists some $\rho : \Omega \rightarrow \mathbb{R}$ such that $F = \nabla \rho$.

(d) If Ω is convex, then we have that $\nabla \times F = 0$.

Theorem 48. (Green's Theorem) Suppose $D \subset \mathbb{R}^2$ is simple bounded by the curve C and F is a vector field. Then

$$\int_C F_1 dx + F_2 dy = \int_D (D_2 F_1 - D_1 F_2) dx dy.$$

Remark 18. Luis didn't prove Theorem 47, but he showed a nice little geometric proof for Theorem 48 that I cannot stand because I know Stokes Theorem exists. Note that in 2-D, we have that

$$D_2 F_1 - D_1 F_2 = \nabla \times F,$$

meaning we have $1 \rightarrow 4$ in Theorem 47. $3 \rightarrow 4$ is equally clear since we show in a pset that $ddf = 0$ and thus $\nabla \times \nabla \rho = 0$.

1.22 Fri, Nov 17: Midterm II

We had another midterm! Thank god I'm god!

Example 1.10. Write explicit examples of the following. No justification is required.

(a) A differentiable function on \mathbb{R} that is not C^1 .

(b) A function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ which has directional derivatives in every direction, but is not differentiable.

Proof. No justification is great.

(a)

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

(b)

$$f(x) = \begin{cases} \frac{x^3 y}{x^4 + y^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

□

Proposition 12. Consider the $C([0, 1], \mathbb{R})$ with the sup-metric, and let A be the subspace

$$A := \{f \in C([0, 1], \mathbb{R}) : f \text{ is differentiable everywhere, } |f| \leq 1, |f'| \leq 1\}.$$

- Is A a closed subset of $C([0, 1])$.
- Is A an open subset of $C([0, 1])$.
- Is the closure of A compact on $C([0, 1])$.

Proof. We give the best proofs of all time.

- **No.** Consider a sequence of $(f_n) \in A$ which uniformly converges to the absolute value function translated to have the crest at $\frac{1}{2}$. Call this f . Then $f_n \rightarrow f$ unif but $f \notin A$. Do a little doodle for this to get full credit.

- **No.** By Baire Category Theorem, we know that the set of continuous but non-differentiable functions is dense in $C([0, 1])$. Thus, for any $f \in A$, any $r > 0$, we have that there exists some g continuous but nondifferentiable such that

$$g \in B_r(f).$$

Thus, A is not open.

- By the generalized Heine Borel (Theorem 26), we have that $\overline{A} \subset C$ is compact if and only if it is equicontinuous, bounded, and closed. Since A is bounded, then \overline{A} is bounded. Obviously \overline{A} is closed. Since $|f'| \leq 1$, then we have that for any $\epsilon > 0$, there exists a $\delta = \epsilon$ such that if $|x - y| < \delta$ for any $x, y \in A$ and $f \in A$, then by MVT,

$$|f(x) - f(y)| = f'(\theta)|x - y| \leq |x - y| < \epsilon.$$

Thus, A is equicontinuous. To show \overline{A} is equicontinuous, let $\epsilon > 0$. It suffices to show that if f is a LP of A , then equicontinuity follows. Thus, suppose $(f_n) \rightarrow f$ with $f_n \in A$. Then we have that

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3},$$

where the first and last terms are bounded by the fact that $f_n \rightarrow f$ uniformly and the middle term is bounded by the equicontinuity of the f_n .

□

Proposition 13. Let α and β be two given continuous functions on $[0, 1]$. Assume that for all $x \in [0, 1]$, $|\alpha(x)| < 1$. prove that there exists a unique continuous function $f : [0, 1] \rightarrow \mathbb{R}$ such that for all $x \in [0, 1]$,

$$f(x) - \int_0^x \alpha(y)f(y)dy = \beta(x).$$

To prove this, used the fixed point theorem.

1.23 Mon, Nov 20: Surface Integrals and the Divergence Theorem

Now that we have talked about line integrals, we can begin talking about surface integrals. Suppose that $S \subset \mathbb{R}^3$ be a surface parameterized by some $\alpha : Q \rightarrow \mathbb{R}^3$, then the **surface integral** of some function f is defined to be

$$\int_S f dA = \int_Q f(\alpha(x, y)) \left| \frac{\partial \alpha}{\partial x} \times \frac{\partial \alpha}{\partial y} \right| dx dy.$$

If $F : S \rightarrow \mathbb{R}^3$ is a vector field and $S \subset \mathbb{R}^3$ is parameterized by α , then the **flux** of F is

$$\int_S F \cdot ndA = \int_Q F(\alpha(t)) \cdot \left(\frac{\partial \alpha}{\partial x} \times \frac{\partial \alpha}{\partial y} \right) dx dy = \int_Q \det \left(F(\alpha), \frac{\partial \alpha}{\partial x}, \frac{\partial \alpha}{\partial y} \right) dx dy. \quad (3)$$

Theorem 49. The Flux of (3) is independent of parametrization.

The proof for this carries a similar flavor to Theorem 45.

Theorem 50. The flux in (3) can be written by

$$\int_S F \cdot ndA = \int_Q F_1 dy dz - F_2 dz dx + F_3 dx dy$$

Remark 19. To sketch a proof for this, we parameterize the surface by $\alpha(x, y) = (x, y, g(x, y))$ and proceed as in Theorem 45!

Theorem 51. (Divergence Theorem) Suppose $D \subset \mathbb{R}^3$ is a simple region. Then if F is a vector field, then

$$\int_{\partial D} F \cdot ndA = \int_D \nabla \cdot F dx dy dz$$

1.24 Fri, Nov 22: Differential Forms

Luis had a terrible treatment of differential forms, so we differ to Pugh for his..., wait his stuff is also terrible?? We take heavy inspiration from Hubbard and Hubbards *Vector Calc, Linear Algebra, and Differential Forms* Chapter 6.

Definition 44. A **k-form** on \mathbb{R}^n is a function φ that takes k vectors in \mathbb{R}^n and returns a number $\varphi(v_1, \dots, v_k)$ such that φ is multilinear and alternating and antisymmetric.

Remark 20. We usually denote 1-forms as dx_1 , etc. A two form as $dx_1 \wedge dx_2$, etc. We provide meaning to his wedge operator later.

Example 1.11.

$$dx_1 \wedge dx_2 \left(\begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 1 \\ 2 \end{bmatrix} \right) = \det \begin{bmatrix} 1 & 3 \\ 2 & -2 \end{bmatrix} = -8$$

Note that the two form gives the signed area of the parallelogram they span. Similarly, $dx_1 \wedge dx_2 \wedge dx_4$ gives the (x_1, x_2, x_4) component of signed volume of the parralelipied spanned by three vectors.

Remark 21. A zero form is just a number in \mathbb{R}^n .

Remark 22. $dx_i \wedge dx_i = 0$. Just look at the determinant!

Remark 23.

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

More generally, on k form, it is the sign of the permutations needed to get to the elementary form.

Definition 45. An **elementary k-form** on \mathbb{R}^n is an expression of the form

$$dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

where $1 \leq i_1 < \dots < i_k \leq n$ which when evaluated on the vectors v_1, \dots, v_k , gives the determinant of the $k \times k$ matrix obtained by selecting rows i_1, \dots, i_k of the matrix who's columns are v_1, \dots, v_k . The only elementary 0-form is just 1.

The following theorem is quite important. Also very cool.

Theorem 52. The space of k -forms in \mathbb{R}^n forms the vector space $A^k(\mathbb{R}^n)$. I.e, we can add k -forms and multiply them by scalars. Moreover, the elementary k -forms form a basis for $A^k(\mathbb{R}^n)$. I.e, any k -form in $A^k(\mathbb{R}^n)$ can be uniquely written by

$$\varphi = \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

where

$$a_{i_1, \dots, i_k} dx_{i_1} = \varphi(e_{i_1}, \dots, e_{i_k}).$$

The dimension of $A^k(\mathbb{R}^n)$ is

$$\binom{n}{k}$$

Definition 46. Let φ be a k -form and ω be an ℓ -form, both on \mathbb{R}^n . The **wedge product** $\varphi \wedge \omega$ is a $(k + \ell)$ -form. It is defined as

$$\varphi \wedge \omega(v_1, \dots, v_{k+\ell}) = \sum_{\sigma \in \text{Perm}(k, \ell)} \varphi(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \omega(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)})$$

It is fairly painful to prove the following properties:

Proposition 14. The wedge product satisfies the following:

- (a) Distributivity: $\varphi \wedge (\omega_1 + \omega_2) = \varphi \wedge \omega_1 + \varphi \wedge \omega_2$.
- (b) Associativity: $\varphi \wedge (\omega_1 \wedge \omega_2) = (\varphi \wedge \omega_1) \wedge \omega_2$.
- (c) Skew commutativity: $\varphi \wedge \omega = (-1)^{k\ell}(\omega \wedge \varphi)$.

Definition 47. A **k-differential form** on an open subset $U \subset \mathbb{R}^n$ is a function that takes k -vectors v_1, \dots, v_k anchored at a point $x \in \mathbb{R}^n$ (i.e, a parallelogram $P_x(v_1, \dots, v_k)$) and returns a number. It is a multilinear and antisymmetric a function of \mathbf{v} .

Example 1.12.

$$\cos(xz)dx \wedge dy \left(P \left(\begin{bmatrix} 1 \\ 2 \\ \pi \end{bmatrix} \right) \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right) \right) = \cos(1 \cdot \pi) \det \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} = -2$$

The entire reason of differential forms is to integrate them, and so we define how to do it with the change of variables formula.

Definition 48. Let $[\gamma(U)]$ be a parametrized domain in \mathbb{R}^n . The **integral of the k-differential form** φ over $[\gamma(U)]$ is

$$\int_{[\gamma(U)]} \varphi = \int_U \varphi (P_{\gamma(U)} (D\gamma(\mathbf{u}))) |d\mathbf{u}|$$

Example 1.13. Let $k = 1$ and $n = 2$ and $\gamma(u) = \begin{pmatrix} R \cos u \\ R \sin u \end{pmatrix}$, with U being the interval $(0, \alpha)$. Integrating $xdy - ydx$ over $[\gamma(U)]$, then we have

$$\int_{[\gamma(U)]} xdy - ydx = \int_0^\alpha xdy - ydx \left(P \left(\begin{bmatrix} R \cos u \\ R \sin u \end{bmatrix} \right) \begin{bmatrix} -R \sin u \\ R \cos u \end{bmatrix} \right) du = \int_0^\alpha (R^2 \cos^2 u + R^2 \sin^2 u) du = R^2 \alpha.$$

Example 1.14. Integrate $dx \wedge dy + ydx \wedge dz$ over $[\gamma(S)]$ where

$$\gamma \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} s+t \\ s^2 \\ t^2 \end{pmatrix}, \quad S = \left\{ \begin{pmatrix} s \\ t \end{pmatrix} \mid s \in [0, 1], t \in [0, 1] \right\} :$$

$$\begin{aligned}
\int_{[\gamma(S)]} dx \wedge dy + ydx \wedge dz &= \int_S dx \wedge dy + ydx \wedge dz \left(P \left(\begin{pmatrix} s+t \\ s^2 \\ t^2 \end{pmatrix} \right) \left(\begin{bmatrix} 1 \\ 2s \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2t \end{bmatrix} \right) \right) dsdt \\
&= \int_0^1 \int_0^1 \det \begin{bmatrix} 1 & 1 \\ 2s & 0 \end{bmatrix} + s^2 \det \begin{bmatrix} 1 & 1 \\ 0 & 2t \end{bmatrix} dsdt \\
&= \int_0^1 \int_0^1 (-2s + 2s^2t) dsdt \\
&= \int_0^1 \left(-1 + \frac{2}{3}t\right) dt \\
&= \frac{-2}{3}
\end{aligned}$$

Going over to vector calculus, we see how amazing differential forms really are.

Definition 49. The **work form** W_F of a vector field $F = \begin{bmatrix} F_1 \\ \vdots \\ F_n \end{bmatrix}$ is the 1-form defined by

$$W_F(P_x(v)) = F(x) \cdot v = F_1 dx + \dots F_n dx_n.$$

Example 1.15. Any differential 1form is the work form of a vector field. The one form $ydx - xdy$ is the work form of $F = \begin{bmatrix} -y \\ x \end{bmatrix}$

Remark 24. Work forms should be integrated over oriented curves.

Definition 50. The **work** of a vector field F along an oriented curve C is

$$\int_C W_F$$

Example 1.16. What is the work of $F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} y \\ -x \\ 0 \end{bmatrix}$ over the helix oriented by the tangent vector field

$$\mathbf{t} = \begin{bmatrix} \sin t \\ \cos t \\ 1 \end{bmatrix}$$

and parameterized by

$$\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}$$

for $t \in (0, 4\pi)$

$$\int_0^{4\pi} ydx_1 - xdx_2 \left(P \left(\begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix} \right) \begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix} \right) dt = \int_0^{4\pi} -\sin^2(t) - \cos^2(t) dt = -4\pi$$

Definition 51. The **flux-form** is the differential 2-form

$$\Phi_F(P_x(v, w)) = \det[F(x), v, w] = F_1 dy \wedge dz - F_2 dz \wedge dx + F_3 dx \wedge dy$$

Remark 25. The flux form of a vector field associates to a parallelogram the flow of the vector field through it.

Definition 52. The **flux** of a vector field F over an oriented surface S is $\int_S \Phi_F$.

Example 1.17. The flux of the vector field $F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} x \\ y^2 \\ z \end{bmatrix}$ through the parameterized domain $\begin{pmatrix} u \\ v \end{pmatrix} \rightarrow \begin{pmatrix} u^2 \\ uv \\ v^2 \end{pmatrix}$ for $0 \leq u, v \leq 1$ is:

$$\begin{aligned} \int_0^1 \int_0^1 x dy \wedge dz - y^2 dz \wedge dx + z dx \wedge dy & \left(P \begin{pmatrix} u^2 \\ uv \\ v^2 \end{pmatrix} \left(\begin{bmatrix} 2u \\ v \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ u \\ 2v \end{bmatrix} \right) \right) dudv = \\ &= \int_0^1 \int_0^1 2u^2 v^2 - 4u^3 v^3 + 2u^2 v^2 \\ &= \frac{7}{36} \end{aligned}$$

Definition 53. The **mass form** M_f is the 3-form defined by

$$M_f(P_x(v_1, v_2, v_3)) = f(x) \det[v_1, v_2, v_3] = f dx \wedge dy \wedge dz$$

We are ready to talk about the exterior derivative!

Definition 54. The **exterior derivative** d of a k -form φ , denoted $d\varphi$, takes a $k+1$ parallelogram and returns a number, i.e.,

$$d\varphi(P_x(v_1, \dots, v_{k+1})) = \lim_{h \rightarrow 0} \frac{1}{h^{k+1}} \int_{\partial P_x(hv_1, \dots, hv_{k+1})} \varphi$$

Proposition 15. Let φ be a k -form:

- (a) The exterior derivative defines a $(k+1)$ -form.
- (b) The exterior derivative is linear over \mathbb{R} . That is,

$$d(a\varphi + b\omega) = ad\varphi + bd\omega.$$

- (c) The exterior derivative of a constant form is 0.
- (d) Let f be a zero form, then

$$df = \sum_{i=1}^n (Df(e_i)) dx_i.$$

- (e) If f is a function, then

$$d(f dx_{i_1} \wedge \dots \wedge dx_{i_k}) = df \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

Example 1.18. Compute the exterior derivative of the following 2-form on \mathbb{R}^4 :

$$\psi = x_1x_2dx_2 \wedge dx_4 - x_2^2dx_3 \wedge dx_4.$$

We have

$$d\psi = d(x_1x_2dx_2 \wedge dx_4) - d(x_2^2dx_3 \wedge dx_4) = x_2dx_1 \wedge dx_2 \wedge dx_4 - 2x_2dx_2 \wedge dx_3 \wedge dx_4$$

Theorem 53. For any k -form φ :

$$d(d\varphi) = 0$$

Theorem 54. The exterior derivative of wedge product is

$$d(\varphi \wedge \omega) = d\varphi \wedge \omega + (-1)^k \varphi \wedge d\omega.$$

Recall from vector calculus class:

Definition 55. Let $U \subset \mathbb{R}^3$ be an open set, $f : U \rightarrow \mathbb{R}$ be C^1 and F be a C^1 vector field on U .

- The **gradient** of f is given by

$$\nabla f = \begin{bmatrix} Df(e_1) \\ Df(e_2) \\ Df(e_3) \end{bmatrix}$$

- The **curl** of F is given by

$$\nabla \times F = \nabla \times F = \begin{bmatrix} DF_3(e_2) - DF_2(e_3) \\ DF_1(e_3) - DF_3(e_1) \\ DF_2(e_1) - DF_1(e_2) \end{bmatrix}$$

- The **divergence** of F is given by

$$\nabla \cdot F = DF_1(e_1) + DF_2(e_2) + DF_3(e_3).$$

We can now relate these concepts to differential forms:

Theorem 55. Let f be a function and F be a vector field. Then

(a)

$$df = W_{\nabla f} = D_1f dx_1 + D_2f dx_2 + D_3f dx_3$$

(b)

$$dW_F = \Phi_{\nabla \times F}$$

(c)

$$d\Phi_F = M_{\nabla \cdot F}$$

We use the previous theorem and Theorem 21 to show two consequences.

Corollary 5.

$$\nabla \times \nabla f = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\nabla \cdot (\nabla \times F) = 0$$

We now get to the most beautiful theorem in Real Analysis:

Theorem 56. (Generalized Stokes Theorem) Let X be a compact piece-with-boundary of a k -dimensional oriented manifold $M \subset \mathbb{R}^n$. Give the boundary ∂X of X the boundary orientation, and let φ be a $(k-1)$ -form defined on an open set containing X . Then

$$\int_{\partial X} \varphi = \int_X d\varphi$$

We can intuit two corollaries from this

Corollary 6. (Green's Theorem) Let S be a bounded region in \mathbb{R}^2 bounded by some curve C . Let F be a vector field defined on a neighborhood of S . Then

$$\int_S dW_F = \int_C W_F.$$

Remark 26. To get the usual form of Green's theorem, just let $F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$, then

$$W_F = F_1 dx + F_2 dy, \quad dW_F = dF \wedge dx + dF_2 \wedge dy = (D_1 F_2 - D_2 F_1) dx \wedge dy.$$

Thus, we have that

$$\int_C F_1 dx + F_2 dy = \int_S (D_1 F_2 - D_2 F_1) dx \wedge dy$$

Corollary 7. (Stokes Theorem) Let S be an oriented surface in \mathbb{R}^3 , bounded by a curve C that is given the boundary orientation. Let φ be a 1-form defined on a neighborhood of S , then

$$\int_S d\varphi = \int_C \varphi.$$

Remark 27. To get the usual form of the divergence theorem, remind yourself that if we write $\phi = W_F$, then we get that

$$\int_S dW_F = \int_S \Phi_{\nabla \times F}.$$

Note that the classical version is

$$\int \int_S (\nabla \times F) \cdot N dx dy = \int_C F \cdot T dx.$$

Corollary 8. (Divergence Theorem) Let X be a bounded domain in \mathbb{R}^3 . Let its boundary ∂X be a union of surfaces S_i . Let φ be a 2-form defined on a neighborhood of X , then

$$\int_X d\varphi = \sum \int_{S_i} \varphi.$$

Remark 28. To extract the classical version, let $\varphi = \Phi_F$, then we have that $d\varphi = M_{\nabla \cdot F}$. We then rewrite

$$\int_X M_{\nabla \cdot F} = \int \int \int_X \nabla \cdot F dx dy dz = \sum \int_S F \cdot N dx dy$$

Theorem 57. (Fundamental Theorem of Line Integrals) Suppose $F : U \rightarrow \mathbb{R}^n$ for $U \subset \mathbb{R}^n$ open, then the following are equivalent:

(a)

$$\int_{\gamma([a,b])} W_{\nabla F} = f(\gamma(b)) - f(\gamma(a))$$

(b)

$$F = \nabla f$$

for some f .

(c)

$$\nabla \times F = 0$$

and U is convex.

Mon, Dec 2: k -cells and pullbacks

We talked about a lot of which I wrote about in the previous section. Here, we more formally define the k -cell, that which we integrate forms over.

Definition 56. A **k-cell** is a parametrization $\varphi : I^k \rightarrow \mathbb{R}^n$ and we can integrate over a k -cell by

$$\int_{\varphi} \omega = \int_{I^k} \omega(\varphi(u)) \left(\frac{\partial \varphi}{\partial u_1}, \dots, \frac{\partial \varphi}{\partial u_k} \right) du_1 du_2 \dots du_k.$$

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a transformation.

Definition 57. The **pushforward** of φ , is where $\varphi : Q \rightarrow \mathbb{R}^n$, is defined as

$$T_*\varphi = T \circ \varphi : Q \rightarrow \mathbb{R}^m.$$

The more important one to us is the pullback, which is defined as

Definition 58. The **pullback** of a k -form in \mathbb{R}^m is a k -form denoted by $T^*\omega$ and defined by

$$\int_{\varphi} T^*\omega = \int_{T_*\varphi} \omega.$$

Remark 29. Thus, we have that $T^*(x)(v_1, \dots, v_k) = \omega(x)(DT_x v_1, \dots, DT_x v_k)$.

1.25 Wed, Dec 4: Pullbacks and Stokes' Theorem

Proposition 16. (a)

$$T^*(\omega_1 \wedge \omega_2) = T^*\omega_1 \wedge T^*\omega_2.$$

(b)

$$T^*(d\omega) = d(T^*\omega)$$

Proof. We don't prove the first because fuck you. We verify for when ω is a zero form ($\omega(y) = f(y)$.) Then we have that by the chain rule:

$$\begin{aligned} T^*\omega &= f(T(x)) \implies d(T^*(\omega)) \\ &= \frac{\partial f \circ T}{\partial x_1} dx_1 + \cdots + \frac{\partial f \circ T}{\partial x_n} dx_n \\ &= \sum_{j=1}^n \left(\sum_{i=1}^m \frac{\partial f_i}{\partial y_i} \frac{\partial T_i}{\partial x_j} \right) dx_j \end{aligned}$$

The other equality we get straight from the definition of the exterior derivative:

$$d\omega = \frac{\partial f}{\partial y_1} dy_1 + \cdots + \frac{\partial f}{\partial y_n} dy_n,$$

and so plugging in the definition of $T^*\omega$ yields the above. For a general k -form, induct. \square

Now we get to Stoke's Theorem, but we must handwave the fuck out of some shit beforehand. Let Q be a cube in \mathbb{R}^n and let

$$d\omega = \sum f_i \hat{dx}_i, \quad \hat{dx}_i = dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_n.$$

Observe that if $\omega = f_1(x) dx_2 \wedge dx_3 \cdots dx_n$, then by FTC:

$$f_1(1, x_2, \dots) - f_1(0, x_2, \dots, x_n) = \int_0^1 \frac{\partial f_1}{\partial x_1}(t, x_2, \dots, x_n) dt,$$

and thus

$$\int_{x_1=1} f_1(x) dx_2 \wedge dx_3 \cdots dx_n - \int_{x_1=0} f_1(x) dx_2 \wedge dx_3 \cdots dx_n = \int_Q \frac{\partial f_1}{\partial x_1} dx_2 \wedge \cdots \wedge dx_n.$$

Theorem 58. Let M be a k -orientable (the k -cells don't cancel out) in \mathbb{R}^n and ω be a $k-1$ form, then

$$\int_{\partial M} \omega = \int_M d\omega.$$

Proof. This is not right, Luis is smoking something. For each k -cell that makes up the manifold, we have that

$$\int_{\partial \varphi} \omega = \int_{\partial \varphi} \varphi^*(\omega)$$

and

$$\int_M d\omega = \int_Q \varphi^* d\omega = \int_Q d(\varphi^*\omega)$$

\square

1.26 Fri, Dec 6: Translating Stokes, Hairy Ball Theorem, and Brower's Theorem

r/mildlyinteresting

1.27 Thu, Dec 13: Final

This final was ok. If only Luis had graded mine correctly!

Write explicit examples of the following. One is not possible. No justification is required.

- (a) A metric space that is not complete.

$$\mathbb{R} \setminus \{0\}$$

- (b) A finite connected set with more than one point.

Impossible

- (c) A compact subset $C([0, 1])$ that has infinitely many elements.

$$\mathcal{F} = \{f_n : [0, 1] \rightarrow \mathbb{R} : f_n(x) = \frac{1}{n} \sin(x)\}, \quad \overline{\mathcal{F}}$$

- (d) An ODE of the form $x'(t) = f(x(t))$ with $x(0) = x_0$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ a C^1 function that has a solution in some interval $t \in (-a, b)$ but not on the full real line $t \in \mathbb{R}$.

$$x'(t) = (x(t))^2$$

- (e) A 1-form ω in $\mathbb{R}^2 \setminus \{(0, 0)\}$ such that $d\omega = 0$. However, there is no 0-form ω_1 such that $d\omega_1 = \omega$.

$$r = \sqrt{x^2 + y^2}, \quad \omega = \frac{-y}{r^2} dx + \frac{x}{r^2} dy$$

Proposition 17. Let K be a compact set and $f : K \rightarrow K$ continuous. Assume that $d(f(x), f(y)) < d(x, y)$ for all $x, y \in K$. Prove that f has a fixed point in K . Show an example in which the conclusion fails if the set K is closed but not necessarily compact.

Proposition 18. Suppose that a sequence of continuous functions $f_n : [a, b] \rightarrow \mathbb{R}$ converges monotonically down to zero. Prove that the convergence is uniform.

Proposition 19. We say that two curves γ_0 and γ_1 with the same endpoints are homotopic in some open set $\Omega \subset \mathbb{R}^n$ if there exists a 2-cell $\varphi : [0, 1]^2 \rightarrow \mathbb{R}^n$ such that

- (a) For all $s \in [0, 1]$, $\gamma_0(0) = \gamma_1(0) = \varphi(0, s)$.

- (b) For all $s \in [0, 1]$, $\gamma_0(1) = \gamma_1(1) = \varphi(1, s)$.

- (c) For all $t \in [0, 1]$, $\gamma_0(t) = \varphi(t, 0) = \gamma_1(t) = \varphi(t, 1)$.

Assume that ω is a 1-form in Ω such that $d\omega = 0$ and γ_1 is homotopic to γ_0 . Prove that

$$\int_{\gamma_1} \omega = \int_{\gamma_0} \omega$$