MAT257 Lecture Notes

Arky!! :3c

'24 Fall & '25 Winter Semester

Contents

§1 Day 1: Course Administrative Details and Equivalent Norms (Sep. 4, 2024)

Course administrative details!

- The prerequisites to this class are calculus and linear algebra; specifically, this class is about calculus on manifolds.
- The main three theorems this class will cover are the inclusive function theorem, change of variable in integral, and Stokes' theorem (generalization of FTC).
- The main techniques we will study this class are manifolds, theory of integration, and differential forms.

To give an example of how linear algebra can be used in defining differentiability, we start by recalling that for a function $f : \mathbb{R} \to \mathbb{R}$ to be differentiable in 1 variable, we need

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} =: c = f'(a)$$

to exist for any $a \in \mathbb{R}$. Specifically, this may be rewritten as

$$\lim_{h \to 0} \frac{f(a+h) - (f(a) + ch)}{h} = 0,$$

where we have $f(a) + f'(a)h$ as the best local linear approximation of $f(a+h)$ at $h = 0$, and $h \mapsto ch$ is a linear transformation $\mathbb{R} \to \mathbb{R}$. For the higher dimensional analogue of the above, let us take $f : \mathbb{R}^n \to \mathbb{R}^m$, and have

$$\begin{aligned} a &= (a_1, \dots, a_n), \\ x &= (x_1, \dots, x_n), \\ h &= (h_1, \dots, h_n). \end{aligned}$$

If we let $y = f(x)$, then $y \in \mathbb{R}^m$, and we can write it as a column vector with components $y_1, \dots, y_m$;

$$f(x) = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix}.$$

In this way, each y_i for $1 \leq i \leq n$ can be viewed as a function in n variables. We say f is differentiable at $a \in \mathbb{R}^n$ if there exists a linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ such that[1]

$$\lim_{h \to 0} \frac{|f(a+h) - (f(a) + \lambda h)|}{|h|} = 0$$

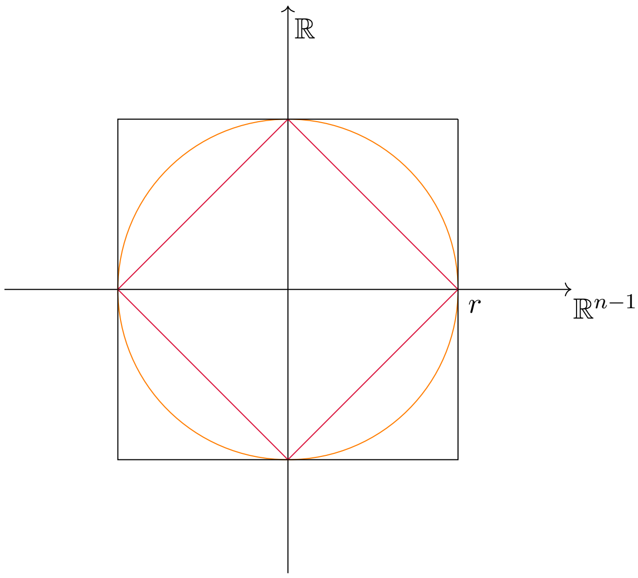
In this way, we have that $f(a) + \lambda h$ is the best linear approximation of $f(a+h)$ at $h = 0$. Notice that compared to the $\mathbb{R} \to \mathbb{R}$ differentiability condition, we impose a norm on $f(a+h) - (f(a) + \lambda h)$ and h, since we cannot divide by $\mathbb{R}^n, \mathbb{R}^m$ vectors.

[1] i think we need a norm in the numerator for the limit below; bierstone didn't do it in class though, so... idk. also lambda is jacobian?

This also brings the topic of what sort of norms we could be looking at; for example, we have[2]

$$
\begin{aligned}
|x| &= \sqrt{x_1^2 + \cdots + x_n^2}, \\
\|x\| &= \max\{|x_1|, \ldots, |x_n|\}, \\
|||x||| &= |x_1| + \cdots + |x_n|.
\end{aligned}
$$

We say that two norms p, q are equivalent to each other if we may find a constant C where $p(x) \leq Cq(x)$ (and vice versa). In fact, the norms $|x|, \|x\|, |||x|||$ are all equivalent to each other; here is a visual proof,



where each figure is the level sets of their respective norms. In particular, $|x| \leq \sqrt{n}\,\|x\|$, and $|||x||| \leq \sqrt{n}\,|x|$. To prove the latter analytically, we may write the norm as an inner product, $|||x||| = \langle x, u(x) \rangle < |x|\,|u(x)| \leq \sqrt{n}r$, where u_i is the sign of each component of x.[3]

[2] these are (in order) euclidean, maximum, and taxicab norms; or $\ell^2, \ell^\infty, \ell^1$

[3] i kinda get where this proof is going, but at the same time i dont really know how he defined u in the first place, so yeah. alternate proof would be to give Cauchy-Schwarz

§2 Day 2: Review of MAT247, Metric Spaces (Sep. 6, 2024)

Course administrative details!

- Fall Office Hours will be held on Mondays from 10:30 to 11:30AM (likely in Bahen 6114).

We start with a review on functions and continuity. Given $\mathbb{R}^n$ with elements $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, we define the following addition and scalar multiplication,

$$\begin{aligned} x + y &= (x_1 + y_1, \ldots, x_n + y_n), \\ \lambda x &= (\lambda x_1, \ldots, \lambda x_n), \lambda \in \mathbb{R}. \end{aligned}$$

We may also equip it with the Euclidean inner product and norm,

$$\begin{aligned} \langle x, y \rangle &= x_1 y_1 + \cdots + x_n y_n, \\ |x| &= (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}} = \sqrt{\langle x, x \rangle}, \end{aligned}$$

of which it has the properties

(a) Non-negativity; $|x| \geq 0$, and is equal to 0 if and only if $x = 0$.

(b) Absolute Homogeneity; $|\lambda x| = |\lambda| \, |x|$ for any scalar $\lambda \in \mathbb{R}$.

(c) Triangle Inequality; $|x + y| \leq |x| + |y|$ for any vectors $x, y \in \mathbb{R}^n$.

We may check that the properties above hold for the Euclidean norm;

(a) Trivial, since the square of each component is non-negative.

(b) Not discussed in class, but we may directly expand the norm to get the equality,

$$|\lambda x| = \left(\sum_{i=1}^{n} (\lambda x_i)^2 \right)^{\frac{1}{2}} = \left(\lambda^2 \sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} = |\lambda| \, |x| \, .$$

(c) Start by observing that from inner product properties, we have $|\langle x, y \rangle| \leq |x| \, |y|$, with equality if and only if x, y are linearly dependent. If x, y are LI, observe $\lambda y - x \neq 0$ for all scalars $\lambda \in \mathbb{R}$, and we may write

$$\begin{aligned} 0 < |\lambda y - x|^2 &= \langle \lambda y - x, \lambda y - x \rangle \\ &= \lambda^2 \langle y, y \rangle - 2\lambda \langle x, y \rangle + \langle x, x \rangle \\ &= |y|^2 \lambda^2 - 2 \langle x, y \rangle \lambda + |x|^2 \, . \end{aligned}$$

Reading the above as a polynomial in λ, we observe that it must not have any real roots, and so the discriminant is necessarily negative, i.e. $4 \langle x, y \rangle^2 - 4 |x|^2 |y|^2 < 0$, which implies $|\langle x, y \rangle| < |x| \, |y|$. Returning to the triangle inequality, let us start by squaring both sides to obtain

$$|x + y|^2 = \langle x, x \rangle + 2 \langle x, y \rangle + \langle y, y \rangle \leq |x|^2 + 2 |x| \, |y| + |y|^2 \, .$$

This leaves us with $2 \langle x, y \rangle \leq 2 |x| \, |y|$, which, from the above, we find equality if x, y are linearly dependent, and "$<$" if LI.

Consider the linear transformation, $T : \mathbb{R}^n \to \mathbb{R}^m$, equipped with standard bases $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$, and $\{f_1, \ldots, f_m\}$ of $\mathbb{R}^m$; let the matrix representation of T be given by the below matrix A,

$$y = T(x) \implies \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \underbrace{\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

In terms of basis vectors, we may visualize T by

$$T(e_i) = \sum_{j=1}^{m} a_{ji} f_j.$$

Now, suppose $S : \mathbb{R}^m \to \mathbb{R}^\ell$ be a linear transformation, with corresponding matrix representation $B \in M_{\ell\times m}(\mathbb{R})$. Then the composition $S \circ T : \mathbb{R}^n \to \mathbb{R}^\ell$ has matrix $B \cdot A \in M_{\ell\times n}(\mathbb{R})$.

Here is an example from class; if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map, then $|Tx| \leq M\,|x|$ for some $M > 0$. To see this, let us start by writing

$$Tx = T\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i T(e_i).$$

This gives us

$$\begin{aligned} |Tx| \leq \sum_{i=1}^{n} |x_i T(e_i)| &= \sum_{i=1}^{n} |x_i|\,|T(e_i)| \\ &= \left(\sum_{i=1}^{n} a_{ji}^2\right)^{\frac{1}{2}} \\ &\leq \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}} =: |A|\,. \end{aligned}$$

Then we have

$$|Tx| \leq \left(\sum_{i=1}^{n} |x_i|\right)|A| \leq \sqrt{n}\,|A|\,|x|\,.$$

We simply pick $M \geq \sqrt{n}\,|A|$ and we are done. □

§2.1 Topology of $\mathbb{R}^n$

Let us equip $\mathbb{R}^n$ with the Euclidean norm; this is an example of a metric space (i.e., a set equipped with a distance function $d(x, y)$). The function $d : X \times X \to \mathbb{R}$ has the following properties,

(a) Symmetry, $d(x, y) = d(y, x)$.

(b) Non-negativity, $d(x, y) \geq 0$, with $d(x, y) = 0$ if $x = y$.

(c) Triangle inequality, $d(x, y) \leq d(x, z) + d(z, y)$.

§3 Day 3: Topology in $\mathbb{R}^n$ (Sep. 9, 2024)

What are some n-dimensional analogues of closed intervals $[a, b] \in \mathbb{R}$? We have

- The closed rectangle $[a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$,
- The closed ball $\{x \in \mathbb{R}^n \mid |x - a| \leq r\}$.

For n-dimensional analogues of open intervals, we have open rectangles and balls, i.e.

- $(a_1, b_1) \times \cdots \times (a_n, b_n)$,
- $\{x \in \mathbb{R}^n \mid |x - a| < r\}$.

We say a subset $U \subseteq \mathbb{R}^n$ is open (two definitions) if:

- For any $a \in U$, we may pick $\varepsilon > 0$ such that the ball $B(a, \varepsilon) \subseteq U$.
- For any $a \in U$, there exists an open rectangle R such that $a \in R \subseteq U$.

We say a subset $C \subseteq \mathbb{R}^n$ is closed if $\mathbb{R}^n \setminus C$ is open. Here are some examples of closed sets,

- $\emptyset, \mathbb{R}^n$;
- Closed rectangles and balls as per earlier;
- Finite sets.

Let us take the closed subset $A \subseteq \mathbb{R}^n$, and have A contain all rationals in $(0, 1)$. We claim that $[0, 1] \subseteq A$. To see this, consider $x \in \mathbb{R}^n \setminus A$. Since A is closed, $\mathbb{R}^n \setminus A$ is open, which means there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq A$. By density of $\mathbb{Q}$ in $\mathbb{R}$, we may always find a rational in $B(x, \varepsilon)$, meaning $B(x, \varepsilon) \cap A = \emptyset$, and $x \notin [0, 1]$. $\square$

A few remarks;

- Any union of open sets is open.
- The finite union of closed sets is closed.
- Arbitrary unions of closed sets are not necessarily closed; observe

$$\bigcup_{n \in \mathbb{N}} \left[\frac{1}{n} - 1, 1 - \frac{1}{n}\right] = (-1, 1).$$

Now, consider $A \subseteq \mathbb{R}^n$, and $x \in \mathbb{R}^n$; there are 3 possibilities:

1. There exists an open ball B such that $x \in B \subset A$ (i.e. the interior of A, $\operatorname{int} A$).
2. There exists an open ball B such that $x \in B \subset \mathbb{R} \setminus A$ (i.e. the exterior of A, $\operatorname{ext} A$).
3. For all open balls B such that $x \in B$, there exists $y_1 \in A$ and $y_2 \in \mathbb{R}^n \setminus A$ such that $y_1, y_2 \in B$ (i.e. the boundary of A, $\operatorname{bdry} A$).

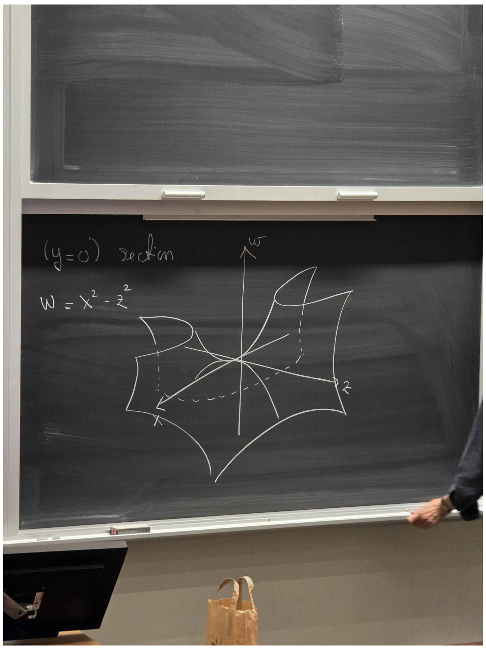
For example, we may consider $A \subset \mathbb{R}^n$ to be the rationals in $(0, 1)$; then the interior of A is $\emptyset$, the exterior of A is $\mathbb{R}^n \setminus [0, 1]$, and the boundary of A is $[0, 1]$.

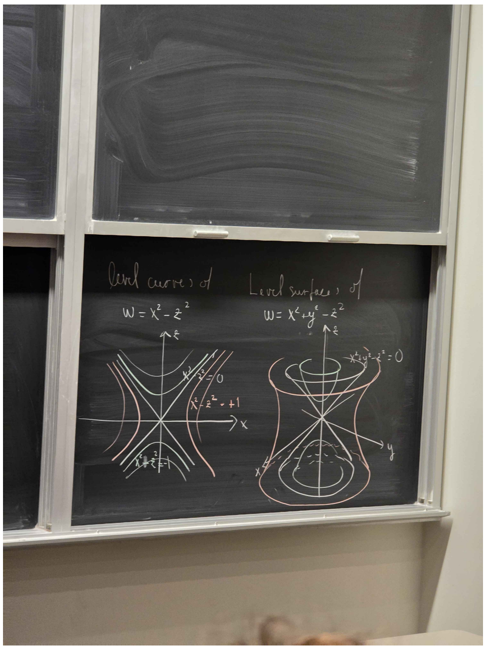
Let $A \subset \mathbb{R}^n$; given the function $f : A \to \mathbb{R}^m$ then we define the graph of f,

$$\operatorname{graph} f = \{(x, f(x)) \mid x \in A\} \subseteq A \times \mathbb{R}^m.$$

§4 Day 4: Graphing Example; Operations on Functions, Definition of Continuity by Open Sets (Sep. 11, 2024)

We graphed $w = x^2 + y^2 - z^2$ in class today. Since I don't really know how to LaTeX these kinds of graphs, I'll just drop these pictures in;





The main idea was to start by considering $z = 0$ and observing that $w = x^2 + y^2$ is really a parabola ($w = x^2$, $w = y^2$) rotated about the w-axis; setting x or y to 0, we get the left picture. If we examine the level sets of $w = x^2 - z^2$, we may combine the two to obtain the level surfaces of $w = x^2 + y^2 - z^2$ (as per the rightmost diagram on the blackboard).

Let $A \subset \mathbb{R}^n$, and consider a function $f : A \to \mathbb{R}^m$ (i.e., a function $f : \mathbb{R}^n \to \mathbb{R}^m$ with domain restricted onto A). If we want to create another function $g : B \to \mathbb{R}^p$ to be composed with A, we implicitly ask that $f(A) \subset B$; with this, we may write

$$(g \circ f)(x) = g(f(x))$$

where $\mathrm{dom}(g \circ f) = f^{-1}(B)$. Now, let us consider the inner product $\langle\, , \rangle : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. Let f, g be functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. Then we may construct

$$(f, g)(x) : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^m,$$
$$f \cdot g = \langle\, , \rangle \circ (f \cdot g),$$

which we may indeed check sends $\mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ as expected of the definition of dot product.

Returning to earlier, let us consider $f : (A \subset \mathbb{R}^n) \to \mathbb{R}^m$, and let us consider $f(x)$ written in its components, $f(x) = (f_1(x), \ldots, f_m(x))$. If we wish to be specific, observe that we may write each f_i as the following composition,

$$f_i = \pi_i \circ f \text{ where } \pi_i : \mathbb{R}^m \to \mathbb{R},$$

where π_i is the mapping $(x_1, \ldots, x_m) \mapsto x_i$.

On the topic of limits, recall from MAT157 that $\lim_{x\to a} f(x) = b$ means that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if x is in a δ-ball of a, then $f(x)$ is in an ε-ball of b. We may extend this idea to define continuity; we say f is continuous at a if

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

and that f is a continuous function if it is continuous for all $a \in A$. If we want to define continuity in its topological notion, though, we have that $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous if and only if $f^{-1}(U)$ is open in $\mathbb{R}^n$ for all $U \subset \mathbb{R}^m$.

Theorem 4.1 (Spivak 1.8)**.** We say $f : (A \subset \mathbb{R}^n) \to \mathbb{R}^m$ is continuous if and only if, for all $U \subset \mathbb{R}^m$, $f^{-1}(U) = A \cap V$ for some open set $V \subset \mathbb{R}^n$.[4]

($\Rightarrow$) Consider any open set $U \subset \mathbb{R}^m$. If $a \in f^{-1}(U)$, we have $b := f(a) \in U$. Since U is open, there exists some $\varepsilon > 0$ such that $B(a, \varepsilon) \subset U$; using the fact that f is continuous at a, we may construct a δ-ball $B(a, \delta(a))$ about a such that

$$A \cap B(a, \delta(a)) \overset{f}{\mapsto} B(f(a), \varepsilon).$$

With this, we may take the union of all such balls $B(a, \delta(a))$ and observe that

$$f^{-1}(U) = A \cap \underbrace{\left(\bigcup_{a\in f^{-1}(U)} B(a, \delta(a)) \right)}_{:=V},$$

where we may note V is open (since the arbitrary union of open sets is open). We may note that $A \cap V$ indeed covers $f^{-1}(U)$; if it did not, then we would be able to pick a new $a \in f^{-1}(U)$ and repeat the above process, contradicting the definition of V.

($\Leftarrow$) For any $a \in A$, let us have $U = B(f(a), \varepsilon)$ for any $\varepsilon > 0$. Then $f^{-1}(U) = A \cap V$ for some open set $V \subset \mathbb{R}^n$; by definition of open sets, we may find a ball about a contained in V; let it be $B(a, \delta)$. Then this fulfills the $\varepsilon - \delta$ definition of continuity, and we are done. □

[4] do note that this proof is different from lecture, since i either mis-transcribed or bierstone made a mistake concerning balls for all $a \in A$ sending to b.

§5 Day 5: Composition of Continuous Functions; Uniform Continuity, Distance Metric, Compactness (Sep. 13, 2024)

We start with a few examples:

- Is $f(x, y) = \frac{x^2-y^2}{x^2+y^2}$ continuous? The answer is no, since the limit on $(x, y) \to 0$ on ℓ_x (read: X-axis) is equal to -1, while the limit on ℓ_y is equal to 1. This means there are two conflicting limits approaching $(x, y) \to 0$, meaning that f is not continuous at 0.
- Is $f(x, y) = e^{-\frac{|x-y|}{x^2-2xy+y^2}}$ continuous? Observe that
$$e^{-\frac{|x-y|}{x^2-2xy+y^2}} = e^{-\frac{1}{|x-y|}};$$
since $|x - y|$ and $e^{-\frac{1}{x}}$ are continuous, we see that the composition is continuous as well, and so $f(x, y)$ is continuous.

While the composition of continuous functions property has been proved in 327, we provide a 257 variant of the proof;

Theorem 5.1 (Composition of Continuous Functions is Continuous)**.** Let $f : (A \subset \mathbb{R}^n) \to \mathbb{R}^m$ and $g : (B \subset \mathbb{R}^m) \to \mathbb{R}^p$ be continuous, with $f(A) \subset B$. Then we have that $g \circ f$ is continuous.

Let U be open in $\mathbb{R}^p$; then $g^{-1}(U) = B \cap V$ where V is open in $\mathbb{R}^m$ (as per our definition of continuity). Furthermore, we also have $(g \circ f)^{-1}(U) = f^{-1}(B \cap V) = f^{-1}(B) \cap f^{-1}(V)$; let $f^{-1}(V) = A \cap W$ for some open $W \subset \mathbb{R}^n$, then we have

$$f^{-1}(B) \cap f^{-1}(V) = \underbrace{f^{-1}(B) \cap A}_{f^{-1}(B) \supset A} \cap W = A \cap W,$$

which is as desired. □

We now provide examples of continuity.

1. Let us have a linear function $T : \mathbb{R}^n \to \mathbb{R}^m$ such that T is uniformly continuous. Then
$$|T(x) - T(y)| = |T(x - y)| \leq C\,|x - y|$$
for some scalar C. In particular, when constructing an $\varepsilon - \delta$ proof for continuity here, for any $\varepsilon > 0$ we may pick $\delta < \frac{\varepsilon}{C}$.

2. Let $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$; we say that f is continuous if and only if each of f_i for $1 \leq i \leq m$ are continuous as well.

   ($\Rightarrow$) By contrapositive; if any f_i is discontinuous at any point a, then f is discontinuous at a as well.

   ($\Leftarrow$) If every f_i is continuous, there exists δ_i for each f_i such that $|f_i(x) - f_i(a)| < \frac{\varepsilon}{\sqrt{n}}$ whenever $|x - a| < \delta_i$. Then we may set $\delta = \min\{\delta_1, \ldots, \delta_m\}$ to see
$$|f(x) - f(a)|^2 = \sum_{i=1}^{m} |f_i(x) - f_i(a)| < n\left(\frac{\varepsilon}{\sqrt{n}}\right)^2 = \varepsilon^2,$$
   yielding $|f(x) - f(a)| < \varepsilon$ whenever $|x - a| < \delta$. □

In terms topological definitions though, we may write

($\Rightarrow$) If f is continuous, then we may use the composition of continuous functions to see that the projections $f_i = \pi_i \circ f$ are indeed continuous.

($\Leftarrow$) Take any open subset $U \subset \mathbb{R}^m$. Then let R be an open rectrangle in U, and let us consider the union of all possible rectangles,

$$f^{-1}(U) = \bigcup_{R \in U} f^{-1}(R).$$

Since each rectangle is defined as the cartesian product of open intervals on each respective f_i, we see that $f^{-1}(R)$ is given by

$$f^{-1}(R) = \bigcap_{i=1}^{m} f_i^{-1}(R_i),$$

where R_i are the said respective open intervals. Since each f_i is continuous and the intersection of open sets is open, we see $f^{-1}(R)$ is open for all R, and so $f^{-1}(U)$ is open as well. $\square$

3. Let $X \subset \mathbb{R}^n$, and define the metric $d(x, X) = \inf_{a \in X} |x - a|$, i.e. the smallest distance from x to some point $a \in X$. We want to show that $f(x) = d(x, X)$ is uniformly continuous on $\mathbb{R}^n$.

   To start, let us consider $|f(x) - f(y)| = |d(x, X) - d(y, X)| \leq |x - y|$; we wish to prove the inequality. To do this, start by taking $d(x, X) - |x - y|$, and consider

$$\begin{aligned} d(x, X) - |x - y| &\leq d(x, X) + \underbrace{|y - a| - |x - a|}_{\text{Triangle Ineq.}} \\ &\leq |y - a| && (d(x, X) - |x - a| \leq 0) \\ &\leq d(y, X). \end{aligned}$$

   In this way, we get $|x - a| \leq |x - y| + |y - a|$, which yields $d(x, X) - d(y, X) \leq |x - y|$. $\square$

We also briefly touched on compactness at the end of the class. We call a subset $X \subset \mathbb{R}^n$ *compact* if every open covering $\mathcal{O}$ of X has a finite subcovering (i.e., a subset of $\mathcal{O}$ that covers X). Here are some examples,

- $\mathbb{R}$ (equipped with the standard topology) is not compact. If we let $\mathcal{O}$ be an open covering given by
$$\mathcal{O} = \{(a, a + 1) \mid a \in \mathbb{R}\},$$
we have that $\mathcal{O}$ covers $\mathbb{R}$, but there does not exist a finite subcovering. Thus, $\mathbb{R}$ cannot be compact.

- The open interval $(0, 1)$ (once again, equipped with the standard topology) is not compact. This time, let
$$\mathcal{O} = \left\{ \left(\frac{1}{n}, 1 - \frac{1}{n} \right) \mid n \in \mathbb{N} \right\}.$$
Clearly, $\mathcal{O}$ covers X, but it does not admit a finite subcovering.

- Any topology on a finite set X is compact. In particular, any covering is necessarily finite, since there are finitely many elements in $\mathcal{P}(X)$; this means any subcovering, even if it is the same as the covering, is finite as well.

§6 Day 6: Compactness (Sep. 16, 2024)

We start by giving some properties on compactness on $\mathbb{R}^n$.

Theorem 6.1 (Compactness $\iff$ Closed and Bounded)**.** A subset $X \subset \mathbb{R}^n$ is compact if and only if it is also closed and bounded.

We prove both directions now.

($\Rightarrow$) Suppose $X \subset \mathbb{R}^n$ is compact. Then consider an open cover of X by open balls of radius 1; then we may find a finite subcover (by definition of compactness). Since there are a finite number of open balls of radius 1 covering X, all points are finitely close to each other.

We now show that X is closed, i.e. $\mathbb{R}^n \setminus X$ is open. Take any $a \in \mathbb{R}^n \setminus x$. We wish to find $\delta > 0$ such that $B(a, \delta) \subset \mathbb{R}^n \setminus X$; then we may consider the closed balls of radius $\frac{1}{k}$ given by $\overline{B}(a, \frac{1}{k})$ for $k = 1, 2, \ldots,$ and denote

$$U_k = \mathbb{R}^n \setminus \overline{B}\left(a, \frac{1}{k}\right).$$

Clearly, $\bigcup_k U_k = \mathbb{R}^n \setminus \{a\}$, and so U_k forms an open covering of X, and so there exists a finite subcover of it. Thus, $X \subset U_k$ for some k, where $B(a, \frac{1}{k}) \subset \mathbb{R}^n \setminus X$. We conclude that X is closed.

($\Leftarrow$) Will be done next lecture (?)

Theorem 6.2 (Continuous Image of Compact is Compact)**.** Let $X \subset \mathbb{R}^n$ be compact, and consider a continuous function $f : X \to \mathbb{R}^n$. Then $f(x)$ is compact.

To start, let $\mathcal{O} = \{U\}$ (read: multiple open sets U) be an open cover of $f(X)$. For every $U \in \mathcal{O}$, consider by continuity we have

$$f^{-1}(U) = X \cap V_U,$$

where V_U is some open set in $\mathbb{R}^n$. Then $\{V_U\}_{U \in \mathcal{O}}$ is an open covering of X. Since X is compact, we may write

$$X \subset V_{U_1} \cup \cdots \cup V_{U_k}$$

for some open sets $U_1, \ldots, U_k$ in the covering $\mathcal{O}$. Therefore, $f(X) \subset U_1 \cup \cdots \cup U_k$, which is indeed a finite covering. □

Theorem 6.3 (Extreme Value Theorem)**.** A continuous function $f : X \to \mathbb{R}$ on a compact $X \subset \mathbb{R}^n$ takes on a minimum and maximum value.

Since $f(X)$ is compact, we know it is closed and bounded; let $M = \sup\{f(x) \mid x \in X\} < \infty$. If $M \notin f(X)$, then there is an open interval around M outside $f(X)$ (since the complement of $f(X)$ is open), contradicting that M is the supremum; thus, f attains M at some point, and we may consider $-f(X)$ to obtain the infimum / minimum. □

Theorem 6.4 (ε-neighborhood Theorem)**.** If we have a compact X in an open set $U \subset \mathbb{R}^n$, then there is $\varepsilon > 0$ such that the ε-neighborhood of X in $\mathbb{R}^n$ lies in U; specifically, the ε-neighborhood of X can be defined as

$$\{y \in \mathbb{R}^n \mid d(y, X) < \varepsilon\} = \bigcup_{x \in X} B(x, \varepsilon). \qquad (\textit{Left as exercise})$$

To prove this, let $f(x) = d(x, \mathbb{R}^n \backslash U)$ be continuous (continuity has been proven previously I think). Then $f(x) > 0$ for all $x \in X$, because we may always pick a ball centered at x in U. Since X is compact, f has a minimum value at $\varepsilon > 0$; this means the ε-neighborhood of X lies in U. $\square$

§7 Day 7: Uniform Continuity, Hard Direction of Heine-Borel (Sep. 18, 2024)

Theorem 7.1 (Cont. Function on Compact Set is Uniformly Cont.)**.** A continuous function $f : (X \subset \mathbb{R}^n) \to \mathbb{R}^n$, where X is compact, is uniformly continuous. [5]

Recall that uniform continuity means that for all $\varepsilon > 0$, there exists some one-size-fits-all $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$ for any $x, y \in X$.

Let us have $X \times X \subset \mathbb{R}^n \times \mathbb{R}^n$. To start, the diagonal $\Delta = \{(x, x) \mid x \in X\}$ is compact, because Δ is the image of X under the map $x \mapsto (x, x)$, which is a continuous function. Thus, we have that $g : X \times X \to \mathbb{R}$ where $g(x, y) = |f(x) - f(y)|$ is continuous, as per the composition of continuous functions.

Given $\varepsilon > 0$, consider $g^{-1}((-\varepsilon, \varepsilon)) = (X \times X) \cup U$ where U is an open set in $\mathbb{R}^n \times \mathbb{R}^n$; clearly, $\Delta \subset U$, since Δ is compact and U is open. By the ε-neighborhood theorem, there exists $\delta > 0$ such that the δ-neighborhood of Δ is in U. Then consider x, y such that $|x - y| < \delta$, and observe we have

$$|(x, x) - (x, y)| \le |x - y| + |y - y| < \delta$$

by the triangle inequality. This means (x, y) is in a δ-neighborhood of Δ, and so is in U. By construction, we see that $(x, y) \in g^{-1}((-\varepsilon, \varepsilon))$, and we conclude $g((x, y)) = |f(x) - f(y)| < \varepsilon$. $\square$

Theorem 7.2 (Closed Interval is Compact)**.** A closed interval $X = [a, b] \subset \mathbb{R}$ is compact.

Consider an open cover $\mathcal{O}$ of $[a, b]$, and let A be the set of all $x \in [a, b]$ such that $[a, x]$ can be covered by finitely many sets in $\mathcal{O}$. Then we want to show that $a, b \in A$, and that A is bounded above by b.

Let $\alpha = \sup A$; we start by showing that $\alpha \in A$. First, observe that $\alpha \in [a, b]$, meaning $\alpha \in U$ for some $U \in \mathcal{O}$. Since U is open, we may find a δ-ball around α in U, i.e. $(\alpha - \delta, \alpha + \delta) \subset U$. Since α is the supremum, there must exist some x in the interval to the left of α such that $[a, x]$ is covered by $U_1, \dots, U_k \in \mathcal{O}$. This means $[a, \alpha]$ is covered by $U \cup \bigcup_{i=1}^k U_i$.

Now, we show that $\alpha = b$. In the opposite direction to the above, suppose we pick $x' \in (\alpha, \alpha + \delta) \cap [a, b]$ (i.e., to the right of α). Since we know $(a - \delta, a + \delta)$ is covered by U and $[a, x]$ is covered by finitely many sets in $\mathcal{O}$, $[a, x']$ is covered by $U \cup \bigcup_{i=1}^k U_i$ as well, which would contradict that α is the supremum of A. The only situation in which there is no contradiction is if $\alpha = b$, since x' would be at most b in this case. $\square$

Lemma 7.3 (Closed Rectangles in $\mathbb{R}^n$ are Compact)**.** Closed rectangles $R = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ are compact.

We may prove this by inducting on n. As per earlier, we have that a closed interval on R is compact. See Day 8 for the complete proof; class was interrupted by a fire alarm :c

[5] In general, this works on any $f : X \to Y$ if X is a compact metric space. for here, we let them both be subsets of $\mathbb{R}^n$

§8 Day 8: Finishing Heine-Borel; Differentiation (Sep. 20, 2024)

Theorem 8.1 (Finishing Heine-Borel; Closed and Bounded implies Compact)**.** We now extend our proof that $X \subset \mathbb{R}^n$ is compact if it is closed and bounded to $n > 1$.

As a preliminary case, start by considering the rectangle $R = [a_1, b_1] \times \cdots \times [a_n, b_n] = Q \times [a_n, b_n]$. Let Q be a closed set in $\mathbb{R}^{n-1}$; continuing on our inductive proof from last time, we have that Q is compact. For any $t \in [a_n, b_n]$, consider $Q_t = \{(y, t) \mid y \in Q\}$ (where we may let y be the rest of the $n-1$ components in Q).

Now, consider $\mathcal{O}$ to be an open cover of R. Then

$$Q_t \subset U_1 \cup \cdots \cup U_k =: U$$

where $U_1, \ldots, U_k \in \mathcal{O}$, then by the ε-neighborhood theorem, there is $\varepsilon > 0$ such that $Q_t \times (t - \varepsilon, t + \varepsilon) \subset U$. Let us consider the set $\{(t - \varepsilon_t, t + \varepsilon_t) \mid t \in [a_n, b_n]\}$. This is an open cover of the closed interval $[a_n, b_n]$; by compactness of $[a_n, b_n]$, there is a finite subcover by the open intervals $(t - \varepsilon_t, t + \varepsilon_t)$. This means $Q \times (t - \varepsilon_t, t + \varepsilon_t)$ for finitely many $t \in [a_n, b_n]$ covers R, and we are done.

Returning to Heine-Borel, now, consider any $X \subset \mathbb{R}^n$. Since X is bounded, we may enclose $X \subset R$ where R is a closed rectangle construction as per above. Then let $\mathcal{O}$ be an open cover of R. Since $\mathbb{R}^n \setminus X$ is open, we have that $\mathcal{O} \cup \{\mathbb{R}^n \setminus X\}$ is an open cover of R, meaning that as per above, $R \subset U_1 \cup \cdots \cup U_k \cup (\mathbb{R}^n \setminus X)$ where $U_i \in \mathcal{O}$. Since $R \supset X$, we conclude that $U_1, \ldots, U_k$ is a finite subcover of $\mathcal{O}$ for X. □

We now cover differentiation. Suppose $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^m$ with U being open. We say that f is differentiable at a if there is a linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - \lambda h}{|h|} = 0.$$

Specifically, $f(a+h) - f(a) - \lambda(h) = o(|h|)$; i.e., it is equal to a function $\varphi(h)$ where $\lim_{h \to 0} \frac{y(h)}{|h|} = 0$.

Lemma 8.2 (Differentiability implies Continuity)**.** If f is differentiable at a, then f is continuous at a.

To see this, take $h \to 0$ in $f(a+h) - f(a) - \lambda(h) = o(|h|)$.

Lemma 8.3. If f is differentiable at a, then there is a unique affine function $h \mapsto c - \lambda(h)$ such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - \lambda(h)}{|h|} = 0.$$

Let $c = f(a)$ by continuity. So we have to show that if $\lambda, \mu : \mathbb{R}^n \to \mathbb{R}^m$, with both satisfying the above limit, then we have $\lambda = \mu$. Observe that if we write

$$\lim_{x \to 0} \frac{\lambda(x) - \mu(x)}{|x|} = 0,$$

then we may take $x = ty$ and take $t \to 0$ to get

$$\lim_{t \to 0} \frac{\lambda(t_y) - \mu(t_y)}{|t_y|} = 0,$$

so $\frac{\lambda(y) - \mu(y)}{|y|} = 0$. This means $\lambda = \mu$. □

With this, we say that λ is the derivative of f at a. We then may write $Df(a)$ or $f'(a)$ or $\partial_a f$. If f is differentiable at every point of U, then we say it is differentiable on the open set U.

§9 Day 9: Partial Differentiation, Jacobians (Sep. 23, 2024)

Let us have a function[6] $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^m$. Then the matrix $D_f(a)$ or $f'(a)$ w.r.t. the standard bases of $\mathbb{R}^n, \mathbb{R}^m$ is called the Jacobian matrix of f at a.

We now give some examples.

(a) For an example, consider $g : \mathbb{R} \to \mathbb{R}$ differentiable at all $a \in \mathbb{R}$. Let $f(x, y) = g(x)$, where $(x, y) \in \mathbb{R}^2$. Then f is differentiable at (a, b), for any $b \in \mathbb{R}$ and $D_f(a, b) : (h, k) \mapsto g'(a)h$. Writing the derivative out, we have

$$\frac{f(a+h, b+k) - f(a, b) - g'(a)h}{|(h, k)|} \to 0$$

as $(h, k) \to 0$, meaning the above is equal to

$$\frac{g(a+h) - g(a) - g'(a)h}{|h|} \cdot \frac{|h|}{|(h, k)|} = 0.$$

Thus, the Jacobian is $D_f(a, b) = (g'(a), 0)$.

(b) Let $f(x, y) = \sqrt{|xy|}$. Is it differentiable at 0? To check this, we want to either find or disprove the existence of λ, μ such that

$$\frac{\sqrt{|hk|} - 0 - (\lambda h + \mu k)}{|(h, k)|} \to 0$$

as $(h, k) \to 0$. Now, suppose $h = k$. Then we have

$$\frac{|h| - (\lambda + \mu)h}{\sqrt{2}\,|h|} = \frac{1}{\sqrt{2}} - \frac{\lambda + \mu}{\sqrt{2}} \cdot \frac{h}{|h|}.$$

If $\lambda + \mu = 0$, then $\frac{1}{\sqrt{2}} \not\to 0$. If $\lambda + \mu \neq 0$, then the limit approaches 2 instead as $h \to 0$. Thus, we conclude that f is not differentiable at 0.

Let the directional derivative of f at a along a vector v be given by

$$D_v f(a) = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}.$$

Then we define the ith partial derivative of f at a, for $i = 1, \dots, n$, to be

$$\frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a) = \lim_{h \to 0} \frac{f(a_1, \dots, a_i + h, \dots, a_n) - f(a_1, \dots, a_n)}{h}.$$

Lemma 9.1 (Differentiability at Point implies Directional Derivatives Exist)**.** If f is differentiable at a, then all directional derivatives (i.e., for all v) at a exist, and $D_v f(a) = Df(a)v$.

To start, we know that

$$\frac{f(a + tv) - f(a) - D_f(a)(tv)}{|tv|} \to 0$$

as $t \to 0$. If $t \to 0^+$, then $|tv| = t\,|v|$ and we may multiply the LHS by $|v|$ to get

$$\frac{f(a + tv) - f(a)}{t} - Df(a)(v) \xrightarrow[t \to 0^+]{} 0$$

In the other way, if $t \to 0^-$, then $|tv| = -t\,|v|$; multiply the LHS by $-|v|$, and we get the same thing. $\square$

[6] wanted to clarify this isn't correct notation, it just looks correct to me so i do it

§10 Day 10: Differentiation (Sep. 25, 2024)

Today we will go over differentiation. We start with a few examples;

(a) Let $f(x,y) = \sin(x\cos y)$. Then

$$\frac{\partial f}{\partial y} = \cos(x\sin y)x\cos y.$$

Note that in the partial derivative above, we hold x as a constant.

(b) Let $f(x,y) = x^{x^{x^{x^y}}} = e^{(logx)\cdot x^{x^{x^y}}}$. Then

$$D_2(1,y) = x^{x^{x^{x^y}}}; \quad \log x \frac{\partial}{\partial y}\left(x^{x^{x^y}}\right) = 0.$$

We now discuss higher order derivatives; consider $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^n$. Suppose $D_i f : U \to \mathbb{R}$ exists for all i, and we consider that $D_j(D_i f)$ may also be written as

$$D_{ij}f(x), \quad f_{x_i x_j}(x), \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(x).$$

However, we must note that order of differentiation is important; observe the example below (with corresponding link),

(c) Let $f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$.

Then we have $D_2 f(x,0) = x$, $D_2 f(0,y) = -y$ by symmetry, and $D_{12}f(0,0) = -1$ since $D_{21}f(0,0) = 1$.

Note that second order mixed partial derivatives at a are equal if they are both defined, and they are continuous on an open set containing a. For clarity, we introduce multi-index notation, given as below,

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} = \frac{\partial^{a_1+\cdots+a_n} f}{\partial x_1^{\alpha_1}\ldots \partial x_n^{a_n}},$$

i.e. $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$. A relevant point worth bringing up is Schwarz's theorem, which discusses when the orders of partials commute (note that this was not covered in class).

Theorem 10.1 (Chain Rule)**.** Let $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^m$, and $g : (V \subset \mathbb{R}^m) \to \mathbb{R}^n$ (where $f(U) \subset V$). If f is differentiable at $a \in U$, and g is differentiable at $f(a)$, then $g \circ f : U \to \mathbb{R}^n$ is differentiable at a, and $D(g\circ f)(a) = D_g(f(n)) \circ Df(a)$, i.e. $(g\circ f)'(a) = g'(f(a)) \cdot f'(a)$ (note that this denotes matrix multiplication).

To start, we're given that $f(x) - f(a) - f'(a)(x-a) = \varphi(x)$ from $\lim_{x\to a} \frac{\varphi(x)}{|x-a|} = 0$, and $g(y) - g(b) - g'(b)(y-a) = \psi(y)$ from $\lim_{y\to b}\frac{\psi(y)}{|y-b|} = 0$. We want to show that

$$\frac{(g\circ f)(x) - g(b) - g'(b)f'(a)(x-a)}{|x-a|} \xrightarrow[x\to a]{} 0.$$

To start, we reduce the numerator as folows,

$$\begin{aligned} &(g\circ f)(x) - g(b) - g'(b)f'(a)(x-a) \\ &= \underbrace{(g\circ f)(x) - g(b) - g'(b)(f(x) - f(a)}_{\psi(f(x))} - \varphi(x)) \\ &= \psi(f(x)) + g'(b)\varphi(x). \end{aligned}$$

Then observe that the latter half of the reduced expression yields,

$$\lim_{x\to a} \frac{g'(b)\varphi(x)}{|x-a|} = g'(b) \lim_{x\to a} \frac{\varphi(x)}{|x-a|} = 0.$$

We now claim that the former half satisfies

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

For all $\varepsilon > 0$, we have $|\psi(f(x))| < \varepsilon\,|f(x) - b|$ if $|f(x) - b| < \delta'$ for some $\delta' = \delta'(\varepsilon)$, which happens when $|x-a| < \delta$ for some $\delta = \delta(\delta')$ by continuity. Therefore, if $|x-a| < \delta$, then

$$\begin{aligned} |\psi(f(x))| &< \varepsilon\,|f(x) - b| \\ &= \varepsilon\left|f'(a)(x-a) + \varphi(x)\right| \\ &\le \varepsilon M\,|x-a| + \varepsilon\,|\varphi(x)|\,, \end{aligned}$$

and so $\frac{|\psi(f(x))|}{|x-a|} \le \varepsilon M + 0$ as $x \to a$, which means we may conclude the fraction goes to 0 as $x \to 0$ as desired. □

§11 Day 11: Computations of Derivatives (Sep. 27, 2024)

We start with a few useful facts: for the following examples, we let $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^m$.

(a) If f is a constant function $f = b$, then $Df(a) = 0$.

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - 0}{|h|} = \lim_{h\to 0} \frac{b-b}{|h|} = 0.$$

(b) If f is a linear transformation, then $Df(a) = f$.

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - f(h)}{|h|} = \lim_{h\to 0} \frac{f(a+h) - f(a+h)}{|h|} = 0.$$

(c) If $f = (f_1, \ldots, f_m)$, f is differentiable at a if and only if each f_i is differentiable at a, and $Df(a) = (Df_1(a), \ldots, Df_m(a))$. We call $Df(a)$ to be the Jacobian (read: $m \times n$ matrix) where the ith row is given by $Df_i(a)$. [7]

($\Rightarrow$) Observe that if f is differentiable at a, then $f_i = \pi_i \circ f$, where π_i is the projection from $\mathbb{R}^m \to \mathbb{R}$, which is differentiable by the chain rule.

($\leftarrow$) Now suppose each f_i is differentiable at a; write $\lambda = (Df_1(a), \ldots, Df_m(a))$, and consider

$$\begin{aligned} &\lim_{h\to 0} \frac{f(a+h) - f(a) - \lambda(h)}{|h|} && \text{(Expand component-wise)} \\ &= \lim_{h\to 0} \frac{(f_1(a+h) - f_1(a) - Df_1(a), \ldots, f_m(a+h) - f_m(a) - Df_m(a))}{|h|} = 0 \end{aligned}$$

by the triangle inequality. □

(d) Suppose $S : \mathbb{R}^2 \to \mathbb{R}$, and $S(x,y) = x + y$. Then we may observe that S is linear, meaning $D_S(a,b) = S$ by (b).

(e) Suppose $P : \mathbb{R}^2 \to \mathbb{R}$, and $P(x,y) = xy$. Then[8]

$$D_P(a,b)(x,y) = bx + ay = \begin{pmatrix} b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Then we may write (using $|hk| \le \sqrt{h^2 + k^2}$)

$$\lim_{(h,k)\to 0} \frac{|P(a+h, b+k) - P(a,b) - (bh + ak)|}{|(h,k)|} = \lim_{(h,k)\to 0} \frac{|hk|}{\sqrt{h^2+k^2}} \le \sqrt{h^2+k^2}.$$

[7] check this... i'm not so sure..? nvm jacobian wikipedia link

[8] yes i used binom i'm a lazy fuck

Theorem 11.1 (Addition, Product, Quotient Rule)**.** If $f, g : (U \subset \mathbb{R}^n) \to \mathbb{R}^m$ are differentiable at $a \in U$, then

(a) $D(f+g)(a) = Df(a) + Dg(a)$.

(b) If $m = 1$, then $D(fg)(a) = g(a)Df(a) + f(a)Dg(a)$.

(c) If $m = 1$ and $g(x) \neq 0$, then $D\left(\frac{1}{g}\right)(a) = -\frac{1}{g(a)^2}Dg(a)$.

The proofs simply follows from our previous examples.

(a) $f + g = S \circ (f+g)$;

(b) $fg = P \circ (f, g)$;

(c) $\frac{1}{g} = \frac{1}{y} \circ g$.

Theorem 11.2 (Jacobian Construction)**.** If $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^m$ is differentiable at a, then $\frac{\partial f_i}{\partial x_j}(a)$ exists for all $1 \leq i \leq n$ and $1 \leq j \leq m$. Then $Df(a)$ is given by the matrix (which we will call the *Jacobian*),

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}(a)\right).$$

To prove this, observe that when $m = 1$, $f : U \to \mathbb{R}$, and $\frac{\partial f}{\partial x_j}(a) = (f \circ h)'(a_j)$, where $h(x) = (a_1, \dots, a_{j-1}, x, a_{j+1}, \dots)$. Then

$$\frac{\partial f}{\partial x_j}(a) = (f \circ h)(a_j) = f'(a) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} j,$$

and so $\frac{\partial f}{\partial x_j}(a)$ is the jth component of $Df(a)$.

For general choices of m, $f = (f_1, \dots, f_m)$; since each f_i is differentiable at a, and $f'(a) = (f_1'(a), \dots, f_m'(a))$, we have that $f_i'(a)$ is the ith row of the matrix $f'(a)$. $\square$

For example, let $f(x, y) = g(x + y)$. Then $Df(c, d)(x, y) = D_g(c + d)(x + y)$, of which the LHS may be expanded as

$$(D_g \circ s)(c, d) \circ \underbrace{D_s(c, d)}_{S}(x, y) = D_g(c + d) \circ s(x, y) = D_g(c + d)(x + y)$$

for more clarity.

§12 Day 12: Computations of Derivatives II (Sep. 30, 2024)

We start with a few examples.

(a) Let $f(x,y) = g(x+y)$ be differentiable. Then

$$D_f(c,d)(x,y) = D_g(c+d)(x+y).$$

(b) Let $f(x,y) = g(xy) = (g \circ p)(x,y)$ be differentiable.

$$D_f(c,d)(x,y) = D_g \circ p(c,d) \circ D_p(c,d)(x,y) = D_g(cd)(dx+cy).$$

(c) Let $f(x,y) = \int_a^{x+y} g$ with g continuous. Then $D_f(c,d)(x,y) = g(c+d)(x+y)$.

(d) Let $f(x,y) = \int_a^{xy} g$ with g continuous. Then $D_f(c,d)(x,y) = g(cd)(dx+cy)$.

(e) Let $F(x) = f(g_1(x),\dots,g_m(x))$. Let $g_1,\dots,g_m : \mathbb{R}^n \to \mathbb{R}$ be differentiable at a. Then $f : \mathbb{R}^m \to \mathbb{R}$ is differentiable at $(g_1(a),\dots,g_m(a))$.

$$\frac{\partial F}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial f}{\partial y_j}(g_1(a),\dots,g_m(a))\frac{\partial g_i}{\partial x_i}(a)$$

Using $F = f \circ g$, where $g = (g_1,\dots,g_m)$ and $F'(a) = Df(g(a)) \cdot D_g(a)$, we get that

$$\left(\frac{\partial F}{\partial x_1}(a),\dots,\frac{\partial F}{\partial x_n}(a)\right) = \left(\frac{\partial f_1}{\partial y_1}(g(a)),\dots,\frac{\partial f_m}{\partial y_m}(g(a))\right)\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(a) & \dots & \frac{\partial g_m}{\partial x_n}(a) \end{pmatrix}.$$

With $y = g(x)$ and $z = f(y)$, we get that

$$\frac{\partial z}{\partial x_i} = \sum_{j=1}^{m} \frac{\partial z}{\partial y_i}\frac{\partial y_i}{\partial x_i}; \qquad \frac{\partial z}{\partial x} = \frac{\partial z}{\partial y}\frac{\partial y}{\partial x}.$$

Recall that whenever f is differentiable at a, then $D_v f(a) = Df(a)(v)$. In particular, $D_v f(a)$ is linear in v. Also recall that f is not necessarily differentiable at a, even if all the directional derivatives at a exist. For example, let

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

Then letting $v = (h,k)$, we get $D_v f(0,0) = \lim_{t\to 0} \frac{f(th,th)-f(0,0)}{t} = \lim_{t\to 0} \frac{h^2k}{t^2h^4+k^2}$, which is equal to $\frac{h^2}{k}$ if $k \neq 0$ and 0 if $k = 0$. Clearly, the directional derivatives are not commensurate to a single value, and so the derivative does not exist. In fact, f is differentiable on every straight line through 0, but it isn't even continuous at 0.

For example, let $f(x,x^2) = \frac{1}{2}$. This function is not continuous at 0. To see this, consider $f(x,mx)$ with $y = mx$, $\frac{mx}{x^2+y^2}$.[9]

[9] what

Theorem 12.1 (Differentiability Condition)**.** Let $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^n$. If all $\frac{\partial f_i}{\partial x_j}$ exist in an open neighborhood of a and are continuous at a, then $Df(a)$ exists.

Start by assuming $m = 1$. Then let us have $a = (a_1, \ldots, a_n), h = (h_1, \ldots, h_n)$.[10]

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2, \ldots, a_n) - f(a_1, \ldots, a_n) \\ &\quad + f(a_1 + h_1, a_2 + h_2, a_3, \ldots, a_n) - f(a_1 + h_1, a_2, \ldots, a_n) \\ &\quad + f(a_1 + h_1, \ldots, a_n + h_n) - f(a_1 + h_1, \ldots, a_{n-1} + h_{n-1}, a_n) \\ &= h_1 \frac{\partial f}{\partial x_1}(b_1, \ldots, a_n) \\ &\quad + h_i \frac{\partial f}{\partial x_i}(a_1 + h_1, \ldots, b_i, a_n) \\ &\quad + \ldots. \end{aligned}$$

Let's write the above as $\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(c_i)$, with $c_i \to a$ as $h \to 0$.

$$\begin{aligned} &\lim_{h \to 0} \frac{f(a+h) - f(a) - \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot h_i}{|h|} \\ &= \lim_{h \to 0} \frac{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(c_i) - \frac{\partial f}{\partial x_i}(a) \right) \cdot h_i}{|h|} \\ &= 0 \end{aligned}$$

by continuity of the partial at a, and $\frac{|h_i|}{|h|} = 1$. □

[10] no idea what's going on here btw

§13 Day 13: Gradient (Oct. 2, 2024)

What is the different of fastest increase of a differentiable function? Let $f : (U \subset \mathbb{R}^n) \to \mathbb{R}$ be differentiable at $a \in U$. Then

$$D_x f(a) = Df(a)(x) = \underbrace{\left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)}_{\nabla f(a)} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \langle \nabla f(a), x \rangle ,$$

where ∇ denotes the gradient of the function f. Invariably, $Df(a) \in (\mathbb{R}^n)^*$ (dual space), so there exists a unique $y \in \mathbb{R}^n$ such that $Df(a)(x) = \langle y, x \rangle$. We denote y by $\mathrm{grad} f(a)$ or $\nabla f(a)$. Let e be a unit vector; then $D_e f(a)$ is the slope of the curve at a, obtained by intersecting the graph of f with vertical planes through the line $x = a + te$ (for $t \in \mathbb{R}$). Then

$$\begin{aligned} D_e f(a) &= \langle \nabla f(a), e \rangle \\ &= |\nabla f(a)| \cos\theta, \end{aligned}$$

where θ is given by the angle between e and $\nabla f(a)$ (let's just assume it's nonzero); i.e., the directional derivative $D_e f(a)$ attains its largest value when $\theta = 0 \implies \cos\theta = 1$, where e is in the direction of $\nabla f(a)$. For example, if $f(x_1, x_2)$ denotes the temperature of a point (x_1, x_2) in the plane, then a heat-seeking bug will move in the direction of the gradient. If $\nabla f(a) = 0$, then there may still be directions of fastest increase, but they are not necessarily unique; for example, graph $z = x^2 - y^2$; we see that z increases along the x axis in both directions.

We now discuss differentiation under the integral sign. Given that $f(x, y)$ is continuous on $[a, b] \times [c, d]$, we have that $\frac{\partial f}{\partial y}$ is also continuous on the rectangle. Let

$$F(y) = \int_a^b f(x, y)\, dx.$$

Then $F(y)$ is continuously differentiable on $[c, d]$, and $F'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx$. More generally, f being C^n implies F being $C^n = \{f : \mathbb{R}^2 \to \mathbb{R} \mid f \text{ is } n \text{ times cont. diff.}\}$. To prove this, let us consider

$$\begin{aligned} &F(y+h) - F(y) - \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx \cdot h \\ &= \int_a^b f(x, y+h) - f(x, y)\, dx - \frac{\partial f}{\partial y}(x, y) \cdot h\, dx. \end{aligned} \tag{13.1}$$

Then $\frac{\partial f}{\partial y}$ is uniformly continuous on $[a, b] \times [c, d]$, so given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|y_1 - y_2| < \delta \implies \left| \frac{\partial f}{\partial y}(x, y_1) - \frac{\partial f}{\partial y}(x, y_2) \right| < \varepsilon,$$

meaning $f(x, y+h) - f(x, y) = \frac{\partial f}{\partial y}(x, y_n) \cdot h$ for some $y_n \in (y, y+h)$ by MVT. Thus,

$$(13.1) = h \int_a^b \frac{\partial f}{\partial y}(x, y_n - y)\, dx,$$

then $\int_a^b \frac{\partial f}{\partial y}(x, y_n - y)\, dx \leq \varepsilon(b - a)$, if $|h| < \delta$, by uniform continuity. Thus, we have that

$$\frac{F(y+h) - F(y) - \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx}{h} \xrightarrow{h \to 0} 0. \quad \square$$

§14 Day 14: Inverse Function Theorem (Oct. 4, 2024)

Let A be an open set on $\mathbb{R}^n$, and let $f : A \to \mathbb{R}^n$ be a C^r map (where $1 \leq r < \infty$). Let $a \in A$; if $Df_a : \mathbb{R}^n \to \mathbb{R}^n$ is invertible, then there exist an open neighborhood U of a where $U \subset A$ such that V of $f(a)$ in $\mathbb{R}^n$, $f|_u : U \to V$ has a C^r inverse $g = (f|_u)^{-1} : V \to U$. Consider the following,

(a) If g is merely differentiable on V, using $f \circ g = \mathrm{id}_V$, for all $y \in V$, we have $Df_{g(y)} \circ Dg_y = D(\mathrm{id}_V)_y = \mathrm{id}_{\mathbb{R}^n}$.

(b) (a) says that $D_g =$ "Inversion" $\circ D_f \circ f^{-1} \in C^\infty \circ C^{r-1} \circ g$. By induction, we see that g is C^r. Thus, in the inverse function theorem, it suffices to prove that g is differentiable (Neumann series).

(c) The proof is trivial in 1 dimension; it is non-trivial in higher dimensions, involving compactness and the Banach contraction fixed point theorem.

(d) Continuity of the derivative is essential; for example, if $f : \mathbb{R} \to \mathbb{R}$ is such that $f'(0) \neq 0$, but f is not invertible in any neighborhood of zero, then we may pick

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

Then

$$f'(x) = \begin{cases} c + 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & x \neq 0, 0 < c < 1, \\ c & x = 0. \end{cases}$$

But $f'(\frac{1}{n}) = c + 0 - (-1)^n \in \{c-1, c+1\}$.

(e) The theorem gives sufficient (but not necessary) conditions. For example, take $f : \mathbb{R} \to \mathbb{R}$. $f(x) = x^3$ is C^∞, and $f'(0) = 0$. Yet it has a globally continuous inverse $f^{-1} : \mathbb{R} \to \mathbb{R}$, $y \mapsto y^{\frac{1}{3}}$. Of course, f^{-1} is not differentiable at $f(0) = 0$, otherwise it would contradict the chain rule.

(f) The theorem is purely local: consider $f : \mathbb{R}^2 \to \mathbb{R}^2$, and $f(x, y) = (e^x \cos y, e^x \sin y)$. Then

$$f'(x, y) = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix},$$

so $\det f'(x, y) = e^{2x} > 0$, but f is not invertible (it is periodic in y).

We now present some examples.

(a) Let $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(r, \theta) = (r \cos \theta, r \sin \theta)$. Then

$$f'(r, \theta) = \begin{bmatrix} \cos theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} \implies det f'(r, \theta) = r,$$

which is obviously nonzero for $r \neq 0$. A typical restriction of f is to $(0, \infty) \times (\theta_1, \theta_2)$ such that $0 < \theta_2 - \theta_1 \leq 2\pi$; for example, $(0, \infty) \times (-\pi, \pi)$ or $(0, \infty) \times (0, 2\pi)$ are both suitable.

Let $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x, y) = \sqrt{x^2 + y^2}$. Then $(g \circ f)(r, \theta) = g(r \cos \theta, r \sin \theta) = \sqrt{r^2} = r$; i.e., changing coordinates (composing f resp. f^{-1}) can simplify things greatly!

(b) Let $f(r, \theta, \varphi) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta)$, where $r \in (0, \infty)$ specifies the radius, $\theta \in (0, \pi)$ specifies the inclination, and $\varphi \in (0, 2\pi)$ is the angle within the plane. This is the parametrization of a sphere (I'm not sure if we did anything else with this in class...).

§15 Day 15: Inverse Function Theorem, Pt. 2 (Oct. 7, 2024)

Given $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^n$ with U open, if $a \in U$ and $\det f'(a) \neq 0$, then

(a) There exists open sets $V \ni a$, $W \ni f(a)$ such that $f : V \to W$ has a continuous inverse $f^{-1} : W \to V$,

(b) f^{-1} is differentiable, and $(f^{-1})'(y) = \big(f'(f^{-1}(y))\big)^{-1}$.

We now make a few remarks;

(a) If we know that f^{-1} is differentiable, then the formula follows directly from the chain rule,
$$f(f^{-1}(y)) = y \implies f'(f^{-1}(y)) \cdot (f^{-1})'(y) = \mathrm{id}.$$

(b) f^{-1} may exist even if $f'(a) = 0$. For example, let $y = x^3$; then $\frac{dy}{dx}\Big|_{x=0} = 0$. However, if $\det f'(a) = 0$, then f^{-1} is not differentiable at $f(a)$, because if it were, then
$$(f^{-1})'(f(a)) \cdot \underbrace{f'(a)}_{\det = 0} = \mathrm{id},$$
which is contradictory. Also, we can't eliminate the hypothesis that f is continuously differentiable.

(c) It follows from the IFT theorem that f is C^1; we want to show that the entries of the matrix $f'(f^{-1}(y))^{-1}$ are continuous. Let A be invertible, and consider $A^{-1}_{ji} = \pm \frac{\det A^{(ij)}}{\det A}$, where $A^{(ij)}$ is A with its ith row and jth column eliminated. Clearly, we see that the entries of A^{-1} are rational functions of entries in A.

(d) If f is C^r (i.e., all partial derivatives of order up to r exist and are continuous), then f^{-1} is C^r:
$$f \text{ is } C^r \iff f \text{ cts. and } D_f \text{ is } C^{r-1},$$
and so f is C^r implies that f is C^{r-1} and D_f is C^{r-1}, meaning f^{-1} is C^{r-1} by induction on r, i.e.
$$(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}$$
is C^{r-1} by the chain rule. $\square$

We'll need the following consequences of the mean value theorem;

Lemma 15.1. Given us have a continuous function $f : B \to \mathbb{R}^m$, where B is a closed ball in $\mathbb{R}^n$, such that f is differentiable on $\mathrm{Int}(B)$, if $\left|\frac{\partial f_i}{\partial x_j}\right| \leq M$ on $\mathrm{Int}(B)$ for all i, j, then $|f(x) - f(y)| \leq \sqrt{mn} M\, |x - y|$ for $x, y \in B$.

If $f = (f_1, \dots, f_n)$ and $|f(x) - f(y)| \leq \sqrt{m} \max_i |f_i(x) - f_i(y)|$, so it is enough to prove that the lemma for $m = 1$, i.e. $f : B \to \mathbb{R}$, $B \subset \mathbb{R}^n$. Let $g(t) = f((1-t)x + ty)$ for given x, y; then
$$\begin{aligned} f(y) - f(x) &= g(1) - g(0) \\ &= g'(t_0) && \text{(for some } t_0 \in (0,1) \text{ by MVT)} \\ &= D_f(c_0)(y - x) && \text{(where } c_0 = (1-t_0)x + t_0 y) \\ &= \langle \nabla f(c_0), y - x \rangle, \end{aligned}$$
and so we have that $|f(y) - f(x)| \leq |\nabla f(c_0)|\, |y - x| \leq \sqrt{n} \max \left|\frac{\partial f}{\partial x_j}\right|$. $\square$

§16 Day 16: Inverse Function Theorem, Pt. 3 (Oct. 9, 2024)

(Note that this is my formulation) If f is a continuously differentiable function from an open subset $U \subset \mathbb{R}^n$ to $\mathbb{R}^n$, and the derivative $f'(a)$ is invertible at a (i.e., $\det f'(a) \neq 0$), then

(a) There exist open neighborhoods V of a, W of $f(a)$ such that $f(V) \subset W$, $f : V \to W$ is bijective, and $f^{-1} W \to V$ is continuously differentiable.

(b) f^{-1} is given by
$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

To prove the theorem, we can assume that $Df(a) = \mathrm{id}$; let $\lambda = Df(a)$, and $g = \lambda^{-1} \circ f$; and so we have that $D_g(a) = \mathrm{id}$ as well. If the theorem is true for g, then it is true for $f = \lambda \circ g$, i.e.
$$f : V \xrightarrow{g} W \xrightarrow{\lambda} \lambda(W),$$
with inverse $g^{-1} \circ \lambda^{-1}$, which is continuous.

(a) To start, we can't have that $f(x) = f(a)$ if $x \neq a$ is sufficiently close to a:
$$\lim_{h \to a} \frac{f(a+h) - f(a) - \overbrace{Df(a)}^{\mathrm{id}} h}{|h|} = 0.$$

If $f(a+h) = f(a)$, then the limit goes to $\frac{h}{|h|}$, which has norm 1. So $f(a+h) \neq f(a)$ for small enough h, i.e. there exists a closed ball B, centered at a, such that $f(x) \neq f(a)$ whenever $x \neq a$ is in B.

We can also assume that for $x \in B$, we have that $\det f'(x) \neq 0$ and
$$\left| \frac{\partial f_i}{\partial x_j} \right| (x) - \frac{\partial f_i}{\partial x_j}(a) < \frac{1}{2n}$$
for all i, j. With this, let $g(x) = f(x) - x$. Then
$$\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a) = \frac{\partial g_i}{\partial x_j}(x),$$
where we may note $\frac{\partial f_i}{\partial x_j}(a)$ is literally just entries from the identity matrix. By the mean value lemma, we have that
$$\begin{aligned}
|g(x_1) - g(x_2)| &\leq n \frac{1}{2n} |x_1 - x_2|, & (x_1, x_2 \in B) \\
\implies |f(x_1) - x_1 - [f(x_2) - x_2]| &\leq \frac{1}{2} |x_1 - x_2| \\
\implies |x_1 - x_2| &\leq \mathrm{LHS} + |f(x_1) - f(x_2)| \\
&\leq \frac{1}{2} |x_1 - x_2| + |f(x_1) - x_2| \\
&\leq 2 |f(x_1) - f(x_2)|. & (*)
\end{aligned}$$

In particular, f is one to one on B. $\square$

Claim 16.1. $f(\text{Int}(B))$ is open in $\mathbb{R}^n$.

Once we prove this, we get (a) by taking $V = \text{It}(B)$, $W = f(V)$; then f^{-1} is continuous by $(*)$, i.e.

$$\left|f^{-1}(y) - f^{-1}(y_2)\right| \le 2\left|y_1 - y_2\right|.$$

To prove the claim, consider $x_0 \in \text{Int}(B)$, $y_0 \in f(x_0)$. We have to find an open ball containing y_0 in $f(\text{Int}(B))$. Let $d = d(y_0, f(\text{Bdry}(B))$. Let us show that $B(y_0, \frac{d}{2}) \subset f(\text{Int}(B))$; we will define $B_{\frac{d}{2}} := B(y_0, \frac{d}{2})$ from here on. Let $y \in B_{\frac{d}{2}}$; if $x \in \text{Bdry}(B)$, then $|y - y_0| < |y - f(x)|$. Then we just need to find $x \in \text{Int}(B)$ such that $y = f(x)$. Define $h : B \to \mathbb{R}$, where $h(x) = |y - f(x)|^2 = \sum_{i=1}^n (y_i - f_i(x))^2$. h is continuous, and so it attains minimum on B. By $|y - y_0| < |y - f(x)|$, it cannot occur on the boundary of B, and so it is strictly in the interior, at a critical point x, i.e. $\frac{\partial h}{\partial x_j} = 0$ for all j.

Thus, we have that

$$\sum_{i=1}^n 2(y_i - f_i(x)) \frac{\partial f_i}{\partial x_j}(x) = 0,$$

where $\frac{\partial f_i}{\partial x_j}(x)$ are the entries of an invertible matrix. Therefore, $y_i - f_i(x) = 0$ for all i, which concludes the proof. $\square$

§17 Day 17: Inverse Function Theorem, Pt. 4 (Oct. 11, 2024)

If $f : (U \subset \mathbb{R}^n) \to \mathbb{R}^n$ is continuously differentiable, given $a \in U$ such that $\det f'(a) \neq 0$, we have that

(a) There are open neighborhoods V of a, W of $f(a)$, such that $f : V \to W$ has a continuous inverse $f^{-1} : W \to V$ with the property

$$\left|f^{-1}(y_1) - f^{-1}(y_2)\right| \leq 2\left|y_1 - y_2\right|$$

for any $y_1, y_2 \in W$.

(b) Today in class, we will show that f^{-1} is differentiable, and $(f^{-1})'(y) = f'(f^{-1}(y))^{-1}$.

Let $y_0 \in W$, $x_0 = f^{-1}(y_0)$, and $\mu = Df(x_0)$. We want to show that f^{-1} is differentiable at y_0, and $(f^{-1})'(y_0) = \mu^{-1}$. Then observe that $f(x) = f(x_0) + \mu(x - x_0) + \varphi(x - x_0)$, where

$$\lim_{x \to x_0} \frac{\varphi(x - x_0)}{|x - x_0|} = 0.$$

We have that $\mu^{-1}(f(x_0) - f(x)) = x - x_0 + \mu^{-1}\varphi(x - x_0)$, and each $y \in W$ can be written $y = f(x)$, $x \in V$. Write

$$\mu^{-1}(y - y_0) = f^{-1}(y) - f^{-1}(y_0) + \mu^{-1}\varphi(f^{-1}(y) - f^{-1}(y_0)).$$

We have to show that

$$\lim_{y \to y_0} \frac{\mu^{-1}\varphi(f^{-1}(y) - f^{-1}(y_0))}{y - y_0} = 0.$$

It is enough to show that the inside term goes to 0. Write

$$\lim_{y \to y_0} \frac{\varphi(f^{-1}(y) - f^{-1}(y_0))}{y - y_0} = \lim_{y \to y_0} \underbrace{\frac{\varphi(f^{-1}(y) - f^{-1}(y_0))}{\left|f^{-1}(y) - f^{-1}(y_0)\right|}}_{=0} \cdot \underbrace{\frac{\left|f^{-1}(y) - f^{-1}(y_0)\right|}{|y - y_0|}}_{\leq 2}.$$

Notice that the first term goes to 0 as $y \to y_0$ because $f^{-1}(y) \to f^{-1}(y_0)$ by continuity of f^{-1}. The second term is bounded above by 2, as per our lemma in (a). Thus, we see that the limit as a whole goes to 0 as $y \to y_0$. Since we already have $\frac{\varphi(t)}{|t|} \to 0$ as $t \to 0$ (we may just take $t = f^{-1}(y) - f^{-1}(y_0)$), we are done as per our earlier observation. $\square$

At the beginning of the proof, we said that we could assume that $f'(a) = \text{id}$. Let $\lambda = f'(a)$, $g = \lambda^{-1} \circ f$. Then $Dg(a) = \lambda^{-1} \circ Df(a) = \text{id}$.

Claim 17.1. If g satisfies the theorem, then so does f.

We have that $f = \lambda \circ g$, meaning $f : V \xrightarrow{g} W \xrightarrow{\lambda} \lambda(W)$; we also have $f^{-1} = g^{-1} \circ \lambda^{-1}$. If g^{-1} is continuous, we have that f^{-1} is continuous; the same goes with differentiability, where if g^{-1} is differentiable, then so is f^{-1} by the chain rule. We may also observe that W is an open set for g, and $\lambda(W)$ an open set for f. $\square$

Test administrative details! There will be 3 to 4 problems, similar to the ones on the problem set. Make sure you understand how to do all the problems on the past homeworks, even better if textbooks one are understood well too. Material is covered up to more or less a week before the test. The test is supposed to be straightforward

§18 Day 18: Implicit Function Theorem (Oct. 16, 2024)

We start with an example. Let $f(x,y) = x^2 + y^2 - 1$, and consider the region on $\mathbb{R}^2$ in which $f(x,y) = 0$. Can we solve $f(x,y) = 0$ for $y = g(x)$ near $x = a$, with $g(a) = b$? If $a \neq \pm 1$, then there are open intervals $I \ni a$ and $J \ni b$ such that, for all $x \in I$, there is a unique $y \in J$ such that $f(x,y) = 0$. Writing $y = g(x)$, we are done.

In general, we can solve for $y = g(x)$ whenever $\partial_y f(a,b) \neq 0$. We may proceed in the example with implicit differentiation; let $f(x,g(x)) = 0$, and consider

$$\frac{\partial f}{\partial x}(x,g(x)) + \frac{\partial f}{\partial y}(x,g(x))\frac{dg}{dx} = 0.$$

In our example, this is given by $2x + 2y\frac{dg}{dx}$. Then

$$g'(x) = -\frac{\partial_x f(x,g(x))}{\partial_y f(x,g(x))}$$

near $x = a$, which means $g'(x) = -\frac{x}{y}$ on the curve near (a,b) with $b \neq 0$. Of course, $g(x) = \sqrt{1-x^2}$; we may consider $(a,-b)$ and proceed to solve for $g_1(x) = -\sqrt{1-x^2}$.

In general, in a system of n equations, let

$$\begin{aligned} f_1(x_1,\ldots,x_m,y_1,\ldots,y_n) &= 0, \\ &\vdots \\ f_n(x_1,\ldots,x_m,y_1,\ldots,y_n) &= 0, \end{aligned}$$

where $f_1,\ldots,f_n$ are functions in $m+n$ variables, i.e. $f(x,y) = 0$ where $x = (x_1,\ldots,x_m) \in \mathbb{R}^m$, $y = (y_1,\ldots,y_n) \in \mathbb{R}^n$, and $f = (f_1,\ldots,f_n) : \mathbb{R}^{m+n} \to \mathbb{R}^n$, given (a,b), can we find, for all x close to $a \in \mathbb{R}^m$, a unique y near b such that $f(x,y) = 0$? (note that $y = g(x)$)

Consider the linear equation $Ax + By = 0$, where $A \in M_{n\times m}(\mathbb{R})$ and $B \in M_{n\times n}(\mathbb{R})$. If $\det B \neq 0$, then B is invertible, meaning we can solve for y as a function of x, i.e. $y = -B^{-1}Ax$. Then $Ax + By = B(B^{-1}Ax + y)$.

Theorem 18.1 (Implicit Function Theorem)**.** *If U is an open set in $\mathbb{R}^m \times \mathbb{R}^n$, consider $f : U \to \mathbb{R}^n$ and let f be of class C^r, i.e. all partials exist up to order r and are continuous. Then for $(a,b) \in U, f(a,b) = 0$, then let*

$$M = \left(\frac{\partial f_i}{\partial y_j}(a,b)\right) = \frac{\partial(f_1,\ldots,f_n)}{\partial(y_1,\ldots,y_n)}(a,b).$$

If $\det M \neq 0$, then there exist open sets $A \ni a$ in $\mathbb{R}^m$, $B \ni b$ in $\mathbb{R}^n$ such that, for all $x \in A$, there exists a unique $y \in B$ with $f(x,y) = 0$, i.e. $y = g(x)$. Moreover, g is of class C^r.

Let $F : U \to \mathbb{R}^m \times \mathbb{R}^n$, where $F(x,y) = (x, f(x,y))$ with $F(a,b) = (a,0)$. Then

$$F'(a,b) = \left(\begin{array}{c|c} I_m & 0 \\ \hline * & M \end{array}\right) \implies \det F'(a,b) = \det M \neq 0.$$

By the inverse function theorem, there are open sets $U \ni (a,b)$ and $V \ni (a,0)$ such that $F : U \to V$ has C^r inverse $F^{-1} : V \to U$. We can assume U is of the form $A \times B$ by taking a smaller open set. Let $(u, v) = F(x, y)$, which implies that F^{-1} is of the form $(u, h(u, v))$. Then

$$\begin{aligned}
&F(F^{-1}(u,v)) = (u,v)\\
\implies &F(u,h(u,v)) = (u,v)\\
\implies &(u, f(u,h(u,v))) = (u,v)\\
\implies &f(u,h(u,v)) = (u,v)\\
\implies &f(u,h(u,0)) = 0.
\end{aligned}$$

Thus, $f(x, g(x)) = 0$ where $g(x) = h(x, 0)$ is C^r. $\square$

§19 Day 19: Implicit Function Theorem, Pt. 2 (Oct. 18, 2024)

Recall that given $f(x, y) = 0$ where

$$\begin{aligned} f &= (f_1, \dots, f_n), \\ x &= (x_1, \dots, x_m), \\ y &= (y_1, \dots, y_n), \end{aligned}$$

with f being C^r in a neighborhood of (a, b) in $\mathbb{R}^{m+n}$, and $f(a, b) = 0$, if $\det \frac{\partial f_i}{\partial y_j}(a, b) \neq 0$, then we can solve $f(x, y) = 0$ for a C^r function $y = g(x)$ satisfying $g(a) = b$, i.e. $f(x, g(x)) = 0$, i.e. there are open sets $A \ni a$, $B \ni b$ such that, for all $x \in A$, there exists a unique $y \in B$ such that $f(x, y) = 0$. Write $y = g(x)$; moreover, g is C^r.

We may find $g'(x)$ by implicit differentiation, i.e. at $f_i(x, g(x)) = 0$, for all $i = 1, \dots, n$, we have

$$\frac{\partial f_i}{\partial x_j}(x, g(x)) + \underbrace{\sum_{k=1}^{n} \frac{\partial f_i}{\partial y_k}(x, g(x))}_{\text{Matrix w/ invt. entries near } (a,b)} \frac{\partial g_k}{\partial x_j}(x)$$

$$\left(\frac{\partial g_k}{\partial x_j}(x)\right) = -\left(\frac{\partial f_i}{\partial y_k}(x, g(x))\right)^{-1} \left(\frac{\partial f_i}{\partial x_j}(x, g(x))\right).$$

The answer, of course, depends on $g(x)$. Moreover, since the implicit function theorem is more or less a generalization of the inverse function theorem, we prove that ImFT implies IFT. Let $U \subset \mathbb{R}^n$, and given $f : U \to \mathbb{R}^n$, consider that f is C^r in an open set U, and for $a \in U$, $\det f'(a) \neq 0$. Then let $F(x, y) = y - f(x)$, $b = f(a)$, $F(a, b) = 0$, where F is C^r in a neighborhood of (a, b). Then $\det \partial_x F(a, b) = \det(-f'(a))$. By the implicit function theorem, there are open neighborhoods A of G, B of A, such that for all $y \in A$, there exist unique $x \in B$ (using $x = g(y)$) such that $F(g(y), y) = 0$, i.e. $y = f(g(y))$. We may take $V = B \cap f^{-1}(A), W = A$ to see that $f : V \to W$ has C^r inverse g. $\square$

We go over a few examples now. Let $y^2 = x^2(x + 1)$ be a curve in $\mathbb{R}^2$, and let $f(x, y) = y^2 - x^2(x + 1)$; on $f(x, y) = 0$, $\frac{\partial f}{\partial y} = 2y \neq 0$ at every point of the curve where $y \neq 0$, so we can solve for y as a function of x.

$$\partial_x f = -3x^2 - 2x = -x(3x + 2) \neq 0$$

except when $x = 0$ or $-\frac{2}{3}$. IVT allows us to distinguish between smooth points and singularities; if f is smooth everywhere in 1 dimension, we call it a manifold. For example, Whitney's umbrella, $X : x^2 - xy^2 = 0$ is smooth according to the implicit function theorem except on the x axis.

§20 Day 20: Implicit Function Theorem, Pt. 3; Rank Theorem (Oct. 21, 2024)

We start with a few remarks on the implicit function theorem.

(a) Consider $f(x, y, z) = z^2 - xy^2$;

$$\frac{\partial f}{\partial z} = 2z, \qquad \frac{\partial f}{\partial x} = -y^2.$$

For the first partial, observe that we can solve for $z = g(x, y)$ at any point except where $z = 0$. For the second partial, we see that $-y^2 \neq 0$ as long as $y \neq 0$, meaning we may solve $x = h(y, z)$ near any point $(0, b, 0)$ with $b \neq 0$. Near $(a, 0, 0)$, with $a < 0$, we see that x is defined locally by $y = z = 0$, i.e. $f : (U \subset \mathbb{R}^3) \to \mathbb{R}^2$, we have $f(x, y, z) = (y, z)$.

(b) For coordinate changes, recall that if we take $f : \mathbb{R}^{m+n} \to \mathbb{R}^n$ and consider f as a function on (x, y), where $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, if $f(a, b) = 0$, and if

$$\det \frac{\partial f}{\partial y}(a, b) = \det \left(\frac{\partial f_i}{\partial y_j}(a, b) \right) \neq 0,$$

then there exists a function $F(x, y) = (x, f(x, y))$ with an inverse H near (a, b). We note that H is given by $H(u, v) = (u, h(u, v))$, with $F \circ H = \text{id}$, and $(u, F \circ H(u, v)) = (u, v)$. That is, $(F \circ H)(u, v) = v$.

(c) Suppose $f : \mathbb{R}^p \to \mathbb{R}^n$ has rank n at a and $f(a) = 0$, i.e.

$$\left(\frac{\partial f_i}{\partial x_j}(a) \right)$$

has rank n for $1 \leq i \leq n$, $1 \leq j \leq p$. This means we may choose indices $j_1 < \cdots < j_n$ such that $\partial_{j_1, \ldots, y_n}(f_1, \ldots, f_n)(a)$ is invertible. For convenience, let us write $(x_{j_1}, \ldots, x_{j_n}) = (y_1, \ldots, y_n)$ and $(z_1, \ldots, z_{p-n})$ be the other x's. By the implicit function theorem, we can solve for $(y_1, \ldots, y_n)$ as a function of $(z_1, \ldots, z_{p-n})$. Let $x = P(z, y)$, where P is a linear transformation given by a change of coordinates. Then $(f \circ P)(z, y)$ satisfies $\partial_y (f \circ P)(P(n))$ being invertible, which means there necessarily exists a change of coordinates $(z, y) = H(u, v)$ such that $(f \circ P \circ H)(u, v) = v$. $\square$

We now introduce the rank theorem. Let $f : \mathbb{R}^p \to \mathbb{R}^n$ be C^n, and assume that f has rank r at every point m in a neighborhood of a given point a (so $r \leq p, r \leq n$). In (c) as above, $r = n$; in general, when $r = n$ is rank r at a, then it is rank r in some open neighborhood about a by continuity of determinant. Then there exists a coordinate change $x = H(u, v)$ with $v = (v_1, \ldots, v_r)$ and $u = (u_1, \ldots, u_{p-r})$, and some coordinate change K in the target space such that $(K \circ f \circ H)(u, v) = (v, 0)$. Then we may find indices $i_1 < \cdots < i_r$, $j_1 < \cdots < j_r$ such that $\partial_{j_1, \ldots, j_r}(f_{i_1}, \ldots, f_{i_r})(a)$ has rank r. We can assume that $(i_1, \ldots, i_r) = (1, \ldots, r)$ by permuting the coordinates in the target space; then there is a coordinate change $x = H(u, v)$, with $v = (v_1, \ldots, v_r)$ such that $(f_1, \ldots, f_r) \circ H(u, v) = v$, i.e.

$$\underbrace{(f \circ H)}_{=(f_1, \ldots, f_r)}(u, v) = (v, (f_{r+1}, \ldots, f_n) \circ H(u, v)).$$

§21 Day 21: Rank Theorem, Tangent Spaces (Oct. 23, 2024)

Let $f : \mathbb{R}^p \to \mathbb{R}^n$ be C^r in a neighborhood of a. Suppose f has constant rank r in a neighborhood of a; then there exists a C^r coordinate change $x = H(u,v)$ near a, and K near $f(a)$ such that $K \circ f \circ H : (u,v) \to (v,0)$, where u, v are $p-r, r$-tuples, i.e. $u = u_1, \dots, u_{p-r}$, and $v = v_1, \dots, v_r$. We also showed that after a permutation of coordinates in the target space, there exists a coordinate change $x = H(u,v)$ such that

$$\begin{aligned}(f \circ H)(u,v) &= (v, \varphi(u,v)),\\ v &= v_1, \dots, v_r,\\ \varphi &= \varphi_1, \dots, \varphi_{n-r}.\end{aligned}$$

(a) φ follows the index of u, i.e.

$$D(f \circ H) = \left(\begin{array}{c|c} 0 & I_r \\ \hline \partial_{u_i}\varphi_k & \partial_{v_j}\varphi_k \end{array}\right),$$

where $i = 1, \dots, p-r$, $j = 1, \dots, r$, and $k = 1, \dots, n-r$. Thus, rank r implies that $\partial_{u_i}\varphi_k = 0$ for all i, k, and so $\varphi(u,v) = \varphi(v)$; this means we have $(f \circ H)(u,v) = (v, \varphi(v))$ with (v, w) being the coordinates in the target space.

(b) We can make the change of coordinates K in the target space *after* the specific φs vanish, i.e. $(v, w) \mapsto (v, w - \varphi(v))$.

We now move onto tangent spaces.[11]

(a) Let us have a plane curve C, given by $f(x,y) = 0$ of which is C^r, with $\partial_y f(a,b) \neq 0$. By IVT, we can solve this as $y = g(x)$ being C^r near a, where $\varphi(a) = b$. The tangent space at (a,b) is given by $y - b = g'(a)(x-a)$, i.e. $f(x, g(x)) = 0$ and $\partial_x f(a,b) + \partial_y f(a,b) g'(a) = 0$. Expanding $g'(a)$, we get

$$\frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) = 0$$

is a tangent space to C at (a,b). By symmetry, we obtain an equivalent result if $\partial_x f(a,b) \neq 0$: if $\nabla f(a,b) \neq 0$, we can define a tangent line at (a,b) as $\partial_x f(a,b)(x - a) + \partial_y f(a,b)(y-b) = 0$. We say C is C^r smooth at (a,b), or a C^r manifold near (a,b).

(b) Let $M \subset \mathbb{R}^n$ be given by the set of points $f(x_1, \dots, x_n) = 0$, with f being C^r. Suppose $\nabla f(a) \neq 0$, i.e. $\partial_{x_i} f(a) \neq 0$ for some i. For example, if $i = n$, then $x_n = g(x_1, \dots, x_{n-1})$ near $(a_1, \dots, a_{n-1})$, where g is C^r, and $g(a_1, \dots, a_{n-1}) = a_n$. Then the tangent space at a is given by

$$x_n - a_n = g'(a_1, \dots, a_{n-1}) \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_{n-1} - a_{n-1} \end{pmatrix} = \sum_{i=1}^{n-1} \frac{\partial g}{\partial x_i}(a_1, \dots, a_{n-1})(x_i - a_i),$$

where $f(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})) = 0$. Then for all $i = 1, \dots, n-1$, we have

$$\frac{\partial f}{\partial x_i}(a) + \frac{\partial f}{\partial x_n}(a) \frac{\partial g}{\partial x_i}(\tilde{a}) = 0.$$

[11] mb if information isn't the most accurate, i am comprehending it myself atm x3

Then the $x(x_i - a_i)$ terms add up for $i = 1, \ldots, n-1$ to give

$$\frac{\partial f}{\partial x_i}(a)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_i}(a)(x_n - a_n) = 0.$$

Specifically, if $\nabla f(a) \neq 0$, we can define a tangent space at a as the above formula, i.e. $\langle \nabla f(a), x - a \rangle = 0$. We say that M is C^r smooth at a, or a C^r manifold at a of dimension $n-1$.

§22 Day 22: Tangent Spaces Pt. 2 (Oct. 25, 2024)

We start with an example;

(a) Let us find the tangent planes to the ellipsoid

$$X : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

at a point $(x_0, y_0, z_0) \in X$. Since tangent planes are orthogonal to the derivative at (x_0, y_0, z_0), we have

$$\frac{2x_0}{a^2}(x - x_0) + \frac{2y_0}{b^2}(y - y_0) + \frac{2z_0}{c^2}(z - z_0) = 0,$$

or

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} + \frac{z_0 z}{c^2} = 1.$$

In general, for $f : \mathbb{R}^p \to \mathbb{R}^n$, with $p \geq n$ and $M : \{x \in \mathbb{R}^p \mid f(x) = 0\}$, given a point $a \in M$ where f has rank n, we have

$$\det \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}(a) \neq 0.$$

For example, $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, where $\det \partial_y f(a, b) \neq 0$, by the implicit function theorem, we can solve for $y = g(x)$ near (a, b), i.e. $g(a) = b$, $f(x, g(x)) = 0$ near $x = a$. The tangent space at (a, b) is given by $y - b = g'(a)(x - a)$, i.e.

$$\begin{pmatrix} y_1 - b_1 \\ \vdots \\ y_n - b_n \end{pmatrix} = \frac{\partial g_i}{\partial x_j}(a) \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix}.$$

Given $f(x, g(x)) = 0$, we have $\partial_{x_i} f(a, b) + \sum_{j=1}^n \partial_{y_j} f(a, b) \partial_{x_i} g(a) = 0$ for $i = 1, \dots, m$, and $\partial_x f(a, b) + \partial_y f(a, b) g'(a) = 0$. Thus, we have

$$g'(a) = -\frac{\partial_y f(a, b)}{\partial_x f(a, b)},$$

i.e. the tangent space to M at a is $Df(a)(x - a) = 0$. When f has rank n at a, i.e., wherever tangent spaces may be defined, we say M is C^r smooth at a, or a manifold of dimension $p - n$ near a.

(a) Let $f(x, y) = 0$ and $g(x, y) = 0$ be two smooth curves in $\mathbb{R}^2$; we consider them at a common point (a, b) (meaning (a, b) lies on both the curves), i.e. $\nabla f(a) \neq 0$ and $\nabla g(a) \neq 0$. For example, the tangent lines to $f = 0$ at (a, b) are represented by

$$\frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) = 0.$$

The curves are orthogonal at (a, b) if

$$\left(\frac{\partial f}{\partial x} \frac{\partial g}{\partial x} + \frac{\partial f}{\partial y} \frac{\partial g}{\partial y} \right)(a, b) = 0,$$

and tangent at (a, b) if

$$\left(\frac{\partial f}{\partial x} \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial y} \right)(a, b) = 0.$$

(b) How to find a point of a surface $g(x, y, z) = 0$ which has least (or greatest) distance from the origin? i.e., find the extreme values[12] of $f(x, y, z) = x^2 + y^2 + z^2$ on a given surface $g(x, y, z) = 0$. Supposing we may solve $g(x, y, z)$ as $z = h(x, y)$, we now want to find the extreme values of $f(x, y, h(x, y))$; if the functions are C^r for $r \geq 1$, then extreme points occur at critical points, i.e.

$$\begin{aligned} f_x + f_z h_x &= 0, \\ f_y + f_z h_y &= 0, \\ g_x + g_z h_x &= 0, \\ g_y + g_z h_y &= 0. \end{aligned}$$

Then $f_x + \lambda g_x = 0$, $f_y + \lambda g_y = 0$, $f_z + \lambda g_z = 0$, where $\lambda = -\frac{f_z}{g_z}$, and $g(x, y, z) = 0$. There are 4 equations and 4 unknowns, so we may solve this. $\square$

[12] root is order preserving, so we may just look at f as the squares of components without rooting them all after

§23 Day 23: Lagrange's Method (Nov. 4, 2024)

We introduce Lagrange's method for obtaining extreme points on an objective function subject to a constraint.

Theorem 23.1 (Lagrange Method)**.** Let $g : U \to \mathbb{R}^p$ be $\mathcal{C}^r$ and $f : U \to \mathbb{R}$ be differentiable, with a local maximum or mnimum at some point $a \in g^{-1}(0)$. If g has rank p at a, i.e. $g^{-1}(0)$ is $\mathcal{C}^r$ smooth at a or $\nabla g_i(a)$ is linearly independent for p choices of i, then there exists $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ called *multipliers* such that we may construct the system of equations (where $i = 1, \dots, p$)

$$\frac{\partial f}{\partial x_j}(a) + \lambda_i \frac{\partial g_i}{\partial x_j}(a) = 0,$$

as well as $g_i(a) = 0$. In particular, this may be written as $\nabla f(a) + \lambda \nabla g(a) = 0$, where $\lambda = (\lambda_1, \dots, \lambda_p)$.

To prove this, let us have

$$\det \frac{\partial(g_1, \dots, g_n)}{\partial(x_{j_1}, \dots, x_{j_p})} \neq 0$$

for some $j_1 < j_2 < \cdots < j_p$. By permutation of variables, we can assume that $(x_{j_1}, \dots, x_{j_p}) = (x_{n-p+1}, \dots, x_n)$; with this, let us write $x = (u, v)$, where $(u, v) = (u_1, \dots, u_{n-p}, v_1, \dots, v_p)$. By the implicit function theorem, we solve $g(u, v) = 0$ for v as a $\mathcal{C}^1$ function $v = h(u)$, i.e. $g(u, h(u)) = 0$ over $u = (a_1, \dots, a_{n-p}))$; this means $f(u, h(u))$ has a local extremum at $u = (a_1, \dots, a_{n-p})$, i.e. $(a_1, \dots, a_{n-p})$ is a critical point of $\varphi(v) = f(u, h(u)) = f \circ H(u)$, where H is given by $H(u) = (u, h(u))$. Then we may write

$$\varphi'(u) = f'(u, h(u)) \begin{pmatrix} I \\ h'(u) \end{pmatrix},$$

and we have

$$0 = f'(a) \begin{pmatrix} I \\ h'(a_1, \dots, a_{n-p}) \end{pmatrix} = \underbrace{\nabla f(a)}_{\in \ker} \begin{pmatrix} I \\ h'(a_1, \dots, a_{n-p}) \end{pmatrix}.$$

Since $g_i(u, h(u)) = 0$ for $i = 1, \dots, p$, we have that

$$0 = g_i'(a) \begin{pmatrix} I \\ h'(a_1, \dots, a_{n-p}) \end{pmatrix} = \nabla g_i(a) \begin{pmatrix} I \\ h'(a_1, \dots, a_{n-p}) \end{pmatrix};$$

since $\nabla g_i(a)$ is linearly independent and the latter matrix is of rank $n - p$, we conclude that $\dim \ker = p$, i.e. $\nabla f(a)$ is a linear combination of $\nabla g_i(a)$. □

> **Remark 23.2.** To apply Lagrange's method, let us have a constraint function g and objective function f. If $\nabla g \neq 0$ over $g^{-1}(\{0\})$ and the set of all points where $g = 0$ is compact (it is useful to check the continuity of g, then use Heine-Borel here), then we may apply EVT to $f : g^{-1}(\{0\}) \to \mathbb{R}$ to see that it must admit an extreme point.

We now provide some examples.

(a) Suppose we want to find the point on a $\mathcal{C}^1$ hypersurface $g(x_1, \dots, x_n) = 0$ at a local maximum or minimum distance from a fixed point x. We may use the objective function $f(x) = |x - c|^2$.[13] This method may fail if $g^{-1}(0)$ is not smooth at a.

[13] what?

(b) Find the triangle with given perimeter and largest possible area. We use the formula $f(x, y, z) = s(s-x)(s-y)(s-z)$, where the perimeter is given by $2s$, to find the area of the triangle.

Here, we want to find the maximum of f subject to the condition $g(x, y, z) = x + y + z - 2s = 0$, where $x, y, z \geq 0$, $x + y \geq z$, $y + z \geq x$, and $z + x \geq y$. On the boundary of the feasible set, one of the inequalities above becomes an equality, and so we have $f = 0$; thus, the maximum of f occurs at a critical point on the interior. Write

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0,$$

which yields the system

$$\begin{aligned} -s(s-y)(s-z) + \lambda &= 0, \\ -s(s-x)(s-z) + \lambda &= 0, \\ -s(s-x)(s-y) + \lambda &= 0. \end{aligned}$$

Substituting z, we have $x = y$, along with $x = z$, $y = z$ by symmetry. Thus, the area of the triangle is maximized when it is an equilateral.

§24 Day 24: Holder's Inequality, Taylor Expansion (Nov. 6, 2024)

(a) Prove that $uv \leq \frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta$ for $u, v \geq 0$, $\alpha, \beta > 0$ such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Start by substituting

$$u = \frac{u_i}{\left(\sum_{j=1}^n u_j^\alpha\right)^{1/\alpha}}, \qquad v = \frac{v_i}{\left(\sum_{j=1}^n v_j^\beta\right)^{1/\beta}},$$

where $u_i, v_i \geq 0$, and at least one of $u_i, v_i \neq 0$. Summing over $i = 1, \dots, n$, we have

$$\sum_{i=1}^n u_i v_i \leq \left(\sum_{j=1}^n u_j^\alpha\right)^{1/\alpha} \left(\sum_{j=1}^n v_j^\beta\right)^{1/\beta}$$

for all $u_i, v_i \neq 0$ and all $\alpha, \beta > 0$ such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Observe that for any $t > 0$, the inequality holds for u, v if and only if it holds for $U = ut^{1/\alpha}$ and $V = vt^{1/\beta}$, since

$$\begin{aligned} uv &\leq \frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta \\ uvt &\leq \left(\frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta\right)t \\ UV &\leq \frac{1}{\alpha}U^\alpha + \frac{1}{\beta}V^\beta \end{aligned}$$

The inequality is trivial if u or $v = 0$, so we may assume that $uv \neq 0$. Given u, v such that $uv \neq 0$, if we let $t = uv$, we get $UV = 1$. It's enough to prove that the inequality holds when $uv = 1$, i.e.

$$\frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta \geq 1$$

when $u, v \geq 0$ and $uv = 1$. Using Lagrange's method, we want to minimize $f(u, v) = \frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta$ on the curve $uv = 1$. The minimum must occur where the curve is smooth, so by Lagrange's method, the minimum occurs at a critical point where

$$\frac{1}{\alpha}u^{1/\alpha}\frac{1}{\beta}v^\beta - \lambda uv,$$

i.e., said point satisfies

$$\begin{aligned} u^{\alpha-1} - \lambda v &= 0 \\ v^{\beta-1} - \lambda u &= 0 \\ u^\alpha &= \lambda \\ v^\beta &= \lambda. \end{aligned}$$

Also, we have that $uv = 1$, implying we have $\lambda = 1, u = 1, v = 1$. Thus, we see

$$\frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta \geq \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

(b) We may apply the second derivative test in higher dimensions. Recall that

$$f(a+h) = f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(k)}(a)}{k!}h^k;$$

in n variables, we have that the sum of terms are given by

$$\frac{1}{k!}\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(a)h_{i_1} \ldots h_{i_k}.$$

We have that the sum of like terms may be written as follows,

$$\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = k}} \frac{1}{\alpha!}\frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a)h^\alpha.$$

Note that $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\alpha! = \alpha_1! \ldots \alpha_n!$, and $h^\alpha = h_1^{\alpha_1} \ldots h_n^{\alpha_n}$. There are exactly $\frac{k!}{\alpha!}$ different ways to choose the partial in the expression by combinatorics.

§25 Day 25: Taylor Expansion in Several Variables (Nov. 8, 2024)

Let $f(x_1, \ldots, x_n)$ be differentiable at $a = (a_1, \ldots, a_n)$. Then

$$f(a+h) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i + O(|h)|),$$

which is the best linear approximation of f at a. Assuming that f has derivatives up to $k+1$ in a neighborhood of a, let $F(t) = f(a+th)$, with $0 \leq t \leq 1$, be a parameterization of f from a to $a+h$. Then the Taylor expansion of order k of $F(t)$ at $t = 0$ with the Lagrange remainder is given by

$$F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \cdots + \frac{F^{(k)}(0)}{k!}t^k + \frac{F^{(k+1)}(\theta t)}{(k+1)!}t^{k+1},$$

for some $0 \leq \theta \leq 1$. Note that this is a Taylor polynomial of order k. Now, let us set $t = 1$. We have $f(a+h)$ has the Taylor expansion

$$\begin{aligned} f(a) &+ \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_i h_j + \ldots \\ &+ \frac{1}{k!}\sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(a)h_{i_1} \ldots h_{i_k} + R, \end{aligned}$$

where R is the Lagrange remainder term given by

$$R = \sum_{i_1=1}^{n} \cdots \sum_{i_{k+1}=1}^{n} \frac{1}{(k+1)!} \frac{\partial^{k+1} f}{\partial x_{i_1} \ldots \partial x_{i_{k+1}}}(a+\theta h)h_{i_1} \ldots h_{i_{k+1}}.$$

Note that the summations are slightly clunky; we can really just write them as tuples of indices as seen in last lecture; i.e.,

$$f(a+h) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq k}} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a)h^\alpha + R.$$

Of course, we may expand the notation used in here as follows,

$$\begin{aligned} \alpha &= (\alpha_1, \ldots, \alpha_n), \\ |\alpha| &= \alpha_1 + \cdots + \alpha_n, \\ x^\alpha &= x_1^{\alpha_1} \ldots x_n^{\alpha_n}, \\ \alpha! &= \alpha_1! \ldots \alpha_n!, \\ \partial^{|\alpha|} &= \partial^{\alpha_1} \ldots \partial^{\alpha_n}, \\ \partial x^\alpha &= \partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}. \end{aligned}$$

The remainder term R can be written using tuples as well; i.e.,

$$R = \sum_{|\alpha|=k+1} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a+\theta h)h^\alpha.$$

We now give an application: the escond derivative test for local extrema in 2 variables. Let $f(x, y)$ be $\mathcal{C}^2$ with critical point (a, b). Let us consider $f(a+h, b+k) - f(a, b)$, given by

$$\frac{1}{2}\left(f_{xx}(a^*, b^*)h^2 + 2f_{xy}(a^*, b^*) + f_{yy}(a^*, b^*)k^2\right),$$

where a^*, b^* are given by $a + \theta h, b + \theta h$ respectively. Consider the quadratic form $Q(h, k) = Ah^2 + 2Bhk + Ck^2$.

- Q is said to be *definite* if $Q \geq 0$ or $Q \leq 0$ always, with $Q = 0$ if and only if $h = k = 0$, such as $\pm(h^2 + k^2)$.
- We say that Q is *indefinite* if it takes values of a different sign, e.g. $h^2 - k^2$.
- We say that Q is *semidefinite* if it is always ≥ 0 or ≤ 0, but it is zero for some nonzero (h, k), such as h^2.

By completing the square, we can put Q in one of the forms above, and also show that

$$Q \text{ is } \begin{cases} \text{definite} \\ \text{indefinite} \\ \text{semidefinite} \end{cases} \implies \begin{cases} AC - B^2 & > 0, \\ AC - B^2 & < 0, \\ AC - B^2 & = 0. \end{cases}$$

§26 Day 26: Extrema in Two Variables (Nov. 11, 2024)

Let $f(x, y)$ be $\mathcal{C}^2$ in two variables in a neighborhood of a critical point (a, b). We may write,

$$f(a+h, b+k) - f(a, b) = \frac{1}{2}(f_{xx}(a^*, b^*)h^2 + 2f_{xy}(a^*, b^*)hk + f_{yy}(a^*, b^*)k^2),$$

where $(a^*, b^*) = (a + \theta h, b + \theta k)$ for some $0 \leq \theta \leq 1$. Consider the quadratic form $Q(h, k) = Ah^2 + 2Bhk + Ck^2$. We have two cases;

- If $A = C = 0$, then $Q(h, k) = 2Bhk = B(u^2 - v^2)$, and after the change of variables

  $$h = \frac{1}{\sqrt{2}}(u + v); \qquad k = \frac{1}{\sqrt{2}}(u - v),$$

  we get $AC - B^2 < 0$, so we see that Q is in indefinite form.

- If $A \neq 0$, let us complete the square as follows;

  $$Q = A\left((h + \frac{B}{A}k)^2 + \frac{AC - B^2}{A^2}k^2\right).$$

  Then we can have three cases. We have that Q is definite, i.e. it only takes one sign and vanishes only at zero, whenever $AC - B^2 < 0$. However, if $AC - B^2 = 0$, then while it still only takes one sign, it vanishes outside of just zero; specifically, on the line $h - \frac{B}{A}k = 0$. If we write

  $$Q = A(u^2 + v^2)$$

  where $u = h + \frac{B}{A}k$ and $v = \sqrt{\frac{AC-B^2}{A^2}}k$, then $Q = Au^2$ where $u = h + \frac{B}{A}k$ and $v = k$.

  We say that Q is indefinite when $AC - B^2 < 0$; we may write $Q = A(u^2 - v^2)$, where $u = h + \frac{B}{A}k$ and $v = \sqrt{\frac{B^2-AC}{A^2}}k$.

Recall the second derivative test from 157; we now extend it to two variables. For $f(x, y)$ that is $\mathcal{C}^2$ near the critical point (a, b), if, at (a, b), we have

$$f_{xx}f_{yy} - f_{xy}^2 > 0, \tag{i}$$
$$f_{xx}f_{yy} - f_{xy}^2 < 0, \tag{ii}$$
$$f_{xx}f_{yy} - f_{xy}^2 = 0. \tag{iii}$$

Then we have the respective three cases;

(i) There is a local maximum or minimum; if $f_{xx}(a, b) < 0$, it is a maximum; if $f_{xx}(a, b) > 0$, it is a minimum.

(ii) There is no local max. or min.; this is called a "saddle point".

(iii) Indeterminate; for example, $x^2 \pm y^3$, or $x^2 \pm y^n$.

We may prove the above by completing the square; leaving it as an exercise, though. We give another example for now; suppose $f(x, y) = (x - y)^n + (y - 1)^n \geq 0$. Let n be a positive even integer; if $f(x, y) = 0$, then $y = 1, x = y = 1$. We see that $(1, 1)$ is a critical point, and that it attains a local minimum, even if the second derivative test is inconclusive.

§27 Day 27: Manifolds (Nov. 13, 2024)

A subset $M \subset \mathbb{R}^n$ is a $\mathcal{C}^r$ submanifold of $\mathbb{R}^n$ of dimension k if it satisfies any of the equivalent conditions in the following theorem.

Theorem 27.1. The following conditions are equivalent,

(a) For all $a \in M$, there is a $\mathcal{C}^r$ map $f : U \to \mathbb{R}^{n-k}$ where U is an open neighborhood such that $M \cap U = f^{-1}(0)$ and Df has rank $n-k$ at every point of $M \cap U$.

(b) Every point of M has an open neighborhood $U = V \times W \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ such that $M \cap U$ is a graph of a $\mathcal{C}^r$ function. If we write $z = g(y)$ where $y = (y_1, \ldots, y_k) \in V$, $z = (z_1, \ldots, z_{n-k}) \in W$, after a permutation of coordinates, we have

$$(y, z) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

The forwards direction (i.e., (a) implies (b)) is given by the implicit function theorem; for the opposite direction, we see that by taking $f(y, z) = z - g(y)$.

We now give some examples of manifolds.

(a) A point in $\mathbb{R}^n$.

(b) An open subset of $\mathbb{R}^n$.

(c) A smooth curve is a 1-dimensional manifold, a smooth surface is a 2-dimensional manifold.

(d) $S^{n-1} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 = 1\}$, i.e. the $n-1$ sphere. If we let $f(x) = x_1^2 + \cdots + x_n^2 - 1$, then $S^{n-1} = f^{-1}(0)$, and $\nabla f(x) = 2(x_1, \ldots, x_n) \neq 0$ at every point of S^{n-1}. Thus, f has rank 1 at every point of S^{n-1}.

Definition 27.2. A *diffeomorphism* is a $\mathcal{C}^r$ map with a $\mathcal{C}^r$ inverse.

We continue the above theorem.

Theorem 27.3. We have a third equivalent condition to the two above;

(c) For all $a \in M$, there is an open neighborhood U of a, an open subset V of $\mathbb{R}^n$, and a $\mathcal{C}^r$ diffeomorphism $h : U \to V$ such that $h(M \cap U) = V \cap (\mathbb{R}^k \times \{0\})$, which is equal to the set $\{(y_1, \ldots, y_n) \in V \mid y_{k+1} = \cdots = y_n = 0\}$.

To see that (b) implies (c), let $h(y, z) = (y, z - g(y))$, and to see (c) implies (b), let $f(x) = (h_{k+1}(x), \ldots, h_n(x))$.

There is another very important characterization of manifolds;

Theorem 27.4 (Coordinate Charts; Spivak Theorem 5-2)**.** We have a fourth equivalent condition to the three above;

(d) For $a \in M$, there is an open neighborhood U of a, open subset W of $\mathbb{R}^n$, and $\mathcal{C}^r$ map $\varphi : W \to \mathbb{R}^n$ such that

  (i) φ is one-to-one,

  (ii) $\varphi(W) = M \cap U$,

  (iii) φ' has rank k at every point of W,

  (iv) for every open open subset Ω of W, $\varphi(\Omega) = \varphi(W) \cap \omega'$, for some open subset Ω' of $\mathbb{R}^n$.

  Note the topology definition for (iv); we say $\varphi(\Omega)$ is open in $\varphi(W)$ in the "subspace topology", i.e. so $\varphi^{-1} : \varphi(W) \to W$ is continuous.

§28 Day 28: Manifolds, Pt. 2 (Nov. 15, 2024)

Recall that a manifold $M \subset \mathbb{R}^n$ of dimension k satisfies:

(c) For all $a \in M$, there is an open neighborhood U of a, an open subset V of $\mathbb{R}^n$, and a $\mathcal{C}^r$ diffeomorphism $h : U \to V$ such that $h(M \cap U) = V \cap (\mathbb{R}^k \times \{0\})$, which is equal to the set $\{(y_1, \dots, y_k) \in V \mid y_{k+1} = \cdots = y_n = 0\}$.

(d) For $a \in M$, there is an open neighborhood U of a, open subset W of $\mathbb{R}^n$, and $\mathcal{C}^r$ map $\varphi : W \to \mathbb{R}^n$ such that

  (i) φ is one-to-one,

  (ii) $\varphi(W) = M \cap U$,

  (iii) φ' has rank k on W,

  (iv) for every open $\Omega \subset W$, $\varphi(\Omega) = \varphi(W) \cap \tilde{U}$, where $\tilde{U}$ is some open subset of $\mathbb{R}^n$.

To see (c) implies (d), take $\varphi = h^{-1}\big|_{V \cap (\mathbb{R}^n \times \{0\})}$. For (d) implies (c), consider $a \in M$, $b \in W$ such that $\varphi(b) = a$; we can assume that $\partial_{y_1,\dots,y_k}(\varphi_1, \dots, \varphi_k)$ has rank k on W. Define $\psi : W \times \mathbb{R}^{n-k} \to \mathbb{R}^n$ to be the map

$$(y, z) \mapsto (\varphi_1(y), \dots, \varphi_k(y), \varphi_{k+1}(y) + z_1, \dots, \varphi_n(y) + z_{n-k}),$$

and consider that

$$\psi(y, z) = \left(\begin{array}{c|c} \dfrac{\partial(\varphi_1, \dots, \varphi_k)}{\partial(y_1, \dots, y_k)} & 0 \\ \hline * & I_{n-k} \end{array}\right)$$

By the inverse function theorem, there exists an open neighborhood V' of $(k, 0)$, U' of a such that $\psi : U' \to V'$ is $\mathcal{C}^r$.

Now, consider the set $\{\varphi(y) \mid (y, 0) \in V'\} = \varphi(W) \cap U''$, where U'' is open in $\mathbb{R}^n$. Take $U_1 = U_1' \cap U''$, $V_1 = \psi^{-1}(U_1)$, and $h = \psi^{-1}\big|_{U_1}$. Then

$$\begin{aligned} M \cap U_1 &= \{\varphi(y) \mid (y, 0) \in V_1\} \\ &= \{\psi(y, 0) \mid (y, 0) \in V_1\}, \end{aligned}$$

and

$$\begin{aligned} h(M \cap U_1) &= \psi^{-1}(M \cap U_1) \\ &= \psi^{-1}\{\psi(y, 0) \mid (y, 0) \in V_1\} \\ &= \{(y, 0) \mid (y, 0) \in V_1\} \\ &= V_1 \cap (\mathbb{R}^k \times \{0\}). \end{aligned}$$

A remark on (d) from the proof above; let $a = \psi(b) \in M$, and $b \in W$. ψ^{-1} is given near a by the...?? *if someone knows what this is pls lmk lmao*

§29 Day 29: Smooth mappings between manifolds (Nov. 18, 2024)

Let $V \subset \mathbb{R}^m$ be open, and let us consider the function $f : V \to M \subset \mathbb{R}^n$. Today we will talk about functions to a $\mathcal{C}^r$ submanifold M of dimension k.

Claim 29.1. f is smooth as a mapping to $\mathbb{R}^n$ if and only if $f \circ \varphi^{-1}$ is $\mathcal{C}^r$ for all coordinate charts φ; i.e., $f \circ \varphi^{-1}$ is defined on $f^{-1}(\varphi(W))$.

($\Leftarrow$) $\varphi^{-1} \circ f$ is $\mathcal{C}^r$ implies that $f = \varphi(\varphi^{-1} \circ f)$ by the chain rule.

($\Rightarrow$) f is $\mathcal{C}^r$ implies that $\varphi^{-1} \circ f$ is $\mathcal{C}^r$. For all $a \in \varphi(W)$, there exists a $\mathcal{C}^r$ diffeomorphism $h : U \to V$ where U is open in $\mathbb{R}^n$, and V is open in $\mathbb{R}^k \times \mathbb{R}^{n-k}$ such that $h(M \cap U) = V \cap (\mathbb{R}^n \times \{0\})$, i.e. $\varphi = h^{-1}|_{V \cap (\mathbb{R}^n \times \{0\})}$. ☐

If $\varphi_1 : W_1 \to \mathbb{R}^n$ and $\varphi_2 : W_2 \to \mathbb{R}^n$ are two different coordinate charts for M, then $\varphi_2^{-1} \circ \varphi_1$ is $\mathcal{C}^r$ as the map $\varphi_1^{-1}(\varphi_2(W_2)) \to \mathbb{R}^k$ with $\mathcal{C}^r$ inverse on $\varphi_2^{-1}(\varphi_1(W_1))$.

Definition 29.2 (Functions defined on a $\mathcal{C}^r$ submanifold M of $\mathbb{R}^n$)**.** Let $f : M \to \mathbb{R}$ be $\mathcal{C}^r$ at a point if there exists a coordinate chart $\varphi : W \to \mathbb{R}$ for M at a (i.e., $a \in \varphi(W)$) such that $f \circ \varphi$ is $\mathcal{C}^r$ at b, where $a = \varphi(b)$.

This definition is independent of the choice of coordinate chart by the previous remark.

Lemma 29.3. Let $f : M \to \mathbb{R}$ be $\mathcal{C}^r$ at a point a if there is an open neighborhood U of a in $\mathbb{R}^n$ and $\mathcal{C}^r$ function $g : U \to \mathbb{R}$ such that $f|_{M \cap U} = g|_{M \cap U}$.

($\Leftarrow$) Because $f \circ \varphi = g \circ \varphi$ is $\mathcal{C}^r$.

($\Rightarrow$) Take $g = f \circ h^{-1} \circ \mathrm{pr} \circ h$, where pr is a projection from $\mathbb{R}^k \times \mathbb{R}^{n-k}$. h^{-1} on $V \cap (\mathbb{R}^k \times \{0\})(\varphi)$.

We start with some definitions.

Definition 29.4. $f : M \to \mathbb{R}^n$ is $\mathcal{C}^r$ at a if every component f_i of f is $\mathcal{C}^r$, i.e. $f = (f_1, \ldots, f_n)$.

Definition 29.5. $\mathcal{C}^r$ mappings between manifolds $M \subset \mathbb{R}^n$ of dimension k and $N \subset \mathbb{R}^p$ of dimension ℓ. $f : M \to N$ is $\mathcal{C}^r$ if for all coordinate charts $\varphi : W \to \mathbb{R}^n$ for M, $\psi : Z \to \mathbb{R}^p$ for N, we have that $\psi^{-1} \circ f \circ \varphi$ is $\mathcal{C}^r$ and $\varphi^{-1} \circ f \circ \psi$ is too.

§30 Day 30: Tangent Space (Nov. 20, 2024)

A tangent space to $\mathbb{R}^n$ at a point a, $\mathbb{R}^n_a$, is a copy of $\mathbb{R}^n$ "centered at a". We write vectors $v \in \mathbb{R}^n_a$ as v_a or (a, v). $v_a = (v_1, \ldots, v_n) = \sum v_i e_{i,a}$. We can identify $v_a \in \mathbb{R}^n_a$ with directional derivative at a, along v_i tangent vectors operate on $\mathcal{C}^1$ functions.

$$V_a(f) = D_u f(a) = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(a).$$

Subsequently equal to

$$\left(\sum_{i=1}^n v_i \left. \frac{\partial f}{\partial x_i} \right|_a \right)(f)$$

so we identify

$$v_a = \sum v_i e_{i,a} \text{ with } \sum v_i \left. \frac{\partial}{\partial x_j} \right|_a,$$

where $\left. \frac{\partial}{\partial x_j} \right|_a$ is the standard basis of $\mathbb{R}^n_a$, with $dx_i(a)$ dual basis, and

$$\mathbb{R}^a_n = \{\text{tangent vectors } \gamma'(0) \text{ to } \mathcal{C}^1 \text{ curves } \gamma : (-\delta, \delta) \to \mathbb{R}^n, \gamma(0) = a\}.$$

Note that the image of the paths γ need to sit in M as well. In particular, $v_a \in \mathbb{R}^n_a$ is $\gamma'(0)$ where $\gamma(t) = a + tv$. Let M be a $\mathcal{C}^r$ submanifold of $\mathbb{R}^n$ of dimension k.

Definition 30.1. The tangent space M_a or TM_a is given by the above, or $\varphi^{-1} \circ \gamma$ being $\mathcal{C}^1$ for any coordinate chart at a.

Let $m : (-\delta, \delta) \to W$, $m(0) = \alpha$ be a $\mathcal{C}^1$ curve in W, and $\eta(0) = b$ if and only if $\gamma = \varphi \circ \eta$ $\mathcal{C}^1$ curve in M, $\gamma(0) = a$. Then

$$\gamma'(0) = D\varphi(b) \cdot \eta'(0)$$

so M_a is the k-dimensional linear subspace of $\mathbb{R}^n_a$ given by $D\varphi(b)(\mathbb{R}^k_b)(a)$ rank k. In terms of other definitions of manifolds,

(i) $h : (U \ni a) \to V \subset \mathbb{R}^n$ diffeomorphism such that $h(M \cap U) = V \cap (\mathbb{R}^k \times \{0\})$. $TM_a = Dh(a)^{-1}(\mathbb{R}^k \times \{0\})$ subspace of $\mathbb{R}^n_{h(a)}$.

(ii) $M \cap U$ is the graph of $z = g(y)$ of a $\mathcal{C}^r$ map $g : V \to W$, with $a = (b, a)$ from earlier[14] then

$$TM_a = \{(y, z) \in \mathbb{R}^n_{(b,a)} \mid z = Dg(b)y\},$$

$\varphi : y \mapsto (y, g(y))$ is a coordinate chart, and $D\varphi(b) = \binom{I}{Dg(b)}$.

(iii) $M \cap U = f^{-1}(0)$, $f : U \to \mathbb{R}^{n-k}$. $Df(a) = \mathbb{R}^n_a \to \mathbb{R}^{n-k}_{f(a)}$. $TM_a = Df(a)^{-1}(\{0\}) = \ker Df(a)$. To see this, let γ be a $\mathcal{C}^1$ curve $\gamma : (-\delta, \delta) \to M$, $\gamma(0) = a$. $f \circ \gamma = 0$, $Df(a)\gamma'(0) = 0$, i.e. $\gamma'(0) \in \ker Df(a)$, so $TM_a \subset \ker Df(a)$ since both are of dimension k.

Let $f : (M \subset \mathbb{R}^n) \to (N \subset \mathbb{R}^p)$ be a $\mathcal{C}^1$ map of manifolds of dimension k and ℓ, respectively. Then this induces a linear transformation

$$f_{*a} : TM_a \to TN_{f(a)}.$$

For example, if $M = \mathbb{R}^n$ and $N = \mathbb{R}^p$, then f_{*a} is given by $Df(a)$. We define f_{*a} as

$$D\psi(f(a))D(\psi^{-1} \circ f \circ \varphi)(b)D\varphi(b)^{-1},$$

where $a = \varphi(b)$ and $\psi(c) = f(a)$. In a way, $f = \psi \circ (\psi^{-1} \circ f \circ \varphi) \circ \varphi^{-1}$.

[14]???

§31 Day 31: Integration (Nov. 22, 2024)

The integral over a rectangle is given as follows,

- The closed rectangle $A = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$.
- The volume of A is given by $(b_1 - a_1) \dots (b_n - a_n)$.
- A partition of A, $P = \{P_1, \dots, P_n\}$ where P_i is a partition of $[a_i, b_i]$. If each P_i has k subintervals, then A has $k_1, \dots, k_n$ subrectangles. Let $f : A \to \mathbb{R}$ be bounded. Then for all subrectangles S of A,

$$m_S(f) = \inf\{f(x) \mid x \in S\},$$
$$M_S(f) = \sup\{f(x) \mid x \in S\}.$$

- The lower and upper bounds are given by

$$L(f, P) = \sum_S m_S(f)v(S),$$
$$U(f, P) = \sum_S M_S(f)v(S),$$

  where $L(f, P) \leq U(f, P)$. If P' is a refinement of P, i.e., every P'_i is a refinement of P_i, then $L(f, P) \leq L(f, P')$ and $U(f, P) \geq U(f, P')$. This is true since every subrectangle S of P is divided into subrectangles $S_1, \dots, S_m$ of P'. In particular, $m_{S_i}(f) \geq m_S(f)$, and

$$m_S(f)V(s) = m_S(f) \sum_j v(S_j) \leq \sum_j m_{S_j} v(s_j),$$

  so $L(f, P) \leq L(f, P')$ as desired.

- For any two partitions P, P' of A, $L(f, P) \leq U(f, P')$. We can see this by comparing them to a common refinement; then $\sup_P L_(f, P) \leq \inf_P U(f, P)$.

Definition 31.1. A bounded function $f : A \to \mathbb{R}$ is called *integrable* if $\sup_P L(f, P) = \inf_P U(f, P)$. In this case, we write $\int_A f$ for this common valve, e.g. $n = 1$,

$$\int_{[a,b]} f = \int_a^b f.$$

Lemma 31.2. A bounded function $f : A \to \mathbb{R}$ is integrable on the closed rectangle A if $\forall \varepsilon > 0$, there exists a partition P of A such that $U(f, P) - L(f, P) < \varepsilon$.

($\Leftarrow$) Yes.

($\Rightarrow$) Integrable means for all $\varepsilon > 0$, there are partitions P, P' of A such that $U(f, P') - L(f, P) < \varepsilon$, so a common refinement P'' satisfies $U(f, P'') - L(f, P'') < \varepsilon$.

We now give some examples.

(a) If $f(x) = c$ is constant, then $\int_A f = cv(A)$ because for all subrectangles S of partition P of A, $m_S(f) = c = M_S(f)$.

(b) Let $A = [0, 1] \times [0, 1] \subset \mathbb{R}^2$. Then let $f(x, y) = 1_{\mathbb{Q}}(x)$; then f is not integrable on A, since $L(f, P) = \sum_S m_S(f)v(S) = 0$, while $U(f, P) = \sum_S M_S(f)v(S) = 1$.

Some basic properties of integration; suppose $f : A \to \mathbb{R}$ is integrable. Then

(i) $f + g$, $c \cdot f$ integrable, and

$$\int_A (f+g) = \int_A f + \int_A g,$$
$$\int_A cf = c \int_A f.$$

(ii) If $f \le g$, then $\int_A f \le \int_A g$.

(iii) If f is integrable, then $|f|$ integrable and in fact $\left|\int_A f\right| \le \int_A |f|$.

(iv) If f, g are integrable, then $f \circ g$ is integrable.

We prove the third property; consider the three cases.

(a) $f \ge 0$ on a subrectangle S. Then

$$M_S(|f|) = M_S(f),$$
$$m_S(|f|) = m_S(f),$$

so $M_S(|f|) - m_S(|f|) = M_s(f) - m_s(f)$.

(b) Let $f \le 0$ on S.

$$M_S(|f|) = -M_S(f),$$
$$m_S(|f|) = -m_S(f);$$

again, we obtain the same result.

(c) Let $m_S(f) \le 0 \le M_S(f)$.

$$-m_S(f) \le M_S(f),$$
$$M_S(|f|) = M_S(f),$$
$$M_S(|f|)m_S(|f|) \le M_S(f) \le M_S(f) - m_s(f).$$

The other case is... left as an exercise?

In all cases, we have that

$$M_S(|f|) - m_S(|f|) \le M_S(f) - m_S(f).$$

So $|f|$ is integrable, and

$$-\int_A |f| \le \int_A f \le \int_A |f|.$$

§32 Day 32: Integrals over General Bounded Sets (Nov. 25, 2024)

We start with some definitions.

Definition 32.1. We say that $A \subset \mathbb{R}^n$ has n-dimensional *measure zero* if, for all $\varepsilon > 0$, there is a covering $\{U_1, \dots, \}$ by countably many closed rectangles with $\sum_{i=1}^{\infty} v(U_i) < \varepsilon$. Note that we need not use closed rectangles necessarily; open rectangles, open or closed balls would work too.

Here are some examples.

(a) A finite subset of $\mathbb{R}^n$ has measure zero.

(b) A countable subset $\{a_1, a_2, \dots\}$ of $\mathbb{R}^n$ has measure zero. Cover each a_i by a closed rectangle U_i, with $v(U_i) < \frac{\varepsilon}{2^i}$. Then $\sum v(U_i) < \varepsilon$.

Lemma 32.2. Any countable union $A = A_1 \cup A_2 \cup \dots$ of subsets A_i of $\mathbb{R}^n$ with measure zero has measure zero.[15]

Cover each A_i by countably many closed rectangles $\{U_{i1}, U_{i2}, \dots, \}$ with

$$\sum_{j=1}^{n} v(U_{ij}) < \frac{\varepsilon}{2^i}.$$

Then $\{U_{ij}\}$ covers A, and is countable by a diagonal argument. Thus, $\{U_{ij}\}$ can be rearranged as $\{V_k\}$, and we have

$$\sum_{k=1}^{\infty} v(V_k) < \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon. \quad \square$$

Theorem 32.3. Given a closed rectangle $A \subset \mathbb{R}^n$ and a bounded function $f : A \to \mathbb{R}$, f is integrable on A if and only if $\{x \in A \mid f \text{discont. at } a\}$ has measure zero.

This will be proved in the next lecture.

Definition 32.4 (Characteristic Function)**.** For C in $\mathbb{R}^n$, let us define the *characteristic function* of C as follows,

$$\chi_C(x) = \begin{cases} 1 & x \in C, \\ 0 & x \notin C. \end{cases}$$

If $f : C \to \mathbb{R}$ is bounded, then $f \circ \chi_C$ makes sense as a function on $\mathbb{R}^n$. Note that the characteristic function and indicator function are of the same notion, just used in different fields..?

As an example, if $C \subset A$ rectangle, we can define

$$\int_C f = \int_A f \circ \chi_C$$

if the latter exists.

[15] wordsalad definition

Lemma 32.5. $\{x \in \mathbb{R}^n \mid \chi_C \text{ discont. at } x\} = \text{bdry}(C)$.

Consider the following cases.

(i) Let x be in the interior of C. Then $x \in U \subset \text{int}(C)$, where U is an open rectangle. And so $\chi_C = 1$ on U, and so it contains all such x.

(ii) Let x be in the exterior of C. Then $x \in U \subset \text{int}(C)$, where $\chi_C = 0$ on U, and so it contains all such x.

(iii) Let x be on the boundary of C. For any open rectangle $U \ni x$, there exists $x_1 \subset C$, where $x_1 \in U : \chi_C(x_1) = 1$, and $x_2 \subset C$, where $x_2 \notin U : \chi_C(x_2) = 0$. Thus, χ_C is not continuous at x.

Corollary 32.6. If $C \subset A$, where A is a closed rectangle and $f : A \to \mathbb{R}$ is continuous, then f is integrable on C if and only if the boundary of C has measure zero.

(c) $[a, b] \subset \mathbb{R}$ with $a < b$ does not have measure zero. Any cover of $[a, b]$ by open intervals has finite subcover $\{U_1, \dots, U_n\}$ since the interval is compact. Then

$$\sum_{i=1}^{n} v(U_i) \geq b - a.$$

To see this, say $U_i = (a_i, b_i)$. Then the endpoints of all (a_i, b_i) form a partition $\{t_0, \dots, t_k\}$ of a closed interval containing $[a, b]$. Every $[t_{j-1}, t_j] \subset [a_i, b_i]$ for some i; then

$$\sum_{i=1}^{n} v((a_i, b_i)) \geq \sum_{j=1}^{n} v([t_{j-1}, t_j]) \geq b - a.$$

(d) If C is a bounded set of measure zero, does the boundary have measure zero? Not necessarily; let $C = \mathbb{Q} \cap [0, 1]$. C is a countable set, is of measure zero, but $\text{bdry}(C) = [0, 1]$ which is clearly not measure zero as per the previous example.

§33 Day 33: Integrals, Pt. 3 (Nov. 27, 2024)

Theorem 33.1. Given a bounded function $f : A \to \mathbb{R}$ on a closed rectangle $A \subset \mathbb{R}^n$, f is integrable on A if and only if the set of discontinuities $B = \{x \in A \mid f \text{ discont.}\}$ has measure zero.

($\Leftarrow$) Let $\varepsilon > 0$. We can cover B by interiors of countable many closed rectangles $U_1, U_2, \dots$ in $\mathbb{R}^n$ such that $\sum_{i=1}^{\infty} v(U_i) < \varepsilon$ (by definition of measure zero). If $x \in A \setminus B$, then there is a closed rectangle $V_x \subset \mathbb{R}^n$ such that $M_{V_x \cap A}(f) - m_{V_x \cap A}(f) < \varepsilon$, $x \in \operatorname{int}(V_x)$ by continuity. Recall that continuity means $|f(y) - f(x)| < \frac{\varepsilon}{2}$ when $|y - x| < \delta$, i.e. $f(x) - \frac{\varepsilon}{2} < f(y) < f(x) + \frac{\varepsilon}{2}$.

Since A is compact, there exists a finite subcollection of $\{U_i, V_x\}$ covering A. Choose a partition P of A such that every subrectangle S of P lies in some V_x or U_i. If $|f(x)| \leq M$, then

$$\begin{aligned}
U(f,P) - L(f,P) &= \sum_S (M_S(f) - m_S(f))v(S) \\
&\leq \sum_{S \subset V_x} (\dots) + \sum_{S \subset U_i} (\dots) && \text{(for some } i,\ x) \\
&< \sum_{s \subset V_x} \varepsilon v(S) + \sum_{S \subset U_i} 2Mv(S) \\
& && \text{(where } 2M \textstyle\sum_{S \subset S_i} v(S) \leq 2M \sum v(U_i)) \\
&\leq \varepsilon v(A).
\end{aligned}$$

($\Rightarrow$) We will check this later on.

We start by giving some examples. The derivation of a bounded function from continuity is given as follows (???); given $X \subset \mathbb{R}^n$, $f : X \to \mathbb{R}$ bounded, $\delta > 0$ and $a \in X$,

$$\begin{aligned}
M(f, a, \delta) &= \sup\{f(x) \mid x \in X, |x - a| < \delta\}, \\
m(f, a, \delta) &= \inf\{f(x) \mid x \in X, |x - a| < \delta\}.
\end{aligned}$$

Then we say that the *oscillation* of f at a is

$$o(f, a) = \lim_{\delta \to 0} M(f, a, \delta) - m(f, a, \delta).$$

This necesssarily exists, since the difference is monotonically decreasing in δ.

Lemma 33.2. f is continuous at a if and only if $o(f, a) = 0$.

($\Leftarrow$) For all $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ implies $|x - a| < \delta$ where $x \in X$. Thus, $M(f, a, S) - m(f, a, S) \leq 2\varepsilon$, i.e. $f(a) - \varepsilon < f(x) < f(a) + \varepsilon$. This applies for any $\varepsilon > 0$, so $o(f, a) = 0$.

($\Rightarrow$) Essentially the same argument; left as an exercise.

Proposition 33.3. Let $X \subset \mathbb{R}^n$ be closed, and $f : X \to \mathbb{R}^n$ be bounded on X. For any $\varepsilon > 0$, $\{x \in X \mid o(f, x) \geq \varepsilon\}$ is closed.

Let $Y = \{x \in X \mid o(f, x) \geq \varepsilon\}$. To show that $\mathbb{R}^n \setminus Y$ is open, let x be an element. If $x \in X$, then there is an open ball B such that $x \in B$ and $B \subset \mathbb{R}^n \setminus X \subset \mathbb{R}^n \setminus Y$. On the other hand, if $x \in X$, then there exists δ such that $M(f, x, \delta) - m(f, x, \delta) < \varepsilon$. Let B be an open ball centered at x with radius δ. If $y \in B$, then there is $\delta_1 > 0$ such that if $|z - y| < \delta_x$, then $|z - x| < \delta$. So $M(f, y, \delta_1) - m(f, y, \delta_1) < \varepsilon$. Therefore, $o(f, y) < \varepsilon$, and so $B \subset \mathbb{R}^n \setminus Y$ as desired. $\square$

§34 Day 34: Integrals, Pt. 4 (Nov. 29, 2024)

Theorem 34.1. Given a bounded function $f : A \to \mathbb{R}$ on a closed rectangle, we have that f is integrable on A if and only if $\{x \in A \mid f \text{ discont. }\}$ has measure zero.

($\Leftarrow$) Done last lecture.

($\Rightarrow$) For any $\varepsilon > 0$, let $B_\varepsilon = \{x \in A \mid o(f, x) > \varepsilon\}$. B_ε is closed, and is also compact. If we let $B = \{x \in A \mid o(f, x) > 0\}$, it is enough to write $B = B_1 \cup B_{1/2} \cup B_{1/3} \cup \ldots$, and it suffices to write $B_{1/k}$ has measure zero for any k.

Given some $\varepsilon > 0$, there is a partition P of A such that $U(f, P) - L(f, P) < \varepsilon$. Let $\mathcal{S}$ be the set of subrectangles of P such that $S \cap B_{1/k} \neq \emptyset$. Let $\mathcal{S}'$ be the set of $S \in \mathcal{S}$ such that $\inf S \cap B_{1/k} \neq \emptyset$. Then $M_S(f) - m_S(f) \geq \frac{1}{k}$. Cover the boundaries of all $S \in P$ by finitely many closed rectangles of total volume ε. The latter rectangles together with $\mathcal{S}'$ cover $B_{1/k}$. In fact, finitely many cover $B_{1/k}$. Directly compute the volume as follows,

$$\sum_{s \in P} (M_S(f) - m_S(f))\, v(S) < \varepsilon \implies \sum_{S \in \mathcal{S}'} \underbrace{(M_S(f) - m_S(f))}_{\geq \frac{1}{k}} v(S) < \varepsilon,$$

meaning we have $\sum_{s \in \mathcal{S}'} v(S) < k\varepsilon$. Thus, the total volume is less than $\sum_{s \in \mathcal{S}'} v(S) + \varepsilon < (k+1)\varepsilon$. $\square$

Definition 34.2. A bounded subset C of $\mathbb{R}^n$ is *Jordan measurable* if the boundary of C has measure 0.

In this case, we define the n-dimensional volume of C as $\int_C 1$, which is equal to $\int_A \chi_C$ for any rectangle $A \supset C$. For example, 1-dimensional volume is referred to as *length*, 2-dimensional volume as *area*, and so forth. We now give some examples.

(a) A bounded set C of measure zero does not necessarily have to be Jordan measurable. For example, take $C = \mathbb{Q} \cap [0, 1]$.

(b) A bounded open set is not necessarily Jordan measurable. For example, let C be the union of intervals $(a_i, b_i) \subset (0, 1)$ such that each rational in $(0, 1)$ is in some (a_i, b_i), and $\sum(b_i - a_i) < 1$. Then the boundary of this set is given by $[0, 1] \setminus C$. This is clearly not measure zero, since if it were, we could cover it with a countable set of open intervals with volume as small as desired; however, $[0, 1]$ is covered by finitely many of these and the (a_i, b_i); however, the total length is less than $\sum(b_i - a_i) + 1 - \sum_(b_i - a_i) < 1$.

(c) A bounded closed set is not necessarily Jordan measurable. The complement in $[0, 1]$ of our example in (b) suffices.

§35 Day 35: Fubini's Theorem (Dec. 2, 2024)

We start with *Cavalieri's principle*: it is reasonable to expect, for $A = [a,b] \times [c,d]$, that

$$\int_A f = \int_a^b h(x)\,dx = \int_a^b \left(\int_c^d f(x,y)\,dy \right)\,dx.$$

However, even if f is integrable on A, i.e., the set of discontinuities of f lies in $\{x_0\} \times [c,d]$, $\int_c^d f(x_0,y)\,dy$ may not exist. For this region, formulate the integral in a more general way as follows; suppose $z = f(x,y) \geq 0$ is continuous. Then the area of the cross section

$$h(x) = \int_c^d g(x) = \int_c^d f(x,y)\,dy,$$

i.e. $g_x(y) = f(x,y)$ (as in, we're fixing x and considering said cross section). If $f : A \to \mathbb{R}$ is a bounded function on a closed rectangle $A \subset \mathbb{R}^n$, then let us consider the lower and upper integrals of f on A:

$$\begin{aligned} L_A f &= \sup \text{ of all lower sums,} \\ U_A f &= \inf \text{ of all upper sums,} \end{aligned}$$

which always exist.

Theorem 35.1 (Fubini's Theorem)**.** Let $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^p$ be closed rectangles, and let $f : A \times B \to \mathbb{R}$ be integrable.

Given $x \in A$, let $g_x : B \to \mathbb{R}$ be given by $g_x(y) = f(x,y)$. Set

$$\begin{aligned} \mathcal{L}(x) &= L_B g_x = L \int_B f(x,y)\,dy, \\ \mathcal{U}(x) &= U_B g_x U \int_B f(x,y)\,dy. \end{aligned}$$

Then SL,$\mathcal{U}$ are integrable on A, and

$$\begin{aligned} \int_{A\times B} f = \int_A \mathcal{L} &= \int_A \left(L \int_B f(x,y)\,dy \right)\,dx \\ = \int_B U &= \int_A \left(U \int_B f(x,y)\,dy \right)\,dx. \end{aligned}$$

A few remarks;

(i) A similar proof will give

$$\begin{aligned} \int_{A\times B} f &= \int_B \left(L \int_A f(x,y)\,dx \right)\,dy \\ &= \int_B \left(U \int_A f(x,y)\,dx \right)\,dy. \end{aligned}$$

(ii) If g_x is integrable, then for all x, we have that

$$\int_{A\times B} f = \int_A \left(\int_B f(x,y)\,dy \right)\,dx.$$

(iii) If $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ and f is nice enough, then we can interate the formula

$$\int_A f = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \ldots dx_1.$$

We now prove Fubini's Theorem. Let P_A, P_B be partitions of A, B respectively, with subrectangles S_A, S_B. We have that P_A, P_B determine the partition P of $A \times B$. Every subrectangle S of P is of form $S_A \times S_B$. Moreover, every partition P of $A \times B$ has this form,

$$\begin{aligned} L(f, P) &= \sum_S m_S(f) \cdot v(S) \\ &= \sum_{S_A, S_B} m_{S_A \times S_B}(f) v(S_A \times S_B) \\ &= \sum_{S_A} \left(\sum_{S_B} m_{S_A \times S_B}(f) v(S_B) \right) v(S_A) \leq \mathcal{S}(x) \end{aligned}$$

for all $x \in S_A$. If $x \in A$, $m_{S_A \times S_B}(f) \leq m_{S_B}(g_\alpha)$. So

$$\sum_{S_B} m_{S_A \times S_B}(f) v(S_B) \leq \sum_{S_B} m_{S_B}(g_\alpha)(S_B) \leq L_B \int g_x = \mathcal{L}(x).$$

We continue the proof next lecture.

§36 Day 36: Fubini's Theorem, Pt. 2 (Dec. 3, 2024)

Theorem 36.1 (Fubini's Theorem)**.** Let $f : A \times B \to \mathbb{R}$ integrable, where A, B are rectangles in $\mathbb{R}^n, \mathbb{R}^p$ respectively. Let $x \in A$; then $g_x : B \to \mathbb{R}$ with $g_x(y) = f(x, y)$. Set $\mathcal{L}(x) = L\int_B g(x) = L\int_B f(x,y)\,dy$, with $\mathcal{U}(x)$ defined analogously. Then $\mathcal{L},\mathcal{U}$ are integrable on B, and

$$\int_{A\times B} f = \int_A \mathcal{L} = \int_A \left(L\int_B f(x,y)\,dy \right) dx = \int_A \mathcal{U}.$$

We continue the proof; consider partitions P_A, P_B of A, B with subrectangles S_A, S_B respectively. Let P_A, P_B define the partition P of $A \times B$ with subrectangles $S_A \times S_B$. Then

$$\begin{aligned} L(f,P) &= \sum_{S_A,S_B} m_{S_A\times S_B}(f) v(S_A \times S_B) \\ &= \sum_{S_A} \left(\sum_{S_B} \underbrace{m_{S_A\times S_B}(f)}_{\le m_{S_B} g(x) \forall x \in S_A} v(S_B) \right) v_(S_A), \end{aligned}$$

where we may note that $\sum_{S_B} m_{S_A\times S_B}(f)v(S_B) \le L\int_B g_x = \mathcal{S}(x)$ for all $x \in S_A$. Thus,

$$L(f,P) \le L(\mathcal{L}, P_A) \le U(L, P_A) \le U(\mathcal{U}, P_A) \le U(f, P),$$

and we're given $\sup_p L(f,P) = \inf_p U(f,P) = \int_{A\times B} f$. Thus, $\mathcal{L}$ is integrable on A, and $\int_{A\times B} f = \int_A \mathcal{L}$. Running the same argument for $\mathcal{U}$, we would obtain

$$L(f,P) \le L(\mathcal{L}, P_A) \le L(\mathcal{U}, P_A) \le U(\mathcal{U}, P_A) \le U(f,P). \quad \square$$

> **Remark 36.2.** When can we write the iterated integral with $L\int_B$? For example, when g_x is integrable on B except for finitely many x, beacuse $\int_A L$ is unchanged if you change the value of $\mathcal{SL}$ at finitely many L. Nevertheless, we sometimes need the theorem as written. For example, let $A \times B = [0,1] \times [0,1]$, with
>
> $$f(x,y) = \begin{cases} 1 & x, y \notin \mathbb{Q}, \\ 1 - \frac{1}{q} & x, y \in \mathbb{Q}, x = \frac{p}{q}, \gcd(p,q) = 1. \end{cases}$$
>
> Then f is integrable, and $\int_{A\times B} f = 1$. However, if x is irrational, then $\int_B f(x,y)\,dy = 1$. If x is rational, then $\int_B f(x,y)\,dy$ is undefined. Applying Fubini's theorem to ealuate $\int_C f$, where f is integrable, e.g. continuously bounded and C Jordan measurable, then $\int_C f = \int_A f\chi_C$ where $C \subset A$ rectangle.[a]
>
> [a] idk where the hell this went in the end

We now give one last example; let $C = [-1,1] \times [-1,1] \setminus \{(x,y) \mid x^2 + y^2 \le 1\}$, i.e. with the unit disc removed. Then the integral something something wasn't continued in class I think. mb

§37 Day 37: Computation of Integrals (Jan. 6, 2025)

Course administrative information! Our new time for office hours, starting next week, is **3:30** to **4:30pm** on Mondays.

Sometimes, there is an advantage to integrating with respect to one of the variables first; let's start with the following example, where we wish to show

$$\iint_{x^2+y^2\leq 1} x^2y^2\,dx\,dy. \tag{1}$$

Intuitively, we are integrating over the unit disc. Since the integral is symmetric, it doesn't matter if we integrate with respect to x or y first. Let us write as follows,

$$\begin{aligned}\iint_{x^2+y^2\leq 1} x^2y^2\,dx\,dy &= 4\int_0^1 \left(\int_0^{\sqrt{1-x^2}} x^2y^2\,dy\right)dx \\ &= 4\int_0^1 \left(x^2\frac{y^3}{3}\bigg|_0^{\sqrt{1-x^2}}\right)dx \\ &= \frac{4}{3}\int_0^1 x^2(1-x^2)^{\frac{3}{2}}\,dx.\end{aligned}$$

Having put the integral in this form, we observe that it is useful to perform the trig sub $x = \sin\theta$ and $dx = \cos\theta\,d\theta$, which gives us

$$\frac{4}{3}\int_0^1 x^2(1-x^2)^{\frac{3}{2}}\,dx = \frac{4}{3}\int_0^{\frac{\pi}{2}} \sin^2\theta\cos^4\theta\,d\theta.$$

We now evaluate a second integral as an example;

$$\int_0^2\int_{\frac{y}{2}}^1 ye^{-x^3}\,dx\,dy. \tag{2}$$

We evaluate the integral as follows,

$$\begin{aligned}\int_0^2\int_{\frac{y}{2}}^1 ye^{-x^3}\,dx\,dy &= \int_0^1\left(\int_0^{2x} ye^{-x^3}\,dy\right)dx \\ &= \int_0^1\left(\frac{y^2}{2}\bigg|_0^{2x}\cdot e^{-x^3}\right)dx \\ &= 2\int_0^1 x^2e^{-x^3}\,dx \\ &= -\frac{2}{3}\,e^{-x^3}\bigg|_0^1 \\ &= \frac{2}{3}(1-e^{-1}).\end{aligned}$$

For a third integral, we have

$$\int_2^4\int_{\frac{4}{x}}^{\frac{20-4x}{8-x}} (y-4)\,dy\,dx. \tag{3}$$

Start by rewriting the bounds on the integral as follows; if we let $y = \frac{4}{x}$, we have $x = \frac{4}{y}$, and we may write

$$y = \frac{20-4x}{8-x} = 4 - \frac{12}{8-x}; \qquad x = 8 - \frac{12}{4-y}.$$

With these bounds, we may integrate with respect to x first, in which the integrand $y-4$ is constant w.r.t. x, yielding

$$\begin{aligned}\int_2^4 \int_{\frac{4}{x}}^{\frac{20-4x}{3-x}} (y-4)\,dy\,dx &= \int_1^2 \left(\int_{\frac{4}{y}}^{8-\frac{12}{4-y}} (y-4)\,dx \right) dy \\ &= \int_1^2 (y-4)\left(8 - \frac{12}{4-y} - \frac{4}{y}\right) dy,\end{aligned}$$

which is easy to integrate. Now for a fourth integral!

$$\iiint_{\substack{x^2+y^2\le z^2 \\ x^2+y^2+z^2\le 1 \\ z\ge 0}} z\,dx\,dy\,dz.$$

Using symmetry, we may observe as follows,

$$\begin{aligned}\iiint_{\substack{x^2+y^2\le z^2 \\ x^2+y^2+z^2\le 1 \\ z\ge 0}} z\,dx\,dy\,dz &= \iint_{x^2+y^2\le\frac{1}{2}} \left(\int_{\sqrt{x^2+y^2}}^{\sqrt{1-(x^2+y^2)}} z\,dz \right) dx\,dy \\ &= \iint_{x^2+y^2\le\frac{1}{2}} \big(1-(x^2+y^2)-(x^2+y^2)\big)\,dx\,dy \\ &= \frac{1}{2}\cdot\frac{\pi}{2} - \iint_{x^2+y^2\le\frac{1}{2}} (x^2+y^2)\,dx\,dy.\end{aligned}$$

We now change to polar coordinates, with $x = r\cos\theta$ and $y = r\sin\theta$. Notice that

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}; \qquad \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| = r.$$

Then we may perform the following substitution,

$$\begin{aligned}\frac{\pi}{4} - \int_0^{2\pi} \left(\int_0^{\frac{1}{\sqrt{2}}} r^2 r\,dr \right) d\theta &= \frac{\pi}{4} - \int_0^{2\pi} d\theta\, \frac{r^4}{4}\bigg|_0^{\frac{1}{\sqrt{2}}} \\ &= \frac{\pi}{4} - 2\pi\cdot\frac{1}{16} = \frac{\pi}{8}.\end{aligned}$$

§38 Day 38: Partition of Unity (Jan. 8, 2025)

This lecture, we will build up to the idea of a *bump function*. Given $C \subset U \subset \mathbb{R}^n$, with C compact and U open, there is a $\mathcal{C}^\infty$ function $f : U \to \mathbb{R}$ such that $0 \le f \le 1$, $f(x) = 1$ for $x \in C$, and $f(x) = 0$ for x outside some compact subset of U. The construction is based on[16]

$$\begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

(i) Let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function such that $g(x) > 0$ on $(-1, 1)$ and $g(x) = 0$ outside $(-1, 1)$. Then $g(x)$ can be given by

$$g(x) = \begin{cases} e^{-\frac{1}{(x-1)^2}} e^{-\frac{1}{(x+1)^2}}, & x \in (-1, 1), \\ 0, & x \notin (-1, 1). \end{cases}$$

(ii) Given $a \in \mathbb{R}^n$, $\delta > 0$, there is a $\mathcal{C}^\infty$ function $g_{a,\delta} : \mathbb{R}^n \to \mathbb{R}$ such that $g_{a,\delta}(x) \ge 0$ on the ball $|x - a| < \delta$ and equal to 0 otherwise. In particular, we may write

$$g_{a,\delta}(x) = g\left(\frac{|x-a|^2}{\delta^2}\right).$$

(iii) Given $C \subset U \subset \mathbb{R}^n$, there is a $\mathcal{C}^\infty$ function $F : \mathbb{R}^n \to \mathbb{R}$ such that $F \ge 0$, with $F(x) > 0$ whenever $x \in C$ and $F(x) = 0$ if x is outside some compact subset of U. Take $F = \sum_k g_{a_k,\delta_k}$.

(iv) Given $\varepsilon > 0$, there is a $\mathcal{C}^\infty$ function $h_\varepsilon : \mathbb{R} \to \mathbb{R}$ such that $0 \le h_\varepsilon \le 1$, where $h_\varepsilon(x) = 0$ if $x \le 0$ and $h_\varepsilon(x) = 1$ if $x \ge \varepsilon$. Let g be a $\mathcal{C}^\infty$ function such that $g(x) \ge 0$ on $(0, \varepsilon)$, and 0 outside of $(0, \varepsilon)$ with the same construction75 from (ii). Take

$$h_\varepsilon(x) = \frac{\int_0^x g}{\int_0^\varepsilon g}.$$

We can now write down the definition of the bump function. Let $f(x) = h_\varepsilon \circ F$, where $\varepsilon = \min_{x \in C} F(x)$.

Theorem 38.1. Given $A \subset \mathbb{R}^n$, $\mathcal{O}$ an open cover of A, there is a countable collection Φ of $\mathcal{C}^\infty$ functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that

(i) $0 \le \varphi \le 1$,

(ii) For all $x \in A$, there is an open neighborhood V of x such that all but finitely many $\varphi \in \Phi$ vanish on V.

(iii) For all $x \in A$, $\sum_{\varphi \in \Phi} \varphi(x) = 1$; note that the sum is finite as per (ii).

(iv) For all $\varphi \in \Psi$, there is $U \in \mathcal{O}$ such that $\varphi(x) = 0$ outside of some compact subset of U.

A collection of functions Φ satisfying the first three properties is called a $\mathcal{C}^\infty$ *partition of unity* for A; if Φ satisfies (iv), it is said to be *subordinate to* $\mathcal{O}$ with *compact support*. This theorem will be proved next lecture. :3c meow

[16] lost here, idk what is going on lol i got to class late

§39 Day 39: Partition of Unity, Pt. 2 (Jan. 10, 2025)

A $\mathcal{C}^\infty$ partition of unity for $A \subset \mathbb{R}^n$ is subordinate to the open cover $\mathcal{O}$ of A if there is a countable collection Φ of $\mathcal{C}^\infty$ functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that

(i) $0 \le \varphi \le 1$;

(ii) For all $x \in A$, there is an open neighborhood V of x such that all but finitely many $\varphi \in \Psi$ vanish on V;

(iii) For all $x \in A$, $\sum_{\varphi \in \Psi} \varphi(x) = 1$.

(iv) For all φ, there is $U \in \mathcal{O}$, such that $\varphi(x) = 0$ outside some compact subset of U.

We start with a few examples to consider. [17]

(a) If A is compact, then finitely many $U \in \mathcal{O}$ are enough to cover A, say, $\mathcal{O}' = \{U_1, \dots, U_n\}$. Then it is enough to find a partition of unity that is subordinate to the collection $\mathcal{O}'$. Now, for all i, let us find compact $D_i \subset U_i$ for each i such that the interiors of D_i also cover A. To see that this is possible, for all $x \in A$, take a closed ball D_x in some U_i; by compactness, we may choose $D_{x_1}, \dots, D_{x_n}$ whose[18] interiors cover A. Now, construct D_i by

$$D_i = \bigcup_{D_{x_j} \subset U_i} D_{x_j}.$$

For all $i = 1, \dots, k$, let ψ_i be nonnegative $\mathcal{C}^\infty$ such that $\psi_i > 0$ on D_i, and equal to 0 outside some compact subset of U_i. Clearly, $\psi_1 + \dots + \psi_k > 0$ on some open set $U \supset A$. To start, we may let

$$\varphi_i(x) = \frac{\psi_i(x)}{(\psi_1 + \dots + \psi_k)(x)}$$

for all $x \in U$; let $f(x)$ be non-negative $\mathcal{C}^\infty$ on which $f = 1$ on A, and 0 outside a compact subset of U. Then $\Psi = \{|\varphi_i|\}$ is indeed a partition of unity as desired. $\square$

(b) Suppose $A = A_1 \cup A_2 \cup \dots$, where $A_i \subset \text{int} A_{i+1}$, and each A_i is compact. We revisit this example later.

(c) Suppose A is an open set, and $A = A_1 \cup A_2 \cup \dots$ with the same constraints as case (b). Let $A_i = \{x \mid |x| \le i\} \cap \{x \in A \mid d(x, \mathbb{R}^n \setminus A) \ge \frac{1}{i}\}$.

(d) If A is arbitrary, let $B = \bigcup_{U \in \mathcal{O}} U$. By the above case, there is a partition of unity for B, and so there is a partition of unity for A.

We now revisit case (b). Let $B_i = A_i \setminus \text{int} A_{i-1}$; by construction, B_i is compact. Let $\mathcal{O}_i$ be the collection of all $U \cap (\text{int} A_{i+1} \setminus A_{i-2})$; this forms an open cover of B_i. Then there is a partition of unity Ψ_i for B_i, subordinate to $\mathcal{O}_i$. Let

$$\sigma(x) = \sum_i \sum_{\varphi \in \Psi_i} \varphi(x).$$

This is a finite sum on an open neighborhood V of x for all $x \in A$. Let

$$\psi(x) = \frac{\varphi(x)}{\sigma(x)}$$

for all $\varphi \in \Psi_i$, for all i. This is the partition of unity as desired. $\square$

[17] i highly recommending reading spivak page 63 onwards for the latter 3 cases, cuz i was not really paying too much attention here

[18] who's? whos? whomstdve? what?

§40 Day 40: Extended Definition of Integral (Jan. 13, 2025)

Recall that a bounded function on a closed rectangle A is integrable if and only if its set of discontinuities has measure zero. Let $C \subset A \subset \mathbb{R}^n$ be bounded, and consider a bounded function $f : C \to \mathbb{R}$. Then $\int_C f$ is defined as $\int_A f \cdot \chi_C$, provided that it exists. So f is integrable on C if and only if the discontinuities of $f\chi_C$ has measure zero. We say that the boundary ∂C is given by the discontinuities of χ_C. The bounded set C is *Jordan measurable* if ∂C has measure zero.

A continuous function on a bounded open set is *not necessarily* integrable. For example, consider x^{-1} on the open unit interval $(0, 1)$.

> **Remark 40.1.** Consider $A \subset \mathbb{R}^n$ open, and let $f : A \to \mathbb{R}$ be locally bounded (i.e., bounded on some neighborhood of any point), and the set of discontinuities of f has measure zero. If f vanishes outside a compact subset C of A, then f is integrable on any bounded subset U of A containing C, and $\int_U f = \int_C f$.

With this, it makes sense to say that f is integrable on A and to define $\int_A f$ as $\int_C f$. We now introduce the extended definition of the integral,

Definition 40.2. Let $A \subset \mathbb{R}^n$ be open, and let $f : A \to \mathbb{R}$ be a locally bounded function whose set of discontinuities has measure 0. Let Φ be a $\mathcal{C}^\infty$ partition of unity for A with compact support, subordinate to $\mathcal{O} = \{A\}$. Then $\int_A f\,|\varphi|$ exists, for all $\varphi \in \Psi$. We have that f is integrable on A if $\sum_{\varphi\in\Phi} \int_A \varphi\,|f|$ converges. In this case, we define

$$\int_A f = \sum_{\varphi\in\Phi} \int_A \varphi \cdot f.$$

Theorem 40.3. Let $A \subset \mathbb{R}^n$ be open, and let $f : A \to \mathbb{R}$ be bounded, whose set of discontinuities has measure zero. Then let Φ be a $\mathcal{C}^\infty$ partition of unity for A with compact support, subordinate to $\{A\}$. Then,

(i) If $\sum_{\varphi\in\Phi} \int_A \varphi\,|f|$ converges and Ψ is another such partition of unity satisfying the above conditions such that $\sum_{\psi\in\Psi} \int_A \psi\,|f|$ converges, then

$$\sum_{\varphi\in\Phi} \int_A \varphi \cdot f = \sum_{\psi\in\Psi} \int_A \psi \cdot f$$

(ii) If A and f are bounded, then $\sum_{\varphi\in\Phi} \int_A \varphi\,|f|$ converges.

(iii) If A is Jordan measurable and f is bounded, then

$$\int_A f = \sum_{\varphi\in\Phi} \int_A \varphi \cdot f,$$

where the left hand side is the original integral.

§41 Day 41: Extended Definition of Integral, Pt. 2 (Jan. 15, 2025)

Given $A \subset \mathbb{R}^n$ open, let $f : A \to \mathbb{R}$ be locally bounded such that its discontinuities are of measure zero. Let Φ be a $\mathcal{C}^\infty$ partition of unity for A with compact support, subordinate to $\{A\}$. We say f is integrable on A if

$$\sum_{\varphi \in \Phi} \int_A \varphi \, |f|$$

converges. If it does, we define $\int_A f = \sum_{\varphi \in \Phi} \int_A \varphi f$.

(i) If $\sum_{\varphi \in \Phi} \int_A \varphi \, |f|$ converges, and Ψ is another partition of unity, then $\sum_{\psi \in \Psi} \int_A \psi \, |f|$ also converges, and

$$\sum_{\psi \in \Psi} \int_A \psi f = \sum_{\varphi \in \Phi} \varphi f.$$

Proof. For any $\varphi \in \Phi$, φf vanishes outside a compact subset of A, and only finitely many $\psi \in \Psi$ are nonzero on a given compact set. So we may write

$$\sum_{\varphi \in \Phi} \int_A \varphi f = \sum_{\varphi \in \Phi} \int_A \sum_{\psi \in \Psi} \psi \varphi f = \sum_{\varphi \in \Phi} \sum_{\psi \in \Psi} \int_A \psi \varphi f,$$

which is a convergent series. Rewriting the above, we have that

$$\sum_{\varphi \in \Phi} \int_A \varphi \, |f| = \sum_{\varphi \in \Phi} \sum_{\psi \in \Psi} \int_A \psi \varphi \, |f| \, ,$$

which means it converges absolutely as well, because $\left| \int_A \psi \varphi f \right| \leq \int_A \psi \varphi \, |f|$. A rearrangement gives the convergence of $\sum_{\psi \in \Psi} \int_A \psi \, |f|$ as desired. □

(ii) If A, f are both bounded, then $\sum_{\varphi \in \Phi} \int_A \varphi \, |f|$ converges, and so f is integrable on A.

Proof. Let $|f| \leq M$ on A. It is enough to show that the partial sums are bounded on $B \supset A$, where B is a closed rectangle. Let Θ be a finite subset of Φ, then

$$\sum_{\varphi \in \Theta} \int_A f \varphi \, |f| \leq \sum_{\varphi \in \Theta} M \int_A \varphi = M \underbrace{\int_A \sum_{\varphi \in \Theta} \varphi}_{\leq 1} \leq M v(B).$$

Hence, f is integrable on A. □

(iii) If A is Jordan measurable and $|f|$ is bounded, then $\int_A f = \sum_{\varphi\in\Phi} \int_A \varphi f$.

Proof. We'll use Spivak problem 3-22. For all $\varepsilon > 0$, there exists a compact Jordan measurable $C \subset A$ such that $\int_{A\setminus C} 1 = \int_A \chi_{A\setminus C} < \varepsilon$. Only finitely many $\varphi \in \Phi$ are nonzero on C;

$$\begin{aligned}
\left| \int_A f - \sum_{\varphi\in\Theta} \int_A \varphi f \right| &\leq \int_A \left| f - \sum_{\varphi\in\Theta} \varphi f \right| \\
&\leq M \int_A \left| 1 - \sum_{\varphi\in\Theta} \varphi \right| \\
&= M \int_A \left(1 - \sum_{\varphi\in\Theta} \varphi \right) \\
&= M \int_A \sum_{\varphi\in\Theta} \varphi \leq M \int_{A\setminus C} 1 < M\varepsilon
\end{aligned}$$

as desired. $\square$

§42 Day 42: Change of Variables (Jan. 17, 2025)

Recall the definition of integration by substitution; let $g : [a, b] \to \mathbb{R}$ be continuously differentiable, and let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then

$$\int_{g(a)}^{g(b)} f = \int_a^b (f \circ g)g'.$$

If g is injective, then

$$\int_{g([a,b])} f = \int_{[a,b]} (f \circ g) \left|g'\right|,$$

where we consider the cases in which g is increasing or decreasing separately.

Theorem 42.1. Given $A \subset \mathbb{R}^n$ open, $g : A \to \mathbb{R}^n$ injective and continuously differentiable, and $\det g'(x) \neq 0$ at every point of A, then $f : [a, b] \to \mathbb{R}$ is integrable if and only if $f \circ g \left|\det g'\right| : A \to \mathbb{R}$ is integrable, i.e.

$$\int_{g(A)} f = \int_A f \circ g \left|\det g'\right|.$$

It's enough to prove that $f : g(A) \to \mathbb{R}$ is integrable implies that $f \circ g \left|\det g'\right| : A \to \mathbb{R}$ is integrable, and the above equality is true; i.e., $F = f \circ g \left|\det g'\right|$ is integrable, then f is integrable. This is because by applying the weaker version of the theorem to g^{-1}, F is integrable implies that $(F \circ g^{-1}) \left|\det g^{-1}\right|$ is integrable, and

$$(F \circ g^{-1}) \left|\det g^{-1}\right| (y) = F(g^{-1}(y)) \frac{1}{\left|\det g'(g^{-1}(y))\right|} = f(y),$$

since $(g^{-1})'(y) = (g'(g^{-1}(y)))^{-1}$ by the inverse function theorem. We can use this function to "compare" definite integrals as follows; consider $A \subset C \subset \mathbb{R}^n$, with A open, C compact, and $C \setminus A$ measure zero. Then A, C are Jordan measurable, since $\partial C, \partial C \setminus A$ are of measure zero. Assume g is injective, continuously differentiable on A, $\det g'(x) \neq 0$, and $x \in A$; then f is integrable on $g(C)$ if and only if $f \circ g \left|\det g'\right|$ is integrable on C, and

$$\int_{g(C)} f = \int_C f \circ g \left|\det g'\right|.$$

This follows from the theorem, since $g(C) \setminus g(A) \subset g(C \setminus A)$, and so $g(C) \setminus g(A)$ is of measure zero. Then we have that

$$\int_{g(C)} f = \int_{g(A)} f = \int_A f \circ g \left|\det g'\right| = \int_C f \circ g \left|\det g'\right|.$$

Lemma 42.2. If $A \subset \mathbb{R}^n$, $g : A \to \mathbb{R}^n$ is continuous and $B \subset A$ is of measure zero, then $g(B)$ is of measure zero.

It is enough to prove that if $C \subset A$ is compact, then $g(B \cap C)$ is of measure zero. A can be covered by an expanding union of compact sets $\{C_i\}$ (with $C_i \subset C_{i+1}$) so that $\{g(C_i)\}$ to form a countable cover of $g(A)$. Then there exists $\delta > 0$ such that $|g(x) - g(y)| \leq c\,|x - y|$ for all $x \in C$ and $|x - y| < \delta$, by uniform continuity of g on a neighborhood of C. Therefore, every ball of radius ε is mapped to a ball of radius ε (which is what we mean when we say measure zero).

§43 Day 43: Examples of Change of Variables (Jan. 20, 2025)

We give some examples of change of variables today.[19]

(a) (*Polar Coordinates*) Let A, C be open, closed rectangles respectively in $\mathbb{R}^2$, such that $A = (0, r) \times (0, \theta)$ and $C = [0, r] \times [0, \theta]$ (with $\theta \leq 2\pi$). Then we may perform a change of coordinates with $(x, y) \mapsto (r\cos\theta, r\sin\theta)$. Note that we have

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \implies \left| \det \frac{\partial(x,y)}{\partial(r,\theta)} \right| = r,$$

and so

$$\iint_B f(x,y)\,dx\,dy = \iint_A f(r\cos\theta, r\sin\theta) r\,dr\,d\theta,$$

where B is the sector of the circle with radius r, angle measure θ obtained from the transformation above.

(b) (*Spherical Coordinates*) Let $(x, y, z) \mapsto (r\cos\theta\sin\varphi, r\sin\theta\cos\varphi, r\cos\theta)$, bounded by $(r, \theta, \varphi) \in [0, \infty) \times [0, \pi] \times [0, 2\pi]$. We then compute the Jacobian and find that $\det = r^2 \sin\theta$, as shown below,

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} = \begin{pmatrix} \cos\varphi\sin\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ sin\varphi\sin\theta & r\sin\varphi\cos\theta & r\cos\varphi\sin\theta \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix},$$

in which we may evaluate $|\det \partial_{r,\theta,\varphi}(x,y,z)| = r^2 \sin\theta$ as desired through cofactor expansion. In particular, by the same construction as previous, we have

$$\iiint_B f(x,y,z)\,dx\,dy\,dz = \iiint_A f(r,\theta,\varphi) r^2 \sin\theta\,dr\,d\theta\,d\varphi.$$

For example, the volume of a ball with radius r is given by the above,

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{r} \rho^2 \sin\theta\,d\rho\,d\theta\,d\varphi = \left.\frac{\rho^3}{3}\right|_0^r \cdot (-\cos\theta)|_0^{2\pi} \cdot 2\pi = \frac{4}{3}\pi r^3.$$

(c) (*Volume of an Ellipsoid of Revolution*) For an example, consider the region given by the following equation,

$$\frac{x^2+y^2}{a^2} + \frac{z^2}{b^2} \leq 1,$$

where the boundary is given by $z = \pm\frac{b}{a}\sqrt{a^2 - x^2 - y^2}$. We may express half the volume of the ellipsoid as

$$\frac{V}{2} = \frac{b}{a}\iint_{x^2+y^2\leq a^2} \sqrt{a^2 - x^2 - y^2}\,dx\,dy,$$

and use polar coordinates to rewrite as follows,

$$\frac{b}{a}\int_0^{2\pi}\int_0^{a} r\sqrt{a^2 - r^2}\,dr\,d\theta = \frac{b}{a}\cdot 2\pi \int_0^a r\sqrt{a^2-r^2}\,dr = \frac{b}{a}\cdot 2\pi \cdot \frac{a^3}{3},$$

where we obtain $V = \frac{4}{3}\pi a^2 b$.

[19] good lord this is tikz hell

(d) (*Volume of a General Ellipsoid*) Consider the region given by the following equation,

$$E : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1.$$

We could use generalized spherical coordinates, i.e.

$$(x, y, z) \mapsto (a \cos\varphi \sin\theta, b \sin\varphi \sin\theta, c\cos\theta),$$

or $(x, y, z) \mapsto (au, bv, cw)$, which takes $B : u^2 + v^2 + w^2 \leq 1$ to the ellipsoid. Then we have

$$\left| \det \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| = abc,$$

and we get

$$\iiint_E 1 \, dx \, dy \, dz = \iiint_B abc \, du \, dv \, dw = abc \cdot \left(\frac{4}{3}\pi \right).$$

§44 Day 44: Change of Variables, Pt. 2 (Jan. 22, 2025)

Given $A \subset \mathbb{R}^n$ open, $g : A \to \mathbb{R}^n$, injective, and continuously differentiable with $\det g'(x) \neq 0$ for $x \in A$, if f is integrable on $g(A)$, then $f \circ g\,|\det g'|$ is integrable on A, and

$$\int_{g(A)} f = \int_A f \circ g \left|\det g'\right|.$$

We provide a correction to a previous lemma;

Lemma 44.1. Given $A \subset \mathbb{R}^n$ open, $\mathcal{C}^1$ $g : A \to \mathbb{R}^m$, if $B \subset A$ has measure zero, then $g(B)$ has measure zero.

This is true if g is $\mathcal{C}^1$, by MVT.

For example, let $g : \mathbb{R}^n \to \mathbb{R}^n$, $g(x) = T(x)$, where T is a linear transformation. If $A \subset \mathbb{R}^n$ is Jordan measurable, i.e. $g'(x) = T$, then $g(A)$ is Jordan measurable and $\text{vol}(g(A)) = |\det T \cdot \text{vol}(A)|$.

(i) If T is invertible, then g takes $\int A$ to $\int B$, and $g(A)$ is Jordan measurable by the lemma,

$$\text{vol}(g(A)) = \int_{g(A)} 1 = \int_A \left|\det g'\right| = |\det T| \int_A 1 = |\det T|\,\text{vol}(A).$$

(ii) If T is not invertible at 0, then let $T = 0$. T is linear, so $g(A)$ lies in a proper linear subspace, and therefore has measure zero.

We now prove the Change of Variables theorem. Consider $U \subset A$ open, Jordan measurable, with closure in A. If f is integrable on $g(A)$, then f is integrable on $g(U)$, which is open, Jordan measurable, has closure in $g(A)$, i.e., is bounded on $g(U)$, since the closure of $g(U)$ is compact. Moreover, its discontinuities are of measure zero.

(i) (*Reduction to Local Statement*) Let $\mathcal{O}$ be an open cover of A by sets U as per above; i.e., for each $U \in \mathcal{O}$ and any integrable f, we have

$$\int_{g(U)} f = \int_U (f \circ g) \left|\det g'\right|.$$

Then the theorem is true on the entirety of A, since g is automatically locally injective for each point. To prove this, observe that $\{g(U) \mid U \in \mathcal{O}\}$ covers $g(A)$. Then let Φ be a compactly supported partition of unity subordinate to this cover; every $\varphi \in \Phi$ vanishes outside a compact subset of $g(U)$ for some U, so $\varphi \circ g$ vanishes outside a compact subset of U, from the local injectivity of g, and so we have

$$\int_{g(A)} \varphi f = \int_{g(U)} \varphi f = \int_U (\varphi f) \circ g \left|\det g'\right| = \int_A (\varphi f) \circ g \left|\det g'\right|$$

as desired.

The other 3 cases from Spivak will be covered in next lecture.

§45 Day 45: Change of Variables, Pt. 3 (Jan. 24, 2025)

Given $A \subset \mathbb{R}^n$ open, let $g : A \to \mathbb{R}$ be injective, $\mathcal{C}^1$, $\det g'(x) \neq 0$ on A. Then if $f : g(A) \to \mathbb{R}$ is integrable, then $(f \circ g)\left|\det' g\right|$ is integrable on A, and

$$\int_{g(A)} f = \int_A (f \circ g)\left|\det g'\right|.$$

We now prove the theorem.

(i) (*Coverings*) It's enough to prove that for all open covers $\mathcal{O}$ of A by Jordan measurable sets U with closure in A, we have, for all $U \in \mathcal{O}$,

$$\int_{g(U)} f = \int_U (f \circ g)\left|\det g'\right|.$$

Let $\{g(U) \mid U \in \mathcal{O}\}$ be an open cover of A by sets as above; then let $\Phi = \{\varphi\}$ be a compactly supported, $\mathcal{C}^\infty$ partition of unity subordinate to $\{g(U)\}$. Then $\{\varphi \circ g\}$ is a partition of unity subordinate to $\mathcal{O}$, and

$$\int_{g(A)} \varphi f = \int_{g(U)} \varphi f = \int_U (\varphi f) \circ g\left|\det g'\right| = \int_A (\varphi f) \circ g\left|\det g'\right|,$$

in which we may note

$$\int_{g(A)} f = \sum_{\varphi \in \Phi} \int_{g(A)} \varphi f = \sum_{\varphi \in \Phi} \int_A \underbrace{(\varphi f) \circ g}_{=(\varphi \circ g)(f \circ g)} \left|\det g'\right|.$$

The same equation holds with $|f|$ in place of f, which shows that $(f \circ g)|\det g'|$ is indeed integrable on A.

> **Remark 45.1.** The theorem also follows from the assumption that
>
> $$\int_V f = \int_{g^{-1}(V)} (f \circ g)\left|\det g'\right|$$
>
> for V in some admissible cover of $g(A)$. This follows from (i) applied to g^{-1}.

(ii) (*Constant Functions*) If $\{V\}$ is an open cover of $g(A)$ by open rectangles with closure in $g(A)$, then for all V,

$$\int_V f = \int_{g(V)} (f \circ g)\left|\det g'\right|$$

for $f = 1$. To start, it is enough to show that the above equation holds for any integrable f; if it holds for $f = 1$, then it holds for any $f = c$, where c is some constant. Let P be a partition of V, and let, for all subrectangles S of P, $f_S = m_S(f)$ a constant function; then write

$$\begin{aligned} L(f, P) = \sum_S m_S(f) v(S) &= \sum_S \int_S f_s \\ &= \sum_S \int_{g^{-1}(\int S)} (f_s \circ g)\left|\det g'\right| \\ &\leq \sum_S \int_{g^{-1}(\int S)} (f \circ g)\left|\det g'\right| \leq \int_{g^{-1}(V)} (f \circ g)\left|\det g'\right|. \end{aligned}$$

We may note that $\int_V f = \sup_P L(f, P)$, which shows

$$\int_V f \leq \int_{g^{-1}(V)} (f \circ g) \left|\det g'\right|$$

as desired; a similar argument with $f_S = M_S(f)$ yields the above inequality in the opposite direction.

(iii) (*Compositions*) If the theorem is true for $g : A \to \mathbb{R}^n$ and $h : g(A) \to \mathbb{R}^n$, then it holds for $h \circ g : A \to \mathbb{R}^n$. Directly write as follows,

$$\begin{aligned}\int_{h \circ g(A)} f &= \int_{g(A)} (f \circ h) \left|\det h'\right| \\ &= \int_A ((f \circ h) \left|\det h'\right|) \circ g \cdot \left|\det g'\right| \\ &= \int_A f \circ (h \circ g) \left|\det(h \circ g)'\right|,\end{aligned}$$

as desired.

(iv) (*Invertible Linear Transformations*) The theorem is true if g is an invertible linear transformation, i.e. $g'(x) = g$.[20] By (i), (ii), it is enough to show that for all open rectangles V with closure in $g(A)$, we have that

$$\int_{g(U)} 1 = \int_U \left|\det g'\right|,$$

i.e., $\text{vol}(g(U)) = |\det g'| \,\text{vol}(U)$. Then g is a composite of elementary functions; by (iii), it is enough to prove that the theorem holds for elementary functions.

We finish the proof next lecture.

[20] watefak

§46 Day 46: Change of Variables, Pt. 4 (Jan. 27, 2025)

Today, we complete the proof of the change of variables theorem. Given a point $a \in A$, we can assume $g'(a) = I$; we may do this since $T = g'(a)$, and consider that

$$(T^{-1} \circ g)'(a) = T^{-1} \circ g'(a) = I,$$

and $g = T \circ T^{-1}(g)$ as desired. We now prove the theorem by inducting on n;

- For $n = 1$, the theorem clearly holds as per integration by substitution.
- Assume that the theorem holds for $n - 1$; it is enough to show that for all $a \in A$, there exists a Jordan measurable open subset U with closure in A such that the theorem holds for U with $f = 1$; we may also assume that $g'(a) = I$. Let's express g *locally* as a composite of two mappings, each of which changes fewer than n coordinates. Let $h : A \to \mathbb{R}^n$, and $h(x) = (g_1(x), \ldots, g_{n-1}(x), x_n)$, with $h'(a) = I$; so there exists an open neighborhood $U' \ni a$ such that $h|_{U'}$ is a $\mathcal{C}^1$ injective function that satisfies $\det(h|_{U'})'(x) \neq 0$.

  Let $k : h(U') \to \mathbb{R}^n$ be given by $y \mapsto (y_1, \ldots, y_{n-1}, g_n(h^{-1}(y)))$, and $g = k \circ h$ on U'. With this, we have expressed g as the composition of two maps, each of which changes fewer than n coordinates, and so there is an open neighborhood V of $h(a)$ in $h(U')$ on which h is $\mathcal{C}^1$ and injective, with $\det h'(y) \neq 0$.

  Take $U = h^{-1}(V)$, and $g = h \circ k$, where $h : U \to \mathbb{R}^n$, $k : V \to h(U) \to \mathbb{R}^n$. By (iii) from last lecture, it is enough to prove the theorem for h, k individually. Given $a \in U$, let D' be a closed rectangle in U with $a \in \operatorname{int} D'$, and $D' = D \times [a_n, b_n]$, where D is a rectangle in $\mathbb{R}^{n-1}$. We want to show that

  $$\int_{h(D')} 1 = \int_{D'} \left|\det h'\right|;$$

  specifically, by Fubini's, we have

  $$\int_{h(D')} 1 = \int_{[a_n, b_n]} \left(\int_{h(D \times \{x_n\})} 1 \, dx_1, \ldots, \, dx_{n-1} \right) dx_n.$$

  Let $h_{x_n}(x_1, \ldots, x_{n-1}) = (g_1(x_1, \ldots, x_n), \ldots, g_{n-1}(x_1, \ldots, x_n)))$. Then $h_{x_n} : D \to \mathbb{R}^{n-1}$, is injective because h is also injective, and we get that the above integral is equal to

  $$\begin{aligned} &= \int_{[a_n, b_n]} \left(\int_D \left|\det h'_{x_n}(x_1, \ldots, x_{n-1})\right| \, dx_1, \ldots, dx_{n-1} \right) dx_n \\ &= \int_{[a_n, b_n]} \left(\int_D \left|\det h'(x_1, \ldots, x_{n-1})\right| \, dx_1, \ldots, dx_{n-1} \right) dx_n \\ &= \int_{D'} \left|\det h'\right|. \quad \square \end{aligned}$$

§47 Day 47: Sard's Theorem (Jan. 29, 2025)

The change of variables theorem is true without the hypothesis that $\det g'(x) \neq 0$ for all $x \in A$.

Theorem 47.1 (Sard's Theorem)**.** Given $A \subset \mathbb{R}^n$, open, $\mathcal{C}^1$ function $g : A \to \mathbb{R}^n$, let $B = \{x \mid A \mid \det g'(x) = 0\}$. Then $g(B)$ has measure zero.

It is enough to prove that if J is a closed cube in A, then $g(J \cap B)$ has measure zero. This is because we can cover A with countably many closed cubes in A. Subdivide J into N^n subcubes with sidelength ℓ/N; given ε, we can choose N such that, for all x, y in some subcube S, we have

$$|Dg(x)(y-x) - g(y) + g(x)| < \varepsilon |x-y|.$$

If $A = (a_{ij})$, then let $|A| = \max_{i,j} |a_{ij}|$. $Dg(x)$ uniformly acts on J, so we can choose N such that $|Dg(x) - Dg(y)| < \frac{\varepsilon}{n^2}$ for all $x, y \in S$. Since a continuous function on a compact set is uniformly continuous, consider $x \in S$, and let $f(z) = Dg(x)(z) - g(z)$, which is a $\mathcal{C}^1$ function on z. Then $Df(z) = Dg(x) - Dg(z)$, so $|Df(z)| < \frac{\varepsilon}{n^2}$, for any $z \in S$. Then we have that

$$|Dg(x)(y-x) - g(y) + g(x)| = |f(y) - f(x)| < \varepsilon |x-y|$$

by MVT as desired. If $S \cap B \neq \emptyset$, then take $x \in B$; since $\det g'(x) = 0$, as the image lies in a proper linear subspace, we have that $Dg(x)(y-x)$ lies in an $n-1$ dimensional linear subspace V of $\mathbb{R}^n$. Let $Dg(x)(y-x) + g(x) \in V + g(x)$; so $g(y) < \sqrt{n}\varepsilon\frac{\ell}{N}$. Therefore, $\{g(y) \mid y \in S\}$ lies within $\varepsilon\sqrt{n}\frac{\ell}{N}$ of $v + g(x)$, and we have that

$$|g(y) - g(x)| < M|x-y| < M\sqrt{n}\frac{\ell}{N}$$

for all $x, y \in S$, by MVT. Therefore, $g(y)$ lies within a cylinder with height $< 2\varepsilon\sqrt{n}\frac{\ell}{N}$ whose base is a ball of radius $< M\sqrt{n}\frac{\ell}{N}$ centered on $g(x)$, in $V + g(x)$. The volume of said cylinder is less than $c\varepsilon\left(\frac{\ell}{N}\right)^n$ for some constant c, and so $g(J \cap B)$ lies in a set of volume less than $c\varepsilon\ell^n$ for all $\varepsilon > 0$, which we may take $\varepsilon \to 0$ to see $g(J \cap B)$ has measure zero. [21]

[21] note to self: revisit this proof. i don't intuitively get it.

§48 Day 48: Integration of Lower Dimensional Subsets (Jan. 31, 2025)

How do we integrate over a lower dimensional subset of $\mathbb{R}^n$, such as over a curve or a manifold?

(i) Let us consider the integral of a continuous function $f(x_1, \dots, x_n)$ over a curve C in $\mathbb{R}^n$. Let γ be a piecewise, $\mathcal{C}^1$, parameterized curve $C : x = \gamma(t), t \in [a, b]$, where γ is continuous except at finitely many points in $[a, b]$. We ask that γ be smooth, injective, and $\gamma'(t) \neq 0$ for any t. For example, consider $\gamma(\theta) = (\cos\theta, \sin\theta)$ on $[0, 2\pi]$, where we may exclude $\theta = 0$. Can we define $\int_C f$ as $\int_a^b f(\gamma(t))\, dt$? This depends on the parameterization; for example, $t : [c, d] \to [a, b]$ with $t(s)$ $\mathcal{C}^1$, $t'(s) > 0$ on (a, b), and $C : x = \delta(s)$, where $\delta(s) = \gamma(t(s))$, we have

$$\int_a^b (f \circ \gamma)(t)\, dt = \int_c^d (f \circ \delta)(s)\delta'(s)\, ds;$$

In a way, we may choose a parameterization that depends only on arclength. Recall that the length of C is given by

$$\ell(c) = \int_a^b \left|\gamma'(t)\right|\, dt.$$

We may define length as the supremum of polygonal approximations to C. Let $P = \{t_i\}$ be a partition of $[a, b]$, and let $\ell_P(c) = \sum_{i=1}^k |\gamma(t_i) - \gamma(t_{i-1})|$. We say that γ is *rectifiable* if $\sup_P \ell_P(c)$ exists; if it does exist, then we define length as such.

Exercise 48.1 (Using Riemann sums)**.** If γ is $\mathcal{C}^1$, then γ is rectifiable and the length is given by $\int_a^b |\gamma'(t)|\, dt$, which we may write as $\int_C f\, ds$. In this manner, we may define the integral of f over C with respect to arclength as

$$\int_C f\, ds = \int_a^b f(\gamma(t)) \left|\gamma'(t)\right|\, dt,$$

which depends only on C.

(ii) For line integrals, let the integral over $C : x = \gamma(t)$ with $t \in [a, b]$, and let us have the function $g(x) = (g_1(x), \dots, g_n(x))$ be continuous. Then we have that

$$\int_C g_1(x)\, dx_1 + \cdots + g_n(x)\, dx_n = \int_a^b (g_1 \circ \gamma)(t)\gamma_1'(t) + \cdots + (g_n \circ \gamma)(t)\gamma_n'(t)\, dt.$$

The index of parameterization is the differential form $g_1(x)\, dx_1 + \cdots + g_n(x)\, dx_n$, i.e., the differential of a $\mathcal{C}^1$ function $f(x)$ where

$$\partial f = \frac{\partial f}{\partial x_1}\, dx_1 + \cdots + \frac{\partial f}{\partial x_n}\, dx_n$$
$$\omega = \sum_{i=1}^n g_i(x)\, dx_i.$$

If $\omega = df$, we say that f is primitive of ω; then $\int_C \omega = f(\gamma(b)) - f(\gamma(a))$, where

$$\int_C \omega = \int_a^b \sum_{i=1}^n \frac{\partial}{\partial x_i}(f \circ \gamma)(t)\gamma_i'(t) = \int_a^b \frac{d(f \circ \gamma)(t)}{dt}\, dt.$$

Recall that the tangent space of $\mathbb{R}^n$ at a point a can be identified with vectors $v = v_a \in \mathbb{R}$ that are the directional derivatives at a;

$$\begin{aligned} v_a(f) = (D_v f)(a) = Df(a)(v) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_j}(a) v_i \\ &= \left(\sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i}\bigg|_a \right)(f), \end{aligned}$$

where we have that $v_a = \sum_{i=1}^{n} v_i e_{i,a}$. In particular, we may identify $\frac{\partial}{\partial x_i}\Big|_a$ with the standard basis $e_{i,a}$ of $\mathbb{R}^n_a$, and we have that $dx_i(a)$ is a dual basis. The dual basis of $\mathbb{R}^n_a$ w.r.t. $e_{i,a}$ are the coordinate functions, if $v_a = \sum_{j=1}^{n} v_j e_{j,a}$, i.e. $v = (v_1, \ldots, v_n)$, then $dx_i(a)(v_a) = v_i$.

§49 Day 49: Multilinear Algebra (Feb. 3, 2025)

The coordinates for $\mathbb{R}^n$, $x = (x_1, \dots, x_n)$ with respect to the standard basis, can be thought of as linear functions $x_i : \mathbb{R}^n \to \mathbb{R}$ where $a \mapsto a_i$, i.e., $x_i = \pi_i$, which is the projection onto the ith component. Then $\mathbb{R}^n_a$ tangent space to $\mathbb{R}^n$ at a, where we may denote $e_{i,a}$ as the standard basis or $\frac{\partial}{\partial x_j}\Big|_a$. Then for $v_a \in \mathbb{R}^n_a$, where $v_a = (v_1, \dots, v_n)$ and $dx_i(a)$ is the dual basis, with $dx_i(a)(v_a) = v_i$. Then the $\mathcal{C}^r$ function $f : \mathbb{R}^n \to \mathbb{R}$ determines the $\mathcal{C}^{r-1}$ 1-form,

$$df : x \mapsto \text{element of dual to tangent space } \mathbb{R}^n_x.$$

We can write $df(a)(v_a) = Df(a)(v) = v_a(f)$. For example, let $f = x_i = pi_i : \mathbb{R}^n \to \mathbb{R}$ is the projection onto the ith coordinate; then

$$\begin{aligned} dx_i(a)(v_a) &= d\pi_i(a)(v_a) \\ &= D\pi_i(a)(v_a) \\ &= v_i. \end{aligned}$$

We also have that

$$dx_i(a)\left(\frac{\partial}{\partial x_j}\Big|_a\right) = \frac{\partial}{\partial x_j}\Big|_a (x_i) = \delta_{ij}.$$

Specifically, $dx_i(a)$ is a dual basis.

Lemma 49.1. $df = \frac{\partial f}{\partial x_1}\, dx_1 + \cdots + \frac{\partial f}{\partial x_n}\, dx_n$.

To see this, simply write

$$df(a)(v_a) = Df(a)(v) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) v_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) dx_i(a)(v_a).$$

We now give some background on multilinear algebra. Let V be an n-dual vector space over $\mathbb{R}$, and let $T : V^k \to \mathbb{R}$. We say that T is a multilinear function, or equivalently, a k-tensor if $T(v_1, \dots, v_k)$ is linear in each v_i. We denote $\mathcal{T}^k(v)$ to be the set of all such tensors. $\mathcal{T}^k(v)$ has the structure of a vector space; specifically,

$$\begin{aligned} (S+T)(v_1, \dots, v_k) &= S(v_1, \dots, v_k) + T(v_1, \dots, v_k), \\ (aT)(v_1, \dots, v_k) &= aT(v_1, \dots, v_k), \end{aligned}$$

For example, $\mathcal{T}^1(V)$ is the same as the dual space V^*. A *tensor product* is the operation $\otimes : \mathcal{T}^k(V) \times \mathcal{T}^\ell(V) \to \mathcal{T}^{k+\ell}(V)$, where

$$(S \otimes T)(V_1, \dots, V_{k+\ell}) = S(v_1, \dots, v_k) T(v_{k+1}, \dots, v_{k+\ell}).$$

Note that $\otimes$ is not commutative, but is distributive and associative, i.e.

$$\begin{aligned} (S_1 + S_2) \otimes T &= S_1 \otimes T + S_2 \otimes T, \\ S \otimes (T_1 + T_2) &= S \otimes T_1 + S \otimes T_2, \\ (aS) \otimes T &= a(S \otimes T), \\ (S \otimes T) \otimes U &= S \otimes (T \otimes U), \end{aligned}$$

where we may simply write $S \otimes T \otimes U$ instead.

Proposition 49.2. Let $v_1, \dots, v_n$ be a basis of V, and let $\varphi_1, \dots, \varphi_n$ be a dual basis. Then $\varphi_{i_1}, \dots, \varphi_{i_k}$, where $1 \le i_1, \dots, i_k \le n$, is a basis of $\mathcal{T}^k(V)$. In particular, $\dim \mathcal{T}^k(V) = n^k$.

The proof is straightforward. Let us write as follows,

$$(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k})(v_{ji}, \dots, v_{jk}) = \delta_{i_1,j_1}\delta_{i_2,j_2} \dots \delta_{i_k,j_k} = \begin{cases} 1 & \text{if } i_1 = j_1, \dots, i_k = j_k, \\ 0 & \text{otherwise.} \end{cases}$$

§50 Day 50: Multilinear Algebra, Pt. 2 (Feb. 5, 2025)

We recap the proposition from class last time.

Proposition 50.1. Let $\{v_1, \ldots, v_n\}$ be a basis of V, $\{\varphi_1, \ldots, \varphi_n\}$ be a dual basis. Then $\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}$, over all combinations $1 \le i_1, \ldots, i_k \le n$, forms a basis of $\mathcal{T}^k(V)$.

Let $T \in \mathcal{T}^k(V)$, and let $w_j = \sum_{i=1}^n a_{ji} v_i$, where $j = 1, \ldots, k$. Then

$$\begin{aligned} T(w_1, \ldots, w_k) &= \sum_{i_1, \ldots, i_k = 1}^{n} a_{1_{i_1}} \ldots a_{k_{i_k}} T(v_{i_1}, \ldots, v_{i_k}) \\ &= \sum_{i_1, \ldots, i_k = 1}^{n} T(v_{i_1}, \ldots, v_{i_k})(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k})(w_1, \ldots, w_k), \end{aligned}$$

where $(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k})(v_{ji}, \ldots, v_{jk}) = 1$ if $(j_1, \ldots, j_k) = (i_1, \ldots, i_k)$, and 0 otherwise. Therefore,

$$T = \sum_{i_1, \ldots, i_k = 1}^{n} T(v_{i_1}, \ldots, v_{i_k})(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}). \quad \square$$

A linear transformation $f : V \to W$ induces the linear transformation $f^* : \mathcal{T}^k(W) \to \mathcal{T}^k(V)$ through $(f^*T)(v_1, \ldots, v_k) = T(f(v_1), \ldots, f(v_k))$. f^* commutes with $\otimes$, i.e., $f^*(S \otimes T) = f^*S \otimes f^*T$.

(i) The inner product on V is a 2-tensor T which is symmetric and positive-definite ($T(v, v) \ge 0$, $= 0$ iff $v = 0$). If T is an inner product on V, there exists an orthonormal basis $v_1, \ldots, v_n$ w.r.t. T i.e. $T(v_i, v_j) = \delta_{ij}$. An orthonormal basis for T provides an isomorphism $f : \mathbb{R}^n \to V$ such that $f^*T = \langle\, , \rangle$.

Proof. By the Gram-Schmidt process, we have $f(x_1, \ldots, x_n) = \sum_{i=1}^n x_i v_i$. $\square$

(ii) Determinant function. Consider an alternating k-tensor $\omega \in \mathcal{T}^k(V)$, where $\omega(v_1, \ldots, v_i, \ldots, v_k)$ has the property that "if we interchange two of the vectors, it changes sign". Specifically, we say that it is alternating and skew-symmetric. Consider,

$$\mathrm{Alt} : \mathcal{T}^k(V) \to \mathcal{T}^k(V),$$

where

$$(\mathrm{Alt}\, T)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),$$

where S_k is the symmetric group on k symbols. Let $\Omega^k(V) \subset \mathcal{T}^k(V)$ be the subspace of alternating tensors.

Theorem 50.2 (Spivak 4-3)**.** $\mathrm{Alt} : \mathcal{T}^k(V) \to \mathcal{T}^k(V)$ is a projection onto $\Omega^k(V)$. Specifically,

(a) $\mathrm{Alt}\, T \in \Omega^k(V)$,

(b) $w \in \Omega^k(V) \implies \mathrm{Alt}\, w = w$,

(c) $\mathrm{Alt}(\mathrm{Alt}\, T) = \mathrm{Alt}\, T$.

Proof. Note that (c) follows (a) and (b). For (a), observe that

$$\begin{aligned}
\operatorname{Alt} T(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) T(v_{\mathcal{T}(\sigma(1))}, \ldots, v_{\mathcal{T}(\sigma(k))}) \\
&= \frac{1}{k!} \underbrace{\operatorname{sgn} \mathcal{T}}_{=-1} \sum_{\sigma \in S_k} (\operatorname{sgn} \mathcal{T}\sigma) T(v_{\mathcal{T}(\sigma(1))}, \ldots, v_{\mathcal{T}(\sigma(k))}) \\
&= -\operatorname{Alt} T(v_1, \ldots, v_k)
\end{aligned}$$

For (b), we have

$$\begin{aligned}
w(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) &= \operatorname{sgn} \sigma \cdot w(v_1, \ldots, v_k) \\
(\operatorname{Alt} w)(v_1, \ldots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \underbrace{w(v_{\sigma(1)}, \ldots, v_{\sigma(k)})}_{\operatorname{sgn} \sigma \cdot w(v_1, \ldots, v_k)} \\
&= w(v_1, \ldots, v_k).
\end{aligned}$$

□

§51 Day 51: Alternating Tensors and Wedge Product (Feb. 7, 2025)

Consider the projection Alt of $\mathcal{T}^k(V)$ onto $\Omega^k(V)$, where

$$\operatorname{Alt} T(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)(v_{\sigma(1)}, \dots, v_{\sigma(k)}).$$

Let us define the wedge product,

$$\wedge : \Omega^k(V) \times \Omega^\ell(V) \to \Omega(k+\ell)(V)$$
$$\omega \wedge \eta = \frac{(k+\ell)!}{k!\ell!} \operatorname{Alt}(\omega \otimes \eta).$$

We have that $\wedge$ enjoys a handful of properties, specifically

$$\begin{aligned}
(\omega_1 + \omega_2) \wedge \eta &= \omega_1 \wedge \eta + \omega_2 \wedge \eta \\
\omega \wedge (\eta_1 + \eta_2) &= \omega \wedge \eta_1 + \omega \wedge \eta_2 \\
(a\omega \wedge \eta) &= a(\omega \wedge \eta) \\
\omega \wedge \eta &= (-1)^{k\ell} \eta \wedge \omega \\
f^*(\omega \wedge \eta) &= f^*\omega \wedge f^*\eta.
\end{aligned}$$

Additionally, $\wedge$ is associative, i.e. $(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta)$, but requires more work to prove.

Lemma 51.1 (Spivak 4-4)**.** The alternating tensor enjoys the following properties,

(i) If $S \in \mathcal{T}^k(V)$, $T \in \mathcal{T}^\ell(V)$, and $\operatorname{Alt} S = 0$, then $\operatorname{Alt}(S \otimes T) = \operatorname{Alt}(T \otimes S) = 0$.

(ii) If $S \in \mathcal{T}^k(V)$, $\eta \in \Omega^\ell(V)$, then $\operatorname{Alt} S \wedge \eta = \frac{(k+\ell)!}{k!\ell!} \operatorname{Alt}(S \otimes \eta)$.

The first part of the lemma is left in a future homework assignment. For the second part, we have that

$$\operatorname{Alt} S \wedge \eta = \frac{(k+\ell)!}{k!\ell!} \operatorname{Alt}(\operatorname{Alt} S \otimes \eta),$$

and so on. (where are we going with this?)

Lemma 51.2 (Spivak 4-4, part (iii))**.** The alternating tensor also enjoys the property,

(iii) Let $\omega \in \Omega^k(V), \eta \in \Omega^\ell(V), \theta \in \Omega^m(V)$, then

$$(\omega \wedge \ell) \wedge \theta = \omega \wedge (\ell \wedge \theta) = \frac{(k+\ell+m)!}{k!\ell!m!} \operatorname{Alt}(\omega \otimes \eta \otimes \theta).$$

To see that this is true, simply plug in as follows,

$$\begin{aligned}
(\omega \wedge \ell) \wedge \theta &= \frac{(k+\ell+m)!}{(k+\ell)!m!} \operatorname{Alt}((\omega \wedge \eta) \otimes \theta) \\
&= \frac{(k+\ell+m)!}{(k+\ell)!m!} \frac{(k+\ell)!}{k!\ell!} \operatorname{Alt}(\omega \otimes \eta \otimes \theta).
\end{aligned}$$

Proposition 51.3 (Spivak 4-5)**.** Let $v_1, \dots, v_n$ be a basis of V, and let $\varphi_1, \dots, \varphi_n$ be a dual basis. Then $\varphi_{i1} \wedge \cdots \wedge \varphi_{ik}$, over $1 \le i_1 < \cdots < i_k \le n$, forms a basis of $\Omega^k(V)$. In particular $\dim \Omega^k(V) = \binom{n}{k}$.

Proof. Linear independence is just like before (Spivak 4-1); for the span, let $\omega \in \Omega^k(V)$. Then

$$\omega = \sum_{i_1,\ldots,i_k=1}^{n} a_{i_1,\ldots,i_k} \varphi i_1 \otimes \cdots \otimes \varphi_{i_k}.$$

Then

$$\omega = \operatorname{Alt}\omega = \sum_{i_1,\ldots,i_k=1}^{n} a_{i_1,\ldots,i_k} \operatorname{Alt}(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}).$$

Since each $\operatorname{Alt}(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k})$ is a constant multiple of $\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}$, these elements span $\Omega^k(V)$. □

Proposition 51.4 (Spivak 4-6)**.** Let $\omega \in \Omega^n(V)$, $n = \dim V$, and $v_1, \ldots, v_n \in V$. Let $w_i = \sum a_{ij} v_j$, where $i = 1, \ldots, n$. Then

$$\omega(w_1, \ldots, w_n) = \det(a_{ij})\omega(v_1, \ldots, v_n),$$

and so $\omega(v_1, \ldots, v_n) = 0$ if $v_1, \ldots, v_n$ are linearly dependent.

Proof. Define $\eta \in \mathcal{T}^n(\mathbb{R}^n)$ by

$$\eta((a_{11}, \ldots, a_{1n}), \ldots, (a_{n1}, \ldots, a_{nn})) = \omega\left(\sum_{j=1}^{n} a_{ij} v_j, \ldots, \sum_{j=1}^{n} a_{nj} v_j\right).$$

Then $\eta \in \Omega^n(\mathbb{R}^n)$, so $\eta = \lambda \cdot \det$ for some $\lambda \in \mathbb{R}$ and $\lambda = \eta(e_1, \ldots, e_n) = \omega(v_1, \ldots, v_n)$. □

§52 Day 52: Orientation and Volume (Feb. 10, 2025)

Previously, we showed that if $\omega \in \Omega^n(V)$ where $\dim V = n$, and $v_1, \ldots, v_n \in V$, we have that

$$w_i = \sum_{j=1}^{n} a_{ij} v_j., \qquad (i = 1, \ldots, n)$$

or more specifically,

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

then we have $\omega(w_1, \ldots, w_n) = \det(a_{ij})\omega(v_1, \ldots, v_n)$. This can be regarded as the "change of basis" analogue with tensors. This means that if $\dim V = n$, and $\omega \in \Omega^n(V)$ is nonzero, then ω divides up the bases of V into two disjoint classes,

$$\omega(v_1, \ldots, v_n) \begin{array}{l} > 0, \text{ or} \\ < 0. \end{array}$$

In particular, $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_n)$ are in the same class if and only if $\det(a_{ij}) > 0$, where $w_i = \sum_{j=1}^n a_{ij} v_j$. These classes are called *orientations* for V. The orientation to which a basis $\{v_1, \ldots, v_n\}$ of V belongs to is denoted $[v_1, \ldots, v_n]$, and is specifically an equivalence class, with the opposite orientation being denoted $-[v_1, \ldots, v_n]$. In $\mathbb{R}^n$, we write $[e_1, \ldots, e_n]$ as the standard orientation.

Suppose V admits an inner product T, and consider two orthonormal bases $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_n\}$ with respect to T; then

$$w_i = \sum_{j=1}^{n} a_{ij} v_j,$$

where $\det(a_{ij}) = \pm 1$, which we see from

$$\underbrace{T(w_i, w_j)}_{=\delta_{ij}} = T\left(\sum_{k=1}^{n} a_{ik} v_k, \sum_{\ell=1}^{n} a_{j\ell} v_\ell\right) = \sum_{k,\ell} a_{ik} a_{j\ell} \underbrace{T(v_k, v_\ell)}_{=\delta_{k\ell}} = \sum_{k} a_{ik} a_{jk},$$

and since $AA^T = I$, we have $\det A = \pm 1$. If $\omega \in \Omega^n(V)$, and $\omega(v_1, \ldots, v_n) = 1$, it follows from Spivak 4-6 that $\omega(w_1, \ldots, w_n) = \pm 1$. Therefore, given some inner product T and orientation μ for V, there exists a unique alternating n-tensor over V (which we will call ω) such that $\omega(v_1, \ldots, v_n) = 1$ when $\{v_1, \ldots, v_n\}$ are orthonormal, and $[v_1, \ldots, v_n] = \mu$. Such a unique ω is called the *volume element* defined by inner product T and orientation μ.

As an example, det is the volume element of $\mathbb{R}^n$ equipped with the standard inner product and usual orientation.

For another example, consider $v_1, \ldots, v_{n-1} \in \mathbb{R}^n$, with φ defined by

$$\varphi(w) = \det \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ w \end{pmatrix}.$$

Then $\varphi \in \Omega^1(\mathbb{R}^n)$; therefore, there exists some unique $z \in \mathbb{R}^n$ such that $\langle w, z \rangle = \varphi(w)$, which is more specifically denoted $z = v_1 \times \cdots \times v_{n-1}$, where $\times$ is the cross product. For example, if we are working in $\mathbb{R}^3$ with $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ already given as v_1, v_2 above, we obtain that

$$\det \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} = \langle u \times v, w \rangle \, .$$

Note: Bierstone introduced a tiny bit of vector fields and differential forms in this lecture, but I'm going to move everything to the next day's notes for cohesion purposes.

§53 Day 53: Vectors Fields and Differential Forms (Feb. 12, 2025)

A vector field F is a function on an open set $U \in \mathbb{R}^n$ such that $F(x) \in \mathbb{R}^n_a$, for some $a \in U$, i.e.

$$F(a) = \sum_{i=1}^n F_i(a)e_{i,a} = (F_1(a), \ldots, F_n(a))_a.$$

We say that F is $\mathcal{C}^r$ if each component function F_i is $\mathcal{C}^r$. Here are a few examples.

(i) (x, y) may be written as $xe_{1,(x,y)} + ye_{2,(x,y)}$,

(ii) $(x, -y)$ may be written as $xe_{1,(x,y)} - ye_{2,(x,y)}$,

(iii) The gradient vector field, $F(a) = \nabla f(a)$, is a $\mathcal{C}^r$ vector field (given that f is $\mathcal{C}^r$).

Operators on a vector field are induced by pointwise operations, such as

$$\begin{aligned}
(F + G)(p) &= F(p) + G(p), \\
\langle F, G\rangle (p) &= \langle F(p), G(p)\rangle , \\
(f \cdot F)(p) &= f(p)F(p).
\end{aligned}$$

If $F_1, \ldots, F_{n-1}$ are vector fields in $\mathbb{R}^n$, we can also define their cross product,

$$(F_1 \times \cdots \times F_{n-1})(p) = F_1(p) \times \cdots \times F_{n-1}(p).$$

We now introduce some more definitions.

Definition 53.1 (Divergence)**.** The divergence of F, denoted $\operatorname{div} F$, where F has components $(F_1, \ldots, F_n)$, is given by

$$\operatorname{div} F = \sum_{i=1}^n D_i F_i = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i} = \langle \nabla, F\rangle .$$

Note that $\nabla = \sum_{i=1}^n D_i \cdot e_i$.

Definition 53.2 (Curl)**.** Similarly, the vector field $\nabla \times F$ is called $\operatorname{curl} F$. In $\mathbb{R}^3$, we have that

$$(\nabla \times F)(a) = (D_2F_3 - D_3F_2)(a)e_{1,a} + (D_3F_1 - D_1F_3)(a)e_{2,a} + (D_1F_2 - D_2F_1)(a)e_{3,a}.$$

Expanding the notion of a vector field to a submanifold $M \subset \mathbb{R}^n$, we may denote $F : x \mapsto F(x) \in TM_x$, which is a function assigning vectors in the tangent space to each point on the submanifold M (which we will call a *vector field on* M). If $\varphi : W \to M$ is a $\mathcal{C}^r$ coordinate system for M, then there exists a unique (differentiable) vector field J on W such that $\varphi_{*a}(G(a)) = F(\varphi(a))$, where φ_{*a} is an injective linear transformation $TM_a \to \mathbb{R}^n_x$ (where $f(x) = a$).[22]

We say that F is $\mathcal{C}^r$ is G is $\mathcal{C}^r$ for any coordinate system. The definition is independent of the choice of coordinate system on a manifold of class $\mathcal{C}^{r+1}$. Now, consider a vector field H on V such that $\psi_{*b}(H(b)) = F(\psi(b))$. If $\varphi(a) = \psi(b)$, i.e. $b = (\psi^{-1} \circ \varphi)(a)$, then

$$H(b) = (\psi^{-1} \circ \varphi)_{*a}(J(a)) = D(\psi^{-1} \circ \varphi)(a) \cdot J(a).$$

We now proceed to define k-forms.

[22] note: we are skipping ahead from page 88 to page 115. i don't know what bierstone was writing but the book explains it better, so read there tbh

Definition 53.3 (Differential k-form)**.** A differential k-form on an open subset U of $\mathbb{R}^n$ is a function ω where $(a \in U) \mapsto \omega(a) \in \Omega^k(\mathbb{R}^n_a)$. Specifically, it is a function that assigns each point $a \in U$ an alternating k-tensor on the tangent space at $x \in M$, where $x = f(a)$.

In particular, we may write

$$\omega(a) = \sum_{i_1 < \cdots < i_k} \omega_{i_1,\ldots,i_k}(a) \cdot (\varphi_{i_1}(a) \wedge \cdots \wedge \varphi_{i_k}(a)),$$

where $\varphi_i(a)$ for $i = 1, \ldots, n$ is the dual basis w.r.t. to $(e_i)_a$. We say that ω is $\mathcal{C}^r$ if each $\omega_{i_1,\ldots,i_k}$ is also $\mathcal{C}^r$. Operations such as $\omega + \eta, f \cdot \omega, \omega \wedge \eta$ are defined in the usual way, and a function f is said to be a 0-form, where $f \cdot \omega$ is also written $f \wedge \omega$.

§54 Day 54: Differential Forms, Pt. 2 (Feb. 14, 2025)

Let $f : \mathbb{R}^n \to \mathbb{R}^p$, and consider the map $f_{*a} : \mathbb{R}^n_a \to \mathbb{R}^p_{f(a)}$. We define f_{*a} to be the map given by

$$f_{*a}(v_a) = (Df(a)(v))_{f(a)};$$

specifically, $f_{*a}(e_{i,a}) = (Df(a)(e_i))_{f(a)} = \sum_{j=1}^p \frac{\partial f_j}{\partial x_i}(a)(e_j)f(a)$. We can also write

$$f_{*a}\left(\left.\frac{\partial}{\partial x_i}\right|_a\right) = \sum_{j=1}^p \frac{\partial f_j}{\partial x_i}(a)\left.\frac{\partial}{\partial y_j}\right|_{f(a)},$$

or $f_{*a}(\frac{\partial}{\partial x_i}) = \sum_{j=1}^n \frac{\partial f_j}{\partial x_i}\frac{\partial}{\partial y_j}$. f_{*a} induces $f_a^* : \Omega^k(\mathbb{R}^p_{f(a)}) \to \Omega^k(\mathbb{R}^n_a)$, where

$$(f_a^*\omega)(v_1, \dots, v_k) = \omega(f_{*a}v_1, \dots, f_{*a}v_k),$$

where $v_i \in \mathbb{R}^n_a$. In particular, f^* takes k-forms on $\mathbb{R}^p$ and send them to k-forms on $\mathbb{R}^n$, i.e. $\omega \mapsto f^*\omega$. For example,

$$(f^*\omega)(a)(v_1, \dots, v_k) = \omega(f(a))(f_{*a}v_1, \dots, f_{*a}v_k).$$

A 0-form on $\mathbb{R}^p$ is a function g on $\mathbb{R}^p$, where $(f^*g)(x) = g(f(x))$, i.e. $f^*g = g \circ f$.

Proposition 54.1 (Spivak 4-8)**.** We have four properties:

(i) $f^*(dy_i) = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}dx_j = df_i = d(y_i \circ f)$,

(ii) $f^*(\omega + \eta) = f^*\omega + f^*\eta$,

(iii) $f^*(g\omega) = (g \circ f)f^*\omega$,

(iv) $f^*(\omega \wedge \eta) = f^*\omega \wedge f^*\eta$.

Proof. Directly write as follows,

$$\begin{aligned}(f^*dy_i)(a)(v_a) &= dy_i(f(a))(f_{*a}v_a) \\ &= dy_i(f(a))\left(\sum_{j=1}^n \frac{\partial f_1}{\partial x_j}(a)v_j, \dots, \sum_{j=1}^n \frac{\partial f_p}{\partial x_j}(a)v_j\right) \\ &= \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a)v_j = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a)dx_j(a)(v_a).\end{aligned}$$

For example,

$$f^*dg = f^*\left(\sum_{i=1}^n \frac{\partial g}{\partial y_i}dy_i\right) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f)(f^*(dy_i)) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f)\left(\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}dx_j\right),$$

which is equal to $d(g \circ f) = d(f^*g)$. ∎

Proposition 54.2 (Spivak 4-9)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be $\mathcal{C}^r$, and let $y = f(x)$. Then

$$f^*(g\,dy_1 \wedge \dots \wedge dy_n) = (g \circ f)\det g'\,dx_1 \wedge \dots \wedge dx_n.$$

§55 Day 56: Closed and Exact Forms (Feb. 26, 2025)

We say that a differential form is *closed* if $d\omega = 0$ (where ω is $\mathcal{C}^1$), and *exact* if $\omega = dy$ for some $\mathcal{C}^1$ y. In particular, since $d^2 = 0$ is exact, we have that every exact differential form is necessarily closed. For example, let $\omega = P(x,y)\,dx + Q(x,y)\,dy$ be a $\mathcal{C}^1$ 1-form on some open set $U \subset \mathbb{R}^2$. Then

$$d\omega = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy;$$

in the case that ω is closed, we have that $\partial_x Q = \partial_y P$. This is sufficient for ω exact if $U = \mathbb{R}^2$, or if it some open disk or rectangle. We may define

$$f(x,y) = \int_0^x P(t,0)\,dt + \int_0^y Q(x,t)\,dt.$$

Note that this looks just like the integrating factor as shown in MAT267! Note that, per construction of f, we indeed have

$$\begin{aligned}
\frac{\partial f}{\partial y}(x,t) &= Q(x,y), \\
\frac{\partial f}{\partial x}(x,y) &= P(x,0) + \int_0^y \frac{\partial Q}{\partial x}(x,t)\,dt \\
&= P(x,0) + \int_0^y \frac{\partial P}{\partial y}(x,t)\,dt \\
&= P(x,0) + P(x,y) - P(x,0) = P(x,y).
\end{aligned}$$

Specifically, this means we have $df = \omega$ (and so ω is exact). For another example, let us consider $U = \mathbb{R}^2 \setminus \{0\}$ to be the punctured plane[23], and let

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

We have that ω is closed (which we may indeed check by differentiating), but ω is not exact. Transforming to polar coordinates with $f(x,y) = (r\cos\theta, r\sin\theta) \in (0,\infty)\times[0,2\pi]$, we obtain that $\omega = d\theta$. However, suppose that $\omega = df$ where f is $\mathcal{C}^1$ on $U = \mathbb{R}^2 \setminus \{0\}$. Then

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial \theta}{\partial x}, \\
\frac{\partial f}{\partial y} &= \frac{\partial \theta}{\partial y},
\end{aligned}$$

meaning $f = \theta + C$, where C is some constant, meaning that θ cannot be extended to a continuous function on the punctured plane, and so such an f cannot exist. Thus, ω is not exact.

Theorem 55.1 (Spivak 4-11; Poincaré Lemma)**.** Let $A \subset \mathbb{R}^n$ be a strictly convex open set. Then every closed $\mathcal{C}^1$ form ω is exact.[24]

[23] this is exercise 4-21 in Spivak, also discussed page 93 as the classic example of if ω is defined on only a subset of $\mathbb{R}^2$, then such a function for $\omega = dy$ may not necessarily exist

[24] more generally, this is true for any star shaped convex set w.r.t. origin

Recall the fundamental theorem of calculus, where we have that

$$\int_a^b f(t)\,dt = f(b) - f(a).$$

We may rewrite the above as follows,

$$\int_{[a,b]} f = \int_{\partial[a,b]} f.$$

As a notational digression, more generally, let ω be a $\mathcal{C}^0$ 1-form on an open $U \subset \mathbb{R}^n$, and let $\gamma : [0,1] \to U$ be a $\mathcal{C}^1$ curve. Then

$$\omega = \sum_{i=1}^{n} \omega_i(x)\,dx_i$$

means we have that

$$\int_\gamma \omega := \int_\gamma \omega_1(x)dx_1 + \cdots + \omega_n(x)dx_n,$$

where we may substitute in $x = \gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$ to yield $dx_i = \gamma_i'(t)\,dt$ to obtain

$$\int_0^1 \big(\omega_1(\gamma(t))\gamma_1'(t) + \cdots + \omega_n(\gamma(t))\gamma_n'(t)\big)\ dt.$$

We now return to the lemma. If ω is exact, then $\omega = df$, where f is $\mathcal{C}^1$. Then

$$\int_\gamma \omega = \int_\gamma df = \int_{[0,1]} \gamma^*(df) = \int_{[0,1]} d(\gamma^* f) = \int_{[0,1]} d(f \circ \gamma) = f(\gamma(1)) - f(\gamma(0)) \overset{\text{def}}{=} \int_{\partial\gamma} f.$$

We complete the other case of the proof next lecture.