

MAT257 Final Portfolio

Ethan Hua

Contents

1	Big List Problems	4
1.1	#1	4
1.2	#2 - Incomplete	5
1.3	#3	6
1.4	#4	6
1.5	#5	8
1.6	#6	9
1.7	#7	12
1.8	#8	12
1.9	#9	13
1.10	#10	14
1.11	#11	15
1.12	#12*	17
1.13	#13 - Unfinished	18
1.14	#16*	19
1.15	#17*	21
1.16	#18	22
1.17	#19 - Incomplete	22
1.18	#20*	23
1.19	#22*	24
1.20	#23	26
1.21	#24	28
1.22	#25	29
1.23	#26	30
1.24	#27*	33
1.25	#28	35
1.26	#29	37
1.27	#30	41
1.28	#31	43
1.29	#32*	45
1.30	#33	47
1.31	#34*	49
1.32	#35*	51
1.33	#36	54
1.34	#37	56
1.35	#38	58
1.36	#39	62
1.37	#40	65
1.38	#41	67
1.39	#42*	69
1.40	#43	71
1.41	#44	72
2	Handout Exercises	72
2.1	Written by Me	72
2.1.1	Exercise 4.10.	72
2.1.2	Exercise 6.33.	73
2.1.3	Lemma 6.44	73
2.1.4	Exercise 7.16. Chain Rule	74
2.1.5	Exercise 8.12 (c) (iii)	75
2.1.6	Exercise 9.10	76
2.1.7	Exercise 11.8	77

2.1.8	Fubini's Theorem + Exercise 12.15	78
2.1.9	Exercise 13.14	79
2.1.10	Exercise 15.14	81
2.1.11	Exercise 16.7. Tangent space to a graph	82
2.1.12	Exercise 18.15	82
2.2	Not Written by Me	83
2.2.1	Exercise 0.6.	83
2.2.2	Exercise 2.21	84
2.2.3	Exercise 3.12.	85
2.2.4	Exercise 4.24.	85
2.2.5	Exercise 6.28	86
2.2.6	Exercise 9.21	86
2.2.7	14.8	86
3	Original Work	88
3.1	Antiderivatives	90
3.2	Precompactness	97
3.3	Generalizing the Inverse Function Theorem	98
4	Major Theorems	101
4.1	Sequential Compactness and Compactness	101
4.2	Equivalence of Norms in Finite Dimensional Vector Spaces	101
4.3	Differentiability Theorem	102
4.4	Chain Rule	103
4.5	Inverse Function Theorem and Implicit Function Theorem	104
4.6	Condition for Local Minimum	106
4.7	Change of Variables	108

1 Big List Problems

Problem Number	Completed?
#1	Complete
#2	Incomplete
#3	Complete
#4	Complete
#5	Complete
#6	Complete
#7	Complete
#8	Complete
#9	Complete
#10	Complete
#11	Complete
#12	Complete (Specially Selected)
#13	Unfinished
#14	Unattempted
#15	Unattempted
#16	Complete (Specially Selected)
#17	Complete (Specially Selected)
#18	Complete
#19	Incomplete
#20	Complete (Specially Selected)
#21	Unattempted
#22	Complete (Specially Selected)
#23	Complete
#24	Complete
#25	Complete
#26	Complete
#27	Complete (Specially Selected)
#28	Complete
#29	Complete
#30	Complete
#31	Complete
#32	Complete (Specially Selected)
#33	Complete
#34	Complete (Specially Selected)
#35	Complete (Specially Selected)
#36	Complete
#37	Complete
#38	Complete
#39	Complete
#40	Complete
#41	Complete
#42	Complete (Specially Selected)
#43	NA
#44	NA

1.1 #1

Let A_1, A_2, A_3, \dots be a sequence of countable sets. Prove that $\bigcup_{i \geq 1} A_i$ is countable.

My previous submission did not take into account the fact that the sets could have been not disjoint. This resubmission should resolve that issue.

Changelog:

1. Fixed the argument to address the non-disjoint case.

Proof. Notice that

$$\bigcup_{i \geq 1} A_i = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cap A_2)) \cup \dots$$

We define a new collection of sets $\{B_i\}$ as follows:

$$B_1 = A_1$$
$$B_k = A_k \setminus \left(\bigcap_{i=1}^{k-1} A_i \right), k > 1$$

Since A_i is countable, denote A_{ij} to be the j th element of set A_i . Noting that $\bigcup_{i \geq 1} A_i = \bigcup_{i \geq 1} B_i$, Define $f : \bigcup_{i \geq 1} B_i \rightarrow \mathbb{N}$ as

$$f(A_{ij}) = 2^i 3^j$$

We will show that f is injective. Let $A_{pq}, A_{rs} \in \bigcup_{i \geq 1} B_i$. Suppose that $f(A_{pq}) = f(A_{rs})$. Then

$$2^p 3^q = 2^r 3^s$$

Since every integer has a unique prime factorisation, it follows that $p = r, q = s$. Thus $A_{pq} = A_{rs}$.

Now, define the injection $g : \mathbb{N} \rightarrow \bigcup_{i \geq 1} A_i$ to be $g(n) = A_{1n}$.

Therefore by the Schröder–Bernstein theorem, $|\bigcup_{i \geq 1} A_i| = |\mathbb{N}|$. Thus the set is countable. □

1.2 #2 - Incomplete

This problem is incomplete for a variety of reasons. Firstly, I did not show that ϕ is a linear mapping. Secondly, ϕ is not well-defined because based on the supposed codomain of ϕ , $\phi(T)$ should only take in one input, whereas I defined $\phi(T)$ by inputting two inputs. Finally, the basis argument I made when showing that ϕ is surjective was unnecessary, and I could have just directly defined $T(x, y) = U(x)(y)$.

Let X, Y, Z be three vector spaces. Prove that $L^2(X, Y; Z)$ is isomorphic to $L(X, L(Y, Z))$.

(For two vector spaces X, Y , we use $L(X, Y)$ to denote the space of linear mappings from X to Y . For three vector spaces X, Y, Z , and a function $\beta : X \times Y \rightarrow Z$, we let $\beta(\cdot, y_0)$ denote the function $X \rightarrow Z$ given by $x \mapsto \beta(x, y_0)$, where $y_0 \in Y$ is some fixed vector; this is called the y_0 -**slice** of β . The x_0 -**slice** $\beta(x_0, \cdot)$ is defined similarly.)

Proof. For a bilinear map $T \in L^2(X, Y; Z)$, $(x, y) \in X \times Y$, Define $\phi : L^2(X, Y; Z) \rightarrow L(X, L(Y, Z))$ such that

$$\phi(T)(x, y) = T(x, \cdot)(y)$$

where $T(x, \cdot)$ is the x -slice of T . We claim that this transformation is an isomorphism.

First, let $\phi(T) = 0$. Then $\forall x \in X, T(x, \cdot)(y) = 0$, from which it follows that $T(x, y) = 0$, meaning ϕ is injective.

Next, fix $U \in L(X, L(Y, Z))$. Let β be the basis for X . Let $T \in L^2(X, Y; Z)$ be the linear transformation such that $T(x, \cdot) = U(x)$, for all $x \in \beta$. We see that $\forall x, y \in X \times Y$,

$$\phi(T)(x, y) = T(x, \cdot)(y) = U(x)(y)$$

making ϕ surjective.

Thus ϕ is an isomorphism, and we get that $L^2(X, Y; Z) \cong L(X, L(Y, Z))$ as desired. □

1.3 #3

Let $I = (a, b)$ and $J = (c, d)$ be two open intervals on the real line. Let $f : I \rightarrow J$ be an increasing function such that $f(I)$ is dense in J . Prove that f is continuous.

(For two sets $D, S \subseteq \mathbf{R}$ we say that D is **dense** in S if $D \cap (s - \varepsilon, s + \varepsilon) \neq \emptyset$ for all $s \in S$ and all $\varepsilon > 0$.)

Proof. Fixing an $a \in I$, let $\varepsilon > 0$. We can assume without loss of generality that ε is small enough that $(f(a) - \varepsilon, f(a) + \varepsilon) \subseteq J$. Since $f(I)$ is dense in J , we can always find an $y_1 \in (f(a) - \varepsilon, f(a))$ and $y_2 \in (f(a), f(a) + \varepsilon)$ such that $y_1, y_2 \in f(I)$, which means $y_1 = f(x_1)$ and $y_2 = f(x_2)$ for some $x_1, x_2 \in I$.

Take $\delta = \min\{|a - x_1|, |a - x_2|\}$. Let $x \in I$. Suppose that $|x - a| \leq \delta$. If $x = a$, clearly $|f(x) - f(a)| < \varepsilon$. Consider when $x < a$.

We see that due to the choice of δ , we have that $x_1 < x < a$. Using the fact that f is increasing, we obtain

$$f(a) - \varepsilon < y_1 = f(x_1) < f(x) < f(a) \implies -\varepsilon < f(x) - f(a) < 0 \implies f(a) - f(x) = |f(x) - f(a)| < \varepsilon$$

The argument for the case when $x < a$ is almost the exact same, except for the use of x_2 and y_2 instead of x_1 and y_1 , as well as the inequalities being swapped.

With this, we can conclude that f is continuous. □

1.4 #4

- (a) Prove that there exists an infinitely differentiable function $\alpha : \mathbf{R} \rightarrow \mathbf{R}$ such that $\alpha(t) = 0$ for all $t \leq 0$, and $\alpha(t) > 0$ for all $t > 0$.

Proof. We define $\alpha(t) = \begin{cases} 0, & \text{if } t \leq 0; \\ e^{-\frac{1}{t}}, & \text{if } t > 0. \end{cases}$

Trivially $\alpha(t) = 0$ if $t \leq 0$ and $\alpha(t) > 0$ if $t > 0$. It remains to show that α is infinitely differentiable.

Since 0 is infinitely differentiable, and $e^{-\frac{1}{x}}$ is infinitely differentiable for $x > 0$, it suffices to show that derivatives of all orders of α are continuous at $t = 0$.

We will continue by proving a lemma.

Lemma. $\forall n \in \mathbb{N}$, for $t > 0$, $\alpha^{(n)}(t) = Q(t)\alpha(t)$, where $Q(t)$ is a linear combination of nonpositive integer powers of t .

We will show this using induction.

Proof. When $n = 1$, fixing $t > 0$, we have

$$\alpha'(t) = \frac{1}{t^2} e^{-\frac{1}{t}} = \frac{1}{t^2} \alpha(t)$$

We let $Q(t) = t^{-2}$ and we are done.

Now suppose that the lemma holds for all $i \leq k$, for some $k \in \mathbb{N}$.

Then for $t > 0$,

$$\alpha^{(k)}(t) = P(t)\alpha(t), \text{ where } P(t) \text{ is a linear combination of nonpositive integer powers of } t$$

Taking the derivative of both sides with respect to t , we obtain

$$\alpha^{(k+1)}(t) = P'(t)\alpha(t) + P(t)\alpha'(t)$$

The lemma holds for $i = 1$, therefore for some $R(t)$,

$$P'(t)\alpha(t) + P(t)\alpha'(t) = P'(t)\alpha(t) + P(t)R(t)\alpha(t) = (P'(t) + P(t)R(t))\alpha(t)$$

$P'(t) + P(t)R(t)$ is a linear combination of nonpositive integer powers of t , therefore letting $Q(t) = P'(t) + P(t)R(t)$, we get our desired conclusion.

By the principle of induction, the lemma holds true for all $n \in \mathbb{N}$. □

We now continue in proving that derivatives of all orders of α are continuous at 0. We show this by proving that

$$\lim_{t \rightarrow 0} \alpha^{(n)}(t) = 0, \text{ where } n \in \mathbb{N}$$

We will only worry about the right hand limit, as the left hand limit always evaluates to 0.

Let $n \in \mathbb{N}$. Since we only deal with positive t , by our lemma,

$$\alpha^{(n)}(t) = Q(t)\alpha(t)$$

Where $Q(t)$ is a linear combination of nonpositive integer powers of t . So

$$\lim_{t \rightarrow 0^+} \alpha^{(n)}(t) = \lim_{t \rightarrow 0^+} Q(t)\alpha(t) = \lim_{t \rightarrow 0^+} \sum_{i=0}^k a_i t^{-i} \alpha(t)$$

Where a_i are real constants and $k \in \mathbb{N}$.

Consider an arbitrary term $a_i t^{-i} \alpha(t) = a_i t^{-i} e^{\frac{-1}{t}}$. We want to show that $\lim_{t \rightarrow 0^+} a_i t^{-i} e^{\frac{-1}{t}}$ exists and is equal to 0.

First, we will perform the substitution $x = \frac{1}{t}$. Then the limit becomes

$$\lim_{x \rightarrow \infty} a_i x^i e^{-x} = 0$$

The proof for this fact is omitted, but applying L'Hopital's rule i times produces the same result. Thus,

$$\lim_{t \rightarrow 0^+} \alpha^{(n)}(t) = \lim_{t \rightarrow 0^+} \sum_{i=0}^k a_i t^{-i} \alpha(t) = \sum_{i=0}^k \lim_{t \rightarrow 0^+} a_i t^{-i} \alpha(t) = \sum_{i=0}^k 0 = 0$$

Thus α is infinitely differentiable everywhere. □

- (b) Prove that there exists an infinitely differentiable function $\beta : \mathbf{R} \rightarrow \mathbf{R}$ such that $\beta(t) = 1$ for all $t \geq 1$, and $\beta(t) = 0$ for all $t \leq 0$.

Proof. Define

$$\beta(t) = \frac{\alpha(t)}{\alpha(t) + \alpha(1-t)}$$

If $t \geq 1$, we also have $1-t \leq 0$. Then

$$\beta(t) = \frac{\alpha(t)}{\alpha(t)} = 1$$

As well, if $t \leq 0$, we have

$$\beta(t) = 0$$

Since $\alpha(t)$ is infinitely differentiable and β is composed of α , it follows that β is also infinitely differentiable. □

- (c) Prove that there exists an infinitely differentiable function $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ such that $\varphi(t) = 1$ for all $t \in [2, 3]$, and $\varphi(t) = 0$ for $t \in \mathbf{R} \setminus (1, 4)$.

Proof. Define $\phi(t) = \beta(t-1)\beta(4-t)$. Since β is infinitely differentiable, ϕ is as well. For $t \in [2, 3]$, $t-1 \geq 1$ and $4-t \geq 1$. Thus

$$\phi(t) = \beta(t-1)\beta(4-t) = 1$$

If $t \in \mathbb{R} \setminus (1, 4)$, then $t-1 \leq 0$ or $4-t \leq 0$. In each case $\beta(t-1) = 0$ or $\beta(4-t) = 0$, respectively, thus we have

$$\phi(t) = \beta(t-1)\beta(4-t) = 0$$

□

1.5 #5

Let $S \subseteq \mathbf{R}^n$. Consider the following three statements:

- S is a bounded subset of $(\mathbf{R}^n, \|\cdot\|_1)$.
- S is a bounded subset of $(\mathbf{R}^n, \|\cdot\|_2)$.
- S is a bounded subset of $(\mathbf{R}^n, \|\cdot\|_{\max})$.

Among these statements, determine which implications are true and which are false. There are six implications to investigate. Supply proof or counterexample as appropriate. Include pictures.

Note. My previous submission did not consider when $d(x, y) = 0$ when proving symmetry. As well, I made a typo when writing the max-norm. Finally, the last paragraph was not written out mathematically. This resubmission should fix those issues.

Changelog.

1. Added an argument when proving symmetry to address division by 0
2. Fixed max-norm typo
3. Rewrote last explanation in a more mathematical fashion

Claim. Each statement implies all the other statements.

Proof. Lemma. Strong equivalence is an equivalence relation.

Proof. Let d_1, d_2, d_3 be metrics on some set X . We proceed by proving each property of equivalence relations.

1. Reflexivity: Let $\alpha = \beta = 1$. Clearly $d_1(x, y) \leq d_1(x, y) \leq d_1(x, y)$. Thus the relation is reflexive.
2. Symmetry: Suppose $\hat{\alpha}d_1(x, y) \leq d_2(x, y) \leq \hat{\beta}d_1(x, y)$. Let $\alpha = \frac{1}{\hat{\beta}}, \beta = \frac{1}{\hat{\alpha}}$. If $x = y$, then $d_1(x, y) = d_2(x, y) = 0$, so the inequality trivially holds. Otherwise, $d_1(x, y) > 0, d_2(x, y) > 0$, so we see that

$$\hat{\alpha} \leq \frac{d_2(x, y)}{d_1(x, y)} \leq \hat{\beta} \implies \frac{1}{\hat{\beta}} \leq \frac{d_1(x, y)}{d_2(x, y)} \leq \frac{1}{\hat{\alpha}} \implies \alpha d_2(x, y) \leq d_1(x, y) \leq \beta d_2(x, y)$$

Thus the relation is symmetric.

3. Transitivity: Suppose that $\hat{\alpha}d_1(x, y) \leq d_2(x, y) \leq \hat{\beta}d_1(x, y)$ and $\tilde{\alpha}d_2(x, y) \leq d_3(x, y) \leq \tilde{\beta}d_2(x, y)$. Let $\alpha = \hat{\alpha}\tilde{\alpha}, \beta = \tilde{\beta}\hat{\beta}$. We have

$$\begin{aligned} \alpha d_1(x, y) &= \hat{\alpha}\tilde{\alpha} \leq \tilde{\alpha}d_2(x, y) \leq d_3(x, y) \leq \tilde{\beta}d_2(x, y) \leq \tilde{\beta}\hat{\beta}d_1(x, y) = \beta d_1(x, y) \\ &\implies \alpha d_1(x, y) \leq d_3(x, y) \leq \beta d_1(x, y) \end{aligned}$$

Thus the relation is transitive.

Therefore, the relation is an equivalence relation.

□

We want to show that $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_{\max}$ are strongly equivalent to each other. First, we will show that $\|\cdot\|_1 \sim \|\cdot\|_{\max}$ and $\|\cdot\|_2 \sim \|\cdot\|_{\max}$. We have that

$$\frac{1}{n} \sqrt{\sum_{i=1}^n |x_i - y_i|} \leq \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \leq \max_{1 \leq i \leq n} \{|x_i - y_i|\} \leq \sqrt{\sum_{i=1}^n |x_i - y_i|} \leq \sum_{i=1}^n |x_i - y_i|$$

From this, we have the inequalities

$$\begin{aligned} \frac{1}{n} \sqrt{\sum_{i=1}^n |x_i - y_i|} &\leq \max_{1 \leq i \leq n} \{|x_i - y_i|\} \leq \sqrt{\sum_{i=1}^n |x_i - y_i|} \text{ and } \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \leq \max_{1 \leq i \leq n} \{|x_i - y_i|\} \leq \sum_{i=1}^n |x_i - y_i| \\ \implies \frac{1}{n} \|x - y\|_2 &\leq \|x - y\|_{\max} \leq \|x - y\|_2 \text{ and } \frac{1}{n} \|x - y\|_1 \leq \|x - y\|_{\max} \leq \|x - y\|_1 \\ \implies \|\cdot\|_2 &\sim \|\cdot\|_{\max} \text{ and } \|\cdot\|_1 \sim \|\cdot\|_{\max} \end{aligned}$$

Using transitivity and symmetry we can also conclude that $\|\cdot\|_1 \sim \|\cdot\|_2$. Thus every metric is strongly equivalent to one another. It follows that for any $i, j \in \{1, 2, \max\}$, there exists $c_{ij} > 0$ such that for all $x, y \in \mathbb{R}^n$, $d_i(x, y) \leq c_{ij} \cdot d_j(x, y)$, meaning that for $S \subseteq \mathbb{R}^n$, if S is bounded with respect to the 1-norm, 2-norm, or the max-norm, it is bounded with respect to the other norms as well, from which our claim follows immediately after. \square

1.6 #6

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed vector spaces. A linear mapping $T : X \rightarrow Y$ is called **bounded** if there exists a constant $M \geq 0$ such that

$$\|T(x)\|_Y \leq M \|x\|_X \quad \text{for all } x \in X.$$

Let $B(X, Y)$ denote the set of these bounded linear operators. The **operator norm** on $B(X, Y)$, denoted by $\|\cdot\|_{\text{op}}$, is defined as follows:

$$\|T\|_{\text{op}} = \sup\{\|T(x)\|_Y : x \in X \text{ and } \|x\|_X \leq 1\}.$$

(a) Prove that $B(X, Y)$ is a linear subspace of $L(X, Y)$.

Proof. The 0-transformation $\|Z(x)\| = 0 \leq \|x\|_X$, $\forall x \in X$, because of the definition of a norm. Thus $0 \in B(X, Y)$. Let $T, U \in B(X, Y)$, $c \in \mathbb{R}$. Then for all $x \in X$, there exist $M, N \geq 0$ such that

$$T(x) \leq M \|x\|_X \text{ and } U(x) \leq N \|x\|_X \implies T(x) + U(x) \leq (M + N) \|x\|_X$$

Which implies that $T + U$ is a member of $B(X, Y)$. As well, from the first inequality, depending on if c is positive or negative, we have

$$T(x) \leq M \|x\|_X \implies cT(x) \leq cM \|x\|_X \text{ or } -cT(x) \leq -cM \|x\|_X$$

Note that if $c = 0$, $cT = 0$. Regardless, this implies that cT is a member of $B(X, Y)$.

Since $0 \in B(X, Y)$ and $B(X, Y)$ is closed under addition and scalar multiplication, $B(X, Y)$ is a subspace of $L(X, Y)$. \square

(b) Prove that $\|\cdot\|_{\text{op}}$ is a norm on $B(X, Y)$.

Proof. To prove that the operator norm is a norm, we first verify that $\|T\|_{op} = 0 \iff T = 0$.

We denote the set of elements x in X such that $\|x\|_X \leq 1$ as X' .

Fix $T \in B(X, Y)$ and suppose that $T = 0$. Then for all $x \in X'$, $\|T(x)\|_Y = \|0\|_Y = 0$. Thus $\|T\|_{op} = 0$.

Now suppose the converse, that $\|T\|_{op} = 0$. Then

$$\forall x \in X', \|T(x)\|_Y \leq \|T\|_{op} = 0.$$

But by the definition of the norm in Y ,

$$0 \leq \|T(x)\|_Y$$

It follows that $\|T(x)\| = 0 \implies T(x) = 0$.

To show nonnegativity, we note that for $x \in X'$,

$$\|T\|_{op} \geq \|T(x)\|_Y \geq 0$$

To show homogeneity, let $T \in G(X, Y)$, $c \in \mathbb{R}$. Then

$$\|cT\|_{op} = \sup\{\|cT(x)\|_Y : x \in X'\} = \sup\{|c| \|T(x)\|_Y : x \in X'\} = |c| \sup\{\|T(x)\|_Y : x \in X'\} = |c| \|T\|_{op}$$

Now we show that the triangle inequality holds with respect to the operator norm.

Fix $T, U \in B(X, Y)$. Let $x \in X'$. By definition,

$$T(x) \leq \sup T(X') \text{ and } U(x) \leq \sup U(X')$$

Adding both together obtains

$$T(x) + U(x) \leq \sup T(X') + \sup U(X')$$

We see that $\sup T(X') + \sup U(X')$ is an upper bound for $T(x) + U(x)$. By the definition of the least upper bound,

$$\sup\{T(X') + U(X')\} \leq \sup T(X') + \sup U(X') \implies \|T + U\|_{op} \leq \|T\|_{op} + \|U\|_{op}$$

Thus the operator norm is, indeed, a norm. □

- (c) Let $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ be the linear mapping given by $T(x, y) = (x + y, x)$. Find, with proof, the exact value of $\|T\|_{op}$. (Here, \mathbf{R}^2 is equipped with the usual norm.)

Claim. $\|T\|_{op} = \sqrt{\frac{3+\sqrt{5}}{2}}$

Proof. We will show that $\|T(x, y)\|_2$ is bounded above by this value, and that equality is possible.

Let $(x, y) \in \mathbb{R}^2$ such that $\sqrt{x^2 + y^2} \leq 1 \implies y \leq \pm\sqrt{1 - x^2} \leq \sqrt{1 - x^2}$. For such (x, y) ,

$$\|T(x, y)\|_2 = \|(x + y, x)\|_2 = \sqrt{(x + y)^2 + x^2} = \sqrt{2x^2 + 2xy + y^2}$$

By the monotonicity of the squareroot,

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

We attempt to maximize this function for $x \in [0, 1]$ (The interval is the set of all x that satisfy the constraint $x^2 + y^2 \leq 1$). Maximizing this function is synonymous to maximizing $f(x) = x^2 + 2x\sqrt{1 - x^2} + 1$ on the interval $[0, 1]$. Taking the derivative, we get

$$f'(x) = 2x + 2\sqrt{1 - x^2} - \frac{2x^2}{\sqrt{1 - x^2}} = \frac{2x\sqrt{1 - x^2} + 2 - 4x^2}{\sqrt{1 - x^2}}$$

Now, we find every critical point. When f' is undefined, $x = 1$. Now, let $f'(x) = 0$, $x \neq 1$. Then through a series of calculations I really don't want to type out,

$$\frac{2x\sqrt{1-x^2} + 2 - 4x^2}{\sqrt{1-x^2}} = 0 \implies 5x^4 - 5x^2 + 1 = 0 \implies x^2 = \frac{1}{2} \pm \frac{1}{2\sqrt{5}} \implies x = \sqrt{\frac{1}{2} \pm \frac{1}{2\sqrt{5}}}$$

We disregard the negative solution since we want $x \in [0, 1]$. Now we evaluate f at the endpoints, as well as at every point we found:

$$f(0) = 1$$

$$f(1) = 2$$

$$f\left(\sqrt{\frac{1}{2} + \frac{1}{2\sqrt{5}}}\right) = \frac{3 + \sqrt{5}}{2}$$

$$f\left(\sqrt{\frac{1}{2} - \frac{1}{2\sqrt{5}}}\right) = \frac{3}{2} + \frac{3}{2\sqrt{5}}$$

It is not too hard to see that f achieves the maximum at $x = \sqrt{\frac{1}{2} + \frac{1}{2\sqrt{5}}}$. Then $\sqrt{x^2 + 2x\sqrt{1-x^2} + 1}$ also achieves a maximum at $x = \sqrt{\frac{1}{2} + \frac{1}{2\sqrt{5}}}$, which is $\sqrt{\frac{3+\sqrt{5}}{2}}$.

In summary, we have for all $(x, y) \in \mathbb{R}^2$,

$$\|T(x, y)\|_2 \leq \sqrt{\frac{3 + \sqrt{5}}{2}}$$

Thus $\sqrt{\frac{3+\sqrt{5}}{2}}$ is an upper bound for $\|T(x, y)\|_2$.

To show that $\sqrt{\frac{3+\sqrt{5}}{2}}$ is the least upper bound, it suffices to show that $\|T(x, y)\|_2$ can achieve that value. Indeed, if we let $x = \sqrt{\left(\frac{1}{2} + \frac{1}{2\sqrt{5}}\right)}$, $y = \sqrt{\left(\frac{1}{2} - \frac{1}{2\sqrt{5}}\right)}$ we see that

$$\begin{aligned} \|T(x, y)\| &= \sqrt{\left(\sqrt{\left(\frac{1}{2} + \frac{1}{2\sqrt{5}}\right)} + \sqrt{\left(\frac{1}{2} - \frac{1}{2\sqrt{5}}\right)}\right)^2 + \left(\sqrt{\left(\frac{1}{2} + \frac{1}{2\sqrt{5}}\right)}\right)^2} \\ &= \sqrt{2\left(\frac{1}{2} + \frac{1}{2\sqrt{5}}\right) + 2\sqrt{\left(\frac{1}{2} + \frac{1}{2\sqrt{5}}\right)}\sqrt{\left(\frac{1}{2} - \frac{1}{2\sqrt{5}}\right)} + \left(\frac{1}{2} - \frac{1}{2\sqrt{5}}\right)} = \sqrt{\frac{3 + \sqrt{5}}{2}} \end{aligned}$$

Thus $\|T\|_{op} = \sup\{\|T(x, y)\|_2 : \|(x, y)\|_2 \leq 1\} = \sqrt{\frac{3+\sqrt{5}}{2}}$

□

(d) Find, with proof, an example of an unbounded linear operator.

Proof. Define ℓ^0 to be the set of all sequences that are eventually 0. Consider the metric spaces $(\ell^0, \|\cdot\|_{\ell^\infty})$ and $(C[0, 1], \|\cdot\|_{C^\infty})$. Here, we denote $\|\cdot\|_{\ell^\infty}$ as the sup norm on ℓ^∞ and $\|\cdot\|_{C^\infty}$ as the sup norm on $C[0, 1]$.

Let $T : \ell^0 \rightarrow C[0, 1]$ be defined by

$$T((a_n)_n) = \sum_{i=0}^k a_i i^x, \text{ where } k \text{ is the last index where } a_k \neq 0$$

First, we will show that T is a linear transformation. Fix $(a_n)_n, (b_n)_n \in \ell^0, c \in \mathbb{R}$. Let $k = \max\{k_a, k_b\}$, where k_a, k_b are the last index where a_{k_a} and b_{k_b} are non-zero, respectively. Then

$$T(c(a_n) + (b_n)) = \sum_{i=0}^k (ca_i + b_i)i^x = c \sum_{i=0}^k a_i i^x + \sum_{i=0}^k b_i i^x = c \sum_{i=0}^{k_a} a_i i^x + \sum_{i=0}^{k_b} b_i i^x = cT((a_n)) + T((b_n))$$

This verifies that T is a linear transformation.

Now we show that T is unbounded. Fix $M \geq 0$. Let $(a_n)_n \in \ell^0$ such that $a_i = 1$ if $i = M + 1$ and 0 otherwise. We have that

$$\|T((a_n))\|_{C^\infty} = \|(M + 1)^x\|_{C^\infty} = M + 1 > M = M \| (a_n) \|_{\ell^\infty}$$

Thus T is an unbounded linear operator. □

1.7 #7

Let $S \subseteq C[0, 1]$. Consider the following two statements:

- S is an open subset of $(C[0, 1], \|\cdot\|_1)$.
- S is an open subset of $(C[0, 1], \|\cdot\|_\infty)$.

Determine if the first statement implies the second, and vice-versa. Supply proof or counterexample as appropriate.

Proof. We claim that the first statement implies the second, but not the converse.

Suppose that S is an open subset of $(C[0, 1], \|\cdot\|_1)$. For any $g \in S$, there is an open ball with respect to the 1-norm centered around g with radius ε such that $B_1(g, \varepsilon) \subseteq S$. We proceed to show that $B_\infty(g, \varepsilon) \subseteq B_1(g, \varepsilon)$. Let $f \in B_\infty(g, \varepsilon)$. Then $\|f - g\|_\infty < \varepsilon$. Thus we have

$$\|f - g\|_1 = \int_0^1 |f - g| \leq \int_0^1 \sup\{|f - g|\} = \|f - g\|_\infty < \varepsilon$$

Thus $B_\infty(g, \varepsilon) \subseteq B_1(g, \varepsilon) \subseteq S$. Thus S is an open subset of $(C[0, 1], \|\cdot\|_1)$.

Now we show that the converse is not necessarily true. Let $S = B_\infty(0, 1)$. This is an open subset of $(C[0, 1], \|\cdot\|_\infty)$. Consider $f(x) = 0 \in B_\infty(0, 1)$. For every $\varepsilon > 0$, we can always find $n \in \mathbb{N}$ such that $n > \frac{1}{\varepsilon}$. Let $g(x) = x^{n-1}$. Since $\int_0^1 g(x)dx = \frac{1}{n} < \varepsilon$, $g(x) \in B_1(0, \varepsilon)$. But $g(1) = 1$, which means that $g(x) \notin B_\infty(0, 1)$. Thus f is not an interior point of $B_\infty(0, 1)$ with respect to the 1-norm, which means that $B_\infty(0, 1)$ is not an open subset of $(C[0, 1], \|\cdot\|_1)$. □

1.8 #8

Let X be any set. The **diagonal** of $X \times X$ is the following set:

$$\Delta = \{(x, x) : x \in X\}.$$

Prove that if (X, d) is a metric space, then Δ is a closed subset of $X \times X$ (with respect to the product metric).

Proof. We define d as the metric in X and d_X to be the product metric of $X \times X$. Suppose that (X, d) is a metric space. We want to show that every limit point of Δ is a member of Δ .

Let $(a, b) \in X \times X$ be a limit point of Δ and suppose for contradiction that (a, b) is not an element of X . This implies that $d(a, b) > 0$. Let $r = d(a, b)$, and consider the open ball $B((a, b), r)$. For all $(x, x) \in X$, by the triangle inequality,

$$d_X((a, b), (x, x)) = d(a, x) + d(b, x) \geq r$$

This means that no element lies within this open ball centered around (a, b) , contradicting the fact that (a, b) is a limit point of X . Thus $(a, b) \in X$. □

1.9 #9

- (a) Let $C^\infty[0, 1]$ denote the set of infinitely differentiable functions $f : [0, 1] \rightarrow \mathbf{R}$. Prove that $C^\infty[0, 1]$ is *not* a closed subset of $(C[0, 1], \|\cdot\|_\infty)$.

Proof. To show that this set is not closed, we just need to find a limit point that is not an element of the set. Let $f(x) = |x - \frac{1}{2}|$. Clearly $f \notin C^\infty[0, 1]$. To prove f is a limit point of $C^\infty[0, 1]$, first let $\varepsilon > 0$ and consider the open ball $B(f, \varepsilon)$. There exists an $n \in \mathbb{N}$ such that $n > \frac{1}{\varepsilon^2} \implies \frac{1}{\sqrt{n}} < \varepsilon$.

Let $g_n(x) = \sqrt{(x - \frac{1}{2})^2 + \frac{1}{n}}$. Since $x^2 + \frac{1}{n}$ is always positive and the squareroot function is infinitely differentiable for all positive real numbers, g_n is infinitely differentiable for all $x \in [0, 1]$. Our goal is to find an upper bound for $|f - g_n|$. For convenience, we perform the change of variables $t = x - \frac{1}{2}$. Define a new function in $C[-\frac{1}{2}, \frac{1}{2}]$ as $h(t) = f(t) - g_n(t)$. We find the global maximum and minimum of h . Taking the derivative with respect to t , we obtain

$$h'(t) = \frac{|t|}{t} - \frac{t}{\sqrt{t^2 + \frac{1}{n}}}$$

We see that h has a critical point at $t = 0$ since $h'(0)$ is undefined. Note that $\pm t \leq |t| \leq \sqrt{t^2 + \frac{1}{n}} \implies \sqrt{t^2 + \frac{1}{n}} \pm t > 0$. If $t > 0$,

$$h'(t) = 1 - \frac{t}{\sqrt{t^2 + \frac{1}{n}}} = \frac{\sqrt{t^2 + \frac{1}{n}} - t}{\sqrt{t^2 + \frac{1}{n}}} > 0$$

If $t < 0$,

$$h'(t) = -1 - \frac{t}{\sqrt{t^2 + \frac{1}{n}}} = \frac{-\sqrt{t^2 + \frac{1}{n}} - t}{\sqrt{t^2 + \frac{1}{n}}} < 0$$

Therefore h achieves a local minimum at $t = 0$. We evaluate h at its critical points and endpoints:

$$h\left(-\frac{1}{2}\right) = \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{n}}$$

$$h(0) = -\frac{1}{\sqrt{n}}$$

$$h\left(\frac{1}{2}\right) = \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{n}}$$

By the monotonicity of the squareroot, $-\frac{1}{\sqrt{n}} < \sqrt{\frac{1}{4} + \frac{1}{n}} < \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{n}} < 0$. Thus $-\frac{1}{\sqrt{n}}$ is the global minimum of h and $\frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{n}}$ is the global maximum of h . Note that $|h(0)| > |h(\frac{1}{2})|$. Thus

$$\|f - g_n\|_\infty = \sup_{0 \leq x \leq 1} \{|f(x) - g_n(x)|\} = \sup_{-\frac{1}{2} \leq t \leq \frac{1}{2}} \{|h(t)|\} = \frac{1}{\sqrt{n}} < \varepsilon$$

Thus f is a limit point of C^∞ , but is not a member of the set. Thus C^∞ is not closed. □

- (b) Let C be the set of **convergent** sequences of real numbers. Prove that C is a closed subset of $(\ell^\infty, \|\cdot\|_\infty)$.

Proof. Let $(a_n)_{n \geq 1}$ be a limit point of C . We want to show (a_n) is convergent. We will do this by showing that (a_n) is a Cauchy sequence.

Let $\varepsilon > 0$. Using the fact that (a_n) is a limit point of C , there exists a convergent sequence $b_n \in B((a_n), \frac{\varepsilon}{3}) \implies \forall k \in \mathbb{N}, |a_k - b_k| < \|(b_n) - (a_n)\|_\infty < \frac{\varepsilon}{3}$. Since (b_n) converges, there is a $N' \in \mathbb{N}$ such that $m, n \geq N \implies |b_m - b_n| < \frac{\varepsilon}{3}$.

Let $N = N'$. Then for all $m, n \geq N$, we have

$$|a_m - a_n| = |(a_m - b_m) + (b_m - b_n) + (b_n - a_n)| \leq |a_m - b_m| + |b_m - b_n| + |b_n - a_n| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

Thus (a_n) is a Cauchy sequence, and therefore converges, which means that $(a_n) \in C$. Thus C is closed. □

- (c) Let $(X, \|\cdot\|)$ be a normed vector space, and let M be a **linear subspace** of X . Prove that M is an open set if and only if $M = X$.

Proof. First, suppose that $M = X$. Then for every $x \in X$, any open ball centered around x is obviously a subset of X . Thus M is open.

Next, suppose that M is an open. To show that $X = M$, it suffices to show that $X \subseteq M$, since we already have $M \subseteq X$ from M being a subspace of X .

Since M is a subspace, we know that $0 \in M$. Thus there exists $\varepsilon > 0$ such that $B(0, \varepsilon) \subseteq M$.

For every $\vec{x} \in X$, $\left\| \frac{\varepsilon \vec{x}}{2\|\vec{x}\|} \right\| = \frac{\varepsilon \|\vec{x}\|}{2\|\vec{x}\|} = \frac{\varepsilon}{2} < \varepsilon$, which means that $\frac{\varepsilon \vec{x}}{2\|\vec{x}\|} \in M$. Since M is a subspace, it follows that $\vec{x} \in M \implies X \subseteq M$. Thus $M = X$ and we are done. □

1.10 #10

- (a) Prove that every bounded sequence in $(\mathbb{R}^d, \|\cdot\|_2)$ has a convergent subsequence.

Proof. We take the Bolzano-Weierstrass Theorem in \mathbb{R} for granted and use this to prove it for \mathbb{R}^d . We will do this using induction on d . When $d = 1$, it follows trivially from the theorem in \mathbb{R} .

Now suppose that the claim is true for some $d = k - 1$, for some $k \in \mathbb{N}$. Let $(a_n)_{n \geq 1}$ be a bounded sequence in $(\mathbb{R}^k, \|\cdot\|_2)$. Define another sequence $(b_n)_{n \geq 1}$ in \mathbb{R}^{k-1} such that b_i is the first $k - 1$ components of a_i . Since (a_n) is bounded, (b_n) is also bounded. From our assumption, b_i has a convergent subsequence $(b_{n_i})_{i \geq 1}$. Consider another sequence $(c_{n_i})_{i \geq 0}$ in \mathbb{R} , where c_{n_i} is equal to the last component of a_{n_i} . By the Theorem in \mathbb{R} , $(c_{n_i})_{i \geq 1}$ has a convergent subsequence $(c_{n_{m_i}})_{i \geq 1}$. Since $(b_{n_i})_{i \geq 1}$ converges, we know that $(b_{n_{m_i}})_{i \geq 1}$ converges to the same value as well. Suppose that $(b_{n_{m_i}})_{i \geq 1}$ converges to $B = (B_1, B_2, \dots, B_{k-1})$ and $(c_{n_{m_i}})_{i \geq 1}$ converges to C . We claim that $(a_{n_{m_i}})_{i \geq 1}$ is our desired subsequence, which converges to $(B_1, B_2, \dots, B_{k-1}, C)$.

Let $\varepsilon > 0$. Since $(B_1, B_2, \dots, B_{k-1})$ and $(c_{n_{m_i}})_{i \geq 1}$ converge,

$$\exists N_b, N_c > 0 \text{ such that } n_b > N_b \implies \|b_{n_b} - B\|_2 < \frac{\varepsilon}{2} \text{ and } n_c > N_c \implies |c_{n_c} - C| < \frac{\varepsilon}{2}$$

Let $N = \max\{N_b, N_c\}$. Let $n \in \mathbb{N}$, $n > N$. Then

$$\|a_n - (B_1, B_2, \dots, B_{k-1}, C)\|_2 = \sqrt{(a_1 - B_1)^2 + (a_2 - B_2)^2 + \dots + (a_{k-1} - B_{k-1})^2 + (a_k - C)^2}$$

Using an inequality that I don't know the name of, we have

$$\sqrt{(a_1 - B_1)^2 + (a_2 - B_2)^2 + \dots + (a_{k-1} - B_{k-1})^2 + (a_k - C)^2}$$

$$\begin{aligned} &\leq \sqrt{(a_1 - B_1)^2 + (a_2 - B_2)^2 + \cdots + (a_{k-1} - B_{k-1})^2} + \sqrt{(a_k - C)^2} = \|b_n - B\|_2 + |c_n - C|_2 \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Thus $(a_{n_{m_i}})_{i \geq 1}$ converges, which means that $(a_n)_{n \geq 1}$ does indeed have a convergent subsequence. By the principle of induction, the Bolzano-Weierstrass Theorem holds in $(\mathbb{R}^d, \|\cdot\|_2)$ and we are done. \square

- (b) Give an example of a normed vector space $(X, \|\cdot\|)$ containing a (bounded?) sequence (x_n) which has no convergent subsequences.

Proof. Consider the normed vector space $(\ell^\infty, \|\cdot\|_\infty)$. Let $(\vec{x}_n)_{n \geq 1}$ be a sequence in $(\ell^\infty, \|\cdot\|_\infty)$ defined

$$\text{by } \vec{x}_{ik} = \begin{cases} 1, & \text{if } k = i; \\ 0, & \text{otherwise.} \end{cases}$$

Clearly $(\vec{x}_n)_{n \geq 1}$ is bounded. Let $(\vec{x}_{n_i})_{i \geq 1}$ be a subsequence of $(\vec{x}_n)_{n \geq 1}$. We will show that $(\vec{x}_{n_i})_{i \geq 1}$ diverges.

For any $(L_k)_{k \geq 1} \in \ell^\infty$ Let $\varepsilon = \frac{1}{2}$. Let $N \in \mathbb{N}$. If there are no values of $p \geq N$ so that $\|(\vec{x}_p)_k - L_k\|_\infty < \varepsilon$, we are done. Now suppose the opposite, that there is a $p \geq N$ such that $\|(\vec{x}_p)_k - L_k\|_\infty < \varepsilon$. Let $n = p + 1$. By the reverse triangle inequality,

$$\begin{aligned} \|(\vec{x}_n)_k - L_k\|_\infty &= \|(\vec{x}_n)_k - (\vec{x}_p)_k + (\vec{x}_p)_k - L_k\|_\infty \geq \| \|(\vec{x}_n)_k - (\vec{x}_p)_k\|_\infty - \|(\vec{x}_p)_k - L_k\|_\infty \| \\ &= |1 - \|(\vec{x}_p)_k - L_k\|_\infty| \geq 1 - \varepsilon = \frac{1}{2} \end{aligned}$$

Thus $(\vec{x}_{n_i})_{i \geq 1}$ diverges. \square

1.11 #11

Let (X, d) be a metric space. A function $f : X \rightarrow X$ is called a **contraction mapping** if there exists a constant $M \in (0, 1)$ such that

$$d(f(x), f(y)) \leq M d(x, y) \quad \text{for all } x, y \in X.$$

- (a) Suppose that (X, d) is a complete metric space, and that $f : X \rightarrow X$ is a contraction mapping. Prove that f has a unique fixed point; i.e. there exists a unique point $x_0 \in X$ such that $f(x_0) = x_0$.

Proof. Let (X, d) be a complete metric space and f be a contraction mapping. In this proof, for $n \in \mathbb{N}$, we denote f^n to be a composition of f n times.

Let $x \in X$. Define the sequence $(x_i)_{i \in \mathbb{N}}$ by $x_i = f^i(x)$. We show that this is a Cauchy sequence.

Let $\varepsilon > 0$. There exists an $N > 0$ such that $\frac{M^N \cdot d(x, f(x))}{1 - M} < \varepsilon$, since $\frac{d(x, f(x))}{1 - M}$ is constant. Then for all $m, n \geq N, m < n$,

$$d(x_m, x_n) = d(f^m(x), f^n(x)) \leq M^m d(x, f^{n-m}(x))$$

By the triangle inequality,

$$\begin{aligned} M^m d(x, f^{n-m}(x)) &\leq M^m (d(x, f(x)) + d(f(x), f^2(x)) + d(f^2(x), f^3(x)) + \cdots + d(f^{n-m-1}(x), f^{n-m}(x))) \\ &\leq M^m (d(x, f(x)) + M d(x, f(x)) + M^2 d(x, f(x)) + \cdots + M^{n-m-1} d(x, f(x))) \\ &= M^m d(x, f(x)) (1 + M + \cdots + M^{n-m-1}) = \frac{M^m (1 - M^{n-m})}{1 - M} d(x, f(x)) < \frac{M^m}{1 - M} d(x, f(x)) \end{aligned}$$

Since $0 < M < 1$ and $M \geq N$, $M^m < M^N$. Thus

$$\frac{M^m}{1-M}d(x, f(x)) < \frac{M^N}{1-M}d(x, f(x)) < \varepsilon$$

All in all, we have that

$$d(x_m, x_n) < \varepsilon$$

which shows that (x_i) is a Cauchy sequence. By the completeness of X , (x_n) converges to some value which will be denoted as x_0 . In fact, x_0 is the fixed point we want. To prove this, we will show that if (x_n) converges to x_0 , then it also converges to $f(x_0)$.

Suppose $(x_n) \rightarrow x_0$. Letting $\varepsilon > 0$, there exists an $N' > 0$ such that $d(f^n(x), x_0) < \varepsilon$ for all $n \geq N'$. Let $N = N' + 1$. For all $n \geq N$,

$$d(f^n(x), f(x_0)) \leq Md(f^{n-1}(x), x_0)$$

$n - 1 \geq N'$ so

$$Md(f^{n-1}(x), x_0) < \varepsilon$$

We have shown that (x_0) converges to both x_0 and $f(x_0)$. By the uniqueness of limits, we have $x_0 = f(x_0)$.

To show uniqueness, suppose that x_0, x_1 are two distinct fixed points of f . Then using the definition of contraction mapping, we have that

$$d(f(x_0), f(x_1)) \leq Md(x_0, x_1)$$

but because x_0, x_1 are fixed points, this implies that

$$d(x_0, x_1) \leq Md(x_0, x_1)$$

Since they are distinct, $d(x_0, x_1) \neq 0$, so

$$1 \leq M$$

which is a contradiction. Thus there is only one fixed point. □

- (b) Give an example of a normed vector space $(X, \|\cdot\|)$ and a contraction mapping $f : X \rightarrow X$ such that f does **not** have a fixed point.

Proof. Let $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \setminus \{0\}$ be defined by $f(x) = \frac{x}{2}$, where $\mathbb{R} \setminus \{0\}$ is endowed with the usual metric in \mathbb{R} . First we verify that f is a contraction mapping. Let $M = \frac{2}{3}$ and $x, y \in \mathbb{R} \setminus \{0\}$. Then indeed,

$$|f(x) - f(y)| = \left| \frac{x}{2} - \frac{y}{2} \right| = \frac{1}{2} |x - y| \leq \frac{2}{3} |x - y| = M |x - y|$$

Suppose for contradiction that f has a fixed point $x_0 \in \mathbb{R} \setminus \{0\}$. Then

$$x_0 = f(x_0) \implies x_0 = \frac{x_0}{2} \implies \frac{x_0}{2} = 0 \implies x_0 = 0$$

But $0 \notin \mathbb{R} \setminus \{0\}$. There is a contradiction. Thus f has no fixed point. □

1.12 #12*

I specially selected this problem because at the time, I did not know how topology could be used in a way to prove theorems in calculus. It was interesting to see how proving the intermediate value theorem just amounts to showing that the image of an interval is an interval.

- (a) A subset $I \subseteq \mathbf{R}$ is called an **interval** if $a, b \in I$ implies $[a, b] \subseteq I$.

Let $I \subseteq \mathbf{R}$. Prove that I is connected (with respect to the usual metric on \mathbf{R}) if and only if I is an interval.

Proof. We prove the equivalent statement I is disconnected if and only if I is not an interval.

Suppose I is disconnected. Then there exist disjoint, relatively open, non-empty sets $A, B \subseteq I$ such that $A \cup B = I$. Take any $a \in A$ and $b \in B$. We can assume without loss of generality that $a < b$.

Let $S = [a, b] \cap A$. Notice that for all elements $x \in S$, $x \leq b$. Thus b is an upper bound for S . Since \mathbb{R} possesses the least upper bound property, this set has a supremum $m \in \mathbb{R}$. By the definition of least upper bound, we see that $a \leq m \leq b$, so $m \in [a, b]$. Now, we consider two cases.

If $m \notin A$, it suffices to show that $m \notin B$. Since m is the supremum of S , we can always find an element in S that is greater than $m - \varepsilon$, so there is always an element in A that is within the ε ball surrounding m . Thus m is a limit point of A , which means that it is impossible for m to be an element of B , which is an open set. Thus we have that $m \notin A \cup B = I$ as desired.

If $m \in A$, then since A is an open subset of I , m is an interior point of A . But we also know that m is greater than or equal to any other element in A . It must be that for some $\varepsilon > 0$, the interval $(m, m + \varepsilon)$ is disjoint from I . Otherwise, it contradicts the fact that A is open in I . Then take any value x from the set $(m, \varepsilon) \cap [a, b]$ and notice that $x \notin I$ but $x \in [a, b]$.

In both cases, we have shown that I is not an interval.

Conversely, suppose that I is not an interval. Then for $p < q \in I$, there is a $c \in [p, q]$ such that $c \notin I$. Define subsets A and B in I as $A = \{x \in I : x < c\}$ and $B = \{x \in I : x > c\}$. A and B are non-empty because $p \in A$ and $q \in B$. The sets are also disjoint by construction.

To show that A is open, first take any $a \in A$. For this value of a , take $\varepsilon = c - a > 0$. For all $x \in B_I(a, \varepsilon)$, note that $x \in I$. As well, if $x < a$, immediately we have $x < a < c \implies x \in A$. If $x > a$, since x is within the open ball surrounding a , $x - a = |x - a| < c - a \implies x < c \implies x \in A$. Thus every element of A is an interior point, so A is open.

We can use a very similar argument for the set B , so the proof is omitted. We conclude that B is open as well. Therefore I is disconnected. □

- (b) Let (X, d_X) and (Y, d_Y) be two metric spaces, and let $f : X \rightarrow Y$ be a continuous function. Prove that if C is a connected subset of X , then $f(C)$ is a connected subset of Y .

Proof. We examine the contrapositive: if $f(C)$ is a disconnected subset of Y , then C is a disconnected subset of X .

Suppose that $f(C)$ is a disconnected subset of Y . Then there exist non-empty, disjoint, relatively open subsets $A, B \subseteq f(C)$ such that $A \cup B = f(C)$. Consider the sets $f^{-1}(A), f^{-1}(B) \subseteq C$. Notice that if $x \in f^{-1}(A)$, $f(x) \in A$. Since A, B are disjoint, $f(x) \notin B$, so $x \notin f^{-1}(B)$. It follows that $f^{-1}(A)$ and $f^{-1}(B)$ are disjoint.

Now, we will show that $f^{-1}(A)$ is non-empty and open.

We know that A is non-empty, so there exists an element $y \in A$. By the definition of image, $y = f(x)$ for some $x \in C$. It follows that $x \in f^{-1}(A)$. Thus $f^{-1}(A)$ is non-empty.

To show that $f^{-1}(A)$ is open, we use the topological definition of continuity of f to conclude that since A is open, $f^{-1}(A)$ is open as well.

We can apply the same argument for $f^{-1}(B)$, so B is non-empty and open as well.

Thus $f^{-1}(A)$, $f^{-1}(B)$ are non-empty, disjoint, and open. It remains to show that $f^{-1}(A) \cup f^{-1}(B) = C$. Immediately, we know that $f^{-1}(A) \subseteq C$ and $f^{-1}(B) \subseteq C$, so $f^{-1}(A) \cup f^{-1}(B) \subseteq C$.

For the other direction, let $x \in C$. We know that $f(x) \in f(C)$, which means that either $f(x) \in A$ or $f(x) \in B$. It follows that $x \in f^{-1}(A)$ or $x \in f^{-1}(B) \implies x \in f^{-1}(A) \cup f^{-1}(B)$. Thus $f^{-1}(A) \cup f^{-1}(B) = C$.

We conclude that C is a disconnected subset of X , which shows that the contrapositive of the original statement is true, completing the proof. \square

- (c) Recall the **Intermediate Value Theorem** from single-variable calculus. “Let $I \subseteq \mathbf{R}$ be an open interval and $f : I \rightarrow \mathbf{R}$ be a continuous function. Suppose that $a, b \in f(I)$ are two numbers such that $a < b$, and suppose that $a < y_0 < b$. Then there exists $x_0 \in I$ such that $f(x_0) = y_0$.”

Prove that this theorem immediately follows from (a) and (b). Thus, (b) is a generalization of the Intermediate Value Theorem.

Proof. Let $a, b, y_0 \in f(I)$ and suppose that $a < y_0 < b$. Since I is an interval, by part (a) I is connected. As well, by continuity of f , it follows from the result of (b) that $f(I)$ is a connected subset of \mathbf{R} . By (a) once again, $f(I)$ is an interval. Thus $[a, b] \subseteq f(I)$. Since $a < y_0 < b$, $y_0 \in [a, b] \subseteq f(I)$ which implies that $y_0 = f(x_0)$ for some $x_0 \in I$, which is what we were looking for. \square

1.13 #13 - Unfinished

Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be two normed vector spaces and let $T : X \rightarrow Y$ be a linear mapping.

- (a) Prove that T is continuous if and only if T is a bounded linear mapping.

Proof. Suppose that T is continuous. Then at $x_0 = 0$, there exists a $\delta > 0$ such that

$$\|x\|_X < \delta \implies \|T(x)\|_Y \leq 1$$

We claim that our $M = \frac{2}{\delta}$. Let $x \in X$. Notice that $\left\| \frac{\delta \cdot x}{2\|x\|_X} \right\|_X \leq \delta$. By the continuity of T we have that

$$\left\| T \left(\frac{\delta \cdot x}{2\|x\|_X} \right) \right\|_Y \leq 1 \implies \frac{\delta}{2 \cdot \|x\|_X} \|T(x)\|_Y \leq 1 \implies \|T(x)\|_Y \leq \frac{2}{\delta} \|x\|_X$$

Thus T is a bounded linear mapping.

Next, suppose that T is a bounded linear mapping. Then there is an $M > 0$ such that for all $x \in X$,

$$\|T(x)\|_Y \leq M\|x\|_X$$

We will show that T is continuous everywhere. Fix $a \in X$. Let $\varepsilon > 0$. Let $\delta = \frac{\varepsilon}{M}$. Let $x \in X$ and suppose that $\|x - a\|_X < \delta$. Then

$$\|T(x) - T(a)\|_Y = \|T(x - a)\|_Y \leq M\|x - a\|_X < \varepsilon$$

and we are done. \square

- (b) Suppose that $(Y, \|\cdot\|_Y) = (\mathbf{R}^n, \|\cdot\|_2)$. Prove that T is continuous if and only if $\ker(T)$ is closed.

Proof. Suppose that T is continuous. Let a be a limit point for $\ker(T)$. We want to show that $T(a) = 0$. By definition, T being continuous implies that

$$\lim_{x \rightarrow a} T(x) = T(a)$$

or more formally, letting ε be arbitrary, there exists a δ such that

$$\|x - a\|_X < \delta \implies \|T(x) - T(a)\|_2$$

Since a is a limit point of $\ker(T)$, there exists $x \in \ker(T)$ such that

$$x \in B(a, \delta) \implies \|x - a\|_X \implies \|T(x) - T(a)\|_2 < \varepsilon \implies \|T(a)\|_2 < \varepsilon$$

By properties of norms $\|T(a)\|_2 \geq 0$, so we must have that $T(a) = 0 \implies a \in \ker(T)$. Thus $\ker(T)$ is closed.

I have no clue how to prove the converse so yeah.

□

1.14 #16*

I specially selected this problem because the proof for this was very unorthodox and unintuitive, which I find pretty cool. This problem has inspired me to try and find creative ways to approach other problems in case I get stumped using conventional methods.

Let $(X, \|\cdot\|)$ be a normed vector space, let $B(X) = B(X, X)$ be the space of bounded linear operators on X , and equip $B(X)$ with the operator norm $\|\cdot\|_{\text{op}}$. Let $\text{GL}(X)$ be the set of invertible bounded linear operators:

$$\text{GL}(X) = \{T \in B(X) : T \text{ is invertible}\}.$$

(The notation GL means “general linear group.”)

Prove that if $(X, \|\cdot\|)$ is complete, then $\text{GL}(X)$ is an open subset of $(B(X), \|\cdot\|_{\text{op}})$.

Proof. First, we prove a couple lemmas.

Lemma 1. For all $S, T \in B(X, X)$,

$$\|ST\|_{\text{op}} \leq \|S\|_{\text{op}}\|T\|_{\text{op}}$$

To prove this, let $x \in X$. If $x = 0$ the result is immediate, so suppose that $0 < \|x\|_X \leq 1$. Then

$$\|S(T(x))\|_X = \left\| S \left(\frac{\|T(x)\|_X}{\|T(x)\|_X} T(x) \right) \right\|_X = \|T(x)\|_X \left\| S \left(\frac{T(x)}{\|T(x)\|_X} \right) \right\|_X$$

Notice that $\left\| \frac{T(x)}{\|T(x)\|_X} \right\|_X = 1$. We obtain that

$$\|S(T(x))\|_X = \|T(x)\|_X \left\| S \left(\frac{T(x)}{\|T(x)\|_X} \right) \right\|_X \leq \|T\|_{\text{op}}\|S\|_{\text{op}}$$

Since $\|T\|_{\text{op}}\|S\|_{\text{op}}$ is an upper bound for $\|S(T(x))\|_X$ when $\|x\|_X \leq 1$, by definition of supremum we can conclude that $\|ST\|_{\text{op}} \leq \|S\|_{\text{op}}\|T\|_{\text{op}}$, proving the first lemma.

Next, the following lemma will help in proving Lemma 3.

Lemma 2. If $T, U \in B(X)$, then for $x \in X$, $\|T(x) - U(x)\|_X \leq \|x\|_X \|T - U\|_{\text{op}}$.

Let $T, U \in B(X)$. Then

$$\|(T - U)(x)\|_X = \left\| \|x\|_X (T - U) \left(\frac{x}{\|x\|_X} \right) \right\|_X = \|x\|_X \left\| (T - U) \left(\frac{x}{\|x\|_X} \right) \right\|_X$$

Since $\frac{x}{\|x\|_X} = 1$,

$$\|x\|_X \left\| (T - U) \left(\frac{x}{\|x\|_X} \right) \right\| \leq \|x\|_X \|T - U\|_{\text{op}}$$

which is what we wanted.

We will use the proceeding lemma to help us prove Lemma 4.

Lemma 3. If X is complete then $B(X, X)$ is complete.

For every Cauchy sequence (T_n) in $B(X, X)$, the sequence $(T_n(x_i))$, where x_i is an arbitrary element in X , is a Cauchy sequence in X . By the completeness of X , $(T_n(x_i))$ converges to some L_i . Let T be a linear operator on X defined by $T(x_i) = L_i$. We can see that this is indeed a linear operator because

$$cT(x_i) + T(x_j) = cL_i + L_j = c \lim_{n \rightarrow \infty} T_n(x_i) + \lim_{n \rightarrow \infty} T_n(x_j) = \lim_{n \rightarrow \infty} cT_n(x_i) + T_n(x_j) = \lim_{n \rightarrow \infty} T_n(cx_i + x_j)$$

X is closed, so $cx_i + x_j$ is equal to some $x_k \in X$. Thus

$$cT(x_i) + T(x_j) = \lim_{n \rightarrow \infty} T_n(cx_i + x_j) = \lim_{n \rightarrow \infty} T_n(x_k) = L_k = T(x_k)$$

which implies that T is a linear operator.

To show that T is bounded, we know that T_n is bounded, so $\|T_n(x)\|_X < M_n \|x\|_X$, for some $M_n > 0$. As well, we have that T_n is uniformly continuous. Thus for a sufficiently large $n \in \mathbb{N}$, for all $x \in X$, using Lemma 2,

$$\begin{aligned} \|T(x)\|_X &= \lim_{m \rightarrow \infty} \|T_m(x)\|_X \leq \lim_{m \rightarrow \infty} \|T_m(x) - T_n(x)\|_X + \|T_n(x)\|_X < \lim_{m \rightarrow \infty} \|T_n - T_m\|_{\text{op}} \|x\|_X + \|T_n(x)\|_X \\ &\leq \|x\|_X + M_n \|x\|_X \end{aligned}$$

We get the last inequality because T_n is Cauchy and bounded as well. Thus $T \in B(X)$. To show that T_n converges to T , fix $\varepsilon > 0$. For all $x \in X$ such that $\|x\|_X \leq 1$, there exists a sufficiently large N such that for all $x \in X$, if $n > N$ then

$$\|T(x) - T_n(x)\|_X < \varepsilon \implies \|T - T_n\|_{\text{op}} < \varepsilon$$

Thus we can conclude that $B(X, X)$ is complete.

Lemma 4. If $T \in B(X, X)$ such that $\|I - T\|_{\text{op}} < 1$, then T is invertible and

$$T^{-1} = \sum_{i=0}^{\infty} (I - T)^i$$

Suppose that $T \in B(X, X)$ and $\|I - T\|_{\text{op}} < 1$. Let $(a_n)_{n \geq 1}$ be a Cauchy sequence in $B(X, X)$ defined by

$$a_n = \sum_{i=0}^n (I - T)^i$$

We verify that this sequence is Cauchy. Let $\varepsilon > 0$. Since $\|I - T\|_{\text{op}} < 1$, there exists natural N such that

$$\frac{\|I - T\|_{\text{op}}^N}{1 - \|I - T\|_{\text{op}}} < \varepsilon$$

Let $m, n > N$ and suppose that $m < n$. Then

$$\|a_n - a_m\|_{\text{op}} = \left\| \sum_{i=0}^n (I - T)^i - \sum_{i=0}^m (I - T)^i \right\|_{\text{op}} = \left\| \sum_{i=m+1}^n (I - T)^i \right\|_{\text{op}}$$

By the triangle inequality and Lemma 1,

$$\left\| \sum_{i=m+1}^n (I - T)^i \right\| \leq \sum_{i=m+1}^n \|(I - T)^i\| \leq \sum_{i=m+1}^n \|I - T\|^i = \|I - T\|^{m+1} \sum_{i=0}^{n-m-1} \|I - T\|^i$$

$$= \|I - T\|^{m+1} \frac{1 - \|I - T\|^{n-m}}{1 - \|I - T\|}$$

$1 - \|I - T\|^{n-m} < 1$, so

$$\|I - T\|^{m+1} \frac{1 - \|I - T\|^{n-m}}{1 - \|I - T\|} < \frac{\|I - T\|^{m+1}}{1 - \|I - T\|} < \frac{\|I - T\|^N}{1 - \|I - T\|} < \varepsilon$$

We see that (a_n) is indeed a Cauchy sequence.

By Lemma 4, $B(X, X)$ is complete, so (a_n) converges, meaning that $\sum_{i=0}^{\infty} (I - T)^i$ exists. Now, notice that

$$T \left(\sum_{i=0}^{\infty} (I - T)^i \right) = (I - (I - T)) \left(\sum_{i=0}^{\infty} (I - T)^i \right) = \sum_{i=0}^{\infty} (I - T)^i - \sum_{i=0}^{\infty} (I - T)^{i+1} = \sum_{i=0}^{\infty} ((I - T)^i - (I - T)^{i+1})$$

This is a telescoping series. As $i \rightarrow \infty$, $(I - T)^i \rightarrow 0$. Thus

$$T \left(\sum_{i=0}^{\infty} (I - T)^i \right) = \sum_{i=0}^{\infty} ((I - T)^i - (I - T)^{i+1}) = I$$

Thus $\sum_{i=0}^{\infty} (I - T)^i$ is the inverse of T , which implies that T^{-1} exists.

Now we can show that $B(X)$ is open.

Let $T \in B(X)$. Consider the open ball $B(T, \frac{1}{\|T^{-1}\|_{\text{op}}})$. For all $S \in B(T, \frac{1}{\|T^{-1}\|_{\text{op}}})$,

$$\|T - S\|_{\text{op}} < \frac{1}{\|T^{-1}\|_{\text{op}}} \implies \|T - S\|_{\text{op}} \|T^{-1}\|_{\text{op}} < 1$$

By Lemma 1, we have that

$$1 > \|T - S\|_{\text{op}} \|T^{-1}\|_{\text{op}} \geq \|(T - S)T^{-1}\|_{\text{op}} = \|I - ST^{-1}\|_{\text{op}}$$

By Lemma 4, ST^{-1} is invertible. If we let $S^{-1} = T^{-1}(ST^{-1})^{-1}$, we see that

$$SS^{-1} = S(T^{-1}(ST^{-1})^{-1}) = (ST^{-1})(ST^{-1})^{-1} = I$$

Thus S is invertible, which means that $S \in B(X)$. Therefore $B(X)$ is open. □

1.15 #17*

I specially selected this problem because it gave me a better understanding of topology and how to prove results in topology. It was my first time working with sequential compactness so everything felt unfamiliar. It was very satisfying solving the problem after thinking about it for a whole day, when everything clicked.

Let (X, d) be a metric space and let $\{U_i\}_{i \in I}$ be an open cover of X ; this means that each U_i is an open subset of X , and that $X = \bigcup_{i \in I} U_i$. A **magic number** for $\{U_i\}_{i \in I}$ is a number $\delta > 0$ with the following property: if $A \subseteq X$ is a set with $\text{diam}(A) < \delta$, then $A \subseteq U_i$ for at least one index $i \in I$.

Suppose that (X, d) is a clustering metric space. Prove that every open cover has a magic number.

Proof. Suppose that (X, d) is a clustering metric space. Suppose for the sake of contradiction that there exists an open cover $\{U_i\}_{i \in I}$ that doesn't have a magic number.

For $n \in \mathbb{N}$, there is $A_n \subseteq X$ with $\text{diam} A_n < \frac{1}{n}$ so that $A_n \not\subseteq U_i$ for all indices $i \in I$. Since $\text{diam} A_n < \frac{1}{n}$, we can cover A_n with an open ball $B(a_n, \frac{1}{n})$, where a_n is some element in X .

Define a sequence $(a_n)_{n \in \mathbb{N}}$ in X such that a_n is equal to the one above.

By the clustering property of X , (a_n) has a convergent subsequence (b_n) . Denote the limit of (b_n) as p .

p is an element of X , so it is contained in some U_i in the open cover. Since U_i is open, we can find $\varepsilon > 0$ such that $B(p, \varepsilon) \subseteq U_i$. As well, since (b_n) converges to p we can find infinitely many entries of the sequence

within the open ball $B(p, \frac{\varepsilon}{2})$. Thus we can find a large enough n such that $n > \frac{2}{\varepsilon}$, which gives $\frac{1}{n} < \frac{\varepsilon}{2}$, and still have that b_n is $\frac{\varepsilon}{2}$ -close to p .

Now, consider the open ball $B(b_n, \frac{1}{n})$. We will show that $B(b_n, \frac{1}{n}) \subseteq B(p, \varepsilon)$.

Let $x \in B(b_n, \frac{1}{n})$. Then $d(x, b_n) < \frac{1}{n} < \frac{\varepsilon}{2}$, so we have

$$d(x, p) \leq d(x, b_n) + d(b_n, p) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \implies x \in B(p, \varepsilon)$$

which is what we wanted.

Recall that the set A_n is covered by $B(b_n, \frac{1}{n})$. Then we have

$$A_n \subseteq B(b_n, \frac{1}{n}) \subseteq B(p, \varepsilon) \subseteq U_i$$

contradicting the fact that $A_n \not\subseteq U_i$. Thus every open cover has a magic number. □

1.16 #18

The sequence space ℓ^1 has two norms: the 1-norm $\|\cdot\|_1$, and the sup norm $\|\cdot\|_\infty$ inherited from ℓ^∞ .

(a) Is $(\ell^1, \|\cdot\|_\infty)$ separable?

(b) Is $(\ell^1, \|\cdot\|_1)$ separable?

Proof. Yes to both. We will show part (b) first and notice that part (a) follows right after.

Let D be the set of all sequences of rational numbers that are eventually 0. This set is obviously a subset of ℓ^1 . We can write D as the countable union of rational sequences that terminate after the n th term, which is countable since there is a bijection to \mathbb{Q}^n . It follows that D is countable. It remains to show that D is dense in ℓ^1 with respect to the 1-norm.

Let $x \in \ell^1$ and $\varepsilon > 0$. Since x is absolutely summable, every partial sum is strictly less than the limit of the series, denoted as L . Also, for all n greater than some $N > 0$,

$$\sum_{i=N+1}^n |x_i| = \sum_{i=0}^n |x_i| - \sum_{i=0}^N |x_i| < L - \sum_{i=0}^N |x_i| < L - \left(L - \frac{\varepsilon}{2}\right) = \frac{\varepsilon}{2}$$

When $1 \leq i \leq N$, by the density of \mathbb{Q} in \mathbb{R} , we can find a $q_i \in \mathbb{Q}$ such that $\|x_i - q_i\| < \frac{\varepsilon}{2N}$.

Let $q = \begin{cases} q_i, & \text{if } 1 \leq i \leq N; \\ 0, & \text{otherwise.} \end{cases}$. We see that

$$\|x - q\|_1 = \sum_{i=0}^{\infty} |x_i - q_i| = \sum_{i=0}^N |x_i - q_i| + \sum_{i=N+1}^{\infty} |x_i| < \frac{N\varepsilon}{2N} + \frac{\varepsilon}{2} = \varepsilon$$

Thus D is dense, implying that ℓ^1 is separable with respect to the 1-norm.

Note that for all $x \in \ell^1$, $\|x\|_\infty \leq \|x\|_1$. it follows immediately that $(\ell^1, \|\cdot\|_\infty)$ is separable as well. □

1.17 #19 - Incomplete

This question is incomplete because when I argued that (K_n) is a compact exhaustion, I did not properly verify that it satisfied the conditions to be a compact exhaustion (i.e., I simply claimed that $K_n \subseteq (K_n)^\circ$ without proving it).

Let X be a metric space and let $A \subseteq X$. A **compact exhaustion** for A is a sequence of compact sets K_1, K_2, K_3, \dots such that $U = \bigcup_{i \geq 1} K_i$ and $K_i \subseteq K_{i+1}^\circ$.

- (a) Let $U \subseteq \mathbf{R}^n$ be a bounded open set. Show that U has a compact exhaustion.
- (b) Now show that every open set $U \subseteq \mathbf{R}^n$ has a compact exhaustion.

Proof. (a):

Let $U \subseteq \mathbf{R}^n$ be a bounded open set. Let N be smallest natural number such that the set $S = \{x \in U : \text{dist}(\{x\}, U^c) \geq \frac{1}{N}\}$ is non-empty. Define the sequence of sets $(S_n)_{n \geq 1}$ by

$$S_n = \left\{ x \in U : \text{dist}(\{x\}, U^c) \geq \frac{1}{n+N} \right\}$$

We will show that S_n is closed. Let s be limit point of S_n . Then for $y \in U^c$,

$$d(s, y) \geq d(x, y) - d(x, s) \implies d(s, y) \geq \sup_{x \in S_n} (d(x, y) - d(x, s))$$

We will show that $\sup_{x \in S_n} (d(x, y) - d(x, s)) = \frac{1}{N+n}$. For all $\varepsilon > 0$, there is an $x \in S_n$ such that $d(x, s) < \varepsilon$. Thus

$$d(x, y) - d(x, s) > \frac{1}{N+n} - \varepsilon$$

which means that $\sup_{x \in S_n} (d(x, y) - d(x, s)) = \frac{1}{N+n}$. Therefore we have that

$$d(s, y) \geq \frac{1}{N+n} \implies s \in S_n$$

so it follows that S_n is closed. Since $S_n \subseteq U$, it is also bounded, so because we are working in \mathbf{R}^n , S_n is compact.

As well, we need to have that $S_n \subseteq S_{n+1}^\circ$. This result is quite quick, as we can notice that

$$S_{n+1}^\circ = \left\{ x \in U : \text{dist}(\{x\}, U^c) > \frac{1}{N+n} \right\}$$

From this, the inclusion follows nicely.

Then, as n tends to infinity, $S_n = \{x \in U : \text{dist}(\{x\}, U^c) > 0\}$ which is exactly $(U^c)^c = U$. Therefore U has compact exhaustion.

(b):

Now, let $U \subseteq \mathbf{R}^n$ be open. Define the sequence of bounded open sets $A_n = U \cap B(\vec{0}, n)$. As n tends to infinity, $\bigcup_{n=1}^\infty A_n = U$. By the previous part, A_n has a compact exhaustion $(K_{nk})_{k \geq 1}$. Take the sequence $(K_k)_{k \geq 1}$ to be $K_k = \bigcup_{n=1}^\infty K_{nk}$. This sequence of sets satisfies the conditions for a compact exhaustion. Now we attempt to prove that the sequence converges to U . We have

$$\bigcup_{k=1}^\infty K_k = \bigcup_{k=1}^\infty \bigcup_{n=1}^\infty K_{nk}$$

We swap the order of union and get that

$$\bigcup_{k=1}^\infty \bigcup_{n=1}^\infty K_{nk} = \bigcup_{n=1}^\infty \bigcup_{k=1}^\infty K_{nk} = \bigcup_{n=1}^\infty A_n = U$$

Thus every open subset of \mathbf{R}^n has a compact exhaustion. □

1.18 #20*

I specially selected this question because I found that it increased my understanding of compactness and the idea of a set being 'large'. This problem was also pretty fun to solve because it utilised alternative definitions of compactness and completeness in a complete ambient space. I find proofs involving open/closed sets are pretty interesting.

Let $x, y \in \ell^\infty$ be two sequences. Let us say that y is **dominated** by x , denoted $x \geq y$, if $|x_n| \geq |y_n|$ for all $n \in \mathbf{N}$. Let D_x denote the set of all sequences which are dominated by x :

$$D_x = \{y \in \ell^\infty : |y_n| \leq |x_n| \text{ for all } n \in \mathbf{N}\}.$$

Prove that D_x is compact if and only if $x_n \rightarrow 0$.

Proof. Suppose that D_x is compact. Suppose for contradiction that $x_n \not\rightarrow 0$. For some $\varepsilon > 0$, $|x_{N_k}| \geq \varepsilon$ for an infinite number of N_k . Consider the open cover $\{B(\vec{y}_i, \frac{\varepsilon}{2})\}_{i \in I}$, which is the collection of $\frac{\varepsilon}{2}$ -balls centered around every $\vec{y}_i \in D_x$. By compactness of D_x , there is a finite subcover $\{B(\vec{y}_i, \frac{\varepsilon}{2})\}_{i \leq m}$. Now, we construct a $y \in D_x$ as follows:

For every sequence \vec{y}_i , let

$$y_{N_i} = \begin{cases} \varepsilon, & \text{if } (\vec{y}_i)_{N_i} < \frac{\varepsilon}{2}; \\ 0, & \text{if } (\vec{y}_i)_{N_i} \geq \frac{\varepsilon}{2}; \end{cases}$$

For all other terms in y , make it 0. Notice that for all $B(\vec{y}_i, \frac{\varepsilon}{2})$,

$$\|y - \vec{y}_i\|_\infty \geq |y_{N_i} - (\vec{y}_i)_{N_i}| \geq \frac{\varepsilon}{2} \implies y \notin D_x$$

which is a contradiction.

Conversely, suppose that $x_n \rightarrow 0$. We will show that D_x is complete and totally bounded, which is equivalent to compactness.

To show completeness, notice that the ambient space ℓ^∞ is complete. Thus if we can show that D_x is closed, it will follow that D_x is complete.

Let $a \notin D_x$. We will show that a is not a limit point of D_x , which means that D_x is closed. We know that there is a $k \in \mathbf{N}$ such that $|a_k| > |x_k|$. Define $\varepsilon = |a_k| - |x_k|$. Fix $y \in D_x$. Then

$$\|a_k - y_k\|_\infty \geq |a_k - y_k| \geq |a_k| - |y_k| \geq |a_k| - |x_k| = \varepsilon$$

which implies that $y \notin B(a_k, \varepsilon)$, so a is not a limit point of D_x . Thus D_x is closed. It follows that D_x is complete.

To show that D_x is totally bounded, first let $\varepsilon > 0$. Since $x_n \rightarrow 0$, there is a large enough N such that for $n > N$, $|x_n| < \frac{\varepsilon}{2}$. For all $y \in D_x$, $|y_n| < |x_n| < \frac{\varepsilon}{2}$. Consider the set of elements in D_x such that their terms are 0 for $n > N$. This set is totally bounded, so there is a ε -ball cover $\{B(y_i, \varepsilon)\}_{i \leq n}$. We show that this collection also covers D_x .

For $y \in D_x$, there is an open ball $B(y_i, \varepsilon)$ such that $\sup_{n \leq N} |y_n - (y_i)_n| < \varepsilon$. But also notice that for $n > N$, $|(y_i)_n - y_n| \leq |(y_i)_n| + |y_n| < |y_n| < \frac{\varepsilon}{2}$. Thus $\|y - y_i\|_\infty < \varepsilon$ so $y \in B(y_i, \varepsilon)$. Thus D_x is totally bounded.

Since D_x is both complete and totally bounded, we can conclude that D_x is compact. \square

1.19 #22*

I specially selected this problem because I felt that it was rewarding to solve it with minimal help. I am not very strong at doing proofs by construction so I think I learned a lot by attempting this problem. It was very satisfying not only finding the right construction, but also proving that it works. Everything was done on my own and the only help I received was a tip on how to show surjectivity.

For a normed vector space $(X, \|\cdot\|)$, let $X^* = B(X, \mathbf{R})$ denote the set of bounded linear mappings from X to \mathbf{R} . Here X^* is equipped with the operator norm $\|\cdot\|_{\text{op}}$, and the normed vector space $(X^*, \|\cdot\|_{\text{op}})$ is called the **topological dual** of X .

Let c_0 denote the set of sequences converging to zero. Prove that $c_0^* \equiv \ell^1$.

Proof. Equip ℓ^1 with the 1-norm and c_0 with the sup-norm. Define the function $f : \ell^1 \rightarrow c_0^*$ by $f(\vec{x})(\vec{a}) = \sum_{i=1}^{\infty} x_i \cdot a_i$. This sum is convergent because for sufficiently large $N > 0$, $|a_n| < 1$ for $n > N$, so

$$\sum_{i=1}^{\infty} x_i \cdot a_i = \sum_{i=1}^N x_i \cdot a_i + \sum_{i=N+1}^{\infty} x_i \cdot a_i < \sum_{i=1}^N x_i \cdot a_i + \sum_{i=N+1}^{\infty} x_i < \infty$$

We claim that this is a bijective isometry.

First, it can be quickly verified that this function is linear. Letting $a_n \in c_0$, for $x, y \in \ell^1$, $k > 0$,

$$f(kx + y)(a_n) = \sum_{i=1}^{\infty} a_i \cdot (kx_i + y_i) = k \sum_{i=1}^{\infty} a_i \cdot x_i + \sum_{i=1}^{\infty} a_i \cdot y_i = kf(x)(a_n) + f(y)(a_n).$$

To show that this function is an isometry, let $x, y \in \ell^1$.

Let $A = \{|f(x)(a) - f(y)(a)| : a \in c_0, \|a\|_{\infty} \leq 1\}$. It will be shown that $\|x - y\|_1 = \sup A$.

Let $a \in c_0$ such that $\|a\|_{\infty} \leq 1$. Then

$$\begin{aligned} |f(x)(a) - f(y)(a)| &= \left| \sum_{i=1}^{\infty} (x_i \cdot a_i - y_i \cdot a_i) \right| = \left| \sum_{i=1}^{\infty} a_i (x_i - y_i) \right| \leq \sum_{i=1}^{\infty} |a_i| |x_i - y_i| \leq \sum_{i=1}^{\infty} |x_i - y_i| \\ &= \|x - y\|_1. \end{aligned}$$

This means that $\|x - y\|_1$ is an upper bound for A . To prove that this is the least upper bound, let $\varepsilon > 0$. There is an $N > 0$ such that $\sum_{i=N+1}^{\infty} |x_i - y_i| < \varepsilon$. Define the sequence (a_n) in c_0 as the sequence of 1's until and including $n = N$ and 0 afterwards. Then

$$\begin{aligned} |f(x)(a) - f(y)(a)| &= \left| \sum_{i=1}^{\infty} a_i (x_i - y_i) \right| = \left| \sum_{i=1}^{\infty} (1 - 1 + a_i) (x_i - y_i) \right| \\ &\geq \left| \sum_{i=1}^{\infty} (x_i - y_i) \right| - \left| \sum_{i=1}^{\infty} (1 - a_i) (x_i - y_i) \right| \\ &= \left| \sum_{i=1}^{\infty} (x_i - y_i) \right| - \sum_{i=N+1}^{\infty} |x_i - y_i| > \|x - y\|_1 - \varepsilon \end{aligned}$$

Thus we have that

$$\|f(x) - f(y)\|_{\text{op}} = \sup A = \|x - y\|_1$$

so f is an isometry.

To prove that f is bijective, it suffices to show that f is surjective, since we already have injectivity from the isometry.

Let $T \in c_0^*$. Construct a sequence (x_n) in ℓ^1 defined by $x_n = T(e_n)$, where (e_n) is the sequence with the only non-zero term being $e_n = 1$.

First, we need to verify that $(x_n) \in \ell^1$. To do this, we bound the absolute sum by a finite number. We have that

$$\sum_{i=1}^{\infty} |x_i| = \sum_{i=1}^{\infty} |T(e_i)|$$

For all non-zero terms, we can rewrite it as $|T(e_i)| \cdot \frac{T(e_i)}{T(e_i)} = T\left(\frac{|T(e_i)|}{T(e_i)} e_i\right)$. Define (b_n) to be a sequence defined by

$$b_n = \begin{cases} 0, & \text{if } T(e_i) = 0; \\ \frac{|T(e_i)|}{T(e_i)} e_i, & \text{otherwise.} \end{cases}$$

Then

$$\sum_{i=1}^{\infty} |T(e_i)| = T(b_n)$$

Notice that $|b_n| \leq 1$ for all $n \in \mathbb{N}$. Thus $\|b_n\|_{\infty} = 1$, so

$$\sum_{i=1}^{\infty} |x_i| = T(b_n) \leq \|T\|_{\text{op}}$$

By monotone bounded convergence, $\sum_{i=1}^{\infty} |x_i|$ converges. Therefore $(x_n) \in \ell^1$.

Now, we want to show that $f(x_n) = T$. For $(a_n) \in c_0$, we have that

$$f(x_n)(a_n) = \sum_{i=1}^{\infty} x_i \cdot a_i = \sum_{i=1}^{\infty} T(e_i) \cdot a_i = T\left(\sum_{i=1}^{\infty} a_i(e_i)\right) = T((a_n))$$

so f is surjective.

Thus f is a bijective isometry, so $c_0^* \equiv \ell^1$. □

1.20 #23

Let S^2 denote the unit sphere in \mathbf{R}^3 . Let $N = (0, 0, 1)$ denote the “north pole”. In this problem, you will show that $S^2 \setminus \{N\}$ is homeomorphic to \mathbf{R}^2 . To do this, we define a function $\Phi : S^2 \setminus \{N\} \rightarrow \mathbf{R}^2$ known as the **stereographic projection**: given a point P in $S^2 \setminus \{N\}$, draw a line between P and N , and let $\Phi(P)$ denote the point where this line intersects the xy -plane in \mathbf{R}^3 .

- (a) Given $P = (x, y, z)$, find an explicit formula for $\Phi(P)$ in terms of x, y, z .
- (b) Deduce that Φ is continuous.
- (c) Prove that Φ is a bijection; in fact, given $p = (s, t) \in \mathbf{R}^2$, find an explicit formula for $\Phi^{-1}(p)$.
- (d) Deduce that Φ is a homeomorphism.

Proof. We will work off the assumption that both metric spaces are equipped with the max-norm. We can do this because all norms on \mathbb{R}^n are equivalent.

(a):

Let $P = (x, y, z)$. First, we find the equation of the line that passes P and N . Consider the equation of the line $L(t) = (tx, ty, (z-1)t + 1)$. Notice that $L(0) = N$ and $L(1) = P$, so L satisfies what we were looking for. Now we find the point where L intersects with the xy -plane. This happens exactly when $(z-1)t + 1 = 0$. Solving for t gives $t = \frac{1}{1-z}$. This value is always defined as $z \neq 1$. As a result, it turns out that

$$L\left(\frac{1}{1-z}\right) = \left(\frac{x}{1-z}, \frac{y}{1-z}, 0\right).$$

Thus

$$\Phi(P) = \frac{1}{1-z}(x, y).$$

(b):

To show that Φ is continuous, fix $(a, b, c) \in \mathbb{R}^3$. Let $\varepsilon > 0$. Then let $\delta = \min\left\{\frac{1-c}{2}, \frac{(1-c)\varepsilon}{4}, \frac{(1-c)^2\varepsilon}{4\max\{|a|, |b|\}}\right\}$. For $(x, y, z) \in \mathbb{R}^3$ such that $\|(x, y, z) - (a, b, c)\|_{\max} < \delta$, notice that

$$z - c < |z - c| \leq \|(x, y, z) - (a, b, c)\|_{\max} < \frac{1-c}{2}$$

Manipulating this, we obtain

$$z < \frac{1+c}{2} \implies 1-z > \frac{1-c}{2} \implies \frac{1}{1-z} < \frac{2}{1-c}.$$

So

$$\begin{aligned} \|\Phi(x, y, z) - \Phi(a, b, c)\|_{\max} &= \left\| \left(\frac{x}{1-z}, \frac{y}{1-z}\right) - \left(\frac{a}{1-c}, \frac{b}{1-c}\right) \right\|_{\max} \\ &= \frac{\|(x(1-c) - a(1-z), y(1-c) - b(1-z))\|_{\max}}{(1-z)(1-c)} \\ &< \frac{2\|(x(1-c) - a(1-z), y(1-c) - b(1-z))\|_{\max}}{(1-c)^2} \end{aligned}$$

Without loss of generality, suppose that $|x(1-c) - a(1-z)| \geq |y(1-c) - b(1-z)|$. Then

$$\begin{aligned} \frac{2\|(x(1-c) - a(1-z), y(1-c) - b(1-z))\|_{\max}}{(1-c)^2} &= \frac{2|x(1-c) - a(1-z)|}{(1-c)^2} \\ &= \frac{2|(x-a)(1-c) + a(z-c)|}{(1-c)^2} \leq \frac{2(|x-a||1-c| + |a||z-c|)}{(1-c)^2} \\ &< \frac{2\left(\frac{(1-c)^2}{4}\varepsilon + \frac{|a|(1-c)^2}{4\max\{|a|, |b|\}}\varepsilon\right)}{(1-c)^2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Therefore $\|\Phi(x, y, z) - \Phi(a, b, c)\|_{\max} < \varepsilon$, so Φ is continuous.

(c):

Let $p = (s, t) \in \mathbb{R}^2$. Our goal is to find $(x, y, z) \in S^2 \setminus \{N\}$ such that $\Phi(x, y, z) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right) = (s, t)$. Immediately, we obtain the following system of equations:

$$\begin{aligned} \frac{x}{1-z} &= s, \\ \frac{y}{1-z} &= t, \\ x^2 + y^2 + z^2 &= 1 \end{aligned}$$

We also have the restriction $z \neq 1$ because $(x, y, z) \neq N$. Isolating for x and y yields

$$\begin{aligned} x &= s(1-z) \\ y &= t(1-z) \end{aligned}$$

Then we substitute this into the third equation and get

$$s^2(1-z)^2 + t^2(1-z)^2 + z^2 = 1 \implies (s^2 + t^2 + 1)z^2 - 2(s^2 + t^2)z + s^2 + t^2 - 1 = 0$$

We can replace the term $t^2 + s^2$ with $\|p\|_2^2$, and the equation becomes

$$(\|p\|_2^2 + 1)z^2 - 2\|p\|_2^2 z + \|p\|_2^2 - 1 = 0$$

Using the quadratic formula:

$$\begin{aligned} z &= \frac{2(\|p\|_2^2) \pm \sqrt{4(\|p\|_2^2)^2 - 4(\|p\|_2^2 + 1)(\|p\|_2^2 - 1)}}{2(\|p\|_2^2 + 1)} \\ &\implies z = \frac{\|p\|_2^2 \pm \sqrt{\|p\|_2^4 - (\|p\|_2^4 - 1)}}{\|p\|_2^2 + 1} \\ &= \frac{\|p\|_2^2 \pm 1}{\|p\|_2^2 + 1} \end{aligned}$$

Notice that we cannot use the positive solution, for then

$$z = \frac{\|p\|_2^2 + 1}{\|p\|_2^2 + 1} = 1$$

Thus it must be true that

$$\begin{aligned} z &= \frac{\|p\|_2^2 - 1}{\|p\|_2^2 + 1} \\ x = s(1-z) &= \frac{2s}{\|p\|_2^2 + 1} \end{aligned}$$

$$y = \frac{2t}{\|p\|_2^2 + 1}$$

It can be verified that these values of x, y, z result in $\Phi(x, y, z) = (s, t)$, which show that Φ is surjective. Additionally, it was also proven that these values are the only ones that work, so Φ is bijective. In fact, using this, we obtain that the formula for Φ^{-1} is

$$\Phi^{-1}(s, t) = \left(\frac{2s}{\|p\|_2^2 + 1}, \frac{2t}{\|p\|_2^2 + 1}, \frac{\|p\|_2^2 - 1}{\|p\|_2^2 + 1} \right)$$

(d):

We know that Φ^{-1} is bijective because its inverse exists. It remains to show that it is continuous. This can be done by verifying pointwise continuity.

Let $(a, b) \in \mathbb{R}^2$. Considering the first component, fix $\varepsilon > 0$, and let $\delta = \frac{\varepsilon}{2}$. Let $(s, t) \in \mathbb{R}^2$ so that $\|(s - a, t - b)\|_2 < \delta$. Consider the case where $\frac{2s}{\|(s, t)\|_2^2 + 1} - \frac{2a}{\|(a, b)\|_2^2 + 1} > 0$. Then

$$\begin{aligned} \left| \frac{2s}{\|(s, t)\|_2^2 + 1} - \frac{2a}{\|(a, b)\|_2^2 + 1} \right| &= \frac{2s}{\|(s, t)\|_2^2 + 1} - \frac{2a}{\|(a, b)\|_2^2 + 1} < \frac{2s(\|(s, t)\|_2^2 + 1)}{\|(s, t)\|_2^2 + 1} - 2a \\ &= 2(s - a) < \varepsilon \end{aligned}$$

The case when $\frac{2s}{\|(s, t)\|_2^2 + 1} - \frac{2a}{\|(a, b)\|_2^2 + 1} \leq 0$ is the exact same but the variables are swapped. Thus the first component is continuous. We can do the same thing for the second component so we can accept that it is continuous as well.

Finally, notice that since norms are continuous, so the third component is a composition of continuous functions, making it continuous. Therefore Φ^{-1} is continuous.

We can conclude that Φ is indeed a homeomorphism and we are done. \square

1.21 #24

Let X be a normed vector space. Prove that the following statements are equivalent.

- (i) X is finite-dimensional.
- (ii) The unit ball $\overline{B}(\vec{0}, 1)$ is compact.
- (iii) X is **locally compact**: each point $p \in X$ is contained in some open set U such that \overline{U} is compact.

Proof. It will be proven that (i) \implies (ii) \iff (iii) \implies (i).

(i) \implies (ii):

Suppose that X has finite dimension n . Then there is a continuous linear isomorphism Φ between X and \mathbb{R}^n . Since $\overline{B}(0, 1)$ is closed and bounded, $\Phi(\overline{B}(0, 1))$ is also closed and bounded in \mathbb{R}^n , so the set is compact. Since homeomorphisms preserve compactness, we can conclude that the closed unit ball in X is compact.

(ii) \implies (iii):

Suppose that the unit ball $\overline{B}(\vec{0}, 1)$ is compact. Let $p \in X$. We claim that $U = B(p, 1)$. Consider $\overline{U} = \overline{B}(p, 1)$. There is an isometry Φ from $\overline{B}(0, 1)$ to $\overline{B}(p, 1)$ defined by $\Phi(x) = p + x$. Since the closed unit ball is compact, it follows that $\overline{B}(p, 1)$ is compact. Thus X is locally compact.

(iii) \implies (i):

Suppose that X is locally compact. Then $\vec{0}$ is contained in an open set U such that \overline{U} is compact. Since U is open, $B(0, \varepsilon) \subseteq U$ for some ε -ball centered around 0. It follows that $\overline{B}(0, \varepsilon)$ is compact, as it is a closed subset of \overline{U} . Since there is a homeomorphism from this closed ball to $\overline{B}(0, 1)$, the closed unit ball is also compact, as desired.

(iii) \implies (i):

Suppose that X is locally compact. We know previously that this implies that the closed unit ball is compact. Firstly, a quick lemma will be proven.

Lemma 1. $\forall x \in X, r > 0, B(x, r) = \{x\} + rB(0, 1)$.

Let $y \in B(x, r)$. Notice that $y - x \in B(0, r)$. Thus $y = x + (y - x)$ so $y \in \{x\} + rB(0, 1)$.

Then, let $y \in \{x\} + B(0, r)$, so $y = x + s$, for some $s \in B(0, r)$. Since $\|y - x\| = \|s\| < r$, it follows that $y \in B(x, r)$ and we are done.

Moving on, we construct a finite set of vectors in the following way:

Since $\overline{B}(0, 1)$ is totally bounded, we can find a finite set of vectors β such that $\bigcup_{x \in \beta} B(x, \frac{1}{2})$.

Lemma 2. $B(0, 1) \subseteq \text{span}(\beta) + 2^{-n}B(0, 1)$, $\forall n \in \mathbb{N}$.

To prove this, let $n \in \mathbb{N}$ and $y \in B(0, 1)$. It follows that $y \in B(x_1, \frac{1}{2})$ for some $x_1 \in \beta$. Using Lemma 1,

$$y \in \{x_1\} + B\left(0, \frac{1}{2}\right) = \{x_1\} + \frac{1}{2}B(0, 1)$$

We can repeatedly use this argument to obtain that

$$y \in \left\{x_1 + \frac{1}{2}x_2 + \frac{1}{2^2}x_3 + \cdots + \frac{1}{2^{n-1}}x_n\right\} + \frac{1}{2^n}B(0, 1)$$

Note that x_i are not necessarily distinct. This simplifies down to

$$y \in \text{span}(\beta) + \frac{1}{2^n}B(0, 1) \implies B(0, 1) \subseteq \text{span}(\beta) + \frac{1}{2^n}B(0, 1)$$

which is what we wanted.

Finally, we claim that β is the basis for X .

Let $v \in X$. Since $\frac{v}{2\|v\|} \in B(0, 1)$, by Lemma 2,

$$\frac{v}{2\|v\|} \in \text{span}(\beta) + \frac{1}{2^n}B(0, 1), \text{ for all } n \in \mathbb{N}$$

Then we construct a sequence $\{v_n\}_{n \geq 1}$ in $\text{span}(\beta)$ where $v_n \in \text{span}(\beta)$ such that $y = v_n + p$, for some $p \in \frac{1}{2^n}B(0, 1)$. Notice that

$$\|y - v_n\| = \|p\| < \frac{1}{2^n}$$

which can become arbitrarily small. This implies that y is a limit point for $\text{span}(\beta)$. However, this set is closed because it is finite dimensional, so we also have that $y \in \text{span}(\beta)$. Therefore $X \subseteq \text{span}(\beta)$, and since X is spanned by a finite set, it is finite dimensional.

These implications are sufficient for proving that (i), (ii), (iii) are equivalent. \square

1.22 #25

Let $\varphi : M_n(\mathbf{R}) \rightarrow M_n(\mathbf{R})$ be the function given by $\varphi(A) = A^2$. For each $A \in M_n(\mathbf{R})$, find a linear approximation $L_A : M_n(\mathbf{R}) \rightarrow M_n(\mathbf{R})$ to φ at A . Give an explicit formula for $L_A(B)$ as a function of B , a proof that L_A is a bounded linear mapping, and a proof that L_A is a linear approximation to φ at A .

Proof. First, we supply a lemma.

Lemma. For all $B \in M_n(\mathbb{R})$, $\|B^2\| \leq K\|B\|^2$, for some positive constant K .

Define the isomorphism $\Phi : M_n(\mathbb{R}) \rightarrow B(\mathbb{R}^n, \mathbb{R}^n)$ as mapping a matrix representation of a linear mapping to the original linear mapping. We define the operator norm on $M_n(\mathbb{R})$ by $\|A\|_{\text{op}} = \|\Phi(A)\|_{\text{op}}$, where the right hand side is the operator norm on $B(\mathbb{R}^n, \mathbb{R}^n)$.

Since all norms are equivalent on $M_n(\mathbb{R})$, there are constants $M, N > 0$ so that for any norm $\|\cdot\|$,

$$M\|A\| \leq \|A\|_{\text{op}} \leq N\|A\|$$

From this, using the subnormality of bounded linear operators, it follows that

$$\|B^2\| \leq \frac{1}{M}\|\Phi(B^2)\|_{\text{op}} = \frac{1}{M}\|\Phi(B) \circ \Phi(B)\|_{\text{op}} \leq \frac{1}{M}\|\Phi(B)\|_{\text{op}}^2 \leq \frac{N^2}{M}\|B\|^2$$

Since $M, N > 0$, we have what we wanted.

We claim that for $A \in M_n(\mathbb{R})$, $L_A(B) = BA + AB$. For $C, D \in M_n(\mathbb{R})$, $k \in \mathbb{R}$,

$$L_A(kC + D) = (kC + D)A + A(kC + D) = k(CA + AC) + DA + AD = kL_A(C) + L_A(D)$$

so L_A is linear. As well, we get that L_A is bounded for free because we are working in a finite dimensional vector space. Finally, we have that

$$\begin{aligned} 0 &\leq \frac{\|\varphi(A+B) - \varphi(A) - L_A(B)\|}{\|B\|} = \frac{\|(A+B)^2 - A^2 - (BA+AB)\|}{\|B\|} \\ &= \frac{\|A^2 + AB + BA + B^2 - A^2 - BA - AB\|}{\|B\|} = \frac{\|B^2\|}{\|B\|} \leq K\|B\| \\ \implies 0 &\leq \frac{\|\varphi(A+B) - \varphi(A) - L_A(B)\|}{\|B\|} \leq K\|B\| \end{aligned}$$

By the Squeeze Theorem, $\lim_{h \rightarrow 0} \frac{\|\varphi(A+B) - \varphi(A) - L_A(B)\|}{\|B\|} = 0$ and we are done. \square

1.23 #26

Let X be a finite-dimensional normed vector space, let U be an open convex subset of X , and let $f : U \rightarrow \mathbf{R}^m$ be a totally differentiable function. (Note: a set $C \subseteq X$ is called **convex** if $tx + (1-t)y \in C$ for all $x, y \in C$ and $t \in [0, 1]$.) Let $f : U \rightarrow \mathbf{R}^m$ be a totally differentiable function.

- (a) Suppose that there exists a constant $C \geq 0$ such that $\|f'(p)\|_{\text{op}} \leq C$ for all $p \in U$. Prove that

$$\|f(p) - f(q)\| \leq C\|p - q\| \quad \text{for all } p, q \in U.$$

Conclude that f is uniformly continuous.

- (b) Prove that $f'(p) = 0$ for all $p \in U$ if and only if f is a constant function.

- (c) Assume $U = X$ and suppose that f is **twice totally differentiable** — meaning that $f' : X \rightarrow B(X, Y)$ itself is differentiable at every point of X , with total derivative $f'' = (f')'$. Show that $f'' = 0$ if and only if f is **affine-linear**: there exists a bounded linear mapping $M : X \rightarrow Y$ and a vector $b \in Y$ such that

$$f(p) = M(p) + b \quad \text{for all } p \in X.$$

(Compare with the formula $y = mx + b$ from single-variable calculus.)

Proof. (a):

Fix $p, q \in U$. If $f(p) = f(q)$, the inequality is trivially true. Otherwise, since U is open, there is some open ball $B(p, \varepsilon_1) \subseteq U$ and $B(q, \varepsilon_2) \subseteq U$. Let $\delta = \frac{\min\{\varepsilon_1, \varepsilon_2\}}{\|q - p\|}$. For all $t \in (-\delta, 1 + \delta)$, we have that

If $t \in [0, 1]$,

$$tq + (1-t)p \in U$$

since U is convex. If $t \in (-\delta, 0)$, then

$$\|(tq + (1-t)p) - p\| = \|t(q - p)\| = |t|\|q - p\| < \varepsilon_1$$

$$\implies tq + (1-t)p \in B(p, \varepsilon_1) \subseteq U$$

If $t \in (1 + \delta)$, then

$$\|(tq + (1-t)p) - q\| = \|(t-1)q - (t-1)p\| = |t-1|\|q - p\| < \varepsilon_2$$

$$\implies tq + (1-t)p \in B(q, \varepsilon_2) \subseteq U$$

Thus we can construct a function $\alpha : (-\delta, 1 + \delta) \rightarrow U$ defined by $\alpha(t) = tq + (1 - t)p$. We will show that $\alpha'(t)(\varphi) = \varphi(q - p)$:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\|((t+h)q + (1-t-h)p) - (tq + (1-t)p) - h(q-p)\|}{\|h\|} \\ = \lim_{h \rightarrow 0} \frac{\|hq - hp - h(q-p)\|}{\|h\|} = 0 \end{aligned}$$

Since α is totally differentiable and f is totally differentiable, by the chain rule, we can say that

$$(f \circ \alpha)'(t) = f'(\alpha(t))(h(q-p)) = hf'(\alpha(t))(q-p)$$

Now, we construct a function from $(-\delta, 1 + \delta)$ to \mathbb{R} by

$$g(t) = (f(q) - f(p)) \cdot f(\alpha(t))$$

Now, we will show that for $t \in (-\delta, 1 + \delta)$,

$$g'(t) = (f(q) - f(p)) \cdot f'(\alpha(t))(q-p)$$

First, we see that

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{|g(t+h) - g(t) - (f(q) - f(p)) \cdot f'(\alpha(t))(q-p)|}{|h|} \\ = \lim_{h \rightarrow 0} \frac{|(f(q) - f(p)) \cdot f(\alpha(t+h)) - (f(q) - f(p)) \cdot f(\alpha(t)) - (f(q) - f(p)) \cdot f'(\alpha(t))(q-p)|}{|h|} \end{aligned}$$

By the linearity of the dot product,

$$\lim_{h \rightarrow 0} \left| (f(q) - f(p)) \cdot \frac{f(\alpha(t+h)) - f(\alpha(t)) - f'(\alpha(t))(q-p)}{|h|} \right| = 0$$

Using this, consider the limit

$$g'(t) = \lim_{h \rightarrow 0} \frac{g(t+h) - g(t)}{h}.$$

Adding and subtracting $h(f(q) - f(p)) \cdot f'(\alpha(t))(q-p)$,

$$= \lim_{h \rightarrow 0} \frac{[g(t+h) - g(t) - h(f(q) - f(p)) \cdot f'(\alpha(t))(q-p)] + h(f(q) - f(p)) \cdot f'(\alpha(t))(q-p)}{h}$$

The first sum is exactly the limit that was computed above and is equal to zero. The limit then reduces to

$$\lim_{h \rightarrow 0} \frac{h(f(q) - f(p)) \cdot f'(\alpha(t))(q-p)}{h} = (f(q) - f(p)) \cdot f'(\alpha(t))(q-p)$$

which shows that g is differentiable on its domain and is equal to $(f(q) - f(p)) \cdot f'(\alpha(t))(q-p)$. In particular, we have continuity on $[0, 1]$ and differentiability on $(0, 1)$. Applying the Mean Value Theorem, there is some $c \in (0, 1)$ so that

$$\begin{aligned} g(1) - g(0) = g'(c) &\implies (f(q) - f(p)) \cdot f(q) - (f(q) - f(p)) \cdot f(p) = (f(q) - f(p)) \cdot f'(\alpha(c))(q-p) \\ &\implies (f(q) - f(p)) \cdot (f(q) - f(p)) = (f(q) - f(p)) \cdot f'(\alpha(c))(q-p) \\ &\implies \|f(q) - f(p)\|^2 = \|f(q) - f(p)\| \|f'(\alpha(c))(q-p)\| \cos(\theta) \end{aligned}$$

where θ is the angle between the vectors $f(q) - f(p)$ and $f'(\alpha(c))(q-p)$. Thus

$$\begin{aligned} \|f(q) - f(p)\|^2 &\leq \|f(q) - f(p)\| \|f'(\alpha(c))(q-p)\| \leq \|f(q) - f(p)\| \|f'(\alpha(c))\|_{\text{op}} \|q-p\| \\ &\implies \|f(q) - f(p)\|^2 \leq \|f(q) - f(p)\| C \|q-p\| \end{aligned}$$

$$\implies \|f(q) - f(p)\| \leq C\|q - p\|$$

which is the desired inequality.

With this, the uniform continuity of f is easy to show.

For $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{C}$. For all $p, q \in X$ so that $\|q - p\| < \delta$,

$$\|f(q) - f(p)\| \leq C\|q - p\| < \varepsilon$$

(b):

Suppose that $f'(p) = 0$ for all $p \in U$. Then $\|f'(p)\|_{\text{op}} \leq 0 = C$, so by part (a), for all $a, b \in U$,

$$\|f(a) - f(b)\| \leq 0 \implies \|f(a) - f(b)\| = 0 \implies f(a) = f(b)$$

so f is constant.

Conversely, suppose that f is a constant function. To show that $f'(p) = 0$, notice that

$$\lim_{h \rightarrow 0} \frac{\|f(p+h) - f(p)\|}{\|h\|} = 0$$

Thus $f'(p) = 0$ for all $p \in U$.

(c):

We know that X is convex because it is a vector space. Suppose that $f'' = 0$. Then by part (b), f' is a constant function. We will denote $f' = L \in B(X, Y)$. We claim that $M = L$ and $b = f(0)$, that is, $f(p) = L(p) + f(0)$ for all $p \in U$.

Define the function $g : U \rightarrow \mathbb{R}^m$ by $g(p) = f(p) - f(0) - L(p)$. For all $p \in U$, we have that

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\|g(p+h) - g(p)\|}{\|h\|} &= \lim_{h \rightarrow 0} \frac{\|(f(p+h) - f(0) - L(p+h)) - (f(p) - f(0) - L(p))\|}{\|h\|} \\ &= \lim_{h \rightarrow 0} \frac{\|f(p+h) - f(p) - L(h)\|}{\|h\|} = 0 \end{aligned}$$

where the last step is from the fact that $f'(p) = L$. We can see that $g'(p) = 0$ for all $p \in U$. From part (b), it follows that g is a constant function. Finally, notice that

$$g(0) = f(0) - f(0) - L(0) = 0,$$

so $g = 0$. Therefore for all $p \in U$,

$$f(p) - f(0) - L(p) = 0 \implies f(p) = L(p) + f(0)$$

as desired.

Conversely, suppose that f is affine-linear. Then for some $M \in B(X, \mathbb{R}^m)$, $b \in \mathbb{R}^m$,

$$f(p) = M(p) + b.$$

We claim that $f'(p) = M$ using the definition. Immediately, we have that

$$\lim_{h \rightarrow 0} \frac{\|(M(p+h) + b) - (M(p) + b) - M(h)\|}{\|h\|} = 0$$

which verifies our claim. Since $f'(p) = M$ for all $p \in U$, f' is constant, and from part (b), this implies that $f'' = 0$, which completes the proof. \square

1.24 #27*

I always found integration fun, because it was similar to solving some kind of puzzle. I have known about Feynmann's trick for a long time, but I never thought about how it ties in with multivariable calculus. Completing this problem was very interesting as I got to see how this trick really worked.

Let $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ be a continuously differentiable function.

(a) Show that the partial function $\mathbf{R} \rightarrow \mathbf{R}$, $t \mapsto f(x, t)$ is integrable (over any bounded interval in \mathbf{R}).

(b) By (a), we can define a function $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ by

$$\varphi(x) = \int_a^b f(x, t) dt.$$

Show that φ is differentiable, and that its (classical) derivative is given by

$$\frac{d\varphi}{dx}(x_0) = \int_a^b \frac{\partial f}{\partial x}(x_0, t) dt.$$

This formula is known as **differentiation under the integral sign**, or **Feynmann's trick**.

(c) Use Feynmann's trick to solve the single-variable integral:

$$\int_0^\infty e^{-t^2} dt$$

Proof. (a):

Since f is continuous, it follows immediately that its partial function is continuous, which implies that it is integrable.

(b):

Let $x_0 \in \mathbf{R}$. We will show that as $h \rightarrow 0$,

$$\frac{1}{h} \left(\int_a^b f(x_0 + h, t) dt - \int_a^b f(x_0, t) dt - h \int_a^b \frac{\partial f}{\partial x}(x_0, t) dt \right) \rightarrow 0,$$

which is equivalent to saying

$$\frac{d\varphi}{dx}(x_0) = \int_a^b \frac{\partial f}{\partial x}(x_0, t) dt.$$

Let $\varepsilon > 0$. By the partial differentiability of f , we obtain a δ so that

$$\left| f(x_0 + h, t) - f(x_0, t) - h \frac{\partial f}{\partial x}(x_0, t) \right| < \frac{|h|\varepsilon}{b-a}$$

for all $0 < |h| < \delta$. Since t is contained in a closed interval $[a, b]$, so the limit converges uniformly and we are guaranteed to be able to find a δ that works for all $t \in [a, b]$.

Fix $h \in \mathbf{R}$ so that $0 < |h| < \delta$. By the linearity of the integral,

$$\begin{aligned} & \left| \frac{1}{h} \left(\int_a^b f(x_0 + h, t) dt - \int_a^b f(x_0, t) dt - h \int_a^b \frac{\partial f}{\partial x}(x_0, t) dt \right) \right| \\ &= \left| \frac{1}{h} \int_a^b (f(x_0 + h, t) - f(x_0, t) - h \frac{\partial f}{\partial x}(x_0, t)) dt \right| \\ &\leq \frac{1}{|h|} \int_a^b \left| f(x_0 + h, t) - f(x_0, t) - h \frac{\partial f}{\partial x}(x_0, t) \right| dt \end{aligned}$$

$$< \frac{1}{|h|} \int_a^b \frac{|h|\varepsilon}{b-a} dt = \varepsilon$$

as desired. Thus we have found an expression for $\varphi'(x_0)$.

(c):

Let $I = \int_0^\infty e^{-t^2} dt$. We will first show that this integral converges. We make the observation that the integrand is continuous and positive. Splitting the integral in two, we see that

$$\int_0^\infty e^{-t^2} dt = \int_0^1 e^{-t^2} dt + \int_1^\infty e^{-t^2} dt$$

The first integral has finite bounds so it is convergent. For $t \geq 1$, $-t^2 \leq -t$. Since $\exp(t)$ is an increasing function,

$$e^{-t^2} \leq e^{-t} \implies \int_1^\infty e^{-t^2} dt \leq \int_1^\infty e^{-t} dt = 1$$

Therefore I converges by the basic comparison test.

Define $\varphi : [0, \infty) \rightarrow \mathbb{R}$ by

$$\varphi(x) = \int_0^\infty \frac{e^{-x^2(t^2+1)}}{t^2+1} dt$$

Notice that $\int_0^\infty \frac{e^{-x^2(t^2+1)}}{t^2+1} dt \leq \int_0^\infty e^{-x^2(t^2+1)} dt$, so φ always converges for similar reasons.

From part (b), we see that

$$\varphi'(x) = -2x \int_0^\infty e^{-x^2(t^2+1)} dt = -2xe^{-x^2} \int_0^\infty e^{-x^2 t^2} dt$$

We perform the substitution $u = xt$. Changing u back to t gives us

$$\varphi'(x) = -2e^{-x^2} \int_0^\infty e^{-t^2} dt$$

Thus we have that

$$\varphi'(x) = -2e^{-x^2} I$$

We can take the definite integral of both sides with respect to x :

$$\int_0^\infty \varphi'(x) dx = \int_0^\infty -2e^{-x^2} I dx \implies \lim_{n \rightarrow \infty} \varphi(n) - \varphi(0) = -2I^2$$

Now we will analyse $\lim_{n \rightarrow \infty} \varphi(n)$ and $\varphi(0)$ separately.

First, we will show that $\lim_{n \rightarrow \infty} \varphi(n) = 0$. Let $\varepsilon > 0$. For $n > N$, where N is a fixed number, we have that $e^{-n^2(t^2+1)} < \frac{2\varepsilon}{\pi}$, so

$$\begin{aligned} |\varphi(n)| &= \left| \int_0^\infty \frac{e^{-n^2(t^2+1)}}{t^2+1} dt \right| \leq \int_0^\infty \left| \frac{e^{-n^2(t^2+1)}}{t^2+1} \right| dt < \frac{2}{\pi} \int_0^\infty \frac{\varepsilon}{t^2+1} dt \\ &\implies \frac{2\varepsilon}{\pi} \arctan(t) \Big|_0^\infty = \varepsilon. \end{aligned}$$

Thus we can conclude that $\lim_{n \rightarrow \infty} \varphi(n) = 0$.

Now, notice that

$$\varphi(0) = \int_0^\infty \frac{1}{t^2+1} dt = \arctan(t) \Big|_0^\infty = \frac{\pi}{2}$$

Applying this to our original equation,

$$0 - \frac{\pi}{2} = -2I^2 \implies 2I^2 = \frac{\pi}{2} \implies I^2 = \frac{\pi}{4}$$

Finally, taking the squareroot of both sides gives us the result:

$$\int_0^\infty e^{-t^2} dt = I = \frac{\sqrt{\pi}}{2}$$

□

1.25 #28

Let U be an open set in a normed vector space X and let $f : U \rightarrow Y$ be a **twice continuously differentiable** function, meaning that the second derivative $f'' : U \rightarrow B(X, B(X, Y))$ exists and is continuous on U . We also say that f is a **C^2 -function**.

- (a) Let $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ be given by $f(x, y) = x^2 - xy + y^2$. Find, with proof, an explicit formula for the linear mapping $f''(2, -1)$. Also, write down the matrix that represents this linear mapping with respect to a suitable “standard” basis.
- (b) Now we investigate the case $X = \mathbf{R}^n$ and $Y = \mathbf{R}$, and let $f : U \rightarrow \mathbf{R}$ be some function defined on an open set $U \subseteq \mathbf{R}^n$. We use the notation $\frac{\partial^2 f}{\partial x_i \partial x_j}$ to refer to the (i, j) **th second partial derivative** of f : this is the i th partial derivative of the j th partial derivative $\frac{\partial f}{\partial x_j}$.
- (i) Show that f is twice continuously differentiable if and only if all second partial derivatives exist and are continuous.
- (ii) Let f be twice continuously differentiable. Let $v \in \mathbf{R}^n$ and let $D_v f : U \rightarrow \mathbf{R}$ be the directional derivative of f along v . Show that $D_v f$ is continuously differentiable.
- (iii) Let f be twice continuously differentiable and let $v \in \mathbf{R}^n$. By (ii), we know that $D_v f$ is C^1 , hence differentiable in every direction $\in \mathbf{R}^n$. Show that the directional derivatives commute:

$$D_v(D_w f) = D_w(D_v f) \quad \text{for all } v, w \in \mathbf{R}^n.$$

- (iv) Deduce **Clairaut’s Theorem**: that the second partial derivatives commute.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \quad \text{for all } i, j \in \{1, \dots, n\}.$$

Proof. (a):

We claim that $f''(2, -1) : \mathbf{R}^2 \rightarrow B(\mathbf{R}^2, \mathbf{R})$ is a bounded linear map given by

$$f''(2, -1)(p, q)(x, y) = (2p - q)x - (p - 2q)y$$

Let $(p, q), (r, s) \in \mathbf{R}^2$ and $c \in \mathbf{R}$. Then for all $(x, y) \in \mathbf{R}^2$,

$$\begin{aligned} f''(2, -1)(cp + r, cq + s)(x, y) &= [2(cp + r) - (cq + s)]x - [(cp + r) - 2(cq + s)]y \\ &= c[(2p - q)x - (p - 2q)y] + [(2r - s)x - (r - 2s)y] = cf''(2, -1)(p, q)(x, y) + f''(2, -1)(r, s) \end{aligned}$$

so $f''(2, -1)$ is linear.

Recall that $f'(p, q)(x, y) = (2p - q)x - (p - 2q)y$. In particular, for $(p, q) = (2, -1)$, $f'(2, -1)(x, y) = 5x - 4y$. Fix $(x, y) \in \mathbf{R}^2$. Then

$$\begin{aligned} &\lim_{h \rightarrow 0} \frac{f'(2 + h_1, -1 + h_2)(x, y) - f'(2, -1)(x, y) - L_p(h)}{\|h\|} \\ &= \lim_{h \rightarrow 0} \frac{[(5 + 2h_1 - h_2)x - (4 + h_1 - 2h_2)y] - [5x - 4y] - L_p(h)}{\|h\|} \\ &= \lim_{h \rightarrow 0} \frac{(2h_1 - h_2)x - (h_1 - 2h_2)y - L_p(h)}{\|h\|} = 0 \end{aligned}$$

which verifies that $f''(2, -1)(p, q)(x, y) = (2p - q)x - (p - 2q)y$.

(b):

Let $f : U \rightarrow \mathbf{R}$ be a function, where U is open in \mathbf{R}^n .

(b)(i):

Suppose that f is twice continuously differentiable. This implies that all its partial derivatives exist.

Consider an arbitrary partial derivative $\frac{\partial f}{\partial x_j}$ for $i \in \{1, \dots, n\}$. It will be shown that their directional derivatives all exist and are continuous.

Let $v \in \mathbb{R}^n$. We see that

$$\lim_{t \rightarrow 0} \frac{1}{t} \left(\frac{\partial f}{\partial x_j}(p + tv) - \frac{\partial f}{\partial x_j}(p) \right)$$

Recall that $f'(q)(v) = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(q)$. Now, notice that

$$\frac{\partial f}{\partial x_j}(p + tv) = f'(p + tv)(x_j) \text{ and } \frac{\partial f}{\partial x_j}(p) = f'(p)(x_j)$$

We make this substitution in the limit expression, so it becomes

$$\lim_{t \rightarrow 0} \frac{1}{t} (f'(p + tv)(x_j) - f'(p)(x_j)) = D_v f'(p)(x_j)$$

by definition. Thus $D_v \left(\frac{\partial f}{\partial x_j}(p) \right) = D_v f'(p)(x_j)$. This is continuous because every directional derivative of f' is continuous by our assumption. Taking $v = x_i$, we conclude that all the second partial derivatives of f exist and are continuous.

Conversely, suppose that all second partial derivatives of f exist and are continuous. Since all first partial derivatives will exist and are continuous, then f is continuously differentiable, meaning that f' exists.

Let $i \in \{1, \dots, n\}$. We claim that $\frac{\partial f'}{\partial x_i}$ exists and is continuous, where x_i represents the i th standard ordered basis vector of \mathbb{R}^n .

Fix $p \in \mathbb{R}^n$. Recall that $f'(p)(\vec{v}) = \sum_{j=1}^n v_j \frac{\partial f}{\partial x_j}(p)$ from the differentiability theorem. For $\vec{v} \in \mathbb{R}^n$, we have that

$$\lim_{h \rightarrow 0} \frac{f'(p + hx_i)(\vec{v}) - f'(p)(\vec{v})}{h} = \lim_{h \rightarrow 0} \sum_{j=1}^n \frac{v_j}{h} \left(\frac{\partial f}{\partial x_j}(p + hx_i) - \frac{\partial f}{\partial x_j}(p) \right)$$

Since all second partial derivatives exist, the limit evaluates to

$$\sum_{j=1}^n v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(p)$$

By definition, this is the i th partial derivative of f' .

Notice that it is also continuous, as it is a finite linear combination of second partial derivatives, which are continuous by assumption.

Since we have that all partial derivatives of f' exist and are continuous, by the differentiability theorem, f' is continuously differentiable, which implies that f is twice continuously differentiable.

Therefore we can conclude that twice continuously differentiable is equivalent to having all second partial derivatives exist and continuous.

(b)(ii):

Suppose that f is twice continuously differentiable. Consider the directional derivative $D_v f$. We will show that for all $p \in U$, $D_v f(p) = f'(p)(v)$.

We can assume that $p + v \in U$ by making $\|v\|$ small enough. Define a function $\alpha : [0, 1] \rightarrow U$ by $\alpha(t) = p + tv$. The derivative is given by $\alpha'(t) = v$. Now, construct a new function $g : \mathbb{R} \rightarrow \mathbb{R}$ defined by $g(t) = (f \circ \alpha)(t)$. By the chain rule, we see that $g'(t) = f'(\alpha(t))(v)$.

Thus we get that

$$D_v f(p) = \lim_{h \rightarrow 0} \frac{f(p + hv) - f(p)}{h} = \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h} = g'(0) = f'(\alpha(0))(v) = f'(p)(v),$$

Now, we can compute the derivative of $D_v f$ directly and obtain that $(D_v f)'(p) = (f')'(p)(\cdot)(v)$. But we know by that f is twice continuously differentiable, so $(D_v f)'$ is continuous.

Thus $D_v f$ is continuously differentiable.

(b)(iii):

We will evaluate $D_v(D_w f)$ from the definition and show that it is equal to $D_w(D_v f)$.

Fix $p \in U$. There exists $\delta > 0$ so that $p + \delta v \in U$. For any such h , construct a function $g : [0, \delta] \rightarrow \mathbb{R}$ defined as

$$g(t) = D_w f(p + tv)$$

g is differentiable on its domain because it is a composition of differentiable functions. By the chain rule, we see that

$$g'(t) = (D_w f)'(p + tv)(v)$$

Using the fact that for all continuously differentiable functions, $D_v h(q) = h'(q)(v)$, we get

$$g'(t) = D_v(D_w f)(p + tv)$$

Now we compute $D_v(D_w f)$:

$$\lim_{h \rightarrow 0} \frac{D_w f(p + hv) - D_w f(p)}{h}$$

For $|h| < \delta$, the limit becomes

$$D_w(D_v f)(p) = \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h} = g'(0) = D_v(D_w f)(p)$$

Therefore the directional derivatives commute.

(b)(iv):

From the previous part, simply take $v = e_i$ and $w = e_j$ and the result follows immediately. \square

1.26 #29

Let $U \subseteq \mathbf{R}^2$ be an open set, and let $f : U \rightarrow \mathbf{R}$ be some differentiable function. This data defines an **explicit surface** in \mathbf{R}^3 , also known as the **graph** of f :

$$S = S(f) = \{(x, y, z) \in \mathbf{R}^3 : (x, y) \in U, z = f(x, y)\}.$$

Colloquially, we often say “Let $z = f(x, y)$ be a surface in \mathbf{R}^3 .”

- (a) Find the equation of the tangent plane to the surface $z = x + xy^2 - y^3$ at the point $p = (2, 1, 3)$, as follows.

(i) First, fix $x = 2$ and set $y = 1 + t$; write z as a function of t , and find $z'(0)$. This is the “slope in the x direction.”

(ii) Similarly, find the “slope in the y direction.”

(iii) Now you have two slopes in orthogonal directions. This gives you two vectors which span a plane. Shift this plane so that it becomes tangent at p . Write the equation of this plane in the form $Ax + By + Cz = D$, where $A, B, C, D \in \mathbf{R}$.

- (b) Find — with proof! — the equations of all planes in \mathbf{R}^3 which (i) are tangent to the surface $z = x + xy^2 - y^3$; (ii) are parallel to the vector $\vec{v} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$; and (iii) pass through the point $p = (-1, -2, 3)$.

- (c) Let $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ be a continuously differentiable function, and define a new function $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ by

$$g(x, y) = f(f(xy, x), f(y, xy)).$$

You are given the following information about f and g :

- $f(3, 1) = 5$ and $\nabla f(3, 1) = (1, 2)$
- $f(3, 3) = 2$ and $\nabla f(3, 3) = (0, -1)$
- $g(1, 3) = 6$ and $\nabla g(1, 3) = (3, 4)$

Find $\nabla f(5, 2)$.

Proof. (a)(i):

Fix $x = 2$ and set $y = 1 + t$. Then

$$\begin{aligned} z_y(t) &= 2 + 2(1+t)^2 - (1+t)^3 \implies z'_y(t) = 4(1+t) - 3(1+t)^2 \\ &\implies z'_y(0) = 1 \end{aligned}$$

(a)(ii):

Fix $y = 1$ and set $x = 2 + t$. Then

$$z_x(t) = (2+t) + (2+t) - 1 = 2t + 3 \implies z'_x(0) = 2$$

(a)(iii):

From the previous two parts gave us two vectors $(1, 0, 2)$ and $(0, 1, 1)$. This spans the plane given by the equation $2x + y - z = 0$. Translate this plane by the vector $(2, 1, 3)$ which yields the equation

$$2(x-2) + (y-1) - (z-3) = 0 \implies 2x + y - z = 2$$

(b):

We claim that there are only two planes that satisfy all three conditions:

$$5x - 14y - z = 20 \text{ and } 5x - 11y - 4z = 5$$

Let $(p, q) \in \mathbb{R}^n$ and suppose that the plane tangent to the surface z at this point has the properties above. We find the equation of this tangent plane like the previous part.

Fix $x = p$ and set $y = q + t$. Then

$$\begin{aligned} z_y(t) &= p + p(q+t)^2 - (q+t)^3 \implies z'_y(t) = 2p(q+t) - 3(q+t)^2 \\ &\implies z'_y(0) = 2pq - 3q^2 \end{aligned}$$

Next, fix $y = q$ and set $x = p + t$. Then

$$z_x(t) = (x+t) + (x+t)q^2 - q^3 \implies z'_x(t) = q^2 + 1$$

This gives two vectors $(0, 1, 2pq - 3q^2)$ and $(1, 0, q^2 + 1)$ that span a plane given by the equation $(q^2 + 1)x + (2pq - 3q^2)y - z = 0$. Translating this plane by $(p, q, p + pq^2 - q^3)$ gives another plane with the equation

$$\begin{aligned} (q^2 + 1)(x-p) + (2pq - 3q^2)(y-q) - (z + q^3 - pq^2 - p) &= 0 \\ \implies (q^2 + 1)x + (2pq - 3q^2)y - z &= p(q^2 + 1) + q(2pq - 3q^2) + (q^3 - pq^2 - p) \\ \implies (q^2 + 1)x + (2pq - 3q^2)y - z &= 2pq^2 - 2q^3 \end{aligned}$$

Now, take the initial vectors $(0, 1, 2pq - 3q^2)$ and $(1, 0, q^2 + 1)$ and compute their cross product to find the normal vector \vec{n} :

$$\vec{n} = (0, 1, 2pq - 3q^2) \times (1, 0, q^2 + 1) = (q^2 + 1, 2pq - 3q^2, -1)$$

This vector is normal to the plane. We only want planes that are parallel to \vec{v} , which means that this normal vector should be normal to \vec{v} as well. We calculate the dot product of this normal vector and \vec{v} :

$$\vec{n} \cdot \vec{v} = 3(q^2 + 1) + (2pq - 3q^2) - 1$$

Set this equal to 0. Simplifying, we have that

$$\begin{aligned} 3(q^2 + 1) + (2pq - 3q^2) - 1 = 0 &\implies 2pq = -2 \\ &\implies pq = -1 \end{aligned}$$

We see that $p, q \neq 0$ and get the relationship $p = -\frac{1}{q}$. Substitute this into our plane equation and we get that

$$\begin{aligned} (q^2 + 1)x + (2pq - 3q^2)y - z &= 2pq^2 - 2q^3 \\ \implies (q^2 + 1)x + (-2 - 3q^2)y - z &= -2q - 2q^3 \\ \implies (q^2 + 1)x - (3q^2 + 2)y - z &= -2q(q^2 + 1) \end{aligned}$$

Finally, recall that our plane should pass through the point $p = (-1, -2, 3)$. Substituting $(x, y, z) = (-1, -2, 3)$, we can solve for q and obtain our answer:

$$\begin{aligned} -(q^2 + 1) + 2(3q^2 + 2) - 3 &= -2q(q^2 + 1) \\ \implies 2q^3 + 5q^2 + 2q &= 0 \end{aligned}$$

We disregard the case where $q = 0$ and end up with the quadratic equation

$$2q^2 + 5q + 2 = 0 \implies q = \frac{-5 \pm \sqrt{25 - 16}}{4} \implies q_1 = -2, q_2 = -\frac{1}{2}$$

Substituting these values into our plane equation, we see that there are two planes that satisfy the above properties:

$$5x - 14y - z = 20 \text{ and } 5x - 11y - 4z = 5$$

(c):

In general,

$$\nabla g(p, q) = \left(\frac{\partial g}{\partial x}(p, q), \frac{\partial g}{\partial y}(p, q) \right)$$

For $a(x, y) = xy$, $b(x, y) = x$, and $c(x, y) = y$, we make the quick note that

$$[a'(x, y)] = [y \quad x], [b'(x, y)] = [1 \quad 0], \text{ and } [c'(x, y)] = [0 \quad 1]$$

Differentiating g at $(1, 3)$, by multiple applications of the chain rule, we get

$$\begin{aligned} [g'(1, 3)] &= [f'(f(3, 1), f(3, 3))] \times \begin{bmatrix} [f'(3, 1)] \times \begin{bmatrix} [a'(1, 3)] \\ [b'(1, 3)] \end{bmatrix} \\ [f'(3, 3)] \times \begin{bmatrix} [c'(1, 3)] \\ [a'(1, 3)] \end{bmatrix} \end{bmatrix} \\ &= [f'(5, 2)] \times \begin{bmatrix} [1 \quad 2] \times \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \\ [0 \quad -1] \times \begin{bmatrix} 0 & 1 \\ 3 & 1 \end{bmatrix} \end{bmatrix} \\ &= [f'(5, 2)] \times \begin{bmatrix} 5 & 1 \\ -3 & -1 \end{bmatrix} \end{aligned}$$

Suppose that $[f'(5, 2)] = [r \quad s]$, for some $r, s \in \mathbb{R}$. Then

$$[g'(1, 3)] = [f'(5, 2)] \times \begin{bmatrix} 5 & 1 \\ -3 & -1 \end{bmatrix}$$

$$\implies \begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} r & s \end{bmatrix} \times \begin{bmatrix} 5 & 1 \\ -3 & -1 \end{bmatrix}$$

$$\implies \begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} 5r - 3s & r - s \end{bmatrix}$$

$$\implies 5r - 3s = 3 \text{ and } r - s = 4$$

Solving the system of equations gives

$$r = -\frac{9}{2} \text{ and } s = -\frac{17}{2}$$

Therefore $\nabla f(5, 2) = \left(-\frac{9}{2}, -\frac{17}{2}\right)$.

□

1.27 #30

Let $U \subseteq \mathbf{R}^n$ be an open set in \mathbf{R}^n , and let K be a compact subset of U . Prove that there exists an *infinitely differentiable* function $\varphi : \mathbf{R}^n \rightarrow [0, 1]$ such that $\varphi(p) = 1$ for all $p \in K$, and $\varphi(p) = 0$ for all $p \in \mathbf{R}^n \setminus U$. This is called a **bump function** supported on U .

(For a function $f : U \rightarrow Y$, the *n th total derivative* $f^{(n)}$ is defined as follows: for $n = 0$, we set $f^{(0)} = f$; for $n \geq 1$, if $f^{(n-1)}$ is totally differentiable, we set $f^{(n)} = (f^{(n-1)})'$. We say that f is **infinitely differentiable** if $f^{(n)}$ exists for all $n \geq 0$.)

Proof. First, we notice that the bump function in Big List #4 can be generalised to arbitrary intervals by simply performing horizontal translations. Now, we show that bump functions for closed rectangles within open rectangles in \mathbf{R}^n can be constructed.

Let $R = \prod_{i=1}^n [a_i, b_i]$ be a closed rectangle that is inside an open rectangle $S = \prod_{i=1}^n (c_i, d_i)$ in \mathbf{R}^n . This implies that for all i , $c_i < a_i \leq b_i < d_i$. Considering this as an interval in \mathbf{R} , we can find a bump function φ_i such that $\varphi_i([a_i, b_i]) = \{1\}$ and $\varphi_i(\mathbf{R} \setminus (c_i, d_i)) = \{0\}$. Notice that this is a function from \mathbf{R} to \mathbf{R} . We define $\alpha_i : \mathbf{R}^n \rightarrow \mathbf{R}$ by $\alpha_i(x) = \varphi_i(x_i)$. It will be shown that the bump function supported in S is

$$\beta(x) = \prod_{i=1}^n \alpha_i(x)$$

First, note that each α_i is infinitely differentiable, so β is infinitely differentiable as well. If $p = (p_1, \dots, p_n) \in R$, then for all $i \in \{1, \dots, n\}$, $p_i \in [a_i, b_i]$, so $\alpha_i(p) = 1$. We have

$$\beta(p) = \prod_{i=1}^n \alpha_i(p) = 1$$

Using a similar argument, if $p \in \mathbf{R}^n \setminus S$, there is at least one component of p such that $p_i \notin (c_i, d_i)$, and so $\alpha_i(p) = 0$, which implies that

$$\beta(p) = 0$$

as desired.

Additionally, if there exists bump functions α_1 and α_2 supported on open sets U_1 and U_2 respectively, where the compact sets are $K_1 \subseteq U_1$ and $K_2 \subseteq U_2$, then there exists a bump function φ supported on $U_1 \cup U_2$ for the compact set $K_1 \cup K_2$, which is defined by

$$\varphi(x) = \alpha_1(x) + \alpha_2(x) - \alpha_1(x)\alpha_2(x).$$

Verifying this, we see that if $x \in K_1 \cup K_2$, then either $x \in K_1$ or $x \in K_2$. Assuming without loss of generality that $x \in K_1$, we have that

$$\varphi(x) = 1 + \alpha_2(x) - \alpha_2(x) = 1$$

As well, if $x \in \mathbf{R}^n \setminus (U_1 \cup U_2)$, then $\varphi(x) = 0$, showing that φ is indeed a bump function.

Next, let \mathbf{R}^n be equipped with the max-norm. We claim that $\prod_{i=1}^n (p - r, p + r) \subseteq B(p, r)$, for $p \in \mathbf{R}^n$ and $r > 0$.

Let $x \in \prod_{i=1}^n (p - r, p + r)$. For all i , $|p_i - x_i| < r$, so $\|x - p\|_{\max} < r$. Hence $x \in B(p, r)$.

Thus $\prod_{i=1}^n (p - r, p + r) \subseteq B(p, r)$, and moreover by taking the closure of both sets,

$$\prod_{i=1}^n [p - r, p + r] \subseteq \overline{B(p, r)}.$$

Now, we can proceed proving the main result.

Let U be an open subset of \mathbf{R}^n , and let $K \subseteq U$ be compact. For every $y_i \in K$, there exists an open ball centered around y_i so that $B(y_i, \delta_i) \subseteq U$. Furthermore, we have that

$$\prod_{i=1}^n \left[y_i - \frac{\delta_i}{2}, y_i + \frac{\delta_i}{2} \right] \subseteq \prod_{i=1}^n (y_i - \delta_i, y_i + \delta_i) \subseteq B(y_i, \delta_i) \subseteq U$$

Let $R_i = \prod_{i=1}^n \left[y_i - \frac{\delta_i}{2}, y_i + \frac{\delta_i}{2} \right]$ and $S_i = \prod_{i=1}^n (y_i - \delta_i, y_i + \delta_i)$. Notice that $\{R_i^\circ\}$ forms an open cover on K , and by compactness, there is a finite subcover $\{R_{i_n}^\circ\}_{n \leq N}$. Note that if we take the closure of each set, we get the closed cover $\{R_{i_n}\}$.

For each closed rectangle R_{i_n} , we can find a bump function supported on S_{i_n} . Since we have a finite number of rectangles, we can “merge” all the bump functions by combining them one at a time using the method outlined above. This eventually obtains a bump function γ supported on $\bigcup_{n=1}^N S_{i_n}$ for the compact set $\bigcup_{n=1}^N K_{i_n}$. We claim that this bump function is also a bump function support on U for the compact set K .

If $x \in K$, then it is also true that $x \in \bigcup_{n=1}^N K_{i_n}$. It follows that $\gamma(x) = 1$.

In a similar fashion, if $x \in \mathbb{R}^n \setminus U$, then $x \in \mathbb{R}^n \setminus \bigcup_{n=1}^N S_{i_n}$, from which it follows that $\gamma(x) = 0$.

Therefore we have found a bump function as desired and the proof is complete. \square

1.28 #31

- (a) Let $A \in M_n(\mathbf{R})$ be a symmetric matrix and let $Q(\vec{x}) = \vec{x}^T A \vec{x}$ be the corresponding quadratic form. Prove that the following two statements are equivalent:

- (i) $Q(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$.
- (ii) All eigenvalues of A are strictly positive.

In this case, we say that Q is a **positive definite** quadratic form, and that A is a **positive definite** matrix.

- (b) Prove the following “stay away” lemma, which you will need for part (c): if $Q : \mathbf{R}^n \rightarrow \mathbf{R}$ is a positive definite quadratic form, then there exists a constant $\eta > 0$ such that $Q(\vec{x}) \geq \eta \|\vec{x}\|^2$ for all $\vec{x} \in \mathbf{R}^n$.
- (c) Let $U \subseteq \mathbf{R}^n$ be an open set, let $f : U \rightarrow \mathbf{R}$ be a twice continuously differentiable function, and let $p_0 \in U$ be a point at which $\nabla f(p_0) = \vec{0}$. Prove that if the Hessian matrix $Hf(p_0)$ is positive definite, then f achieves a **local minimum** at p_0 ; i.e. p_0 has an open neighborhood U_0 , contained in U , such that $f(p) \geq f(p_0)$ for all $p \in U_0$.

Proof. (a):

Suppose that $Q(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$. Since A is symmetric, it is orthogonally diagonalizable, so there exists an orthogonal matrix $P \in M_n(\mathbf{R})$ such that $B = P^T A P$ is diagonal. Letting $Q_B(\vec{x}) = \vec{x}^T B \vec{x}$, we have that

$$Q(\vec{x}) = Q_B \circ P(\vec{x})$$

We will show that $Q_B(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$.

Let $\vec{x} \in \mathbf{R}^n$ so that $\vec{x} \neq \vec{0}$. Since $Q(\vec{x}) > 0$, it follows that $Q_B(\vec{x}) > 0$ as well, which implies that its eigenvalues are strictly positive. Since A and B share the same eigenvalues, we are done.

Conversely, suppose that all the eigenvalues of A are strictly positive. By a similar argument as before, the eigenvalues of B are positive, so $Q_B(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$. The same applies to Q and we are done.

(b):

First, equip \mathbf{R}^n with the 2-norm.

Let Q be a positive definite quadratic form, so $Q(\vec{x}) = \vec{x}^T A \vec{x}$, for some diagonalizable matrix A , where all its eigenvalues λ_i are strictly positive.

Let $\eta = \min_{1 \leq i \leq n} \{\lambda_i\}$. Let B be the diagonal matrix such that $A = P^T B P$ for an orthogonal matrix P . Denote $P(\vec{x}) = (p_1, \dots, p_n)$. We have

$$Q(\vec{x}) = Q_B \circ P(\vec{x}) = (P(\vec{x}))^T B (P(\vec{x})) = \sum_{i=1}^n \lambda_i \cdot p_i^2 \geq \eta \sum_{i=1}^n p_i^2 = \eta \|P(\vec{x})\|_2^2$$

Since P preserves distance, we have that

$$Q(\vec{x}) \geq \eta \|P(\vec{x})\|_2^2 = \eta \|\vec{x}\|_2^2$$

as desired.

(c):

Let $U \subseteq \mathbf{R}^n$ be an open set, let $f : U \rightarrow \mathbf{R}$ be a C^2 function. Let $p_0 \in U$ be a point such that $\nabla f(p_0) = 0$. Suppose that $Hf(p_0)$ is positive definite. We will show that $Q(x) = \frac{1}{2} x^T Hf(p_0) x$ is a quadratic approximation to f at p_0 . That is,

$$\lim_{h \rightarrow 0} \frac{f(p_0 + h) - f(p_0) - Q(h)}{\|h\|^2} = 0$$

Let $\varepsilon > 0$. Since U is open, there is a δ_o so that $p_0 + h \in U$, where $\|h\| < \delta_o$. Additionally, by the continuity of Hf , there exists δ_c so that for $\|h\| < \delta_c$, $|Hf(p_0 + h) - Hf(p_0)| < 2\varepsilon$. Let $\delta = \min\{\delta_o, \delta_c\}$. Fix $h \in \mathbf{R}^n$ such that $0 < \|h\| < \delta$.

Since U is open, there exists $r > 0$ so that $p_0 + th \in U$ for $t \in (-r, 1 + r)$. Define $g(t) = f(p_0 + th)$. Applying Taylor's quadratic remainder formula to g , there exists $\theta \in (0, 1)$ so that

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(\theta) \implies f(p_0 + h) = f(p_0) + f'(p_0)h + \frac{1}{2}f''(p_0 + \theta h)(h)h$$

Since $\nabla f(p_0) = 0$, we have

$$f(p_0 + h) = f(p_0) + \frac{1}{2}h^T Hf(p_0 + \theta h)h$$

Thus we have

$$\begin{aligned} |f(p_0 + h) - f(p_0) - Q(h)| &= \left| \frac{1}{2}h^T Hf(p_0 + \theta h)h - \frac{1}{2}h^T Hf(p_0)h \right| \\ &\leq \frac{1}{2}\|h^T\|_c \|Hf(p_0 + \theta h) - Hf(p_0)\| \|h\| \end{aligned}$$

where $\|\cdot\|_c$ is a norm on row vectors defined by $\|x\|_c = \|x^T\|$. Automatically we have that

$$\begin{aligned} |f(p_0 + h) - f(p_0) - Q(h)| &\leq \frac{1}{2}\|Hf(p_0 + \theta h) - Hf(p_0)\| \|h\|^2 < \varepsilon \|h\|^2 \\ \implies \frac{f(p_0 + h) - f(p_0) - Q(h)}{\|h\|^2} &< \varepsilon \end{aligned}$$

as desired. Thus Q is indeed a quadratic approximation for f at p_0 . It remains to show that f attains a local minimum at p_0 . Suppose that this is not the case. In particular, for every open neighborhood around p_0 there is a point q where $f(q) < f(p_0)$. We write $q = p_0 + s$, where $s = q - p_0$. Since $Hf(p_0)$ is positive definite, Q is positive definite. Using the quadratic approximation, there is a $\delta > 0$ so that for all $0 < \|h\| < \delta$,

$$|f(p_0 + h) - f(p_0) - Q(h)| < \eta \|h\|^2,$$

where η comes from part (b) applied to Q . Find a suitable s such that $f(p_0 + s) < f(p_0)$. By part (a), $Q(s) > 0$. This implies that $f(p_0 + s) - f(p_0) - Q(s) < 0$, so

$$f(p_0) - f(p_0 + s) + Q(s) < \eta \|h\|^2 \implies f(p_0) + Q(s) - \eta \|h\|^2 < f(p_0 + s).$$

From part (b), we know that $Q(s) - \eta \|h\|^2 \geq 0$, so

$$f(p_0) < f(p_0 + s)$$

which is a contradiction.

Therefore we have shown that f attains a local minimum at p_0 and we are done. \square

1.29 #32*

I selected this problem because I found this question to be challenging, but also rewarding to prove. At first, I attempted this problem without knowing about the Implicit Function Theorem. As a result, I could not solve the problem, though I gained some intuition about how to solve it. This intuition was directly related to the application of the Implicit Function Theorem.

Let M be a subset of \mathbf{R}^n , let $p_0 \in M$ be a point, and let $\vec{v} \in \mathbf{R}^n$ be a vector. We say that \vec{v} is a **tangent vector** to M at p_0 if there exists $\delta > 0$ and a C^1 function $\alpha : (-\delta, \delta) \rightarrow M$ such that $\alpha(0) = p_0$ and $\alpha'(0) = \vec{v}$. In other words, \vec{v} is the velocity vector of a curve through M .

- (a) Suppose now that M is the zero set of some C^1 function $f : U \rightarrow \mathbf{R}$, where U is an open set in \mathbf{R}^n : thus

$$M = \{p \in U : f(p) = 0\}.$$

Suppose that $p_0 \in M$ is a point such that $\nabla f(p_0) \neq \vec{0}$, and let $\vec{v} \in \mathbf{R}^n$ be a vector. Show that \vec{v} is a tangent vector to M at p_0 if and only if $\nabla f(p_0) \cdot \vec{v} = 0$.

- (b) Let E be the ellipsoid in \mathbf{R}^3 defined by the following equation:

$$x^2 + yz + y^2 - xy - xz + z^2 = 3.$$

Find the equation of the tangent plane to M at the point $p_0 = (1, 2, 0)$.

Proof. (a):

Suppose that \vec{v} is a tangent vector to M at p_0 . Then there exists a function $\alpha : (-\delta, \delta) \rightarrow M$ so that $\alpha(0) = p_0$ and $\alpha'(0) = \vec{v}$. Define $g : (-\delta, \delta) \rightarrow \mathbf{R}$ by $g(t) = f(\alpha(t))$. For all $t \in (-\delta, \delta)$, $\alpha(t) \in M$, so $g(t) = 0$. It follows that

$$0 = g'(t) = \nabla f(\alpha(t)) \cdot \alpha'(t)$$

Substituting $t = 0$ yields

$$\nabla f(p_0) \cdot \vec{v} = 0$$

as needed.

Conversely, suppose that $\nabla f(p_0) \cdot \vec{v} = 0$. Since $\nabla f(p_0) \neq 0$, $\frac{\partial f}{\partial x_i}(p_0) \neq 0$ for some $i \in \{0, \dots, n\}$. Define the C^1 function $g : \mathbf{R}^n \rightarrow \mathbf{R}$ as the function that swaps the i th and n th coordinate. That is,

$$g(x_1, \dots, x_n) = f(x_1, \dots, x_{i-1}, x_n, x_{i+1}, \dots, x_i)$$

Let p' be the vector in \mathbf{R}^{n-1} whose components are the same as p_0 except that its i th component is p_n . In particular,

$$p' = (p_1, \dots, p_{i-1}, p_n, p_{i+1}, \dots, p_{n-1}).$$

Notice that $\frac{\partial g}{\partial x_n}(p', p_i) = \frac{\partial f}{\partial x_i}(p_0) \neq 0$. Applying the Implicit Function Theorem with $k = 1$, there exists an open set $W \subseteq \mathbf{R}^{n-1}$ that contains p' and a continuously differentiable function $\psi : W \rightarrow \mathbf{R}$ such that for all $x' \in W$, $\psi(p') = p_i$ and

$$g(x', \psi(x')) = f(x_1, \dots, x_{i-1}, \psi(x'), x_{i+1}, \dots, x_{n-1}) = 0$$

Since W is open, there exists $\delta > 0$ so that for all $\|t\| < \delta$, $p' + t\pi_{\mathbf{R}^{n-1}}(\vec{v}) \in W$.

Let \hat{v} be \vec{v} with swapped i th and n th components. Let $\alpha : (-\delta, \delta) \rightarrow M$ be defined by

$$\alpha(t) = (p_1 + tv_1, \dots, p_{i-1} + tv_{i-1}, \psi(p' + tv_i(\pi_{\mathbf{R}^{n-2}}(\hat{v}), -1)), p_{i+1} + tv_{i+1}, \dots, p_n + tv_n).$$

We can see that $\alpha(0) = p_0$. Next, we want to show that $\alpha'(0) = \vec{v}$. We first find the total derivative of $\psi(p' + tv_i(\pi_{\mathbf{R}^{n-2}}(\hat{v}), -1))$. For $j \in \{1, \dots, n-1\}$, we know that the j th partial derivative is given by

$$\frac{\frac{\partial g}{\partial x_j}(p' + tv_i(\pi_{\mathbf{R}^{n-2}}(\hat{v}), -1)) \cdot v_i \cdot \hat{v}_j}{\frac{\partial g}{\partial x_n}(p' + tv_i(\pi_{\mathbf{R}^{n-2}}(\hat{v}), -1)) \cdot (-v_i)} = -\frac{\frac{\partial g}{\partial x_j}(p' + tv_i(\pi_{\mathbf{R}^{n-2}}(\hat{v}), -1)) \cdot \hat{v}_j}{\frac{\partial g}{\partial x_n}(p' + tv_i(\pi_{\mathbf{R}^{n-2}}(\hat{v}), -1))}$$

substituting $t = 0$ into this expression gives us

$$-\frac{\frac{\partial g}{\partial x_j}(p') \cdot \hat{v}_j}{\frac{\partial g}{\partial x_n}(p')} = -\frac{\frac{\partial g}{\partial x_j}(p') \cdot \hat{v}_j}{\frac{\partial f}{\partial x_i}(p_0)}$$

For $j \neq i$,

$$-\frac{\frac{\partial g}{\partial x_j}(p') \cdot \hat{v}_j}{\frac{\partial f}{\partial x_i}(p_0)} = -\frac{\frac{\partial f}{\partial x_j}(p_0) \cdot v_j}{\frac{\partial f}{\partial x_i}(p_0)}.$$

For $j = i$,

$$-\frac{\frac{\partial g}{\partial x_j}(p') \cdot \hat{v}_j}{\frac{\partial f}{\partial x_i}(p_0)} = -\frac{\frac{\partial f}{\partial x_n}(p_0) \cdot v_n}{\frac{\partial f}{\partial x_i}(p_0)}$$

Thus the total derivative is given by

$$-\frac{1}{\frac{\partial f}{\partial x_i}(p_0)} \sum_{j \in \{1, \dots, i-1, i+1, \dots, n\}} \frac{\partial f}{\partial x_j}(p_0) \cdot v_j$$

Recall that $\sum_{j=1}^n \frac{\partial f}{\partial x_j}(p_0) \cdot v_j = \nabla f(p_0) \cdot \vec{v} = 0$. Using this, our expression for the total derivative at $t = 0$ becomes

$$-\frac{1}{\frac{\partial f}{\partial x_i}(p_0)} \cdot \left(-\frac{\partial f}{\partial x_i}(p_0) \cdot v_i \right) = v_i$$

Therefore, we conclude that

$$\alpha'(0) = (v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n) = \vec{v},$$

Which verifies that \vec{v} is indeed a tangent vector to M at p_0 , finishing the proof.

(b):

Let $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ be a function defined by

$$f(x, y, z) = x^2 + yz + y^2 - xy - xz + z^2 - 3$$

Its zero set is exactly M and its gradient is

$$\nabla f(x, y, z) = (2x - y - z, z + 2y - x, y - x + 2z).$$

Substituting p_0 , we get $\nabla f(p_0) = (0, 3, 1) \neq 0$. Let $\vec{v}_1 = (1, 0, 0)$, $\vec{v}_2 = (0, 1, -3)$. Notice that

$$\nabla f(p_0) \cdot \vec{v}_1 = \nabla f(p_0) \cdot \vec{v}_2 = 0$$

By the results of part (a), \vec{v}_1 and \vec{v}_2 are tangent vectors to M at p_0 . The equation of the plane spanned by \vec{v}_1 and \vec{v}_2 is given by

$$3y + z = 0$$

We shift this equation to p_0 , giving us

$$3y + z = 6$$

□

1.30 #33

- (a) Let $g : U \rightarrow \mathbf{R}$ be a C^1 function defined on an open set $U \subseteq \mathbf{R}^n$, and let M be its zero set:

$$M = \{p \in U : g(p) = 0\}.$$

Suppose that we have a C^1 function $f : U \rightarrow \mathbf{R}$, defined on an open set $U \subseteq \mathbf{R}^n$ which contains M , and we wish to find the maximum of f on M . Assume that M is compact, and that f achieves its maximum on M at some point $p_0 \in M$. Prove that there exists a real number $\lambda \in \mathbf{R}$ such that

$$\nabla f(p_0) = \lambda \nabla g(p_0).$$

This number λ is known as the **Lagrange multiplier**.

- (b) Use Lagrange multipliers to solve the following optimization problem: *Find the point(s) on the ellipsoid $x^2 + yz + y^2 - xy - xz + z^2 = 3$ which are **closest** and **furthest** from the origin.*

Proof. (a):

Let $h : U \times \mathbf{R} \rightarrow \mathbf{R}$ be the C^1 function defined by

$$h(x, y) = yg(x) + y - f(x)$$

Notice that $h(p_0, f(p_0)) = f(p_0) \cdot g(p_0) + f(p_0) - f(p_0) = 0$ and $\frac{\partial h}{\partial y}(p_0, f(p_0)) = g(p_0) + 1 = 1 \neq 0$. By the implicit function theorem, there exists an open set $W \subseteq \mathbf{R}^n$ and C^1 function $\psi : W \rightarrow \mathbf{R}$ such that for all $x \in W$,

$$h(x, \psi(x)) = \psi(x)g(x) + \psi(x) - f(x) = 0$$

Taking the derivative of both sides at p_0 with respect to x , we see that

$$\begin{aligned} g(p_0)\nabla\psi(p_0) + \psi(p_0)\nabla g(p_0) + \nabla\psi(p_0) - \nabla f(p_0) &= 0 \\ \implies \psi(p_0)\nabla g(p_0) + \nabla\psi(p_0) - \nabla f(p_0) &= 0 \end{aligned}$$

Also note that $\frac{\partial\psi}{\partial x_i}(p_0) = \frac{\frac{\partial h}{\partial x_i}(p_0, f(p_0))}{\frac{\partial h}{\partial y}(p_0, f(p_0))} = \frac{\partial g}{\partial x_i}(p_0) - \frac{\partial f}{\partial x_i}(p_0)$. Thus our equation becomes

$$(\psi(p_0) + 1)\nabla g(p_0) - 2\nabla f(p_0) = 0 \implies \nabla f(p_0) = \frac{\psi(p_0) + 1}{2}\nabla g(p_0)$$

Therefore the value $\lambda = \frac{\psi(p_0) + 1}{2}$ is the one that we needed.

(b):

Define

$$g(x, y, z) = x^2 + yz + y^2 - xy - xz + z^2 - 3.$$

Its zero set M is defined by the ellipsoid

$$x^2 + yz + y^2 - xy - xz + z^2 - 3 = 0.$$

We will attempt to minimize and maximize $f(x, y, z) = x^2 + y^2 + z^2$ on M . We know that if f achieves a maximum or minimum, by part (a), there exists $\lambda \in \mathbf{R}$ so that

$$\nabla f(p_0) = \lambda \nabla g(p_0) \implies (2x, 2y, 2z) = \lambda(2x - y - z, z + 2y - x, y - x + 2z)$$

We simply evaluate f at these points and check which ones are maximums, minimums, or neither. From the above equality, we end up with the system of equations

$$\begin{aligned} 2(1 - \lambda)x + y + z &= 0 \\ x + 2(1 - \lambda)y - z &= 0 \\ x - y + 2(1 - \lambda)z &= 0 \end{aligned}$$

Consider if $\lambda = \frac{3}{2}$. The equations collect and become

$$x = y + z$$

From the ellipsoid, we have

$$x^2 + yz + y^2 - xy - xz + z^2 = 3 \implies f(x, y, z) = x(y + z) - yz + 3,$$

but at the same time, we have

$$x^2 + yz + y^2 - xy - xz + z^2 = 3 \implies x^2 + yz + y^2 - (y + z)y - (y + z)z + z^2 = 3$$

Expand the brackets and simplify to get

$$x^2 - yz = 3 \implies x^2 - yz + 3 = 6.$$

Thus

$$f(x, y, z) = 6, \text{ when } \lambda = \frac{3}{2}$$

Otherwise, we take the first and second equation in the original system and add them together to obtain

$$\begin{aligned} (2(1 - \lambda) + 1)x + (2(1 - \lambda) + 1)y &= 0 \\ \implies (2(1 - \lambda) + 1)(x + y) &= 0 \end{aligned} \tag{1}$$

Likewise, we can subtract the third equation from the second to get

$$\begin{aligned} (2(1 - \lambda) + 1)y - (2(1 - \lambda) + 1)z &= 0 \\ (2(1 - \lambda) + 1)(y - z) &= 0 \end{aligned} \tag{2}$$

Since $(2(1 - \lambda) + 1) \neq 0$, from equations (1) and (2), we get that $z = y = -x$. We substitute every variable into the original equation for x and see that

$$x^2 + x^2 + x^2 + x^2 + x^2 + x^2 = 3 \implies x = \frac{1}{\sqrt{2}}, y = z = -\frac{1}{\sqrt{2}}$$

so

$$f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \frac{3}{2}$$

Thus the points on the ellipsoid that are furthest from the origin satisfy the equation $x = y + z$, and the point that is closest to the origin is $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$. \square

1.31 #34*

I first struggled on figuring out this problem, despite the statement being so intuitive. Even after a hint from Nigel, I still couldn't figure it out for a while. However, as I played around with it, I realized that the proof was also very intuitive.

Let $\Phi : \mathbf{R}^n \rightarrow \mathbf{R}^m$ be a C^1 mapping.

- (a) Suppose that $n > m = 1$. Show that Φ cannot be injective.
- (b) Suppose that $n < m$. Show that if $K \subseteq \mathbf{R}^n$ is a compact set, then $\Phi(K)$ is a Jordan measurable set, and has Jordan measure zero.

Previous submission was just a complete skill issue in part (b). This new submission hopefully provides a correct proof for part (b).

Changes: all of part (b) lol.

Proof.

(a):

Suppose for contradiction that $n > m = 1$ and Φ is a C^1 injective function. Since Φ cannot be a constant function, by the results of Big List #26, there is a $p \in \mathbf{R}^n$ so that $\nabla \Phi(p) \neq 0$. In particular, we will say that $\frac{\partial \Phi}{\partial x_j} \neq 0$. Define $\alpha : \mathbf{R}^n \rightarrow \mathbf{R}$ by $\alpha(x) = \Phi(x) - \Phi(p)$. Injectivity is translation-invariant, so α is injective. Notice that $\alpha(p) = 0$. We can apply the implicit function theorem to obtain an open set $W \subseteq \mathbf{R}^{n-1}$ that contains $p' = (p_1, \dots, p_{j-1}, p_{j+1}, \dots, p_n)$ and a C^1 function $\Psi : W \rightarrow \mathbf{R}$ such that for all $x = (x_1, \dots, x_{n-1}) \in W$,

$$\alpha(x_1, \dots, x_{j-1}, \Psi(x), x_j, \dots, x_{n-1}) = 0$$

Then, since W is open and contains p' , we can find another distinct point $q \in W$. We have

$$\alpha(p_1, \dots, p_{j-1}, \Psi(p'), p_{j+1}, \dots, p_n) = 0 = \alpha(q_1, \dots, q_{j-1}, \Psi(q), q_j, \dots, q_{n-1})$$

which contradicts the fact that α is injective.

(b):

Since K is compact, and thus bounded, we can enclose it in a closed box $B = [-L, L]^n$, for some positive L . It suffices to show that $\Phi(B)$ has measure 0, as we can apply the monotonicity of measure to conclude that $\Phi(K)$ has measure 0.

First, let $\hat{\Phi} : \mathbf{R}^m \rightarrow \mathbf{R}^m$ be defined by $\hat{\Phi}(x) = \Phi(\pi_{\mathbf{R}^n}(x))$. That is, $\hat{\Phi}$ first projects vectors in \mathbf{R}^m onto \mathbf{R}^n and then composes it with Φ . Let $\hat{B} = B \times \{0\}^{m-n}$. Then note that $\hat{\Phi}(\hat{B}) = \Phi(B)$. Trivially, \hat{B} has measure 0. We now show that $\Phi(B)$ also has measure 0.

Let $\varepsilon > 0$. Since $\hat{\Phi}$ is C^1 , its derivative is continuous. By the extreme value theorem, each component derivative attains a maximum on \hat{B} . Let α a positive number greater than all the maximums. Since \hat{B} has measure 0, we can find a finite number of cubes B_1, \dots, B_k with side length d such that

$$\hat{B} \subseteq \bigcup_{i=1}^k B_i \text{ and } \sum_{i=1}^k \text{vol}(B_i) < \frac{\varepsilon}{m^m \alpha^m}$$

Consider some cube $B_i = \prod_{j=1}^m [a_{ij}, a_{ij} + d]$. For each component $\hat{\Phi}_j : \mathbf{R}^m \rightarrow \mathbf{R}$, it must be true that $\hat{\Phi}_j$ attains a maximum $\hat{\Phi}_j(M_{ij})$ and minimum $\hat{\Phi}_j(m_{ij})$, for some $M_{ij}, m_{ij} \in B_i$. Define $g : [0, 1] \rightarrow \mathbf{R}$ by

$$g(t) = \hat{\Phi}_j(tM_{ij} + (1-t)m_{ij})$$

By the Mean Value Theorem, there exists some $c \in (0, 1)$ such that

$$g(1) - g(0) = g'(c) \implies \hat{\Phi}_j(M) - \hat{\Phi}_j(m) = \hat{\Phi}'_j(cM_{ij} + (1-c)m_{ij})(M_{ij} - m_{ij})$$

Taking the 2-norm of both sides, it follows that

$$\|\hat{\Phi}_j(M_{ij}) - \hat{\Phi}_j(m_{ij})\| \leq \|\hat{\Phi}'_j(cM_{ij} + (1-c)m_{ij})\| \|M_{ij} - m_{ij}\| \leq (\sqrt{m}\alpha)(\sqrt{m}d) = m\alpha d$$

Note that we used the fact that each component of $\hat{\Phi}'_j(cM_{ij} + (1-c)m_{ij})$ is bounded by α and $\|M_{ij} - m_{ij}\|$ can be no bigger than the diagonal of B_i .

Finally, we see that the box $C_i = \prod_{j=1}^m [\hat{\Phi}_j(m_{ij}), \hat{\Phi}_j(M_{ij})]$ covers $\hat{\Phi}(B_i)$, and its volume is given by

$$\prod_{j=1}^m (\hat{\Phi}_j(M_{ij}) - \hat{\Phi}_j(m_{ij})) \leq \prod_{j=1}^m m\alpha d = m^m \alpha^m \text{vol}(B_i)$$

From here, we know that the union $\bigcup_{i=1}^k C_i \supseteq \hat{\Phi}(\hat{B})$ but

$$\sum_{i=1}^k \text{vol}(C_i) < m^m \alpha^m \sum_{i=1}^k \text{vol}(B_i) < \varepsilon.$$

Therefore $\Phi(B)$ has measure 0, so $\Phi(K)$ has measure 0 and we are done. □

1.32 #35*

I chose this problem because I felt that it was quite involved in terms of the smaller statements that I had to prove as well as the caseworking in order to achieve the main result.

Perturbing the roots of a polynomial.

Let $f(x) = \sum_{i=0}^n a_i x^i$ be a **monic** polynomial with **no repeated real roots**. This means that $a_n = 1$, and that all real roots of f have multiplicity 1.

- (a) Let r be a root of $f(x)$. Prove that for all $\varepsilon > 0$, there exists $\delta > 0$ such that: if $g(x) = \sum_{i=0}^n b_i x^i$ is a monic polynomial with coefficients b_i satisfying $|a_i - b_i| < \delta$, then $g(x)$ has at least one root in the interval $(r - \varepsilon, r + \varepsilon)$.

This shows that slight perturbations of the coefficients results in slight perturbations of the roots.

- (b) Suppose that f has fewer than n real roots. Prove that number of real roots of f does not change under small perturbation of the coefficients.

Proof. (a):

First, we prove the following lemma:

Lemma. Let $r \in \mathbb{R}$ be a root of a polynomial p . Then r is a repeated root if and only if $p'(r) = 0$.

Suppose that p has a repeated root. Then we can factor p as $(x - r)^k q(x)$, for some $k > 1$ and $q \in \mathbb{P}(\mathbb{R})$. We can take the derivative of this and get that

$$\begin{aligned} p'(x) &= k(x - r)^{k-1}q(x) + (x - r)^k q'(x) \\ &\implies p'(r) = 0 \end{aligned}$$

Conversely, suppose that $p'(r) = 0$, for some $r \in \mathbb{R}$. We can write

$$p'(x) = (x - r) \sum_{i=0}^m c_i x^i, \text{ for constants } c_0, \dots, c_m$$

We can integrate both sides to get that

$$\begin{aligned} p(x) &= \int (x - r) \sum_{i=0}^m c_i x^i dx = \int x \sum_{i=0}^m c_i x^i - r \sum_{i=0}^m c_i x^i dx \\ &= \sum_{i=0}^m \frac{c_i}{i+2} x^{i+2} - r \sum_{i=0}^m \frac{c_i}{i+1} x^{i+1} + C \end{aligned}$$

In order for r to be a root of p , we must have that

$$\begin{aligned} p(r) &= \sum_{i=0}^m \left(\frac{c_i}{i+2} r^{i+2} - \frac{c_i}{i+1} r^{i+2} \right) + C = 0 \\ &\implies C = \sum_{i=0}^m \left(\frac{c_i}{i+1} r^{i+2} - \frac{c_i}{i+2} r^{i+2} \right) \end{aligned}$$

Therefore (im sorry)

$$\begin{aligned} p(x) &= \sum_{i=0}^m \left(\frac{c_i}{i+2} x^{i+2} - \frac{c_i}{i+2} r^{i+2} + \frac{c_i}{i+1} r^{i+2} - r \frac{c_i}{i+1} x^{i+1} \right) \\ &= \sum_{i=0}^m \left(\left(\frac{c_i}{i+2} \right) (x^{i+2} - r^{i+2}) - \left(\frac{c_i r}{i+1} \right) (x^{i+1} - r^{i+1}) \right) \\ &= \sum_{i=0}^m \left(\left(\frac{c_i}{i+2} \right) (x - r) \sum_{j=0}^{i+1} x^j r^{i-j+1} - \left(\frac{c_i}{i+1} \right) (x - r) \sum_{j=0}^i x^j r^{i-j+1} \right) \end{aligned}$$

$$\begin{aligned}
&= (x-r) \sum_{i=0}^m \left(\frac{c_i}{(i+1)(i+2)} \left((i+1) \left(x^{i+1} + \sum_{j=0}^i x^j r^{i-j+1} \right) - (i+2) \sum_{j=0}^i x^j r^{i-j+1} \right) \right) \\
&= (x-r) \sum_{i=0}^m \left(\frac{c_i}{(i+1)(i+2)} \left((i+1)x^{i+1} - \sum_{j=0}^i x^j r^{i-j+1} \right) \right) \\
&= (x-r) \sum_{i=0}^m \left(\frac{c_i}{(i+1)(i+2)} \sum_{j=0}^i (x^{i+1} - x^j r^{i-j+1}) \right) \\
&= (x-r) \sum_{i=0}^m \left(\frac{c_i}{(i+1)(i+2)} \sum_{j=0}^i x^j (x^{i-j+1} - r^{i-j+1}) \right) \\
&= (x-r) \sum_{i=0}^m \left(\frac{c_i}{(i+1)(i+2)} \sum_{j=0}^i x^j (x-r) \sum_{k=0}^{i-j} x^k r^{i-j-k+1} \right) \\
&= (x-r)^2 \sum_{i=0}^m \left(\frac{c_i}{(i+1)(i+2)} \sum_{j=0}^i \sum_{k=0}^{i-j} x^{k+j} r^{i-j-k+1} \right)
\end{aligned}$$

which implies that r is a repeated root of p .

Define the C^1 function $\Phi : \mathbb{R}^{n+2} \rightarrow \mathbb{R}$ by

$$\Phi(y_0, y_1, \dots, y_n, x) = \sum_{i=0}^n y_i x^i.$$

Let $a = (a_0, \dots, a_n) \in \mathbb{R}^{n+1}$. We have that $\Phi(a, r) = 0$ and as well, by our lemma,

$\frac{\partial \Phi}{\partial x}(a, r) = f'(r) \neq 0$. Applying the Implicit Function Theorem, we get that for some open set $W \subseteq \mathbb{R}^{n+1}$ containing a and C^1 function $\Psi : W \rightarrow \mathbb{R}$ such that $\Psi(a) = r$,

$$\Phi(b, \Psi(b)) = 0, \text{ for all } b \in W.$$

Now, let $\varepsilon > 0$. by the continuity of Ψ at a , there exists a $\delta > 0$ so that for all $b \in W$ with $\|b - a\| < (n+1)\delta$, $|\Psi(b) - r| < \varepsilon$, so $\Psi(b) \in (r - \varepsilon, r + \varepsilon)$. Consider the polynomial

$$g(x) = \sum_{i=0}^n b_i x^i = \Phi(b, x)$$

Where $|a_i - b_i| < \delta$. It follows that

$$(n+1)\delta > \sum_{i=0}^n |a_i - b_i| > \|a - b\|$$

which implies that $\Psi(b) \in (r - \varepsilon, r + \varepsilon)$. But notice that $\Psi(b)$ is actually a root of $g(x)$, as desired.

(b):

From part (a), we know that the number of roots can never decrease under small perturbations. Thus, it suffices to show that it will never increase, either.

First, we will strengthen the statement in part (a) so that for small enough ε , there is exactly one root in the interval $(r - \varepsilon, r + \varepsilon)$. By the continuity of $\frac{\partial \Phi}{\partial x}$, we obtain $\xi > 0$ so that for $(c, s) \in \mathbb{R}^{n+1} \times \mathbb{R}$ such that $\|(c, s) - (a, r)\| < \xi$,

$$\left| \frac{\partial \Phi}{\partial x}(c, s) - \frac{\partial \Phi}{\partial x}(a, r) \right| < \frac{1}{2} \left| \frac{\partial \Phi}{\partial x}(a, r) \right|$$

$$\implies \frac{\partial \Phi}{\partial x}(a, r) - \frac{1}{2} \left| \frac{\partial \Phi}{\partial x}(a, r) \right| < \frac{\partial \Phi}{\partial x}(c, s) < \frac{\partial \Phi}{\partial x}(a, r) + \frac{1}{2} \left| \frac{\partial \Phi}{\partial x}(a, r) \right|$$

Since $\frac{\partial \Phi}{\partial x}(a, r) \neq 0$, it follows that $\frac{\partial \Phi}{\partial x}(c, s) \neq 0$ as well. Let $\varepsilon < \frac{\xi}{2}$. Define δ_a as the value we found in part

(a) for this value of ε . Let g be a polynomial with coefficients b_i such that $|a_i - b_i| < \min \left\{ \delta_a, \frac{\xi}{2(n+1)} \right\}$.

Then $g(x) = \Phi(b, x)$ and $g'(x) = \frac{\partial \Phi}{\partial x}(b, x)$. But notice that for $|x - r| < \varepsilon = \frac{\xi}{2}$, we have that

$$\|(b, x) - (a, r)\| \leq \sum_{i=0}^n |a_i - b_i| + |x - r| < \frac{(n+1)\xi}{2(n+1)} + \frac{\xi}{2} = \xi,$$

so $g'(x) = \frac{\partial \Phi}{\partial x}(b, x) \neq 0$, which implies that g is injective on the interval $(r - \varepsilon, r + \varepsilon)$. We already knew that $g(\Psi(b)) = 0$, now we know that this is the only root in the interval.

For every root r , we have an associated interval $(r - \varepsilon_r, r + \varepsilon_r)$ such that g contains exactly one root within them. Finally, we must show that outside of all these intervals, g does not gain any new roots. To do this, we have to shrink δ even further. Let R be the finite set of all the roots of f . Let $V = \mathbb{R} \setminus \bigcup_{r \in R} (r - \varepsilon_r, r + \varepsilon_r)$. Consider the set

$$S = |f|(V)$$

Notice that $\inf S > 0$. We use the fact that $f \in \Omega(x^n)$ to claim that there exists x_0 such that for all $|x| > x_0$, we have

$$|f(x)| \geq \eta |x|^n, \text{ for some } \eta > 0.$$

Let $m = \max\{1, x_0\}$. Let $\delta = \min \left\{ \delta_a, \frac{\xi}{2(n+1)}, \frac{\eta}{2(n+1)}, \frac{\inf S}{2m^n(n+1)} \right\}$. Let $x \in V$. We have that

$$||f(x)| - |g(x)|| \leq |f(x) - g(x)| = \left| \sum_{i=0}^n a_i x^i - \sum_{i=0}^n b_i x^i \right| \leq \sum_{i=0}^n |a_i - b_i| |x|^i$$

Consider the case where $|x| > m$. We get

$$\begin{aligned} ||f(x) - |g(x)|| &\leq \sum_{i=0}^n |a_i - b_i| |x|^i < \sum_{i=0}^n \frac{\eta}{2(n+1)} |x|^n = \frac{\eta}{2} |x|^n \\ \implies |g(x)| &> |f(x)| - \frac{\eta}{2} |x|^n \geq \eta |x|^n - \frac{\eta}{2} |x|^n > 0 \end{aligned}$$

Otherwise, if $|x| \leq m$,

$$\begin{aligned} ||f(x) - |g(x)|| &\leq \sum_{i=0}^n |a_i - b_i| |x|^i \leq \sum_{i=0}^n \frac{\inf S}{2m^n(n+1)} m^i \leq \sum_{i=0}^n \frac{\inf S}{2m^n(n+1)} m^n \\ \implies |g(x)| &> |f(x)| - \frac{\inf S}{2} \geq \inf S - \frac{\inf S}{2} = \frac{\inf S}{2} > 0 \end{aligned}$$

In either case, we see that g has no roots in V .

Therefore, we see that any g with coefficients δ -close to f has exactly one root close to each root of f , and no other roots, so we can conclude that g exactly the same number of roots as f . \square

1.33 #36

In Handout #7, we defined differentiability for functions on open sets. Now we give a definition that works over arbitrary sets. [For this problem, you will need to read Piazza Post @274 and use Theorem 1.1.](#)

Let $A \subseteq \mathbf{R}^n$ be an arbitrary set, let $f : A \rightarrow \mathbf{R}$ be a function, and let $p \in A$ be a point. We say that f is **differentiable** at p if there exists an open neighborhood U of p and a function $\hat{f} : U \rightarrow \mathbf{R}$ such that \hat{f} is differentiable at p (in the sense of Handout #7) and $\hat{f}|_{U \cap A} = f|_{U \cap A}$.

- Prove that f is differentiable at every point of A if and only if f extends to a differentiable function defined on an open set containing A .
- Suppose further that A is closed. Prove that f is differentiable at every point of A if and only if f extends to a differentiable function on \mathbf{R}^n .

Proof. (a):

Suppose that f extends to a differentiable function \hat{f} on an open set $U \supseteq A$. That is, $\hat{f}|_A = f$. Let $x \in A$. Since U is open, we can find an open ball such that $B(x, \varepsilon) \subseteq U$. Immediately, we get that the function $\hat{f}|_{B(x, \varepsilon)}$ is the desired extension of f at x , as \hat{f} is differentiable at x and $\hat{f}|_{B(x, \varepsilon) \cap A} = f|_{B(x, \varepsilon) \cap A}$.

Conversely, suppose that f is differentiable at every point of A . For all $p_a \in A$, there exists an open neighborhood U_a of p_a and a function $f_a : U_a \rightarrow \mathbf{R}$ that is differentiable at p_a and $f_a|_{U_a \cap A} = f|_{U_a \cap A}$. Notice that $U = \{U_a\}_{a \in I}$ forms an open cover of A . Thus we can find a partition of unity $\{\varphi_a : \mathbf{R}^n \rightarrow [0, 1]\}_{a \in I}$ subordinate to U . In particular,

- $\varphi_a \in C^\infty$,
- $\text{supp}(\varphi_a) \cap U \subseteq U_a$,
- $\{\text{supp}(\varphi_a)\}$ is locally finite,
- $\sum_{a \in I} \varphi_a = 1$.

We claim that $\hat{f} = \sum_{a \in I} \varphi_a \cdot f_a$ is the differentiable function that extends f to $\{U_a\}_{a \in I}$. First, note that this sum is well defined because the partition of unity is locally finite, so the sum is finite at every point, and as well, any point outside the domain of f_a implies that $\varphi_a = 0$, so $\varphi_a \cdot f_a = 0$.

Moreover, for all $x \in U$, since $\{\varphi_a\}_{a \in I}$ is locally finite, for some open neighborhood V of x , for some finite sequence of $a_n \in I$,

$$\hat{f}(y) = \sum_{n=1}^N \varphi_{a_n}(y) \cdot f_{a_n}(y), \text{ where } y \in V.$$

Since each term is differentiable, \hat{f} is differentiable at x . Thus \hat{f} is differentiable on the entirety of U . Additionally, for each a_n , $x \in U_{a_n}$, so by our initial assumption,

$$\sum_{n=1}^N \varphi_{a_n}(x) \cdot f_{a_n}(x) = f(x) \cdot \sum_{n=1}^N \varphi_{a_n}(x) = f(x)$$

which shows that $\hat{f}|_A = f$ as needed.

(b):

Suppose that A is a closed set.

The converse of this statement is exactly the same as the previous part, so we will omit the proof.

Suppose that f is differentiable at every point of A . By the previous part, f extends to a differentiable function \tilde{f} on some open set U that contains A . We want to find an open set V such that $A \subseteq V \subseteq \overline{V} \subseteq U$.

We see that the set $V = \left\{x \in \mathbf{R}^n : \text{dist}_A(x) < \frac{\text{dist}(A, U^c)}{2}\right\}$ is sufficient. It is pretty immediate that $A \subseteq V$, as any element of A has distance 0 from A . As well, if we let $v \in \overline{V}$, since v is a limit point of V we can find

$v' \in V$ such that $\|v - v'\| < \frac{\text{dist}(A, U^c)}{4}$. Furthermore, by the definition of infimum we can find $a \in A$ such that $\|v' - a\| < \text{dist}_A(x) + \frac{\text{dist}(A, U^c)}{4} < \frac{3\text{dist}(A, U^c)}{4}$. Thus we have that

$$\|a - v\| \leq \|a - v'\| + \|v' - v\| < \text{dist}(A, U^c)$$

For all $w \in U^c$,

$$\|w - v\| \geq \|w - a\| - \|a - v\| > \text{dist}(A, U^c) - \text{dist}(A, U^c) = 0$$

which implies that $v \notin U^c$, so $v \in U$ and $\overline{V} \subseteq U$.

Now, notice that A^c is open and $A^c \cup V = \mathbb{R}^n$. Thus $\{A^c, V\}$ form an open cover of \mathbb{R}^n , so we can find a partition of unity $\{\varphi_{A^c}, \varphi_V\}$ subordinate to $\{A^c, V\}$. We claim that our desired function \widehat{f} is given by

$$\widehat{f} = \widetilde{f} \cdot \varphi_V.$$

This function is well defined, as outside of the domain of \widetilde{f} , it is identically equal to 0.

First, we will show that \widehat{f} is differentiable on \mathbb{R}^n . Let $p \in \mathbb{R}^n$, and begin by considering the case when $p \in \overline{V} \subseteq U$. Then for some small enough neighborhood W of p , for all $x \in W$,

$$\widehat{f}(x) = \widetilde{f}(x) \cdot \varphi_V(x)$$

so \widehat{f} is differentiable at p .

If $p \in \mathbb{R}^n \setminus \overline{V}$, then it is also true that $p \notin V$, which is outside of $\text{supp}(\varphi_V)$. Thus, using the fact that $\mathbb{R}^n \setminus \overline{V}$ is open, we get an open neighborhood W' of p such that

$$\widehat{f}(x) = 0, \text{ for } x \in W'.$$

so \widehat{f} is differentiable at p .

Now, we show that it is a valid extension of f . For all $x \in A$, notice that $\varphi_{A^c}|_A \equiv 0$, so it must be true that $\varphi_V|_A \equiv 1$. Then it follows that

$$\widehat{f}|_A = \widetilde{f}|_A = f|_A$$

and the proof is complete. □

1.34 #37

The following set is called the n -**simplex**:

$$\Delta_n := \{\vec{x} = (x_1, \dots, x_n) \in \mathbf{R}^n : x_1, \dots, x_n \geq 0 \text{ and } x_1 + \dots + x_n \leq 1\}.$$

You can assume, without proof, that Δ_n is Jordan measurable.

Find, with proof, an explicit formula for $\mu(\Delta_n)$ in terms of n .

Proof. First, we show that Δ_n is the same as the set

$$S = \left\{ (x_1, \dots, x_n) \in \mathbf{R}^n : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 - x_1, \dots, 0 \leq x_n \leq 1 - \sum_{i=1}^{n-1} x_i \right\}$$

Let $x \in \Delta_n$. We want to show that $0 \leq x_i \leq 1 - \sum_{j=1}^{i-1} x_j$. We get that $x_i \geq 0$ immediately. As well, since $\sum_{j=1}^n x_j \leq 1$ and every component is non-negative, we have that

$$x_i \leq 1 - \sum_{j=1}^{i-1} x_j - \sum_{j=i+1}^n x_j \leq 1 - \sum_{j=1}^{i-1} x_j$$

which shows that $\Delta_n \subseteq S$.

Now, let $x \in S$. We know every x_i is non-negative and additionally

$$x_n \leq 1 - \sum_{i=1}^{n-1} x_i \implies \sum_{i=1}^n x_i \leq 1$$

so $S \subseteq \Delta_n$.

Now, we proceed to find $\mu(S) = \mu(\Delta_n)$. Using Fubini's Theorem, we get

$$\mu(S) = \int_S 1 = \int_0^1 \int_0^{1-x_1} \dots \int_0^{1-\sum_{i=1}^{n-1} x_i} 1 \, dx_n \dots dx_2 \, dx_1$$

Let $I : \mathbb{N} \times [0, 1] \rightarrow \mathbb{R}$ be defined recursively as follows:

$$\begin{aligned} I(1, \alpha) &= \int_0^{1-\alpha} 1 \, dt \\ I(k, \alpha) &= \int_0^{1-\alpha} I(k-1, \alpha+t) \, dt, \end{aligned} \quad \text{for } k > 1.$$

Notice that if we continue applying the definition, we get that

$$I(n, 0) = \mu(S)$$

Now, we will prove using induction on n that for all $\alpha \in [0, 1]$, $I(n, \alpha) = \frac{1}{n!}(1-\alpha)^n$.

Let $n = 1$. Then

$$I(1, \alpha) = \int_0^{1-\alpha} 1 \, dt = 1 - \alpha$$

Now, suppose that $I(k, \alpha) = \frac{1}{k!}(1-\alpha)^k$ holds for all $\alpha \in [0, 1]$ and some $k \in \mathbb{N}$. We want to show that the same holds for $k+1$ as well. For an arbitrary α , we get

$$\begin{aligned} I(k+1, \alpha) &= \int_0^{1-\alpha} I(k, \alpha+t) \, dt \\ &= \int_0^{1-\alpha} \frac{1}{k!}(1-\alpha-t)^k \, dt \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{(k+1)!}(1-\alpha-t)^{k+1}\Big|_0^{1-\alpha} \\
&= \frac{1}{(k+1)!}(1-\alpha)^{k+1}
\end{aligned}$$

as desired. Thus we get that

$$\mu(S) = I(n, 0) = \frac{1}{n!}$$

□

1.35 #38

Basel Problem. Here you will use multivariable calculus to establish the following famous equation:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

To do it, you will evaluate the (improper) double integral $\int_U \frac{1}{1-xy}$ in two ways. Let $f : (0, 1)^2 \rightarrow \mathbf{R}$ be the function given by $f(x, y) = \frac{1}{1-xy}$, and let K_N denote the closed box $\left[\frac{1}{N}, 1 - \frac{1}{N}\right]^2$.

- (a) Evaluate $\int_{K_N} f$ using Fubini's theorem.
- (b) Evaluate $\int_{K_N} f$ using the Change of Variables formula twice: first using the linear diffeomorphism $(x, y) = (u + v, u - v)$, then using the polar coordinates transform.
- (c) Conclude that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Proof. (a):

By Fubini's:

$$\begin{aligned} \int_{K_N} f &= \int_{\frac{1}{N}}^{1-\frac{1}{N}} \int_{\frac{1}{N}}^{1-\frac{1}{N}} \frac{1}{1-xy} dy dx \\ &= \int_{\frac{1}{N}}^{1-\frac{1}{N}} -\frac{1}{x} \ln(1-xy) \Big|_{\frac{1}{N}}^{1-\frac{1}{N}} dx \\ &= \int_{\frac{1}{N}}^{1-\frac{1}{N}} -\frac{1}{x} \left(\ln \left(1 - \left(1 - \frac{1}{N} \right) x \right) - \ln \left(1 - \frac{1}{N} x \right) \right) dx \end{aligned}$$

Notice that $-1 < -\left(1 - \frac{1}{N}\right) < -\frac{1}{N} < 1$, so we can use the power series expansion of $\ln(1+t)$ to get that

$$\begin{aligned} \int_{K_N} f &= \int_{\frac{1}{N}}^{1-\frac{1}{N}} -\frac{1}{x} \left(\sum_{n=1}^{\infty} -\frac{\left(1 - \frac{1}{N}\right)^n x^n}{n} + \sum_{n=1}^{\infty} \frac{\left(\frac{1}{N}\right)^n x^n}{n} \right) dx \\ &= \int_{\frac{1}{N}}^{1-\frac{1}{N}} \sum_{n=1}^{\infty} \frac{x^{n-1}}{n} \left(\left(1 - \frac{1}{N}\right)^n - \left(\frac{1}{N}\right)^n \right) dx \\ &= \sum_{n=1}^{\infty} \frac{x^n}{n^2} \left(\left(1 - \frac{1}{N}\right)^n - \left(\frac{1}{N}\right)^n \right) \Big|_{\frac{1}{N}}^{1-\frac{1}{N}} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^2} \left(\left(1 - \frac{1}{N}\right)^n - \left(\frac{1}{N}\right)^n \right)^2 \end{aligned}$$

(b):

Let $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the diffeomorphism defined by

$$\Phi(x, y) = (x + y, x - y).$$

Notice that $\det J\Phi = \det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2$ and $\Phi^{-1}(x, y) = \left(\frac{x+y}{2}, \frac{x-y}{2}\right)$.

Define

$$E = \left\{ (x, y) \in \mathbb{R}^2 : \frac{1}{N} \leq x \leq 1 - \frac{1}{N}, \left| x - \frac{1}{2} \right| - \frac{1}{2} + \frac{1}{2N} \leq y \leq \frac{1}{2} - \frac{1}{2N} - \left| x - \frac{1}{2} \right| \right\}.$$

We claim that $\Phi(E) = K_N$.

Let $u = \frac{x+y}{2}$, $v = \frac{x-y}{2}$. Recall that our constraints for x and y are $\frac{1}{N} \leq x, y \leq 1 - \frac{1}{N}$. It follows that

$$\frac{1}{N} \leq u \leq 1 - \frac{1}{N}$$

$$\frac{1}{N} \leq u + v \leq 1 - \frac{1}{N} \implies \frac{1}{N} - u \leq v \leq 1 - \frac{1}{N} - u$$

$$\frac{1}{N} \leq u - v \leq 1 - \frac{1}{N} \implies u - 1 + \frac{1}{N} \leq v \leq u - \frac{1}{N}$$

If $\frac{1}{N} \leq u < \frac{1}{2}$, then $u - 1 + \frac{1}{N} < \frac{1}{N} - u$, so the inequality $\frac{1}{N} - u \leq v \leq u - \frac{1}{N}$ suffices. On the other hand, if $\frac{1}{2} \leq u \leq 1 - \frac{1}{N}$, then the inequality $u - 1 + \frac{1}{N} \leq v \leq 1 - \frac{1}{N} - u$ suffices.

Therefore, we have that

$$\begin{aligned} \int_{K_N} f &= \int_E f \circ \Phi \cdot |\det J\Phi| \\ &= 2 \left(\int_{\frac{1}{N}}^{\frac{1}{2}} \int_{\frac{1}{N}-u}^{u-\frac{1}{N}} \frac{1}{1 - (u+v)(u-v)} dv du + \int_{\frac{1}{2}}^{1-\frac{1}{N}} \int_{u-1+\frac{1}{N}}^{1-\frac{1}{N}-u} \frac{1}{1 - (u+v)(u-v)} dv du \right) \\ &= 2 \left(\int_{\frac{1}{N}}^{\frac{1}{2}} \int_{\frac{1}{N}-u}^{u-\frac{1}{N}} \frac{1}{1 - u^2 + v^2} dv du + \int_{\frac{1}{2}}^{1-\frac{1}{N}} \int_{u-1+\frac{1}{N}}^{1-\frac{1}{N}-u} \frac{1}{1 - u^2 + v^2} dv du \right) \end{aligned}$$

By symmetry along $y = 0$, the integral above is equal to

$$4 \left(\int_{\frac{1}{N}}^{\frac{1}{2}} \int_0^{u-\frac{1}{N}} \frac{1}{1 - u^2 + v^2} dv du + \int_{\frac{1}{2}}^{1-\frac{1}{N}} \int_0^{1-\frac{1}{N}-u} \frac{1}{1 - u^2 + v^2} dv du \right)$$

We do both integrals side-by-side:

$$= 4 \left(\int_{\frac{1}{N}}^{\frac{1}{2}} \frac{1}{\sqrt{1-u^2}} \arctan \left(\frac{u - \frac{1}{N}}{\sqrt{1-u^2}} \right) du + \int_{\frac{1}{2}}^{1-\frac{1}{N}} \frac{1}{\sqrt{1-u^2}} \arctan \left(\frac{1 - \frac{1}{N} - u}{\sqrt{1-u^2}} \right) du \right)$$

For the first integral, let $u = \sin \theta$. For the second integral, let $u = \cos \theta$. We have

$$= 4 \left(\int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan \left(\frac{\sin \theta - \frac{1}{N}}{|\cos \theta|} \right) d\theta - \int_{\frac{\pi}{3}}^{\arccos(1-\frac{1}{N})} \arctan \left(\frac{1 - \frac{1}{N} - \cos \theta}{|\sin \theta|} \right) d\theta \right)$$

We remove the absolute values and simplify to get that the integral is equal to

$$4 \left(\int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan \left(\tan \theta - \frac{1}{N} \sec \theta \right) d\theta - \int_{\frac{\pi}{3}}^{\arccos(1-\frac{1}{N})} \arctan \left(\tan \left(\frac{\theta}{2} \right) - \frac{1}{N} \csc \theta \right) d\theta \right)$$

It is quite difficult to find a closed form for these integrals due to the pesky $\frac{1}{N}$, but we will deal with it in part (c).

(c):

Recall from part (a) that $\int_{K_N} f = \sum_{n=1}^{\infty} \frac{1}{n^2} \left(\left(1 - \frac{1}{N}\right)^n - \left(\frac{1}{N}\right)^n \right)^2$. This sum is bounded above by $\sum_{n=1}^{\infty} \frac{1}{n^2}$, so by the dominated convergence theorem as $N \rightarrow \infty$, $\int_{K_N} f \rightarrow \sum_{n=1}^{\infty} \frac{1}{n^2}$. Now, we will show that

$$4 \left(\int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan \left(\tan \theta - \frac{1}{N} \sec \theta \right) d\theta - \int_{\frac{\pi}{3}}^{\arccos(1-\frac{1}{N})} \arctan \left(\tan \left(\frac{\theta}{2} \right) - \frac{1}{N} \csc \theta \right) d\theta \right)$$

converges to

$$\frac{\pi^2}{6}.$$

First, we restrict $N > 2$ to ensure that $\arcsin(\frac{1}{N}) < \frac{\pi}{6}$ and $\arccos(1 - \frac{1}{N}) < \frac{\pi}{3}$. Since $\sec t$ is bounded by $\frac{2}{\sqrt{3}}$ on $[\arcsin(\frac{1}{N}), \frac{\pi}{6}]$, by the monotonicity of $\arctan t$ we have that

$$\arctan \left(\tan \theta - \frac{2}{\sqrt{3}N} \right) \leq \arctan \left(\tan \theta - \frac{1}{N} \sec \theta \right) \leq \arctan \left(\tan \theta + \frac{2}{\sqrt{3}N} \right)$$

We note that $\arctan t$ is uniformly continuous on a closed interval, so for sufficiently large N , we have that for any $\theta \in [\arcsin(\frac{1}{N}), \frac{\pi}{6}]$,

$$\arctan(\tan \theta) - \frac{6\varepsilon}{\pi} < \arctan \left(\tan \theta - \frac{2}{\sqrt{3}N} \right)$$

and

$$\arctan(\tan \theta) + \frac{6\varepsilon}{\pi} > \arctan \left(\tan \theta + \frac{2}{\sqrt{3}N} \right)$$

for some $\varepsilon > 0$ that can be made arbitrarily small. By the monotonicity of the integral,

$$\int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan(\tan \theta) - \frac{6\varepsilon}{\pi} d\theta < \int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan \left(\tan \theta - \frac{1}{N} \sec \theta \right) d\theta$$

and

$$\int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan \left(\tan \theta - \frac{1}{N} \sec \theta \right) d\theta < \int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan(\tan \theta) + \frac{6\varepsilon}{\pi} d\theta$$

It is easy to evaluate the two integrals that bound the original one. We get that

$$\begin{aligned} \int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan(\tan \theta) - \frac{6\varepsilon}{\pi} d\theta &= \int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \theta - \frac{6\varepsilon}{\pi} d\theta \\ &= \frac{1}{2} \left(\frac{\pi^2}{36} - \arcsin^2 \left(\frac{1}{N} \right) \right) - \frac{6\varepsilon}{\pi} \left(\frac{\pi}{6} - \arcsin \left(\frac{1}{N} \right) \right) \\ \int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan(\tan \theta) + \frac{6\varepsilon}{\pi} d\theta &= \int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \theta + \frac{6\varepsilon}{\pi} d\theta \\ &= \frac{1}{2} \left(\frac{\pi^2}{36} - \arcsin^2 \left(\frac{1}{N} \right) \right) + \frac{6\varepsilon}{\pi} \left(\frac{\pi}{6} - \arcsin \left(\frac{1}{N} \right) \right) \end{aligned}$$

We apply Squeeze Theorem to see that

$$\lim_{N \rightarrow \infty} \int_{\arcsin(\frac{1}{N})}^{\frac{\pi}{6}} \arctan \left(\tan \theta - \frac{1}{N} \sec \theta \right) d\theta = \frac{\pi^2}{72}.$$

By a similar argument, we can show that

$$- \lim_{N \rightarrow \infty} \int_{\frac{\pi}{3}}^{\arccos(1-\frac{1}{N})} \arctan \left(\tan \left(\frac{\theta}{2} \right) - \frac{1}{N} \csc \theta \right) d\theta = \int_0^{\frac{\pi}{3}} \frac{\theta}{2} d\theta$$

$$= \frac{\pi^2}{36}$$

Thus our original integral converges to

$$4 \left(\frac{\pi^2}{72} + \frac{\pi^2}{36} \right) = \frac{\pi^2}{18} + \frac{\pi^2}{9} = \frac{\pi^2}{6}$$

Therefore we can conclude that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \lim_{N \rightarrow \infty} \int_{K_N} f = \frac{\pi^2}{6}$$

and we are done. □

1.36 #39

Let $M \subseteq \mathbf{R}^N$ be a set and let $\varphi : U \rightarrow M$ be a smooth parametrization of M , defined on some open set $U \subseteq \mathbf{R}^n$. We say that φ is a **regular parametrization** if φ is a homeomorphism, and if $\varphi'(p)$ has rank n at every point $p \in U$. Note that we require φ^{-1} to be continuous, but not necessarily smooth!

- (a) Prove that if φ is a regular parametrization of M , then φ^{-1} is also smooth.
- (b) Prove that if $\varphi : U \rightarrow M$ and $\psi : V \rightarrow M$ are two regular parametrizations of M , then $\text{vol}_\varphi(M) = \text{vol}_\psi(M)$.
- (c) Use regular parametrizations to find the surface area of a sphere of radius r in \mathbf{R}^4 .

Proof. (a):

First, we show that the inversion map for invertible linear maps $T \mapsto T^{-1}$ is smooth.

Let T be a linear mapping. We will show that the total derivative of the inversion map at T is defined by $h \mapsto -T^{-1}hT^{-1}$. Let $\varepsilon > 0$. Let $\delta = \|T^{-1}\|^{-1}(1 - \frac{\|T^{-1}\|^2}{\varepsilon + \|T^{-1}\|^2})$. Let $\|h\| < \delta$, and recall that for this choice of δ , $T + h$ is still invertible. We have that

$$\begin{aligned} \frac{1}{\|h\|} \|(T+h)^{-1} - T^{-1} + T^{-1}hT^{-1}\| &= \frac{1}{\|h\|} \|(T(I + T^{-1}h))^{-1} - T^{-1} + T^{-1}hT^{-1}\| \\ &= \frac{1}{\|h\|} \|(I + T^{-1}h)^{-1}T^{-1} - T^{-1} + T^{-1}hT^{-1}\| \end{aligned}$$

Notice that $\|I - I - T^{-1}h\|_{\text{op}} = \|T^{-1}h\|_{\text{op}} \leq \|T^{-1}\|_{\text{op}}\|h\|_{\text{op}} < 1$, so by a lemma from long ago, we have that $I + T^{-1}h$ is invertible and

$$(I + T^{-1}h)^{-1} = \sum_{n=0}^{\infty} (I - (I + T^{-1}h))^n = \sum_{n=0}^{\infty} (-1)^n (T^{-1}h)^n$$

Thus we can rewrite our original expression as

$$\begin{aligned} & \frac{1}{\|h\|} \left\| \sum_{n=0}^{\infty} (-1)^n (T^{-1}h)^n T^{-1} - T^{-1} + T^{-1}hT^{-1} \right\| \\ &= \frac{1}{\|h\|} \left\| \left(\sum_{n=0}^{\infty} (-1)^n (T^{-1}h)^n - I \right) T^{-1} + T^{-1}hT^{-1} \right\| \\ &= \frac{1}{\|h\|} \left\| \left(\sum_{n=1}^{\infty} (-1)^n (T^{-1}h)^n \right) T^{-1} + T^{-1}hT^{-1} \right\| \\ &= \frac{1}{\|h\|} \left\| \left(-T^{-1}h + \sum_{n=2}^{\infty} (-1)^n (T^{-1}h)^n \right) T^{-1} + T^{-1}hT^{-1} \right\| \\ &= \frac{1}{\|h\|} \left\| \sum_{n=2}^{\infty} (-1)^n (T^{-1}h)^n T^{-1} \right\| \\ &< \sum_{n=2}^{\infty} \frac{1}{\|h\|} \|(T^{-1}h)^n T^{-1}\| \\ &\leq \sum_{n=2}^{\infty} \frac{1}{\|h\|} \|T^{-1}\|^{n+1} \|h\|^n \\ &= \|T^{-1}\|^3 \|h\| \sum_{n=0}^{\infty} \|T^{-1}\|^n \|h\|^n \\ &= \frac{\|T^{-1}\|^3 \|h\|}{1 - \|T^{-1}\| \|h\|} \quad (\|T^{-1}\| \|h\| < 1) \end{aligned}$$

$$\begin{aligned}
&= \frac{\|T^{-1}\|^2}{1 - \|T^{-1}\|\|h\|} - \|T^{-1}\|^2 \\
&< \varepsilon
\end{aligned}$$

Thus $T^{-1}hT^{-1}$ is the desired derivative. Now, notice that the derivative is also differentiable because it is built from the identity mapping and itself. But this implies that the original mapping is twice differentiable, so the derivative is in fact also twice differentiable. We can continue this process an arbitrary number of times, but hopefully it is convincing enough to show that the inversion mapping is infinitely differentiable. For convenience, define our inversion mapping as V (which i should've done earlier lol).

Suppose φ is a smooth regular parametrization of M . Since φ' has full rank, we can apply the Inverse Function Theorem, we obtain that φ^{-1} is C^1 and

$$(\varphi^{-1})' = (\varphi' \circ \varphi^{-1})^{-1} = V \circ \varphi' \circ \varphi^{-1}$$

We can apply the same argument from before to φ^{-1} to show that it is infinitely differentiable. Since V and φ' are smooth, $(\varphi^{-1})'$ is C^1 , so φ^{-1} is C^2 , and so on. Therefore we can conclude that φ^{-1} is indeed C^∞ .

(b):

Suppose that $\varphi : U \rightarrow M$ and $\psi : V \rightarrow M$ are regular parametrizations of M . We can rewrite

$$\psi = \varphi \circ (\varphi^{-1} \circ \psi).$$

Let $\Phi : V \rightarrow U$ be defined by $\Phi = \varphi^{-1} \circ \psi$. We claim that Φ is a diffeomorphism. ψ is smooth, and by part (a), we know that φ^{-1} is smooth, so Φ is also smooth. Furthermore, $\Phi^{-1} = \psi^{-1} \circ \varphi$, which is smooth for similar reasons. Thus Φ is indeed a diffeomorphism. This being the case, we know that parametrizations are diffeomorphism-invariant (done in class), so we can conclude that

$$\text{vol}_\varphi(M) = \text{vol}_\psi(M).$$

(c):

Let $rS^3 = \{(x, y, z, w) \in \mathbb{R}^4 : \sqrt{x^2 + y^2 + z^2 + w^2} = r\}$. Let $S_w^+ = \{(x, y, z, w) \in \mathbb{R}^4 : \sqrt{x^2 + y^2 + z^2 + w^2} = r, w > 0\}$ be the positive w -hemisphere of rS^3 . We will analyze the volume of S_w^+ , which is half of the total surface area of rS^3 . Let $f(x, y, z) = \sqrt{r^2 - x^2 - y^2 - z^2}$. We define the parametrization $\Phi : B(0, r) \subseteq \mathbb{R}^3 \rightarrow S_w^+$ by

$$\Phi(x, y, z) = (x, y, z, f(x, y, z)).$$

This function is smooth, as the identity function is smooth and f is smooth when $(x, y, z) \neq 0$ and Φ is not defined on 0. We see that

$$J\Phi = \begin{pmatrix} I_3 \\ \nabla f \end{pmatrix}$$

which is guaranteed to be maximal rank 3, so Φ is regular. We have that

$$\begin{aligned}
(J\Phi)^t J\Phi &= \begin{pmatrix} I_3 & (\nabla f)^t \end{pmatrix} \begin{pmatrix} I_3 \\ \nabla f \end{pmatrix} = I_3 + (\nabla f)^t \nabla f \\
&= \begin{pmatrix} 1 + \frac{x^2}{r^2 - x^2 - y^2 - z^2} & \frac{xy}{r^2 - x^2 - y^2 - z^2} & \frac{xz}{r^2 - x^2 - y^2 - z^2} \\ \frac{xy}{r^2 - x^2 - y^2 - z^2} & 1 + \frac{y^2}{r^2 - x^2 - y^2 - z^2} & \frac{yz}{r^2 - x^2 - y^2 - z^2} \\ \frac{xz}{r^2 - x^2 - y^2 - z^2} & \frac{yz}{r^2 - x^2 - y^2 - z^2} & 1 + \frac{z^2}{r^2 - x^2 - y^2 - z^2} \end{pmatrix} \\
&\implies \det((J\Phi)^t J\Phi) = 1 + \frac{x^2 + y^2 + z^2}{r^2 - x^2 - y^2 - z^2} \quad (\text{trust})
\end{aligned}$$

$$= \frac{r^2}{r^2 - x^2 - y^2 - z^2}$$

Therefore the volume is

$$\begin{aligned}
\text{vol}(S_w^+) &= \int_{B(0,r)} \sqrt{\det((J\Phi)^t J\Phi)} \\
&= \int_{-r}^r \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \int_{-\sqrt{r^2-x^2-y^2}}^{\sqrt{r^2-x^2-y^2}} \frac{r}{\sqrt{r^2-x^2-y^2-z^2}} dz dy dx \\
&= 8r \int_0^r \int_0^{\sqrt{r^2-x^2}} \int_0^{\sqrt{r^2-x^2-y^2}} \frac{1}{\sqrt{r^2-x^2-y^2-z^2}} dz dy dx \\
&= 8r \int_0^r \int_0^{\sqrt{r^2-x^2}} \arcsin \left(\frac{z}{\sqrt{r^2-x^2-y^2}} \right) \Big|_0^{\sqrt{r^2-x^2-y^2}} dy dx \\
&= 8r \int_0^r \int_0^{\sqrt{r^2-x^2}} \frac{\pi}{2} dy dx \\
&= 4\pi r \int_0^r \sqrt{r^2-x^2} dx \\
&= 4\pi r \left(\frac{1}{4} \pi r^2 \right) \\
&= \pi^2 r^3
\end{aligned}$$

This is the surface area of one hemisphere, and thus the total surface area of the sphere is $2\pi^2 r^3$. \square

1.37 #40

Let $O_n(\mathbf{R})$ be the set of all $n \times n$ real orthogonal matrices:

$$O_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) : A^t A = I_n\}.$$

Show that O_n is a smooth manifold, and find its dimension.

Proof. First, we will prove the Regular Level-Set Theorem:

Let X, Y be normed vector spaces with dimensions n, m and ordered bases α, β respectively, where $n > m$. Let $F : X \rightarrow Y$ be a smooth function. Define $M = F^{-1}(0_Y)$. If $F'(p)$ is surjective for all $p \in M$, then M is a smooth manifold of dimension $n - m$. That is, if $\phi_\alpha : X \rightarrow \mathbb{R}^n$ is the coordinate isomorphism corresponding to α , then M is a smooth k -manifold if $\phi_\alpha(M)$ is a smooth k -manifold in the usual sense.

This problem reduces to trying to prove that $N := \phi_\alpha(M)$ is a smooth manifold of dimension $n - m$. Notice that $N = \phi_\alpha(F^{-1}(0_Y)) = \phi_\alpha(F^{-1}(\phi_\beta(0_{\mathbb{R}^m})))$. Since ϕ_α, ϕ_β are isomorphisms, we have that $N = \phi_\alpha \circ F^{-1} \circ \phi_\beta(0_{\mathbb{R}^m})$. Let $\hat{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a function defined by $\hat{F} = \phi_\alpha \circ F^{-1} \circ \phi_\beta$, and notice that N is the zero set of \hat{F} . For any $p \in N$, notice that $\hat{F}'(p)$ is surjective, because

$$\hat{F}'(p) = (\phi_\beta \circ F \circ \phi_\alpha^{-1})'(p) = \phi_\beta \circ F'(\phi_\alpha^{-1}) \circ \phi_\alpha^{-1}$$

is just $F'(\phi_\alpha^{-1})$ —a surjective map—composed with linear isomorphisms. This implies that $R(J\hat{F}(p)) = \mathbb{R}^m$ and $\text{rank } J\hat{F}(p) = m$. We can write

$$J\hat{F}(p) = (A \mid B)$$

where A is a $m \times n - m$ matrix and B is a $m \times m$ matrix, and assume without loss of generality that B is invertible, for if not, $J\hat{F}(p)$ is still guaranteed to contain an invertible $m \times m$ submatrix, and we can perform column swaps to move the matrix to the right, which does not change the conclusion of our statement.

Recall that since N is the zero set of \hat{F} , $\hat{F}(p) = 0$. Thus, we write $p = (a, b)$ for $a \in \mathbb{R}^{n-m}$, $b \in \mathbb{R}^m$ and apply the Implicit Function Theorem and obtain an open set $\hat{U} \subseteq \mathbb{R}^{n-m}$ containing a and a C^∞ function $\Phi : \hat{U} \rightarrow \mathbb{R}^m$ so that

$$\hat{F}(x, \Phi(x)) = 0$$

for all $x \in \hat{U}$. We claim that $\varphi : \hat{U} \rightarrow \Phi(\hat{U})$ defined by

$$\varphi(x) = (x, \Phi(x))$$

is our desired smooth regular embedding. It is fairly clear that φ is smooth. Additionally, $J\varphi(x) = \begin{pmatrix} I_{n-m} \\ J\Phi \end{pmatrix}$ is a $(n - m) \times n$ matrix and has at least $n - m$ linearly independent rows, so φ is regular. Finally, if we let $\varphi(x) = \varphi(y)$, we have that $(x, \Phi(x)) = (y, \Phi(y))$, from which we get $x = y$, so φ is injective, and therefore bijective to its image. In addition, we can explicitly find $\varphi^{-1}(y) = \pi_{\mathbb{R}^{n-m}}(y)$, which is continuous because it is a linear map. Therefore φ is a homeomorphism onto its image, so N is a smooth $(n - m)$ -manifold. It follows that $\phi_\alpha^{-1}(N) = M$ is also a smooth manifold of dimension $n - m$ so we are done.

Now, we may proceed to prove that $O_n(\mathbb{R})$ is a smooth manifold of dimension $\frac{1}{2}n(n - 1)$.

We note that $O_n(\mathbb{R})$ is the zero set of the function $f : M_n(\mathbb{R}) \rightarrow S^n$ defined by

$$f(A) = A^t A - I_n$$

where S^n is the set of symmetric $n \times n$ matrices. Notice that f is smooth as it is constructed by smooth functions. Additionally, we show that $f'(X)(h) = X^t h + h^t X$. Indeed,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(X + h) - f(X) - X^t h - h^t X}{\|h\|} &= \lim_{h \rightarrow 0} \frac{(X + h)^t (X + h) - X^t X - X^t h - h^t X}{\|h\|} \\ &= \lim_{h \rightarrow 0} \frac{h^t h}{\|h\|} \\ &= 0 \end{aligned}$$

Next, we want to show that $f'(X)$ is surjective on S^n for all $X \in O_n(\mathbb{R})$.

Let $Y \in S^n$. Let $h = \frac{1}{2}XY$. We see that

$$\begin{aligned} f'(X)(h) &= X^t \left(\frac{1}{2}XY \right) + \left(\frac{1}{2}XY \right)^t X = \frac{1}{2} (X^t XY + Y^t X^t X) \\ &= \frac{1}{2} (Y + Y^t) && (X \text{ is orthogonal}) \\ &= Y && (Y \text{ is symmetric}) \end{aligned}$$

Now, we have the hypotheses needed to apply the Regular Level-Set Theorem, and conclude that $O_n(\mathbb{R})$ is a smooth manifold of dimension $n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1)$. \square

1.38 #41

Let $0 < a < b$. In the xz -plane, draw a circle of radius a centered at the point $(b, 0, 0)$; rotate this circle about the z -axis. The resulting subset of \mathbf{R}^3 is called a **torus**, denoted by $\mathbf{T} = \mathbf{T}_{a,b}$.

- (a) Find a smooth function $f : U \rightarrow \mathbf{R}$, defined on some open set $U \subseteq \mathbf{R}^3$, so that \mathbf{T} is equal to the zero set of f .
- (b) Show that \mathbf{T} is a smooth manifold.
- (c) Find the surface area of \mathbf{T} , in terms of a and b .

Proof. (a):

Notice that in cylindrical coordinates, each θ -slice contains a circle with radius a centered at b , so the torus can be defined by

$$T = \{(r, \theta, z) : (r - b)^2 + z^2 = a^2\}.$$

If we map the polar part of the set back to cartesian coordinates, we see that T is actually the zero set of the function

$$f(x, y, z) = (\sqrt{x^2 + y^2} - b)^2 + z^2 - a^2$$

This function is smooth everywhere except for when $x = 0 = y$, so we let $U = \mathbf{R}^3 \setminus \{(x, y, z) : x = y = 0\}$.

(b):

Notice that for all $p = (x, y, z) \in T$, $f'(p)$ is rank 1, because $f'(p)$ is a 1×3 matrix, so its rank is at most 1, but it cannot be rank 0 because

$$\frac{\partial f}{\partial x}(x, y, z) = 2 \frac{\sqrt{x^2 + y^2} - b}{\sqrt{x^2 + y^2}} = 2 - \frac{2b}{\sqrt{x^2 + y^2}}$$

and

$$\frac{\partial f}{\partial z}(x, y, z) = 2z$$

so being rank 0 implies that

$$f'(p) = \left(2 - \frac{2b}{\sqrt{x^2 + y^2}}, 2 - \frac{2b}{\sqrt{x^2 + y^2}}, 2z \right) = (0, 0, 0) \implies \sqrt{x^2 + y^2} = b \text{ and } z = 0$$

but if this is the case, $f(p) = -a^2 \neq 0$, so $p \notin T$, which is a contradiction.

Thus $Jf(p)$ is always rank 1, so according to the Regular Level-Set Theorem, T is a smooth manifold of dimension $3 - 1 = 2$.

(c):

We split T into 4 quadrants of equal volume, but we will only do the computation for one quadrant. Let Q be the set

$$Q = \{(x, y, z) \in T : x > 0, z > 0\}.$$

We parametrize Q with the function $\varphi : (b - a, b + a) \times (-\frac{\pi}{2}, \frac{\pi}{2})$ defined by

$$\varphi(r, \theta) = (r \cos \theta, r \sin \theta, \sqrt{a^2 - (r - b)^2}).$$

To see that φ is a smooth regular embedding, we see that every component is smooth on Q (the third component is smooth when $r \neq b \pm a$, but $r \in (b - a, b + a)$). As well, the first two components are the polar coordinate transform, so $J\varphi(p)$ is rank 2 for all $p \in (b - a, b + a) \times (-\frac{\pi}{2}, \frac{\pi}{2})$. Finally, notice that $\varphi^{-1}(x, y, z) = \left(\sqrt{x^2 + y^2}, \arctan\left(\frac{y}{x}\right) \right)$ is continuous, so φ is a homeomorphism, thus confirming that φ is a smooth regular embedding. Now, we find an expression for $V(J\varphi(r, \theta))$:

$$V(J\varphi(r, \theta)) = \sqrt{\det((J\varphi(r, \theta))^t J\varphi(r, \theta))}$$

$$\begin{aligned}
&= \sqrt{\det \left(\begin{pmatrix} \cos \theta & \sin \theta & -\frac{r-b}{\sqrt{a^2-(r-b)^2}} \\ -r \sin \theta & r \cos \theta & 0 \end{pmatrix} \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \\ -\frac{r-b}{\sqrt{a^2-(r-b)^2}} & 0 \end{pmatrix} \right)} \\
&= \sqrt{\det \begin{pmatrix} \frac{a^2}{a^2-(r-b)^2} & 0 \\ 0 & r^2 \end{pmatrix}} = \sqrt{\frac{a^2 r^2}{a^2 - (r-b)^2}} \\
&= \frac{ar}{\sqrt{a^2 - (r-b)^2}}
\end{aligned}$$

Now we can calculate the volume of Q :

$$\begin{aligned}
\text{vol}(Q) &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{b-a}^{b+a} V(J\varphi(r, \theta)) \, dr \, d\theta = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{b-a}^{b+a} \frac{ar}{\sqrt{a^2 - (r-b)^2}} \, dr \, d\theta \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-a}^a \frac{a(r+b)}{\sqrt{a^2 - r^2}} \, dr \, d\theta \quad (\text{substitution } r \mapsto r-b) \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-a}^a \frac{ar}{\sqrt{a^2 - r^2}} \, dr \, d\theta + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-a}^a \frac{ab}{\sqrt{a^2 - r^2}} \, dr \, d\theta \\
&= 0 + ab \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \arcsin\left(\frac{r}{a}\right) \Big|_{-a}^a \, d\theta \quad (\text{first integrand is odd}) \\
&= ab \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \pi \, d\theta \\
&= \pi^2 ab
\end{aligned}$$

Finally, we multiply the surface area of the quadrant by 4 to obtain that the total surface area is $4\pi^2 ab$. \square

1.39 #42*

I liked this problem because I found the first part very interesting. While solving this problem, I ran into the issue of trying to get a countable atlas for the manifold. After some work, I realized that all subsets in \mathbb{R}^n permit countable subcovers, which inspired me to do some original work on this concept.

Let $M \subseteq \mathbb{R}^N$ be a smooth n -manifold (with or without boundary!).

- (a) Show that if $n < N$, then M is a Lebesgue null set.
- (b) Show that if $n = N$ and M is closed and its boundary is nonempty, then ∂M coincides with the usual topological boundary (as defined on Handout #2).
- (c) Show that if M is compact and its boundary is nonempty, then M is Jordan measurable.

Proof. (a):

We begin by proving a number of lemmas:

Lemma 1: An open cover of any subset $M \subseteq \mathbb{R}^n$ has a countable subcover.

We know that \mathbb{R}^n is separable, so M is also separable. Let C be a countable dense subset of M . Let \mathcal{U} be an open cover for M . We construct the countable subcover $\hat{\mathcal{U}}$ as follows. For each $q \in C$ and $k \in \mathbb{Q}$, consider the open ball $B(q, k)$. If there exists a $U_{qk} \in \mathcal{U}$ such that $B(q, k) \subseteq U_{qk}$, include it in $\hat{\mathcal{U}}$. Notice that $\hat{\mathcal{U}}$ is at most countable. We claim that it is also an open cover.

Let $x \in M$. Then it is contained in some open set $U \in \mathcal{U}$. As well, we can find an open ball such that $B(x, \delta) \subseteq U$. Since C is dense, we can find $q \in C$ such that $q \in B(x, \frac{\delta}{4})$. Let $k \in \mathbb{Q}$ such that $\frac{\delta}{4} < k < \frac{\delta}{2}$. Then $x \in B(q, k) \subseteq B(x, \delta)$, because for all $y \in B(q, k)$,

$$\|x - y\| \leq \|x - q\| + \|q - y\| < \frac{\delta}{4} + \frac{\delta}{2} < \delta$$

It follows that $B(q, k) \subseteq U$, so it is guaranteed that some U_{qk} from our construction exists. Thus $x \in U_{qk} \in \hat{\mathcal{U}}$ so $\hat{\mathcal{U}}$ is indeed an open cover and we are done.

Lemma 2: A countable union of sets with Jordan measure 0 is a Lebesgue null set.

Let $E = \bigcup_{i \geq 1} E_i$, where $\mu(E_i) = 0$. Let $\varepsilon > 0$. For each E_i , we can find a finite union of boxes B_i such that $B_i \supseteq E_i$ and $\text{vol}(B_i) < \frac{\varepsilon}{2^i}$. We see that $\bigcup_{i \geq 1} B_i$ is a countable union of boxes, $E \subseteq \bigcup_{i \geq 1} B_i$, and

$$\sum_{i=1}^{\infty} \text{vol}(B_i) < \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \frac{\varepsilon}{2(1 - \frac{1}{2})} = \varepsilon$$

as desired.

Lemma 3: If K is a Jordan measurable set with a compact exhaustion K_n , then $\mu(K) = \lim_{n \rightarrow \infty} \mu(K_n)$.

Let $\varepsilon > 0$. Since K is Jordan measurable, we can find a closed polybox $I \subseteq K^\circ$ such that

$$\mu(I) > \mu(K) - \varepsilon$$

Notice that for any $x \in I$, x is contained in some K_n . It follows that $x \in K_{n+1}^\circ$, so I is covered by $\{K_n^\circ\}_{n \in \mathbb{N}}$. Since I is compact, there exists $N \in \mathbb{N}$ such that $I \subseteq K_n$ for all $n > N$. Thus

$$\mu(K_n) > \mu(I) > \mu(K) - \varepsilon \implies |\mu(K) - \mu(K_n)| < \varepsilon$$

as needed.

Now, we prove the problem at hand. Let $M \subseteq \mathbb{R}^N$ be a smooth n -manifold with $n < N$. Let $\{(U_i, \varphi_i)\}_{i \in I}$ be an atlas for M . By Lemma 1, we can assume without loss of generality that the atlas is countable. We can also assume that each U_i is bounded, for if not, we can take a countable union of open balls that cover the unbounded U_i , and restrict the embedding to each ball.

For each chart (U_i, φ_i) , assume that the domain \hat{U} is \mathbb{R}^n or $\overline{\mathbb{H}^n}$, depending on if the chart includes the manifold boundary. Regardless, we can take the compact exhaustion $K_n = B(0, n)$ and see that $\varphi_i(K_n)$ is a compact exhaustion for U_i . Additionally, since φ_i maps from lower dimension to higher dimension, we know

from a previous Big List question that $\mu(\varphi_i(K_n)) = 0$, so by Lemma 3, $\mu(U_i) = 0$. This is true for each chart U_i , so we apply Lemma 2 to conclude that $\cup_{i \geq 1} U_i = M$ is a Lebesgue null set.

(b):

Let p be a point in the topological boundary of M . We will show that $p \in \partial M$. Suppose for contradiction that p is in M° , meaning it is contained in a chart (U, φ) that is diffeomorphic to an open set \hat{U} in \mathbb{R}^N . Note that φ is a diffeomorphism with domain \hat{U} and codomain U . Then it must be true that $\varphi(\hat{U}) = U$ is an open subset of \mathbb{R}^N . But this is a contradiction, as that would imply that the boundary point p is in the interior of M . Therefore $p \in \partial M$.

Next, let $p \in \partial M$ and again suppose for contradiction that p is not in the topological boundary of M . Then it must be true that p is in the topological interior of M . Recall that $p \in \partial M$ implies that it is contained in a chart (U, φ) that is diffeomorphic to \mathbb{H}^n and $p \in \text{bd}(\mathbb{H}^n)$. Since p is in the topological interior of M , we can find an open ball $B(p, r) \subseteq M$ which is also open in \mathbb{R}^N . Then $\varphi^{-1}(B(p, r))$ should also be open in \mathbb{R}^N . But this implies that for small enough δ , $\varphi(p) - (0, \dots, \delta) \in \mathbb{H}^N$, which cannot happen.

Therefore we can conclude that ∂M coincides with the topological boundary of M .

(c):

First, we prove that a compact Lebesgue null set E has Jordan measure 0. It suffices to show that the upper measure $\mu^*(E) = 0$.

Let $\varepsilon > 0$. By definition, we can find a countable union of boxes $B = \bigcup_{i=1}^{\infty} B_i$ such that $E \subseteq B$ and $\text{vol}(B) < \varepsilon$. But since E is compact, it can be covered by finitely many boxes B_{n_i} , $0 < i \leq N$. Thus

$$\text{vol}\left(\bigcup_{i=1}^N B_{n_i}\right) = \sum_{i=1}^N \text{vol}(B_{n_i}) \leq \text{vol}(B) < \varepsilon.$$

Since ε was chosen arbitrarily, we can conclude that $\mu^*(E) = 0$, and E has Jordan measure 0.

Now, suppose that M is compact and its boundary is nonempty. If $\dim M < N$, M is a Lebesgue null set and has Jordan measure 0, and therefore measurable. Otherwise, if $\dim M = N$, since the boundary of M is non-empty, the topological boundary of M is actually a smooth manifold of dimension $(N - 1)$, and therefore a Lebesgue null set. As well, the topological boundary of M is compact, so it is Jordan measure 0, which implies that M is Jordan measurable. \square

1.40 #43

Transverse intersection. Let M, N be two smooth surfaces in \mathbf{R}^3 . We say that M and N **intersect transversally** if $T_p M \neq T_p N$ for all $p \in M \cap N$.

- (a) Prove that if M, N intersect transversally, then $M \cap N$ is a smooth curve in \mathbf{R}^3 .
- (b) Show by example that the conclusion of (a) fails without the assumption of transverse intersection.

Proof. (a):

Suppose that M and N intersect transversally. We will show that $M \cap N$ is a smooth 1-manifold in \mathbf{R}^3 . Let $p \in M \cap N$. Then there is some relatively open neighborhood U of M and V of N that is the zero set of some smooth functions $f, g : \mathbf{R}^3 \rightarrow \mathbf{R}$, that is, $U = Z(f)$ and $V = Z(g)$. Let $\Phi : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ be defined by $\Phi = (f, g)$. We claim that there exists a chart containing p by showing that $U \cap V$ is the zero set of Φ , and $J\Phi(q)$ is surjective for all $q \in U \cap V$. We first start by verifying that $U \cap V$ is relatively open to $M \cap N$. This is quick, as we know that U and V are relatively open to M and N respectively, for each point in $U \cap V$, we can choose two open balls with radii r_1, r_2 that stays within M and N respectively. We then take the lesser of the radii as our radius.

Moving on, we see that Φ is smooth, and $Z(\Phi) = U \cap V$. It remains to show that $J\Phi(q)$ has rank 2 for all $q \in U \cap V$. The Jacobian of Φ is a 2×3 matrix given by

$$J\Phi(q) = \begin{pmatrix} \nabla f(q) \\ \nabla g(q) \end{pmatrix}.$$

It necessarily has rank at most 2 and at least 1 (because $\nabla f, \nabla g \neq 0$). Suppose for contradiction that $\text{rank } J\Phi(q) = 1$. Then $\nabla f(q) = c\nabla g(q)$ for some non-zero constant c . We know that the tangent space $T_q M$ is given by the set of tagged vectors orthogonal to $\nabla f(q)$. Likewise, $T_q N$ consists of tagged vectors orthogonal to $\nabla g(q)$. But notice that for $v \in T_q M$, we have $\nabla f(q) \cdot v = 0$ but also $c\nabla f(q) \cdot v = \nabla g(q) \cdot v = 0$. If we additionally apply the same argument to $u \in T_q N$, we can see that $T_q M = T_q N$ which is a contradiction. Thus $J\Phi(q)$ must have rank 2. From here, it follows that $\Phi^{-1}(\{0\}) = U \cap V$ is a smooth manifold with dimension 1, so we can conclude that $M \cap N$ is a smooth curve.

(b):

Let M be the xy -plane, and let N be the graph of $f(x, y) = x^2 + y^2$. Then $M \cap N$ is simply the origin, which is not a smooth curve.

□

1.41 #44

Suppose that M is a smooth manifold, and let \mathcal{A} be an open cover of M by pairwise consistently oriented charts. Let \mathcal{A}^+ be the collection of all charts on M which are positively oriented with \mathcal{A} ; likewise, let \mathcal{A}^- be the collection of all charts on M which are negatively oriented with \mathcal{A} .

Now suppose that \mathcal{B} is some other open cover of M by charts, such that any two (overlapping) charts in \mathcal{B} are consistently oriented. Prove that if M is connected, then either \mathcal{B} is completely contained in \mathcal{A}^+ , or else it is completely contained in \mathcal{A}^- .

Proof. First, we will prove a form of transitivity for manifold charts.

Lemma. From a manifold M , take two charts (U, φ) , (V, ψ) that are consistently oriented with each other. Then (U, φ) is positively oriented with \mathcal{A} if (V, ψ) is positively oriented with \mathcal{A} .

Suppose that (U, φ) is positively oriented with \mathcal{A} . Consider the chart $(U \cap V, \varphi|_{U \cap V})$ that is consistently oriented with some chart (W, γ) in \mathcal{A} . We will show that (V, ψ) is consistently oriented with (W, γ) . Notice that

$$\psi \circ \gamma^{-1} = (\psi \circ \varphi^{-1}) \circ (\varphi \circ \gamma^{-1})$$

so

$$J(\psi \circ \gamma^{-1}) = J(\psi \circ \varphi^{-1}) \cdot J(\varphi \circ \gamma^{-1}) \text{ and } \det J(\psi \circ \gamma^{-1}) > 0$$

as needed.

Next, we prove the main result. Let \mathcal{B} be an open cover of M by charts. Pick a chart (U, φ) in \mathcal{B} , and say for convenience that it is contained in \mathcal{A}^+ . If the chart picked was in \mathcal{A}^- the argument is analogous. We claim that all other charts are also contained in \mathcal{A}^+ . Let $(V, \psi) \in \mathcal{B}$. Take any two points $p \in U$ and $q \in V$. Since M is connected, we can find a continuous function $\gamma : [0, 1] \rightarrow M$ such that $\gamma(0) = p$ and $\gamma(1) = q$. Consider $C = \gamma[0, 1]$. Since it is compact, we can find a finite number of charts $\{(U_n, \varphi_n)\}_{n \leq N}$ that cover C , such that the open cover includes (U, φ) and (V, ψ) . As well, we can always find a sequence of m charts $(U_{n_k})_{1 \leq k \leq m}$ such that

$$U \cap U_{n_1} \neq \emptyset, U_{n_1} \cap U_{n_2} \neq \emptyset, \dots, U_{n_m} \cap V \neq \emptyset.$$

Suppose if not. Then consider \mathcal{C} , the collection of charts in the finite subcover that can be reached from U using a chain of charts. Let $\bar{\mathcal{C}}$ be the complement; the collection of charts that cannot be reached from U . Then the charts of \mathcal{C} and $\bar{\mathcal{C}}$ are pairwise disjoint from one another, and $V \in \bar{\mathcal{C}}$. We have that $\mathcal{C} \cup \bar{\mathcal{C}}$ cover $\gamma[0, 1]$, but $\bigcup_{C_i \in \mathcal{C}} C_i$ is disjoint from $\bigcup_{D_i \in \bar{\mathcal{C}}} D_i$, which contradicts the fact that $\gamma[0, 1]$ is connected, so such a sequence of charts does indeed exist. Since (U, φ) and (U_{n_1}, φ_{n_1}) are consistently oriented with each other and (U, φ) is positively oriented with \mathcal{A} , by the lemma, (U_{n_1}, φ_{n_1}) is positively oriented with \mathcal{A} . We can apply this argument inductively to conclude that V is also positively oriented with \mathcal{A} and thus $V \in \mathcal{A}^+$.

Therefore, if a single chart in \mathcal{B} is contained in \mathcal{A}^+ , then in fact every chart in \mathcal{B} is contained in \mathcal{A}^+ . The same applies for the case where there is a chart that is contained in \mathcal{A}^- . \square

2 Handout Exercises

2.1 Written by Me

2.1.1 Exercise 4.10.

Solvers: Sanchit, Ethan, Udo

Writeup: Ethan

Let (X, d_X) and (Y, d_Y) be two metric spaces, and let $f : X \rightarrow Y$ be a function. Prove that if f is continuous, then $f^{-1}(U)$ is an open subset of X if U is an open subset of Y .

Proof. Suppose that f is continuous. Let U be an open subset of Y . We will now show that $f^{-1}(U)$ is an open subset of X . Let $x_0 \in f^{-1}(U)$. It follows that $f(x_0) \in U$. We need to find an open ball centered around x_0 such that it is a subset of $f^{-1}(U)$.

Since U is open, there exists an $\varepsilon > 0$ such that $B_Y(f(x_0), \varepsilon) \subseteq U$. By the continuity of f , there is a $\delta > 0$ such that $d_X(x_0, x) < \delta \implies d_Y(f(x_0), f(x)) < \varepsilon$.

Take the open ball $B_X(x_0, \delta)$ in X . We want to show that $B_X(x_0, \delta) \subseteq f^{-1}(U)$.

Let $x \in B_X(x_0, \delta)$. Then $d_X(x_0, x) < \delta$, which implies that $d_Y(f(x_0), f(x)) < \varepsilon$ from continuity. It follows that $f(x) \in B_Y(f(x_0), \varepsilon) \subseteq U$. By definition, this means that $x \in f^{-1}(U)$. Thus $B_X(x_0, \delta) \subseteq f^{-1}(U)$, which means that $f^{-1}(U)$ is an open subset of X . □

2.1.2 Exercise 6.33.

Solvers: Sanchit, Ethan

Writeup: Ethan

Prove that $[0, 1]^2$ is homeomorphic to the closed unit ball $\overline{B}(0, 1)$ in \mathbf{R}^2 .

Proof. It is pretty easy to see that the closed box $[0, 1]^2$ is homeomorphic to $[-1, 1]^2$. We define the function $h : [-1, 1]^2 \rightarrow \overline{B}(0, 1)$ by

$$h(x, y) = \begin{cases} \frac{\|(x, y)\|_{\max}}{\|(x, y)\|_2}(x, y), & \text{if } (x, y) \neq (0, 0); \\ (0, 0), & \text{if } (x, y) = (0, 0). \end{cases}$$

We can verify that this function is indeed well defined because for $(x, y) \in [-1, 1]^2$,

$$\|h(x, y)\|_2 \leq \left\| \frac{\|(x, y)\|_{\max}}{\|(x, y)\|_2}(x, y) \right\|_2 = \frac{\|(x, y)\|_{\max}}{\|(x, y)\|_2} \|(x, y)\|_2 = \|(x, y)\|_{\max} \leq 1$$

which implies that $h(x, y) \in \overline{B}(0, 1)$.

We show that this is a homeomorphism by first showing continuity, and then showing that the inverse is continuous.

Fact! Norms are continuous. Therefore h is continuous when $(x, y) \neq 0$. It remains to show continuity at $(0, 0)$.

Let $\varepsilon > 0$. By the strong equivalence of norms on \mathbb{R}^2 , there is an $M > 0$ such that $\frac{\|(x, y)\|_{\max}}{\|(x, y)\|_2} \leq M$ for all $(x, y) \in \mathbb{R}^2$. Let $\delta = \frac{\varepsilon}{M}$. Let $(x, y) \in \mathbb{R}^2$ such that $\|(x, y)\| < \delta$. Then

$$\|h(x, y)\| = \left\| \frac{\|(x, y)\|_{\max}}{\|(x, y)\|_2}(x, y) \right\| = \frac{\|(x, y)\|_{\max}}{\|(x, y)\|_2} \|(x, y)\| \leq M \|(x, y)\| < \varepsilon$$

Thus h is continuous everywhere. We can explicitly define the inverse h^{-1} as

$$h^{-1}(x, y) = \begin{cases} \frac{\|(x, y)\|_2}{\|(x, y)\|_{\max}}(x, y), & \text{if } (x, y) \neq 0; \\ (0, 0), & \text{if } (x, y) = (0, 0). \end{cases}$$

By a similar argument, we can prove that h^{-1} is continuous. Therefore h is indeed a homeomorphism.

Since $[0, 1]^2 \cong [-1, 1]^2 \cong \overline{B}(0, 1)$, $[0, 1]^2 \cong \overline{B}(0, 1)$ by transitivity. □

2.1.3 Lemma 6.44

Solvers: Ali, Ethan, Samip

Writeup: Ethan

Let X be a vector space with finite dimension n equipped with an arbitrary norm $\|\cdot\|$. Define the linear isomorphism $\Phi : X \rightarrow \mathbb{R}^n$ by

$$\Phi(\vec{x}) = (x_1, \dots, x_n), \text{ where } \vec{x} = \sum_{i=1}^n x_i \vec{b}_i, \{b_1, \dots, b_n\} \text{ is a basis for } X.$$

Define $\|\cdot\|_1$ on X as

$$\|\vec{x}\|_1 = \|\Phi(\vec{x})\|_1$$

Note that the norm on the right hand side is the 1-norm in \mathbb{R}^n . Next, we introduce a lemma that has already been proven.

Lemma 6.43. The closed unit ball of $(X, \|\cdot\|_1)$ is compact.

Corollary. The unit circle in $(X, \|\cdot\|_1)$ is compact.

This follows from the fact that the unit circle is a closed subset of the closed unit ball.

Now, we prove the following lemma:

Lemma 6.44. There exists a constant $m > 0$ such that $m\|\vec{x}\|_1 \leq \|\vec{x}\|$ for all $\vec{x} \in X$.

Proof. Let C' denote the unit circle in $(X, \|\cdot\|_1)$. Define a function $f : C' \rightarrow \mathbb{R}$ by

$$f(\vec{x}) = \frac{\|\vec{x}\|}{\|\vec{x}\|_1}.$$

Since C' is compact, by the generalized EVT, f attains a minimum m . Notice that since norms are positive, $f(\vec{x}) > 0$, so $m > 0$. We claim that this m is the value we are looking for. That is, $m\|\vec{x}\|_1 \leq \|\vec{x}\|$ is true for all $\vec{x} \in X$.

If $\vec{x} = \vec{0}$, then the inequality follows immediately.

Otherwise, for $\vec{x} \in X \setminus \{\vec{0}\}$, notice that $\frac{\vec{x}}{\|\vec{x}\|_1} \in C'$, so

$$f\left(\frac{\vec{x}}{\|\vec{x}\|_1}\right) = \frac{\left\|\frac{\vec{x}}{\|\vec{x}\|_1}\right\|}{\left\|\frac{\vec{x}}{\|\vec{x}\|_1}\right\|_1} \geq m \implies \frac{\frac{1}{\|\vec{x}\|_1}\|\vec{x}\|}{\frac{1}{\|\vec{x}\|_1}} \geq m \implies \|\vec{x}\| \geq m\|\vec{x}\|_1$$

and the proof is complete. □

2.1.4 Exercise 7.16. Chain Rule

Solvers: Ali, Ethan, Ryan

Writeup: Ethan

Let X, Y, Z be normed vector spaces, and $g : X \rightarrow Y$, $f : Y \rightarrow Z$ be functions. For some $p \in X$, suppose that g is totally differentiable at p and f is totally differentiable at $g(p)$. Then the function $f \circ g$ is totally differentiable at p , and $(f \circ g)' = f'(g(p)) \circ g'(p)$.

Proof. Let $L_f = f'(g(p))$ and $L_g = g'(p)$. First, let's prove that $L_f \circ L_g$ is a linear bounded operator. For $a, b \in X$, $c \in \mathbb{R}$,

$$L_f(L_g(ca + b)) = L_f(cL_g(a) + L_g(b)) = cL_f(L_g(a)) + L_f(L_g(b))$$

Thus $L_f \circ L_g$ is linear. Now, since L_f, L_g are bounded linear operators, there exist constants $M_f, M_g > 0$ such that for all $x \in X$, $y \in Y$,

$$\|L_f(y)\|_Z \leq M_f\|y\|_Y \text{ and } \|L_g(x)\|_Y \leq M_g\|x\|_X$$

Then we have that for all $x \in X$,

$$\|L_f(L_g(x))\|_Z \leq M_f\|L_g(x)\|_Y \leq M_f \cdot M_g\|x\|_X$$

so it is bounded as well. Now we can move on to the long part of the proof.

Let $\varepsilon > 0$. By our assumption, there exists positive numbers δ_f, δ_g such that for h_f, h_g with $\|h_f\|_X < \delta_f$, $\|h_g\|_X < \delta_g$,

$$\|f(g(p) + h_f) - f(g(p)) - L_f(h_f)\|_Z \leq \min\left\{\sqrt{\varepsilon}, \frac{\varepsilon}{4M_g}\right\}\|h_f\|_X \quad (1)$$

and

$$\|g(p + h_g) - g(p) - L_g(h_g)\|_Y \leq \min\left\{\frac{\sqrt{\varepsilon}}{4}, \frac{\varepsilon}{2M_f}\right\}\|h_g\|_X \quad (2)$$

As well, since g being totally differentiable at p implies continuity at p , then for all $\|h_c\|_X < \delta_c$,

$$\|g(p + h_c) - g(p)\| < \delta_f \quad (3)$$

Define $\varepsilon_f = \min \left\{ \sqrt{\varepsilon}, \frac{\varepsilon}{4M_g} \right\}$ and $\varepsilon_g = \min \left\{ \frac{\sqrt{\varepsilon}}{4}, \frac{\varepsilon}{2M_f} \right\}$, and let $\delta = \min\{\delta_g, \delta_c, r\}$. For $h \in X$ with $\|h\|_X < \delta$,

$$\begin{aligned} & \frac{\|f(g(p+h)) - f(g(p)) - L_f(L_g(h))\|_Z}{\|h\|_X} \\ & \leq \frac{\|f(g(p+h)) - f(g(p)) - L_f(g(p+h) - g(p))\|_Z + \|L_f(g(p+h) - g(p)) - L_f(L_g(h))\|_Z}{\|h\|_X} \end{aligned}$$

Consider the first term. Because of (3), we can apply (1) with $h_f = g(p+h) - g(p)$ and get that

$$\frac{\|f(g(p+h)) - f(g(p)) - L_f(g(p+h) - g(p))\|_Z}{\|h\|_X} < \frac{\|g(p+h) - g(p)\|}{\|h\|_X} \varepsilon_f$$

Then applying (2) with $h_g = h$,

$$\begin{aligned} & \leq \frac{\|g(p+h) - g(p) - L_g(h)\|_Y + \|L_g(h)\|_Y}{\|h\|_X} \varepsilon_f \leq \frac{\sqrt{\varepsilon}}{4} \varepsilon_f + M_g \varepsilon_f \leq \frac{\sqrt{\varepsilon}}{4} \sqrt{\varepsilon} + M_g \frac{\varepsilon}{4M_g} \\ & = \frac{\varepsilon}{2} \end{aligned}$$

Now, we examine the second term. Using the fact that L_f is a bounded linear operator and (2),

$$\frac{\|L_f(g(p+h) - g(p) - L_g(h))\|_Z}{\|h\|_X} \leq \frac{M_f \|g(p+h) - g(p) - L_g(h)\|_Y}{\|h\|_X} < \frac{M_f \varepsilon}{2M_f} = \frac{\varepsilon}{2}$$

Adding the two together we conclude that

$$\begin{aligned} & \frac{\|f(g(p+h)) - f(g(p)) - L_f(g(p+h) - g(p))\|_Z + \|L_f(g(p+h) - g(p) - L_g(h))\|_Z}{\|h\|_X} \\ & < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

and we are done. □

2.1.5 Exercise 8.12 (c) (iii)

Solvers: Ethan

Writeup: Ethan

Let $U \subseteq \mathbb{R}^n$ be an open set and let $f : U \rightarrow \mathbb{R}$ be a scalar function. Then if all partial derivatives exist and are continuous on U , then f is continuously differentiable and $f'(p)$ is given by $f'(p)(v) = D_v f(p)$.

Recall the statement in (ii): There exists a point $q_k \in U$ such that

$$\frac{f(p_k) - f(p_{k-1})}{h_k} = \frac{\partial f}{\partial x_k}(q_k)$$

We make the additional distinction that q_k is of the form $p_{k-1} + \gamma e_k$, where $|\gamma| < |h_k|$.

Proof. Define $\|\cdot\|$ on U to be the 1-norm.

Let $L_p = f'(p)$. We will show that $\frac{|f(p+h) - f(p) - L_p(h)|}{\|h\|} \rightarrow 0$.

Let $\varepsilon > 0$. Utilising part (i), we define a sequence of points $p_0, \dots, p_n \in X$ by

$$p_0 = p \text{ and } p_i = p_{i-1} + h_i e_i.$$

We know that for $\|h\|$ smaller than some positive δ_1 , $p_i \in U$. By the uniform continuity of $\frac{\partial f}{\partial x_i}$ there is also a δ_2 such that for all $a, b \in U$ such that $\|a - b\| < \delta_2$, $\left| \frac{\partial f}{\partial x_i}(a - b) \right| < \varepsilon$.

Let $\delta = \min\{\delta_1, \delta_2\}$. Let $h \in U$ so that $\|h\| < \delta$. Notice that $p + h = p_n$ and $p = p_0$. We have that

$$\frac{|f(p+h) - f(p) - L_p(h)|}{\|h\|} = \frac{|f(p_n) - f(p_0) - L_p(h)|}{\|h\|}$$

We can expand the numerator by adding and subtracting every term p_i and substituting

$$L_p(h) = D_h f(p) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(p),$$

which yields

$$\frac{\left| \sum_{i=1}^n \left(f(p_i) - f(p_{i-1}) - h_i \frac{\partial f}{\partial x_i}(p) \right) \right|}{\|h\|} \leq \sum_{i=1}^n \frac{\left| f(p_i) - f(p_{i-1}) - h_i \frac{\partial f}{\partial x_i}(p) \right|}{\|h\|}.$$

Now, by part (ii), we can rewrite $f(p_i) - f(p_{i-1})$ as $h_i \frac{\partial f}{\partial x_i}(q_i)$, for some $q_i \in U$, so the expression becomes

$$\sum_{i=1}^n \frac{\left| h_i \frac{\partial f}{\partial x_i}(q_i) - h_i \frac{\partial f}{\partial x_i}(p) \right|}{\|h\|} = \frac{1}{\|h\|} \sum_{i=1}^n |h_i| \left| \frac{\partial f}{\partial x_i}(q_i - p) \right|$$

Note that p_i can also be written as $p + \sum_{j=1}^i h_j e_j$. Thus we can say that $q_i = p_{i-1} + \gamma e_i = p + \gamma e_i + \sum_{j=1}^{i-1} h_j e_j$. We get that

$$\frac{1}{\|h\|} \sum_{i=1}^n |h_i| \left| \frac{\partial f}{\partial x_i}(q_i - p) \right| = \frac{1}{\|h\|} \sum_{i=1}^n |h_i| \left| \frac{\partial f}{\partial x_i} \left(\gamma e_i + \sum_{j=1}^{i-1} h_j e_j \right) \right|$$

We see that the norm of the argument inside the partial derivative is

$$\left\| \gamma e_i + \sum_{j=1}^{i-1} h_j e_j \right\| \leq |\gamma| + \sum_{j=1}^{i-1} |h_j| < \sum_{j=1}^i |h_j| \leq \sum_{j=1}^n |h_j| = \|h\| < \delta_2,$$

so by the continuity of $\frac{\partial f}{\partial x_i}$,

$$\frac{1}{\|h\|} \sum_{i=1}^n |h_i| \left| \frac{\partial f}{\partial x_i} \left(\gamma e_i + \sum_{j=1}^{i-1} h_j e_j \right) \right| < \frac{1}{\|h\|} \sum_{i=1}^n |h_i| \cdot \varepsilon = \frac{\|h\|}{\|h\|} \cdot \varepsilon = \varepsilon$$

and the proof is complete. □

2.1.6 Exercise 9.10

Solvers: Ethan

Writeup: Ethan

Suppose that $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ is a diagonal matrix, and let Q_D be the corresponding quadratic form. Q_D achieves a minimum at the origin if and only if $\lambda_i \geq 0$ for all $i \in \{1, \dots, n\}$.

Proof. First note that $Q_D(x) = x^T D x = (x_1 \ \cdots \ x_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n \lambda_i x_i^2$.

Suppose that every $\lambda_i \geq 0$. It follows that $\lambda_i x_i^2 \geq 0$. Thus

$$Q_D(x) = \sum_{i=1}^n \lambda_i x_i^2 \geq 0$$

Since $Q_D(0) = 0$, Q_D achieves a minimum at the origin.

Conversely, suppose that $\lambda_j < 0$, for some $j \in \{1, \dots, n\}$. $Q_D(0) = 0$, but $Q_D(e_j) = \lambda_j < 0$, so the origin is not a minimum. □

2.1.7 Exercise 11.8

Solvers: Ethan, Ali, Mitya

Writeup: Ethan

Lemma 12.4. Let P be a polybox in \mathbb{R}^n . Then

1. P is a finite disjoint union of boxes.
2. If $\{E_1, \dots, E_k\}$ and $\{F_1, \dots, F_m\}$ are two pairwise disjoint collections of boxes, and $\bigcup_{i=1}^k E_i = \bigcup_{j=1}^m F_j$, then

$$\sum_{i=1}^k \text{vol}(E_i) = \sum_{j=1}^m \text{vol}(F_j)$$

Proof. We start with a proof of part (a). Since P is a polybox, we can write it as a finite union of boxes not necessarily disjoint. We have

$$P = \bigcup_{i=1}^p P_i, \text{ where } P_i \text{ is a box in } \mathbb{R}^n$$

We proceed using induction on p to prove (a). When $p = 1$, P itself is a box and is considered a finite union of disjoint boxes.

Suppose that the claim holds for some $p - 1$. We can rewrite P as

$$P = P_p \cup \bigcup_{i=1}^{p-1} P_i$$

From our assumption, the polybox $\bigcup_{i=1}^{p-1} P_i$ can be expressed as a finite union of disjoint boxes, which we will denote as B_i , so we have

$$P = P_p \cup \bigcup_{i=1}^q B_i = \bigcup_{i=1}^q (B_i \cup P_p)$$

Consider each pair of boxes B_i and P_p . By Lemma 12.2,

1. $P_p \cap B_i$ is a box
2. $P_p \setminus B_i$ and $B_i \setminus P_p$ are finite unions of disjoint boxes.

Additionally, all these sets are disjoint from each other. We can take the union of all the sets above and see that $(P_p \cap B_i) \cup (P_p \setminus B_i) \cup (B_i \setminus P_p) = B_i \cup P_p$ is also a finite union of disjoint boxes. Thus P is a finite union of disjoint boxes.

By induction, part (a) has been proven and we move on to prove part (b).

Let $\{E_1, \dots, E_k\}, \{F_1, \dots, F_m\}$ be pairwise disjoint collections of boxes such that

$$P = \bigcup_{i=1}^k E_i = \bigcup_{j=1}^m F_j$$

For $i \in \{1, \dots, k\}, j \in \{1, \dots, m\}$, define $A_{ij} = E_i \cap F_j$. We will show that $A := \bigcup_{i=1}^k \bigcup_{j=1}^m A_{ij} = P$. It is clear that $A \subseteq P$, so only $P \subseteq A$ will be shown.

Let $x \in P$. Then x must belong to some boxes E_i and F_j , so $x \in A_{ij} \subseteq A$. This shows that A is also a valid disjoint decomposition of P .

Finally, notice that for all E_i, F_j ,

$$E_i = \bigcup_{j=1}^m A_{ij} \text{ and } F_j = \bigcup_{i=1}^k A_{ij}$$

By part (b) of Lemma 12.2, we have

$$\text{vol}(E_i) = \sum_{j=1}^m \text{vol}(A_{ij}) \text{ and } \text{vol}(F_j) = \sum_{i=1}^k \text{vol}(A_{ij})$$

Therefore

$$\sum_{i=1}^k \text{vol}(E_i) = \sum_{i=1}^k \sum_{j=1}^m \text{vol}(A_{ij}) = \sum_{j=1}^m \sum_{i=1}^k \text{vol}(A_{ij}) = \sum_{j=1}^m \text{vol}(F_j)$$

what was what we wanted. □

2.1.8 Fubini's Theorem + Exercise 12.15

Solvers: Ethan Let E be a closed rectangle in \mathbb{R}^2 , so $E = [a, b] \times [c, d]$. Let $f : E \rightarrow \mathbb{R}$ be a continuous function. Then

$$\int_E f = \int_c^d \int_a^b f(x, y) \, dx \, dy$$

Proof. We make the quick note that $f(\cdot, y)$ is integrable because it is continuous. Moreover, it is uniformly continuous because it is defined on the closed interval $[a, b]$. Define $g : [c, d] \rightarrow \mathbb{R}$ by

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

We can argue that g is continuous because for any y ,

$$|g(y+h) - g(y)| = \left| \int_a^b f(x, y+h) \, dx - \int_a^b f(x, y) \, dx \right| < \int_a^b |f(x, y+h) - f(x, y)| \, dx,$$

which we can make arbitrarily small using the uniform continuity of f .

Let P partition $[a, b]$ and Q partition $[c, d]$, and let n, m represent the number of intervals in P and Q respectively. Define

$$M_{ij} = \sup\{f(x, y) : (x, y) \in [p_{i-1}, p_i] \times [q_{j-1}, q_j]\}.$$

We claim that for all $j, y \in [q_{j-1}, q_j]$,

$$\sum_{i=1}^n M_{ij}(p_i - p_{i-1}) \geq U(f(\cdot, y), P).$$

Define $M_{iy} = \sup\{f(x, y) : x \in [p_{i-1}, p_i]\}$. We notice that the set in M_{iy} is a subset of the one in M_{ij} , so we immediately have that $M_{iy} < M_{ij}$. Thus

$$U(f(\cdot, y), P) = \sum_{i=1}^n M_{iy}(p_{i-1}, p_i) \leq \sum_{i=1}^n M_{ij}(p_{i-1}, p_i).$$

From this, we get that for any j ,

$$g(y_j) = \int_a^b f(x, y) \, dx \leq \sum_{i=1}^n M_{ij}(p_{i-1}, p_i), \text{ where } y_j \in [q_{j-1}, q_j],$$

which implies that

$$\sum_{i=1}^n M_{ij}(p_{i-1}, p_i) \geq M_j := \sup\{g(y) : y \in [q_{j-1}, q_j]\}.$$

It follows that

$$\begin{aligned} U(f, (P, Q)) &= \sum_{j=1}^m \sum_{i=1}^n M_{ij}(p_i - p_{i-1})(q_j - q_{j-1}) \geq \sum_{j=1}^m M_j(q_j - q_{j-1}) = U(g, Q) \geq \int_c^d g(y) \, dy, \\ \implies U(f, (P, Q)) &\geq \int_c^d \int_a^b f(x, y) \, dx \, dy. \end{aligned}$$

We can follow the same steps from above but reversing the inequality signs to get that

$$L(f, (P, Q)) \leq \int_c^d \int_a^b f(x, y) \, dx \, dy.$$

Since we did this with arbitrary partitions P, Q , we actually know that

$$L(f) \leq \int_c^d \int_a^b f(x, y) \, dx \, dy \leq U(f),$$

but $L(f) = U(f)$ because of the integrability of f , so we conclude that

$$\int_E f = \int_c^d \int_a^b f(x, y) \, dx \, dy.$$

□

Example: Integrate $f(x, y) = x^y$ on $E = [0, 1]^2$.

Note that f is undefined at $(0, 0)$, but we can construct a function g that is the same as f but with $g(0, 0) = 1$ to solve this issue. It is also continuous, which I invite you to do as an exercise.

Using Fubini's Theorem, we have that

$$\int_E f = \int_0^1 \int_0^1 x^y \, dx \, dy = \int_0^1 \frac{x^{y+1}}{y+1} \Big|_0^1 \, dy = \int_0^1 \frac{1}{y+1} \, dy = \ln|y+1| \Big|_0^1 = \ln 2.$$

Doing this from the definition is kind of painful, so having Fubini in your closet is quite useful.

2.1.9 Exercise 13.14

Recall the statement of Theorem 13.6 and the definition of a simple mapping:

Theorem 13.6. Let $U, V \subseteq \mathbf{R}^n$ be two open sets and let $\Phi : U \rightarrow V$ be a diffeomorphism. Let $E \subseteq U$ be a measurable set. Then $\Phi(E)$ is also measurable, and its measure is

$$\mu(\Phi(E)) = \int_E |\det J\Phi|.$$

Definition 13.12. Let $\Phi : U \rightarrow \mathbf{R}^n$ be a function defined on some set $U \subseteq \mathbf{R}^n$, and let $\varphi_1, \dots, \varphi_n$ be the component functions of Φ :

$$\Phi(p) = (\varphi_1(p), \dots, \varphi_n(p)) \quad \text{for all } p \in U.$$

If the i th component function is just x_i — i.e. $\varphi_i(x_1, \dots, x_n) = x_i$ for some i — then we say that Φ is x_i -**simple**. We say that Φ is **simple** if it is x_i -simple for all but one i .

Exercise 13.14. Show that Theorem 13.6 holds if Φ is simple.

Proof. Let $U, V \subseteq \mathbf{R}^n$ be open, and let $\Phi : U \rightarrow V$ be a simple diffeomorphism. Without loss of generality, let the only non-simple component of Φ be φ_n (I will explain why this reduction is valid below!), and let $E = \prod_{i=1}^n [a_i, b_i] \subseteq U$ be a closed box. Since Φ is a diffeomorphism, this implies that $\Phi(E)$ is measurable, and as well, because Φ is simple,

$$\Phi(E) = \left\{ (x_1, \dots, x_{n-1}, \varphi(x)) \in V : x = (x_1, \dots, x_n) \in \prod_{i=1}^n [a_i, b_i] \right\}$$

Let $E' = \prod_{i=1}^{n-1} [a_i, b_i]$, and define $\Psi : E' \rightarrow \mathbf{R}$ by

$$\Psi(x') = \int_{\varphi(x', [a_n, b_n])} 1 \, dx_n.$$

Φ is a diffeomorphism, so the x' -slice $\varphi(x', \cdot)$ is C^1 and injective. It follows that the image $\varphi(x', [a_n, b_n])$ is an interval. Now, we want to apply single-variable change of variables with $\varphi(x', u) = x_n$. Notice that

$$dx_n = \frac{\partial \varphi}{\partial u}(x', u) du$$

and

$$\det J\Phi = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi}{\partial x_1}(x', u) & \frac{\partial \varphi}{\partial x_2}(x', u) & \cdots & \frac{\partial \varphi}{\partial u}(x', u) \end{vmatrix} = \frac{\partial \varphi}{\partial u}(x', u)$$

so it follows that

$$\Psi(x') = \int_{\varphi(x', [a_n, b_n])} 1 \, dx_n = \int_{a_n}^{b_n} |\det J\Phi|$$

Finally, using Fubini's Theorem, we get that

$$\mu(\Phi(E)) = \int_{\Phi(E)} 1 = \int_{E'} \Psi = \int_{E'} \int_{a_n}^{b_n} |\det J\Phi| = \int_E |\det J\Phi|$$

as needed.

Explanation of Reduction. It, in fact, does not matter in which component Φ is simple in.

Let $\Phi : U \rightarrow V$ be a simple diffeomorphism, where the i th component is non-simple, and we will denote this component as the function $\varphi : U \rightarrow \mathbf{R}$. Let L be the linear map that swaps the i th and n th component. Notice that $L = L^{-1}$ and $\det L = \det L^{-1} = 1$ (this will be important later). We would like to apply the theorem to $L \circ \Phi$, but this is not simple! However, we have a workaround for this issue. Let $\Psi : U \rightarrow V$ be the function defined by $\Psi(x_1, \dots, x_n) = (x_1, \dots, x_{n-1}, \varphi \circ L(x_1, \dots, x_n))$. We claim that $L \circ \Phi(E) = \Psi(E)$.

Let $x \in E$ and consider $L \circ \Phi(x) = (x_1, \dots, x_n, \dots, \varphi(x))$. Let $x' = L(x)$. We have that

$$\Psi(x') = \Psi(x_1, \dots, x_n, \dots, x_i) = (x_1, \dots, x_n, \dots, \varphi(x)) = L \circ \Phi(x)$$

We will not show the converse because it is the same argument. As well, convince yourself that $\det J\Phi = \det J\Psi$. Now, we can apply the statement to the simple function Ψ . Here is everything we need:

- $\mu(L(R)) = |\det L| \mu(R)$: This is Exercise 13.11/12.18
- $\det L = 1$
- $L \circ \Phi(E) = \Psi(E)$
- The exercise above
- $\det J\Psi = \det J\Phi$

We put it all together and see that

$$\mu(\Phi(E)) = \mu(L^{-1} \circ L \circ \Phi(E)) = |\det L^{-1}| \mu(L \circ \Phi(E)) = \mu(\Psi(E)) = \int_E |\det J\Psi| = \int_E |\det J\Phi|$$

□

2.1.10 Exercise 15.14

Solvers: Emerald, Ethan

Writeup: Ethan

Special mention to Alan and Sanchit for helping us out a bit as well!

Let M be a (smooth) n -manifold. Show that each point $p \in M$ has a relatively open neighborhood $U \subseteq M$ such that U homeomorphic to \mathbb{R}^n . In fact, show that there is a smooth regular embedding $\varphi : \mathbb{R}^n \rightarrow U$ such that $\varphi(\vec{0}) = p$.

Proof. Let $p \in M$. By our assumption, we can find a smooth regular embedding $\hat{\varphi} : \hat{V} \rightarrow V$, where $\hat{V} \subseteq \mathbb{R}^n$ is open and V is a relatively open subset of M containing p . Furthermore, there exists $q \in \hat{V}$ such that $\hat{\varphi}(q) = p$ and an open ball $B_{\max}(q, r) \subseteq \hat{V}$. Note that we make use of an open ball with respect to the max-norm, for reasons that will become clear soon. Let $U = \hat{\varphi}(B_{\max}(q, r))$. We are guaranteed that U is open because $\hat{\varphi}$ is a homeomorphism. Now, we define a smooth homeomorphism between $B_{\max}(q, r)$ and \mathbb{R}^n . Let $\Phi : \mathbb{R}^n \rightarrow B_{\max}(q, r)$ be defined by

$$\Phi(\vec{x}) = \left(\frac{2}{\pi r} \arctan(x_1) + q_1, \dots, \frac{2}{\pi r} \arctan(x_n) + q_n \right)$$

Notice that $B_{\max}(q, r)$ is a cartesian product of intervals $\prod_{i=1}^n [q_i - r, q_i + r]$. The function $\Phi_i(x) = \frac{2}{\pi r} \arctan(x) + q_i$ is a well known bijection between \mathbb{R} and $[q_i - r, q_i + r]$, so it is clear that Φ is a bijection with a continuous inverse. Moreover, each component is smooth, so Φ is a smooth homeomorphism. We claim that our desired function $\varphi : \mathbb{R}^n \rightarrow U$ is given by

$$\varphi(x) = \hat{\varphi}(\Phi(x)).$$

φ is a composition of smooth functions, and therefore smooth. $J\Phi$ is a diagonal matrix with non-zero diagonals, so it is rank n . It follows that the Jacobian $J\varphi(x) = J\hat{\varphi}(\Phi(x)) \cdot J\Phi(x)$ is rank n for all $x \in \mathbb{R}^n$. Finally, φ is a homeomorphism since it is a composition of homeomorphisms (so $U \simeq \mathbb{R}^n$, as needed). Thus we can conclude that φ is a smooth regular embedding. As well,

$$\varphi(0) = \hat{\varphi}(\Phi(0)) = \hat{\varphi}(q) = p.$$

and the proof is done.

□

This exercise shows that we can assume WLOG that the domain of every smooth regular embedding is \mathbb{R}^n .

2.1.11 Exercise 16.7. Tangent space to a graph

Ethan

Let $U \subseteq \mathbb{R}^n$ be an open set, and let $f : U \rightarrow \mathbb{R}$ be some function (that is smooth). Let M be the graph of f , so that M is a smooth manifold in \mathbb{R}^{n+1} .

We attempt to find the tangent space $T_p M$ of some point p .

Since M is a graph, we can cover M with a single chart (U, φ) , where $\varphi(x) = (x, f(x))$. Fix $p \in M$ and write $p = \varphi(q)$, for some $q \in U$. From exercise 16.3, a vector v is in the tangent space $T_p M$ if v is in the image of $\varphi'(q)$. That is, the tangent space coincides exactly with the image of $\varphi(q)$. Thus it suffices to find $\varphi(q)(\mathbb{R}^n)$. We have that

$$J\varphi(q) = \left(\frac{I_n}{\nabla f(q)} \right)$$

So for any $h \in \mathbb{R}^n$,

$$\varphi'(q)(h) = (h, \nabla f(q)h)$$

and we can conclude that

$$T_p M = \{ (p, (h, \nabla f(q)h)) \in \mathbb{R}_p^{n+1} : q \in U, p = (q, f(q)), h \in \mathbb{R}^n \}.$$

2.1.12 Exercise 18.15

Recall the following definitions:

- (a) F is **exact** if $F = \nabla f$ for some smooth function $f : U \rightarrow \mathbf{R}$. The function f is called a **scalar potential function** for F .
- (b) F is **conservative** if $\oint_C F \cdot d\vec{x} = 0$ for every loop C contained in U . (A **loop** is the image of a piecewise smooth map $\gamma : [a, b] \rightarrow \mathbf{R}^n$ such that γ is a regular embedding on $[a, b)$, and $\gamma(a) = \gamma(b)$. Or, you could think of it as a smooth, oriented curve C which is closed as a subset of \mathbf{R}^n , and such that $\partial C = \emptyset$.)

In this exercise, we will show that a conservative function is also exact.

Proof. Let $F : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ be conservative. For each connected component U_i of U , we pick an arbitrary $q_i \in U_i$. Define a function $f : U \rightarrow \mathbb{R}$ as follows. For each $p \in U$, it is contained in some U_i . Let $\gamma : [0, 1] \rightarrow U_i$ parametrize a path from p to p_i , that is, γ is smooth, regular, and satisfies $\gamma(0) = p$, $\gamma(1) = p_i$. For this γ , we define

$$f(p) = \oint_{\gamma[0,1]} F \cdot d\vec{x} = \int_0^1 F(\gamma(t)) \cdot \gamma'(t) dt.$$

We assert that this function is well defined, that is to say that $f(p)$ is independent of the choice of the path parametrized by γ . Let $\hat{\gamma} : [0, 1] \rightarrow U$ be a parametrization of another path between p and p_0 . Notice that $\gamma[0, 1] \sqcup \hat{\gamma}[0, 1]$ is a piecewise smooth loop. So from our assumption, we have that

$$\begin{aligned} \oint_{\gamma[0,1] \sqcup \hat{\gamma}[0,1]} F \cdot d\vec{x} &= \oint_{\gamma[0,1]} F \cdot d\vec{x} + \oint_{\hat{\gamma}[0,1]} F \cdot d\vec{x} = \int_0^1 F(\gamma(t)) \cdot \gamma'(t) dt + \int_1^0 F(\hat{\gamma}(t)) \cdot \hat{\gamma}'(t) dt = 0 \\ \implies f(p) &= \int_0^1 F(\gamma(t)) \cdot \gamma'(t) dt = \int_0^1 F(\hat{\gamma}(t)) \cdot \hat{\gamma}'(t) dt \end{aligned}$$

so our choice of path does not matter.

As you can see, the computation above looks wildly different from whatever I was doing in lecture. It turns out that what I did (i.e. parametrizing the loop with a piecewise function) is not the way to evaluate a line integral on a piecewise smooth curve, which is completely my fault. What you should actually do is sum up the line integrals on each individual smooth curve, making sure that each parametrization is oriented correctly. That is why we have one of the integral bounds to be from 1 to 0. We haven't covered a way to formalize the concept of orienting piecewise smooth curves so this argument will have to do for now.

Moving on, we now want to show that $\frac{\partial f}{\partial x_i}(p) = F_i(p)$, which we will do from the limit definition. Let $p \in U$, and take $h \in \mathbb{R}$ small enough so that $p + he_i$ and p lie within the same connected component U_i . Consider two parametrizations: a path from p to p_i , and a path from $p + he_i$ to p_i . We call them $\gamma_1 : [0, 1] \rightarrow U$ and $\gamma_2 : [0, 1] \rightarrow U$ respectively. As well, parametrize the straight line between $p + he_i$ and p with $\varphi(t) = p + the_i$, where $t \in [0, 1]$. Now note that $\gamma_1[0, 1] \sqcup \gamma_2[0, 1] \sqcup \varphi[0, 1]$ form a piecewise smooth closed loop. We have that

$$\oint_{\gamma_1[0,1] \sqcup \gamma_2[0,1] \sqcup \varphi[0,1]} F \cdot d\vec{x} = \int_0^1 F(\gamma_1(t)) \cdot \gamma_1'(t) dt - \int_0^1 F(\gamma_2(t)) \cdot \gamma_2'(t) dt + \int_0^1 F(\varphi(t)) \cdot \varphi'(t) dt = 0$$

and thus

$$\begin{aligned} \frac{1}{h}(f(p + he_i) - f(p)) &= \frac{1}{h} \left(\int_0^1 F(\gamma_2(t)) \cdot \gamma_2'(t) dt - \int_0^1 F(\gamma_1(t)) \cdot \gamma_1'(t) dt \right. \\ &\quad \left. + \int_0^1 F(\varphi(t)) \cdot \varphi'(t) dt - \int_0^1 F(\varphi(t)) \cdot \varphi'(t) dt \right) \\ &= \frac{1}{h} \int_0^1 F(\varphi(t)) \cdot \varphi'(t) dt \\ &= \frac{1}{h} \int_0^1 F(p + the_i) \cdot he_i dt \\ &= \int_0^1 F_i(p + the_i) dt \\ &= \frac{1}{h} \int_0^h F_i(p + ue_i) du \quad (u = th) \end{aligned}$$

If we let $h \rightarrow 0$, we notice that this expression is the derivative of $I(x) = \int_0^x F_i(p + ue_i) du$ at $x = 0$, so

$$\frac{\partial f}{\partial x_i}(p) = \lim_{h \rightarrow 0} \frac{f(p + he_i) - f(p)}{h} = I'(0) = F_i(p)$$

as needed. □

2.2 Not Written by Me

2.2.1 Exercise 0.6.

Solvers: Ethan, Hana, Ryan

Writeup: Ryan

Let $K = [a, b]$ be a closed, bounded interval and let $f : K \rightarrow \mathbf{R}$ be a continuous function. Prove that $f(K)$ is also a closed, bounded interval.

Proof. To show that $f(K)$ is a bounded interval, recall that for an interval to be bounded, $\forall x \in K$, there must exist some $c, d \in \mathbf{R}$ such that $c \leq f(x) \leq d$.

As the function f is a continuous function on K from the question statement, we use the Extreme Value Theorem, and determine that f has a minimum m and a maximum M on K . Thus, we have it that $\forall x \in K, m \leq f(x) \leq M$. Thus, $f(K)$ is bounded.

To prove that $f(K)$ is a closed interval, begin by assuming that $f(K)$ is an open interval. This means that there is no $y \in f(K)$ such that either $f(a) = y$ or $f(b) = y$. (Note that the y value specified can be different for $f(a), f(b)$, respectively). Since $f(K)$ is bounded, we know that $f(K)$ has both a minimum and a maximum. Call the minimum m and the maximum M . Define x_{\min}, x_{\max} such that $f(x_{\min}) = m$ and $f(x_{\max}) = M$.

Let $y \in f([a, b])$ (or $f(K)$). Since we know that $f(K)$ is continuous and $f(x_{\min}) \leq y \leq f(x_{\max})$, we can use Intermediate Value Theorem (IVT) and conclude that $\exists x \in K$ such that $f(x) = y$. This indicates that every y value in $f(K)$ has atleast one associated x such that $f(x) = y$.

This means, that for some $y_a, y_b \in f(K)$, $f(a) = y_a$ and $f(b) = y_b$. This contradicts our earlier statement, of which there was no y value in $f(K)$ such that $f(a) = y$ or $f(b) = y$.

Thus, $f(K)$ must be a closed interval. □

2.2.2 Exercise 2.21

Solvers: Sanchit, Udo, Ethan

Writeup: Sanchit

What can be said about $\overline{A \cup B}$?

Note: For notation we denote $\{1, \dots, n\} = [n]$

Claim 1. Let (X, d) be a metric space, then for $X_i \subseteq X$ for $i \in \mathbf{N}$ we have $\overline{\bigcup_{i \in \mathbf{N}} X_i} = \bigcup_{i \in \mathbf{N}} \overline{X_i}$.

Proof. First we show $\overline{\bigcup_{i \in \mathbf{N}} X_i} \subseteq \bigcup_{i \in \mathbf{N}} \overline{X_i}$.

Let $x \in \overline{\bigcup_{i \in \mathbf{N}} X_i}$ be arbitrary, then for all $\varepsilon > 0$, $B(x, \varepsilon) \cap \bigcup_{i \in \mathbf{N}} X_i \neq \emptyset$. Then we see that for all $\varepsilon > 0$, $\bigcup_{i \in \mathbf{N}} B(x, \varepsilon) \cap X_i \neq \emptyset$ as intersection distributes over union. However this implies there exists an $k \in \mathbf{N}$ such that for all $\varepsilon > 0$, $B(x, \varepsilon) \cap X_k \neq \emptyset$ and thus $x \in \overline{X_k}$ and so $x \in \bigcup_{i \in \mathbf{N}} \overline{X_i}$. Thus $\overline{\bigcup_{i \in \mathbf{N}} X_i} \subseteq \bigcup_{i \in \mathbf{N}} \overline{X_i}$.

Next we show $\bigcup_{i \in \mathbf{N}} \overline{X_i} \subseteq \overline{\bigcup_{i \in \mathbf{N}} X_i}$.

Let $x \in \bigcup_{i \in \mathbf{N}} \overline{X_i}$ be arbitrary, then there exists $k \in \mathbf{N}$ such that $x \in \overline{X_k}$. Then by definition we see that for all $\varepsilon > 0$, $B(x, \varepsilon) \cap X_k \neq \emptyset$. However this implies that for all $\varepsilon > 0$, $B(x, \varepsilon) \cap \bigcup_{i \in \mathbf{N}} X_i \neq \emptyset$. However this implies $x \in \overline{\bigcup_{i \in \mathbf{N}} X_i}$. Thus $\bigcup_{i \in \mathbf{N}} \overline{X_i} \subseteq \overline{\bigcup_{i \in \mathbf{N}} X_i}$.

By double subset inclusion we conclude that $\overline{\bigcup_{i \in \mathbf{N}} X_i} = \bigcup_{i \in \mathbf{N}} \overline{X_i}$. □

Claim 2. Let (X, d) be a metric space and let $\Omega \subseteq \mathcal{P}(X)$ be a collection of subsets of X . Then we have $\bigcup_{X' \in \Omega} \overline{X'} \subseteq \overline{\bigcup_{X' \in \Omega} X'}$.

The difference between this claim and the result from the previous claim is that the previous claim only holds for finitely many sets. However this claim holds for an arbitrary amount of sets.

Proof. Let $x \in \bigcup_{X' \in \Omega} \overline{X'}$ be arbitrary, then there exists $X' \in \Omega$ such that $x \in \overline{X'}$. Then by definition we see that for all $\varepsilon > 0$, $B(x, \varepsilon) \cap X' \neq \emptyset$. However this implies that for all $\varepsilon > 0$, $B(x, \varepsilon) \cap \bigcup_{X' \in \Omega} X' \neq \emptyset$. However this implies $x \in \overline{\bigcup_{X' \in \Omega} X'}$. Thus $\bigcup_{X' \in \Omega} \overline{X'} \subseteq \overline{\bigcup_{X' \in \Omega} X'}$. □

Claim 3. Let (X, d) be a metric space, then for $X_i \subseteq X$ for $i \in \mathbf{N}$ we do **NOT** have $\overline{\bigcup_{i \in \mathbf{N}} X_i} \subseteq \bigcup_{i \in \mathbf{N}} \overline{X_i}$.

Proof. Consider the following counter example:

Let $(X, d) = (\mathbf{R}, |\cdot|)$ and define $X_i = [0, 1 - \frac{1}{i}]$. Now we see that $\overline{X_i} = [0, 1 - \frac{1}{i}]$ and thus

$$\bigcup_{i \in \mathbf{N}} \overline{X_i} = \bigcup_{i \in \mathbf{N}} X_i = [0, 1)$$

However then one sees that

$$\overline{\bigcup_{i \in \mathbf{N}} X_i} = \overline{[0, 1)} = [0, 1]$$

One trivially verifies $[0, 1] \not\subseteq [0, 1)$. □

2.2.3 Exercise 3.12.

Solvers: Ethan, Jaohar, Ali.

Writeup: Ali.

Lemma: \mathbf{Q}^n is dense in \mathbf{R}^n (with respect to $\|\cdot\|_2$), for any $n \in \mathbf{Z}^+$.

Proof. Let $n \in \mathbf{Z}^+$ be arbitrary. To show that \mathbf{Q}^n is dense in \mathbf{R}^n , it is sufficient to show that $\forall \vec{x} \in \mathbf{R}^n, \forall \epsilon > 0, B(x, \epsilon) \cap \mathbf{Q}^n \neq \emptyset$. Let $\vec{x} \in \mathbf{R}^n$ be arbitrary

We can express \vec{x} as an n-tuple (x_1, \dots, x_n) . For every $i \in \mathbf{N}$ with $1 \leq i \leq n$, let $q_i \in (x_i - \epsilon/\sqrt{n}, x_i + \epsilon/\sqrt{n}) \cap \mathbf{Q}$ be arbitrary, by the density of \mathbf{Q} in \mathbf{R} . Let $\vec{q} = (q_1, \dots, q_n)$. $\vec{q} \in \mathbf{Q}^n$. Then $\|\vec{q} - \vec{x}\|_2 = \sqrt{\sum_{i=1}^n (q_i - x_i)^2} < \sqrt{\sum_{i=1}^n \frac{\epsilon^2}{n}} = \sqrt{n \cdot \frac{\epsilon^2}{n}} = \epsilon$. Therefore $\vec{q} \in B(x, \epsilon) \cap \mathbf{Q}^n$, as needed.

\mathbf{Q}^n is dense in \mathbf{R}^n (with respect to $\|\cdot\|_2$) □

Let $x \in \mathbf{R}^d$. Prove that there exists a sequence (q_n) of points in \mathbf{Q}^d such that $q_n \rightarrow x$ (with respect to $\|\cdot\|_2$).

Proof. Let $x \in \mathbf{R}^d$ be arbitrary. By the lemma, we know that $B(x, 1/n) \cap \mathbf{Q}^d$ is non empty, for any point $x \in \mathbf{R}^d$ and any positive real number n . Define q_n as any arbitrary element in the non-empty set $B(x, 1/n) \cap \mathbf{Q}^d$.

Now, we show that (q_n) converges to x .

Let $\epsilon > 0$ be arbitrary. Let $N \in \mathbf{N}$ be chosen such that $1/N < \epsilon$. Then $\forall n \geq N, q_n \in B(x, 1/n)$. Thus $\|q_n - x\|_2 < 1/n \leq 1/N < \epsilon$, as needed.

We have shown that $\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n \in \mathbf{N}, (n \geq N) \implies (\|q_n - x\|_2 < \epsilon)$. Therefore $q_n \rightarrow x$. □

2.2.4 Exercise 4.24.

Solvers: Sanchit, Ethan

Writeup: Sanchit

Let X, Y be metric spaces and let $f : X \rightarrow Y$ be a continuous function. If $A \subseteq X$ is compact, does it follow that $f(A)$ is also compact? If $B \subseteq Y$ is compact, does it follow that $f^{-1}(B)$ is also compact?

$f(A)$ is also compact.

Proof. Let $\{U_n\}_{n \in \mathbf{N}}$ be an open cover of $f(A)$. We see as f is continuous we have that the preimage of an open set must be open and so the preimage of U_n must be open for each $n \in \mathbf{N}$. We then see that $\{f^{-1}(U_n)\}_{n \in \mathbf{N}}$ forms a collection of open sets. Now as for each $x \in A$ we have $f(x) \in f(A) \subseteq \bigcup_{n \in \mathbf{N}} U_n$,

we then have $f(x) \in U_k$ for some $k \in \mathbf{N}$ and so $x \in f^{-1}(U_k)$. We then have $A \subseteq \bigcup_{n \in \mathbf{N}} f^{-1}(U_n)$ and so

$\{f^{-1}(U_n)\}_{n \in \mathbf{N}}$ is an open cover of A . Now we have by the compactness of A that it permits a finite subcover $\{f^{-1}(U_n)\}_{n \in I}$ where $I \subseteq \mathbf{N}$ is a finite set. We then see that $f(A) \subseteq \bigcup_{n \in I} U_n$ and so $\{(U_n)\}_{n \in I}$ is a finite subcover of $f(A)$. □

$f^{-1}(B)$ is not compact. Consider the following counter example.

Counter Example. Consider $f : \mathbf{R} \rightarrow \mathbf{R}$ defined by $f(x) = \sin(x)$

It is known that $[-1, 1]$ is compact and we have that $f^{-1}([-1, 1]) = \mathbf{R}$, however \mathbf{R} is not compact as it is unbounded in \mathbf{R} . □

Unfortunately we see that even if we restrict f to being uniformly continuous the above counter example still holds.

Consider the follow up where we restrict X to be compact.

Proof. Let B be compact in Y , then we see that B is closed. Now as f is continuous we have $f^{-1}(B)$ is closed in X . Now by the theorem that a closed subset of a compact space is compact we have that $f^{-1}(B)$ is compact. \square

2.2.5 Exercise 6.28

Solvers: Sanchit, Ethan, Udo

Writeup: Sanchit

Show that if X is a finite set, then any two metrics on X are topologically equivalent.

Proof. Let d_1 and d_2 be two metrics on X . Now let U be an open set in (X, d_1) , we show it is open in (X, d_2) . We have that for all $x \in U$, there exists $\varepsilon > 0$ such that $B_1(x, \varepsilon) \subseteq U$. Fix $k = \min\{d_2(x, y) \mid y \in X \setminus \{x\}\}$. We see that k exists by the finiteness of X , moreover it is non zero. Now we see that $B_2(x, k) = \{x\}$. We see that $\{x\}$ is open as $B_2(x, k) \subseteq \{x\}$ and so its only point x is an interior point. We see then that $B_2(x, k) \subseteq B_1(x, \varepsilon) \subseteq U$. Thus x is an interior point of U in (X, d_1) . Thus we see that if U is open in (X, d_1) then it is open in (X, d_2) . The proof for the reverse direction follows similarly. Thus we see that any two metrics on X are topologically equivalent. \square

2.2.6 Exercise 9.21

Solvers: Ethan, Udo

Writeup: Udo

prop: find the δ such that $|f(p_0 + h) - f(p_0) - Q(h)| < \epsilon$ if $\|h\| < \delta$.

PROOOOOF. Since Hf is continuous, choose δ_H such that

$$\|Hf(p_0 + r) - Hf(p_0)\| < \epsilon$$

for $\|r\| < \delta_H$.

Now, let $\delta = \min\{\delta_H, \sqrt{\frac{\epsilon}{2}}\}$ for all $\epsilon > 0$. Suppose that $\|h\| < \delta$, then with 9.20, choose $\theta \in (0, 1)$ such that

$$|f(p_0 + h) - f(p_0) - Q(h)| \leq \frac{1}{2} \|Hf(p_0 + \theta h) - Hf(p_0)\|_{\text{op}} \|h\|^2$$

But $\|\theta h\| < \delta_H$, by hypothesis, so

$$\frac{1}{2} \|Hf(p_0 + \theta h) - Hf(p_0)\|_{\text{op}} \|h\|^2 < \frac{1}{2} \times 2 \|h\|^2 = \|h\|^2 < \delta^2 < \frac{\epsilon}{2} < \epsilon$$

as was desired.

qed qed qed qed qed

2.2.7 14.8

Writeup: Ali

Solvers: Ethan, Ryan, Ali

14.8 Theorem *invariance of volume under reparametrization* Let M be a parametrizable set in \mathbf{R}^N , and

let $\alpha : U \rightarrow M$ and $\beta : V \rightarrow M$ two parametrizations of M , where $U, V \subseteq \mathbf{R}^n$ are two open sets. Suppose that $\Phi : U \rightarrow V$ is a diffeomorphism such that $\beta \circ \Phi = \alpha$. Then

$$\text{vol}_\alpha(M) = \text{vol}_\beta(M).$$

Proof.

$$\text{vol}_\alpha(M) = \int_U V(J\alpha)$$

By definition of vol

$$\int_U V(J\alpha) = \int_{\Phi^{-1}(V)} V(J(\beta \circ \Phi))$$

By definition of Φ

$$\int_{\Phi^{-1}(V)} V(J(\beta \circ \Phi)) = \int_{\Phi^{-1}(V)} V((J\beta \circ \Phi) \circ J\Phi)$$

By application of the multivariate chain rule on $J(\beta \circ \Phi)$

$$\int_{\Phi^{-1}(V)} V((J\beta \circ \Phi) \circ J\Phi) = \int_{\Phi^{-1}(V)} \sqrt{\det(((J\beta \circ \Phi) \circ J\Phi)^t((J\beta \circ \Phi) \circ J\Phi))}$$

By the definition of V , the volume of a parallelepiped

$$\int_{\Phi^{-1}(V)} \sqrt{\det(((J\beta \circ \Phi) \circ J\Phi)^t((J\beta \circ \Phi) \circ J\Phi))} = \int_{\Phi^{-1}(V)} \sqrt{\det((J\Phi)^t(J\beta \circ \Phi)^t(J\beta \circ \Phi)(J\Phi))}$$

By property of transpose $(AB)^t = B^t A^t$ for matrices A and B

$$\int_{\Phi^{-1}(V)} \sqrt{\det((J\Phi)^t(J\beta \circ \Phi)^t(J\beta \circ \Phi)(J\Phi))} = \int_{\Phi^{-1}(V)} \sqrt{\det(J\Phi)^t \det((J\beta \circ \Phi)^t(J\beta \circ \Phi)) \det(J\Phi)}$$

Because $J\Phi$ is a square matrix and by the properties of determinant

$$\int_{\Phi^{-1}(V)} \sqrt{\det(J\Phi)^t \det((J\beta \circ \Phi)^t(J\beta \circ \Phi)) \det(J\Phi)} = \int_{\Phi^{-1}(V)} \sqrt{\det((J\beta \circ \Phi)^t(J\beta \circ \Phi)) \det(J\Phi)^2}$$

Because $\det A^t = \det A$ for any square matrix A

$$\int_{\Phi^{-1}(V)} \sqrt{\det((J\beta \circ \Phi)^t(J\beta \circ \Phi)) \det(J\Phi)^2} = \int_{\Phi^{-1}(V)} \sqrt{\det((J\beta \circ \Phi)^t(J\beta \circ \Phi))} \cdot |\det J\Phi|$$

Because $\sqrt{x^2} = |x|$

$$\int_{\Phi^{-1}(V)} \sqrt{\det((J\beta \circ \Phi)^t(J\beta \circ \Phi))} \cdot |\det J\Phi| = \int_{\Phi^{-1}(V)} V(J\beta \circ \Phi) \cdot |\det J\Phi|$$

By the definition of V

$$\int_{\Phi^{-1}(V)} V(J\beta \circ \Phi) \cdot |\det J\Phi| = \int_{\Phi^{-1}(V)} (V(J\beta) \circ \Phi) \cdot |\det J\Phi|$$

Because $V(J\beta \circ \Phi) = V \circ J\beta \circ \Phi = V(J\beta) \circ \Phi$

$$\int_{\Phi^{-1}(V)} (V(J\beta) \circ \Phi) \cdot |\det J\Phi| = \int_V V(J\beta)$$

By the Change of Variables Theorem from Unit 13

Finally, we have shown that

$$\text{vol}_\alpha(M) = \int_U V(J\alpha) = \int_V V(J\beta) = \text{vol}_\beta(M)$$

As needed. □

3 Original Work

All the statements in this section are some results I proved in no particular order.

3.1 Theorem. *Let (X, d) be a separable metric space. Then $A \subseteq X$ is separable with respect to the metric d .*

Proof. Since X is separable, it is pre-totally bounded. We will show that a metric subspace A is separable by showing that it is pre-totally bounded.

Let $\varepsilon > 0$. Since X is pre-totally bounded, there exists a finite collection of open balls $\{B(x_i, \frac{\varepsilon}{2})\}_{i \leq n}$ that covers X . We define the desired collection of open balls in A with the following method:

If $B(x_i, \frac{\varepsilon}{2})$ contains some element $a_i \in A$, we add $B_A(a_i, \varepsilon)$ to the collection. Notice for $a \in B(x_i, \frac{\varepsilon}{2}) \cap A$,

$$d(a, a_i) \leq d(x, a) + d(x, a_i) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which implies that $B(x_i, \frac{\varepsilon}{2}) \subseteq B_A(a_i, \varepsilon)$.

Continuing on, the number of open balls in this collection is at most the number of open balls in the original collection in X , so $\{B_A(a_i, \varepsilon)\}$ is finite.

It remains to show that the collection covers A . Let $a \in A$. Then $a \in B(x_i, \frac{\varepsilon}{2})$ for some i . It follows that $a \in B(x_i, \frac{\varepsilon}{2}) \cap A \subseteq B_A(a_i, \varepsilon)$. Thus A is covered by this collection, so A is pre-totally bounded.

Thus we can conclude that A is separable. □

3.2 Theorem. *Let X, Y be metric spaces let and $f: C \rightarrow X$ be a continuous function. If $A \subseteq X$ is seperable, then $f(A)$ is seperable.*

Proof. Define d_X and d_Y to be metrics on X and Y respectively. Suppose that A is seperable. We will show that $f(A)$ is seperable by equivalently showing that it is pre-totally bounded.

Let $\varepsilon > 0$. By the continuity of f , for every $a \in A$, there exists $\delta > 0$ such that for all $x \in A$, $d_X(x, a) < \delta \implies d_Y(f(x), f(a)) < \varepsilon$. Keep this value of δ .

Since A is seperable, A is pre-totally bounded. By definition, A is covered by a countable subcover $\{B_X(a_i, \delta)\}_{i \in \mathbb{N}}$.

Consider the countable collection $\{B_Y(f(a_i), \varepsilon)\}_{i \in \mathbb{N}}$. We will show that this collection covers $f(A)$. Let $y \in f(A)$. Then $y = f(x)$ for some $x \in A \subseteq \{B_X(a_i, \delta)\}_{i \in \mathbb{N}}$. This implies that x is in some open ball $B_X(a_k, \delta) \implies d_X(x, a_k) < \delta$. By the continuity of f , this implies that $d_Y(y, f(a_k)) = d_Y(f(x), f(a_k)) < \varepsilon \implies y \in B_Y(f(a_k), \varepsilon) \subseteq \{B_Y(f(a_i), \varepsilon)\}_{i \in \mathbb{N}}$.

We see that $f(A)$ is pre-totally bounded, which implies that $f(A)$ is seperable. □

3.3 Lemma. $\forall x \in X, r > 0, \overline{B}(0, 1) \simeq \overline{B}(x, r)$.

Proof. Let $x \in X, r > 0$. Define the homeomorphism $f: \overline{B}(0, 1) \rightarrow \overline{B}(x, r)$ by

$$f(y) = ry + x$$

First, we will show continuity. Let $p \in \overline{B}(0, 1)$. Let $\varepsilon > 0$. Let $\delta = \frac{\varepsilon}{|r|}$. Let $y \in \overline{B}(0, 1)$ such that $\|y - p\| < \delta$. Then

$$\|f(y) - f(p)\| = \|ry - rp\| = |r|\|y - p\| < \varepsilon$$

We can show that f is bijective by showing it has a left and right inverse.

Let $f^{-1}(z) = \frac{1}{r}(z - x)$. Notice that

$$f^{-1}(f(y)) = \frac{1}{r}(f(y) - x) = y$$

and

$$f(f^{-1}(z)) = rf^{-1}(z) - x = z$$

Thus f is bijective and its inverse is given by f^{-1} , which is continuous, which shows that f is a homeomorphism.

□

3.4 Proposition. *Let X be a normed vector space. Then $\overline{B}(0, 1)$ is compact if and only if X has a compact exhaustion.*

Proof. Suppose that the closed unit ball is compact. Then we can construct the compact exhaustion $K_n = \overline{B}(0, n)$. These sets are compact because compactness is preserved by the homeomorphism in the lemma. Notice that $(K_n)^\circ \subseteq K_{n+1}$, as

$$x \in (K_n)^\circ \implies \|x\| < n \leq n+1 \implies x \in K_{n+1}$$

Trivially, $\bigcup_{i=1}^{\infty} K_i \subseteq X$. Now, let $x \in X$. Then there exists $j \in \mathbb{N}$ such that $\|x\| \leq j \implies x \in K_j \subseteq \bigcup_{i=1}^{\infty} K_i$. Thus $X = \bigcup_{i=1}^{\infty} K_i$, so X has a compact exhaustion.

Conversely, assume that X has a compact exhaustion $(K_n)_{n \geq 1}$. Take the first non-empty set K_j . Since $K_j \subseteq (K_{j+1})^\circ$, then for $v \in K_j$, there exists an open ball such that $B(v, \varepsilon) \subseteq (K_{j+1})^\circ$. It follows that $\overline{B}(v, \varepsilon) \subseteq K_{j+1}$. By the lemma, it follows that $\overline{B}(v, \varepsilon) \simeq \overline{B}(0, 1)$, which implies that $\overline{B}(0, 1)$ is compact.

□

3.5 Corollary. *Let X be a normed vector space. Then X has a compact exhaustion if and only if X is finite dimensional.*

Proof. This follows from the previous proposition and a Big List problem.

□

3.6 Proposition. *Equip \mathbb{R}^n with any norm $\|\cdot\|_a$. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function given by $f(x) = \|x\|$, for any norm $\|\cdot\|$ on \mathbb{R}^n . Then f is not differentiable at $x = 0$.*

Proof. Suppose for contradiction that f is differentiable at $x = 0$. Then it has a linear approximation at 0. That is, for some $L \in B(\mathbb{R}^n, \mathbb{R})$,

$$\lim_{h \rightarrow 0} \frac{f(h) - f(0) - L(h)}{\|h\|_a} = 0 \implies \lim_{h \rightarrow 0} \frac{\|h\| - L(h)}{\|h\|_a} = 0$$

Since norms are equivalent in \mathbb{R}^n , there exists an $N > 0$ such that

$$N\|h\|_a \leq \|h\|, \text{ for all } h \in \mathbb{R}^n$$

Using the limit definition, for $h \in \mathbb{R}^n$ such that $\|h\|$ is smaller than some δ ,

$$|\|h\| - L(h)| < N\|h\|_a$$

We can also take the negative of this h and use the linearity of L to see that

$$|\|h\| + L(h)| < N\|h\|_a$$

We see that for all h , $\|h\| < \delta$,

$$\begin{aligned} \|h\| &= \frac{1}{2} \|(h + L(h)) + (h - L(h))\| \leq \frac{1}{2} (\|h + L(h)\| + \|h - L(h)\|) < \frac{1}{2} (N\|h\|_a + N\|h\|_a) \\ &\implies \|h\| < N\|h\|_a \end{aligned}$$

which is a contradiction.

□

3.7 Theorem. *Let Y and K be metric spaces with K compact. Let $f : K \rightarrow Y$ be a continuous function. Then f is uniformly continuous.*

Proof. Let $\varepsilon > 0$. For each $p \in K$, we can find an open ball $U_p \subseteq K$ such that for all $x \in U_p$,

$$d_Y(f(p), f(x)) < \frac{\varepsilon}{2}$$

Notice that $\{U_p\}_{p \in K}$ forms an open cover of K , so by compactness, we can find a Lebesgue number δ for this open cover. We claim that this choice of δ suffices for uniform continuity. Let $x, y \in K$ with $d_K(x, y) < \delta$. Then $y \in B(x, \delta)$. This open ball is contained in some open set U_p of our open cover. It follows that

$$d_Y(f(x), f(y)) \leq d_Y(f(p), f(x)) + d_Y(f(p), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as needed. □

3.8 Theorem. (*Fubini's Theorem*) Let $E \subseteq \mathbb{R}^c, F \subseteq \mathbb{R}^d$ be two measurable sets. Let $f : E \times F \rightarrow \mathbb{R}$ be an integrable function such that the x -slice $f(x, \cdot)$ is integrable for all $x \in E$. Define $\int_F f : E \rightarrow \mathbb{R}$ by $(\int_F f)(x) = \int_F f(x, \cdot)$. Then

$$\int_{E \times F} f = \int_E \int_F f$$

Proof. Let P, Q be partitions for E and F respectively. Consider the subbox P_i of P . For each subbox Q_j in Q , define $M_{ij} = \sup\{f(x, y) : (x, y) \in P_i \times Q_j\}$. For any $x_0 \in P_i$, we have

$$\int_F f(x_0, \cdot) \leq U(f(x_0, \cdot), Q) = \sum_{j=1}^m \sup\{f(x_0, y) : y \in Q_j\} \text{vol}(Q_j) \leq \sum_{j=1}^m M_{ij} \text{vol}(Q_j)$$

$$\implies \sup \left\{ \int_F f(x, \cdot) : x \in P_i \right\} \leq \sum_{j=1}^m M_{ij} \text{vol}(Q_j)$$

$$\implies U \left(\int_F f, P \right) \leq \sum_{i=1}^n \sum_{j=1}^m M_{ij} \text{vol}(Q_j) \text{vol}(P_i) = U(f, (P, Q))$$

$$\implies \int_E \int_F f \leq U(f)$$

By an analogous argument, we can additionally show that $\int_E \int_F f \geq L(f)$. Thus

$$L(f) \leq \int_E \int_F f \leq U(f)$$

and since $L(f) = U(f) = \int_{E \times F} f$, we can conclude that

$$\int_{E \times F} f = \int_E \int_F f.$$

□

3.1 Antiderivatives

In this section, we look into a possible generalisation of function antiderivatives. We create some definitions and prove results that help us solve specific problems. In particular, we will use the concept of antiderivatives to prove that multivariate polynomials are closed under differentiation and integration.

To begin, we are interested in functions that “look like” the derivative of another function. The following proposition characterizes this.

3.9 Proposition. Exact equations. Let $M, N : \mathbb{R}^2 \rightarrow \mathbb{R}$ be totally differentiable functions. Then there exists a twice differentiable function $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that $\frac{\partial F}{\partial x} = M$ and $\frac{\partial F}{\partial y} = N$ if and only if $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$.

Proof. Suppose that such F does indeed exist and is C^2 . By Clairaut's theorem we have

$$\frac{\partial M}{\partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial N}{\partial x}$$

Conversely, suppose that $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. We define $\int f \, dx$ to be the antiderivative of $f(., y)$ with additional constant $C = 0$, and likewise for $\int f \, dy$. Since N, M are continuous, they are integrable with respect to both x and y . Let F be defined by

$$F(x, y) = \frac{1}{2} \int M(x, y) \, dx + \frac{1}{2} \int N(x, y) \, dy$$

Taking the partial derivative with respect to x , with the application of Feynmann's technique we see that

$$\begin{aligned} \frac{\partial F}{\partial x}(x, y) &= \frac{1}{2} M(x, y) + \frac{1}{2} \int \frac{\partial N}{\partial x}(x, y) \, dy \\ &= \frac{1}{2} M(x, y) + \frac{1}{2} \int \frac{\partial M}{\partial y}(x, y) \, dy \\ &= \frac{1}{2} M(x, y) + \frac{1}{2} M(x, y) \\ &= M(x, y) \end{aligned}$$

Since F is symmetric, taking the partial derivative with respect to y will similarly give us $\frac{\partial F}{\partial y}(x, y) = N(x, y)$, and the proof is complete. □

It is important to note that this is not true in general, namely for an arbitrary domain $U \subseteq \mathbb{R}^n$, this statement may not hold. For instance, take $f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2$ to be

$$f(x, y) = \left(\frac{y}{x^2 + y^2}, \frac{-x}{x^2 + y^2} \right)$$

Then

$$\begin{aligned} \frac{d}{dy} \left(\frac{y}{x^2 + y^2} \right) &= \frac{x^2 + y^2 - 2y^2}{(x^2 + y^2)^2} = \frac{x^2 - y^2}{(x^2 + y^2)^2} \\ \frac{d}{dx} \left(\frac{-x}{x^2 + y^2} \right) &= \frac{-(x^2 + y^2) + 2x^2}{(x^2 + y^2)^2} = \frac{x^2 - y^2}{(x^2 + y^2)^2} \end{aligned}$$

so the partial derivatives of each component function match. Following the formula we found above, the supposed function is given by

$$F(x, y) = \frac{1}{2} \int \frac{y}{x^2 + y^2} \, dx + \frac{1}{2} \int \frac{-x}{x^2 + y^2} \, dy = \frac{1}{2} \arctan \left(\frac{x}{y} \right) - \frac{1}{2} \arctan \left(\frac{y}{x} \right).$$

Each partial derivative of F does indeed match the components of f , but the issue is that F is undefined when either $x = 0$ or $y = 0$, which is within the domain of f . I have not found a sufficient restriction on the domain to make the implication true, so we need the condition that the domain is \mathbb{R}^n . Luckily, this is enough to prove results about polynomials, as their domain is \mathbb{R}^n .

3.10 Corollary. *Suppose we have a differential equation of the form*

$$M(x, y) + N(x, y) \frac{dy}{dx} = 0$$

where $M, N : \mathbb{R}^2 \rightarrow \mathbb{R}$ are C^1 functions in terms of x and y , $N \neq 0$, and $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. Then this differential equation has a solution.

Proof. By the proposition above, we have that there exists some C^2 function $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ with $\frac{\partial F}{\partial x} = M$ and $\frac{\partial F}{\partial y} = N \neq 0$. In particular, for some $x_0, y_0 \in \mathbb{R}$,

$$\frac{\partial F}{\partial y}(x_0, y_0) \neq 0.$$

Let C be an arbitrary constant, and set $F(x, y) = C$. By the Implicit Function Theorem, we have that there exists an open set $U \subseteq \mathbb{R}$ containing x_0 and a function $\Psi : U \rightarrow \mathbb{R}$ so that

$$F(x, \Psi(x)) = C.$$

We claim that a solution to the differential equation is given by Ψ . We see that by taking the derivative of the equation above, we immediately get the original differential equation:

$$M(x, \Psi(x)) + N(x, \Psi(x))\Psi'(x) = 0.$$

□

It may seem tricky to find the Ψ function, but the work involved amounts to isolating y in the equation $F(x, y) = C$, which we will illustrate with the following example.

3.11 Example. Find the solution to the differential equation

$$x \frac{dy}{dx} + y = 2xe^x + 6x^2.$$

Let $M(x, y) = 2xe^x + 6x^2 - y$, $N(x, y) = -x$. We see that

$$\frac{\partial M}{\partial y}(x, y) = -1 = \frac{\partial N}{\partial x}(x, y)$$

By the proposition above, we can find a function $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ so that

$$\frac{\partial F}{\partial x}(x, y) = M(x, y) \text{ and } \frac{\partial F}{\partial y}(x, y) = N(x, y).$$

In particular, by following the explicit formula, we have

$$F(x, y) = 2xe^x - 2e^x + 2x^3 - xy.$$

Set $2xe^x - 2e^x + 2x^3 - xy = C$, for some constant $C \in \mathbb{R}$. We claim the solution to the differential equation is $y = 2e^x - \frac{2}{x}e^x + 2x^2 - \frac{C}{x}$. Indeed, we have that

$$\frac{dy}{dx} = 2e^x + \frac{2}{x^2}e^x - \frac{2}{x}e^x + 4x + \frac{C}{x^2}$$

so

$$x \frac{dy}{dx} + y = x \left(2e^x + \frac{2}{x^2}e^x - \frac{2}{x}e^x + 4x + \frac{C}{x^2} \right) + 2e^x - \frac{2}{x}e^x + 2x^2 - \frac{C}{x} = 2xe^x + 6x^2$$

as desired.

We can generalize this result to multiple variables:

3.12 Theorem. Let $N_1, \dots, N_n : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable functions. Then there exists a twice differentiable function $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that $\frac{\partial F}{\partial x_i} = N_i$ if and only if $\frac{\partial N_i}{\partial x_j} = \frac{\partial N_j}{\partial x_i}$ for all $i, j \in \{1, \dots, n\}$.

Proof. Suppose such F does exist. For each i, j , we can apply Clairaut's Theorem to each component function of N_i, N_j to see that

$$\frac{\partial N_i}{\partial x_j} = \frac{\partial F}{\partial x_j \partial x_i} = \frac{\partial F}{\partial x_i \partial x_j} = \frac{\partial N_j}{\partial x_i}.$$

Conversely, suppose that $\frac{\partial N_i}{\partial x_j} = \frac{\partial N_j}{\partial x_i}$ for all $i, j \in \{1, \dots, n\}$. Similar to the smaller case, define

$$F(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n \int N_i(x_1, \dots, x_n) dx_i$$

For $k = 1, 2, \dots, n$, we see that

$$\begin{aligned} \frac{\partial F}{\partial x_k} &= \frac{1}{n} \sum_{i=1}^n \frac{\partial N_i}{\partial x_k}(x_1, \dots, x_n) dx_i \\ &= \frac{1}{n} \left(N_k(x_1, \dots, x_n) + \sum_{i \neq k} \int \frac{\partial N_i}{\partial x_k} dx_i \right) \\ &= \frac{1}{n} \left(N_k(x_1, \dots, x_n) + \sum_{i \neq k} \int \frac{\partial N_k}{\partial x_i}(x_1, \dots, x_n) dx_i \right) \\ &= \frac{1}{n} \left(N_k(x_1, \dots, x_n) + \sum_{i \neq k} N_k(x_1, \dots, x_n) \right) \\ &= \frac{1}{n} (n N_k(x_1, \dots, x_n)) \\ &= N_k(x_1, \dots, x_n) \end{aligned}$$

□

We can use this notion to define a multivariate antiderivative as follows:

3.13 Definition. Let $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a function of the form $f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))$, where U is open and each $f_i : U \rightarrow \mathbb{R}^m$ are differentiable functions. An **antiderivative** of f , if it exists, is a function $F : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that $\frac{\partial F}{\partial x_i} = f_i$ for all i .

We can also define repeated antiderivatives recursively.

3.14 Definition. Let $k \in \mathbb{N}$, $f : \mathbb{R}^n \rightarrow \mathbb{R}^{n^k m}$. The k th antiderivative of f is

- The antiderivative of f if $k = 1$ (in the sense of the definition above),
- The antiderivative of the $k - 1$ th antiderivative of f if $k > 1$.

3.15 Example. Let $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by

$$f(x, y, z) = (y^2 \sin z, 2xy \sin z, xy^2 \cos z).$$

We first verify that f has an antiderivative. We compare the required partial derivatives:

$$\begin{aligned} \frac{\partial f_x}{\partial y}(x, y, z) &= 2y \sin z \\ \frac{\partial f_y}{\partial x}(x, y, z) &= 2y \sin z \\ \frac{\partial f_x}{\partial z}(x, y, z) &= y^2 \cos z \\ \frac{\partial f_z}{\partial x}(x, y, z) &= y^2 \cos z \\ \frac{\partial f_y}{\partial z}(x, y, z) &= 2xy \cos z \\ \frac{\partial f_z}{\partial y}(x, y, z) &= 2xy \cos z \end{aligned}$$

Since the pairs match, we confirm that the antiderivative of f does indeed exist and is given by

$$F(x, y, z) = \frac{1}{3} \left(\int y^2 \sin z \, dx + \int 2xy \sin z \, dy + \int xy^2 \cos z \, dz \right) = xy^2 \sin z$$

A motivation for defining antiderivatives in this way is to be able to express higher order derivatives in a more condensed and understandable way, as we will see below. First, we define a way to transform functions between arbitrary finite dimensional vector spaces over \mathbb{R} into functions between \mathbb{R}^n and \mathbb{R}^m .

3.16 Definition. Let V, W be finite dimensional vector spaces with dimensions n, m respectively and let $f : V \rightarrow W$ be a function. Let α, β be ordered bases for V, W , and let $\phi_\alpha : V \rightarrow \mathbb{R}^n$ be the standard coordinate isomorphism to \mathbb{R}^n corresponding to α , and define ϕ_β similarly. A **coordinate representation** for f with respect to α and β is the function $\hat{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by

$$\hat{f} = \phi_\beta \circ f \circ \phi_\alpha^{-1}.$$

3.17 Theorem. If $F : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the antiderivative of $f = (f_1, \dots, f_n) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, then

$$JF(p) = \begin{pmatrix} f_1(p) \\ \vdots \\ f_n(p) \end{pmatrix}^t$$

for all $p \in U$.

Proof. Let $F = (F_1, \dots, F_m)$, where $F_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are functions. As well, write

$$f = (f_1, \dots, f_n)$$

where $f_i = (f_{i1}, \dots, f_{in})$. Then note that

$$(JF(p))_{ij} = \frac{\partial F_i}{\partial x_j}(p)$$

but since F is the antiderivative of f , we know that $\frac{\partial F}{\partial x_j} = f_j$, so

$$\frac{\partial F_i}{\partial x_j}(p) = f_{ji} = (f_1(p) \quad \cdots \quad f_n(p))_{ji} = \begin{pmatrix} f_1(p) \\ \vdots \\ f_n(p) \end{pmatrix}_{ij}^t.$$

□

It follows that f is the coordinate representation for $(JF)^t$ with respect to the standard bases of \mathbb{R}^n and $M_{n \times m}(\mathbb{R})$. We define this unique function to be the **function representation** of JF or F' . It is not hard to see that the converse holds as well:

3.18 Proposition. If f is the function representation of JF , then F is an antiderivative of f .

Proof. Reverse the steps from the proof above. □

This concept may not seem too useful when only working with first derivatives, but quickly becomes enticing when used for higher order derivatives.

3.19 Theorem. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a smooth function, and let $f : \mathbb{R}^n \rightarrow \mathbb{R}^{k \cdot m}$ be the function representation of $J^k F$. Then the function representation of f' is also the function representation of $J^{k+1} F$.

Proof. We denote $f = (f_1, \dots, f_{n^k m})$, where $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are functions. Let g be the function representation of f' . Write $g = (g_1, \dots, g_n)$, where $g_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n^{k+1} m}$ are functions. Then

$$Jf = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}^t.$$

But since f is the function representation of $J^k F$,

$$J^k F = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}^t.$$

We take the Jacobian of both sides to obtain that

$$J^{k+1} F = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{n^k m}}{\partial x_1} & \dots & \frac{\partial f_{n^k m}}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1}^t & \dots & \frac{\partial f}{\partial x_n}^t \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}^t$$

Recall that f is an antiderivative of g , so

$$J^{k+1} F = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}^t$$

so g is indeed the function representation for $J^{k+1} F$. \square

We will use these concepts to prove a theorem about polynomials. First, we do some preparation by listing definitions and lemmas.

3.20 Definition. A **monomial** in n variables is a function $m : \mathbb{R}^n \rightarrow \mathbb{R}$ of the form $m(x_1, \dots, x_n) = x_1^{i_1} \cdots x_n^{i_n}$, where $i_1, \dots, i_n \geq 0$. The degree of a monomial is $\deg(m) = i_1 + \dots + i_n$. A **polynomial** is a linear combination of monomials, and its degree is defined to be the degree of the monomial with the highest degree. Additionally, we define a **k -polynomial** to be a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ such that each component is a polynomial in n variables. Likewise, its degree is defined to be the degree of the highest degree component. Note that using this definition, a polynomial in the sense of the above definition is considered a 1-polynomial.

3.21 Proposition. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ be a k -polynomial in n variables. Then the characteristic function of Jf is a nk -polynomial in n variables.

Proof. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^{nk}$ be the characteristic function of Jf . By Theorem 4.15, we know that each component function of g is a partial derivative of some component function of f . Denote this component as f_i . We claim that $\frac{\partial f_i}{\partial x_j}$ is a polynomial. \square

3.22 Lemma. The partial derivative of a monomial is a monomial. In particular, if $f(x_1, \dots, x_n) = x_1^{j_1} \cdots x_n^{j_n}$, then

$$\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = j_i x_1^{j_1} \cdots x_i^{j_i-1} \cdots x_n^{j_n}$$

Proof. For any fixed point $p = (p_1, \dots, p_n) \in \mathbb{R}^n$, let $g_i(t) = f(p + te_i) = p_1^{j_1} \cdots (p_i + t)^{j_i} \cdots p_n^{j_n}$. We have that

$$\frac{\partial f}{\partial x_i}(p) = g_i'(0) = j_i p_1^{j_1} \cdots p_i^{j_i-1} \cdots p_n^{j_n}$$

as required. \square

The following corollary follows immediately and generalizes the above result for an arbitrary k -polynomials.

3.23 Corollary. *The partial derivative of any k -polynomial is a k -polynomial.*

The next result relates the degree of a monomial to the degree of its partial derivatives.

3.24 Lemma. *Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be a monomial in n variables, so $g(x_1, \dots, x_n) = x_1^{j_1} \cdots x_n^{j_n}$. For each $i \in \{1, \dots, n\}$, $\deg\left(\frac{\partial g}{\partial x_i}\right) = \deg(g) - 1$.*

Proof. Fix $p = (p_1, \dots, p_n) \in \mathbb{R}^n$. Let $i \in \{1, \dots, n\}$. Notice that

$$\frac{\partial g}{\partial x_i}(x_1, \dots, x_n) = j_i x_1^{j_1} \cdots x_i^{j_i-1} \cdots x_n^{j_n}$$

and $\deg\left(\frac{\partial g}{\partial x_i}\right) = j_1 + \cdots + (j_i - 1) + \cdots + j_n = \deg(g) - 1$. □

This extends easily to k -polynomials, as the degree of every monomial in a polynomial reduces by one, immediately implying the result, which implies the same thing for k -polynomials.

The next result will show that polynomials are “closed” under differentiation.

3.25 Theorem. *If $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a m -polynomial with degree d , then the function representation \hat{f} of Jf is a nm -polynomial with degree $d - 1$.*

Proof. This result is fairly quick, as we know that the components of \hat{f} are the partial derivatives of f , where are polynomials themselves. Since \hat{f} has nm components, we can conclude that it is a nm -polynomial. As well, since the degree of each component of \hat{f} is one less than some component from f , we can also conclude that $\deg(\hat{f}) = d - 1$. □

Now, we can prove the main result we sought out to prove.

3.26 Theorem. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ be a polynomial in n variables. There exists $k \in \mathbb{N}$ such that $f^{(k)} = 0$. As well, if f is a function such that $f^{(k)} = 0$ for some $k \in \mathbb{N}$, then f is necessarily a polynomial.*

Proof. Let d be the degree of f . Take $k = d + 1$. Let g_m be the function representation of $f^{(m)}$. We show using induction that $\deg(g_m) = d - m$, for $m \in \{0, \dots, d\}$.

Let $m = 0$. Then $g_m = f$, so $\deg(g_m) = d$.

Now, for $m < d$, suppose that $\deg(g_m) = d - m$. Then the function representation of Jg_m has degree $d - m - 1$, but we know that this is also the function representation of $f^{(m+1)}$, which is g_{m+1} , so $\deg(g_{m+1}) = d - m - 1$.

Thus g_d has degree 0, but this simply means that g_d is a constant function, from which it follows that g_{d+1} is the zero function. But this implies that $f^{(d+1)} = 0$.

Now, we prove the second half using induction on k .

Let $k = 1$. Then $f' = 0$, which implies that f is a constant function, which is a polynomial of degree 0.

Now, suppose that the statement is true for some $k \in \mathbb{N}$. Suppose that $f^{(k+1)} = 0$. Let \hat{f} be the function representation of f' . Notice that the function representation g of $f^{(k+1)}$ is the function representation of $\hat{f}^{(k)}$. We have that

$$(f)^{(k+1)} = 0 \iff g = 0 \iff \hat{f}^{(k)} = 0$$

By our assumption, we know that \hat{f} is a polynomial, but it is also the function representation of f' , which means that f is the antiderivative of \hat{f} . If we write $\hat{f} = (\hat{f}_1, \dots, \hat{f}_n)$, where $\hat{f}_i : \mathbb{R}^n \rightarrow \mathbb{R}^p$ are functions. Then we have that

$$f = \frac{1}{n} \sum_{i=1}^n \int \hat{f}_i dx_i$$

but each $\int \hat{f}_i dx_i$ is a polynomial, and a linear combination of polynomials is also a polynomial, so we can conclude that f is a polynomial, which completes the proof. □

3.2 Precompactness

We know that compactness can play a big role in proving results, such as the existence of partitions of unity. Here, we look into a slightly weaker version of compactness and see what results can be proved with it.

3.27 Definition. A metric space X is *precompact* if any open cover permits a *countable* subcover.

While it can be cumbersome to directly find examples of precompact metric spaces, we can show that precompactness is equivalent to certain metric space properties, and thus show that all metric spaces with those properties are also precompact.

3.28 Proposition. *A metric space is precompact if and only if it is separable.*

Proof. Let X be a metric space. Suppose that X is precompact. For $n \in \mathbb{N}$, let $U_n = \{B(x_{ni}, \frac{1}{n})\}_{i \in I}$ be an open cover of X , and suppose without loss of generality that U_n is countable. It follows that $U = \bigcup_{n=1}^{\infty} U_n$ is countable as well. Let $C = \{x_{ni}\}_{i \in I}$ be the collection of all the open ball centers in U . This set is countable, so it only remains to show that C is dense in X . Let $x \in X$, $\varepsilon > 0$. Take n such that $\frac{1}{n} < \varepsilon$. Then x is contained in some open ball $B(x_i, \frac{1}{n})$. But this implies that $x_i \in B(x, \varepsilon)$, so C is dense. Thus X is separable.

Conversely, suppose that X is separable. Let $\{U_i\}_{i \in I}$ be an open cover for X . Let $C \subseteq X$ be a separable subset of X and write it as $C = \{q_1, q_2, \dots\}$. We construct the countable subcover U as follows. For all $n, k \in \mathbb{N}$, consider the open ball $B(q_n, \frac{1}{k})$. For each of these balls, if there exists an open set U_{nk} inside the cover such that $B(q_n, \frac{1}{k}) \subseteq U_{nk}$, we include U_{nk} inside the subcover. Notice that U is countable by construction.

Let $x \in X$. There is some open set U_i , not necessarily in U that contains x . Take the open ball $B(x, \delta) \subseteq U_i$. By the separability of X , there exists some $q_n \in C$ so that $q_n \in B(x, \frac{\delta}{2})$. Let $k \in \mathbb{N}$ such that $\frac{1}{k} < \frac{\delta}{2}$ and consider another open ball $B(q_n, \frac{1}{k}) \subseteq B(x, \delta)$. We are guaranteed that U_{nk} exists, because $B(q_n, \frac{1}{k}) \subseteq U_i$, so $B(q_n, \frac{1}{k})$ is of at least one set in the open cover. Thus $x \in U_{nk} \in U$, so U is indeed a countable open cover so X is precompact. \square

3.29 Corollary. \mathbb{R}^n is precompact.

Proof. This is because \mathbb{R}^n is separable. \square

3.30 Corollary. *Every subset of a precompact metric space is also precompact.*

Proof. We know that precompactness is equivalent to separability, and we also know that any subset of a separable metric space is also separable, and thus precompact. \square

Now, we investigate whether the union, intersection, or cartesian product of precompact metric spaces are precompact.

3.31 Proposition. *Let X, Y be precompact metric spaces. Then the following are true:*

- (a) $X \cup Y$ is precompact,
- (b) $X \cap Y$ is precompact,
- (c) $X \times Y$ endowed with the product metric is precompact.

Proof.

(a):

Let \mathcal{U} be an open cover for $X \cup Y$. We can find countable subcovers $\mathcal{U}_1, \mathcal{U}_2$ for X and Y respectively, then note that $\mathcal{U}_1 \cup \mathcal{U}_2$ is a countable subcover of $X \cup Y$.

(b):

$X \cap Y$ is a subset of X , so it is also precompact.

(c):

X, Y being precompact implies that they are separable. Take countable dense subsets $C \subseteq X, D \subseteq Y$. We will show that the countable set $C \times D$ is dense in $X \times Y$.

Fix $(x, y) \in X \times Y$, and let $\varepsilon > 0$. We can find $c \in C$, $d \in D$ such that $d_X(x, c) < \frac{\varepsilon}{2}$ and $d_Y(y, d) < \frac{\varepsilon}{2}$ so therefore

$$d((x, y), (c, d)) = d_X(x, c) + d_Y(y, d) < \varepsilon.$$

as needed. \square

For this section, we focus mainly on \mathbb{R}^n rather than arbitrary precompact metric spaces.

3.32 Proposition. *Every set with a compact exhaustion is precompact.*

Proof. Let A be a set with a compact exhaustion K_n . Let $\{U_i\}_{i \in I}$ be an open cover of A . Notice that this is also an open cover for each K_n , which induces a finite subcover \mathcal{U}_n for each compact set. We take the countable union of these sets and claim that is the desired countable subcover.

Let $x \in A$. Then x is contained in some K_n . But this implies that x is in some open set U_i such that $U_i \in \mathcal{U}_n$, so this choice is indeed a countable subcover. \square

3.33 Theorem. *Let $M \subseteq \mathbb{R}^N$ be a manifold. Then M has a countable atlas.*

Proof. Let $\{(U_i, \varphi_i)\}_{i \in I}$ be an atlas for M . Since $\{U_i\}$ constitutes an open cover of M , we can find a countable subcover $\{U_{i_n}\}_{n \in \mathbb{N}}$ by the precompactness of M . Then $\{(U_{i_n}, \varphi_{i_n})\}_{n \in \mathbb{N}}$ is still an atlas over M , but is countable, so we are done. \square

This result is useful for when we want to apply countable additivity on manifold charts. For instance, this theorem is used to show that a lower dimensional manifold is a Lebesgue null set by showing that we can find a countable atlas where each chart has Jordan measure 0, and then applying countable additivity of the Jordan measure.

3.3 Generalizing the Inverse Function Theorem

Below is the statement for the Inverse Function Theorem:

3.34 Theorem. *Let $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a C^1 function. If $p \in U$ such that $Jf(p)$ is invertible, then*

- *There exists an open set $U_0 \subseteq U$ such that $\Phi = f|_{U_0}$ is injective.*
- *$V_0 = f(U_0)$ is open.*
- *$\Phi^{-1}(q) = (\Phi' \circ \Phi^{-1}(q))^{-1}$*
- *Φ^{-1} is C^1 .*

As well, here is the statement of the Implicit Function Theorem:

3.35 Theorem. *Let $f : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$ be a C^1 function. Let $(a, b) \in \mathbb{R}^{n+k}$, where $a \in \mathbb{R}^n, b \in \mathbb{R}^k$ such that $f(a, b) = 0$. Write*

$$Jf(a, b) = (A \mid B)$$

If B is invertible, then there is an open neighborhood $U_0 \subseteq \mathbb{R}^k$ containing b and a C^1 function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ such that for all $x \in U_0$,

$$f(x, \Phi(x)) = 0 \text{ and } \Phi(a) = b.$$

It seems that these two theorems are a special case for some more general theorem. In fact, these statements are actually equivalent. We have shown that the Implicit Function Theorem follows from the Inverse Function Theorem. We will show that the converse is also true.

Proof. Suppose that the Implicit Function Theorem has been proven. We will show that the Inverse Function Theorem is also true. Let $f : U \rightarrow \mathbb{R}^n$ be a C^1 function. Let $p \in U$ such that $Jf(p)$ is invertible. Let $g : f(U) \times U \rightarrow \mathbb{R}^n$ be defined by

$$g(x, y) = f(y) - x$$

We see that the Jacobian is of the form

$$Jg(x, y) = (I \mid Jf(y))$$

Since $g(f(p), p) = 0$ and $Jf(y)$ is invertible, by the Implicit Function Theorem, we can find an open neighborhood $V_0 \subseteq f(U)$ containing $f(p)$ and a C^1 function $\Psi : V_0 \rightarrow U$ such that for all $x \in V_0$,

$$g(x, \Psi(x)) = 0 \implies f(\Psi(x)) = x$$

We see that f acts as a left inverse for Ψ , so Ψ is injective. We set $U_0 = \Psi(V_0)$ and see that $\Phi = f|_{U_0}$ is the inverse of Ψ , or in other words $\Psi = \Phi^{-1}$ (statement 2 and 4). By the continuity of Φ and the fact that V_0 is open, we have that U_0 is open since $\Phi(U_0) = V_0$ (statement 1). Focussing on the equation above, we can take the derivative of both sides to obtain that for $q \in V_0$,

$$\Phi'(\Psi(q)) \circ \Psi'(q) = I \implies \Psi'(q) = (\Phi' \circ \Psi)^{-1}(q) \quad (\text{statement 3})$$

The fact that $\Phi'(\Psi(q))$ is invertible is swept under the rug here, but the reasoning is that we can shrink U_0 so that it is invertible by using the fact that $\text{GL}(\mathbb{R}^n)$ is open, and reassign $V_0 = \Phi(U_0)$. V_0 stays open because $\Psi(V_0) = U_0$.

Thus we have proven every statement in the Inverse Function Theorem. \square

In order to take a step closer to finding the general case, let us consider other special cases. For instance, let $f : U \subseteq \mathbb{R}^k \rightarrow \mathbb{R}^{n+k}$ be a C^1 function. We write $f = (f_1, f_2)$, where $f_1 : \mathbb{R}^k \rightarrow \mathbb{R}^n$ and $f_2 : \mathbb{R}^k \rightarrow \mathbb{R}^k$ are functions. Then the Jacobian is of the form

$$Jf(p) = \begin{pmatrix} Jf_1(p) \\ Jf_2(p) \end{pmatrix}$$

Suppose that $Jf_2(p_0)$ is invertible, for some $p_0 \in U$. By the Inverse Function Theorem, f_2 has an inverse $f_2^{-1} : V_0 \rightarrow U_0$ in an open neighborhood U_0 of p_0 . Define a function $\Phi : \mathbb{R}^n \times V_0 \rightarrow U_0$ by

$$\Phi(y_1, \dots, y_{n+k}) = f_2^{-1}(y_{n+1}, \dots, y_{n+k}).$$

Notice that for all $x \in U_0$,

$$\Phi(f(x)) = \Phi(f_1(x), f_2(x)) = x$$

So we have the case for when the dimension of domain is the same, less than, and greater than that of the codomain. For the case when the domain is higher dimension, the statement involves composing some function inside f . When the domain is lower dimension, we compose outside of f . We conjecture that the general case will involve composing both sides of f . Additionally, these are cases where $Jf(p)$ has full rank, but what if $\text{rank } Jf(p)$ is less than k ? These questions are answered in the following generalisation of the Inverse Function Theorem:

3.36 Theorem. *Let $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a C^1 function, and suppose that $\text{rank } Jf(p) = k \leq \min\{n, m\}$ for all $p \in U$. Then there exists an open neighborhood $U_0 \subseteq U$ containing p_0 and two diffeomorphisms $\Psi : f(U_0) \rightarrow V_0 \subseteq \mathbb{R}^m$, $\Phi : W_0 \subseteq \mathbb{R}^n \rightarrow U_0$ such that for all $(x_1, \dots, x_n) \in W_0$,*

$$\Psi \circ f \circ \Phi(x_1, \dots, x_n) = (x_1, \dots, x_k, 0, \dots, 0)$$

Proof. Suppose the Jacobian of f at $p_0 = (p_1, p_2)$ is of the form

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right)$$

where $A \in M_{k \times k}(\mathbb{R})$. Suppose without loss of generality that A is invertible. If not, we know that $Jf(p)$ must have k linearly independent rows and columns, so we can apply swaps to both the inner and outer components of f to move them into the top left corner, which does not change the conclusion of the theorem, as we are composing f with diffeomorphisms on both sides.

Write $f = (f_1, f_2)$, where $f_1 : U \rightarrow \mathbb{R}^k$, $f_2 : U \rightarrow \mathbb{R}^{m-k}$ are C^1 functions. Define $g : \mathbb{R}^k \times \mathbb{R}^{n-k} \rightarrow \mathbb{R}^n$ by

$$g(x, y) = (f_1(x, y), y)$$

The Jacobian of g at p_0 is

$$Jg(p_0) = \left(\begin{array}{c|c} A & B \\ \hline O & I_{n-k} \end{array} \right)$$

Since A is invertible, $Jg(p_0)$ is invertible, so by the Inverse Function Theorem, we can find an open neighborhood U_0 containing p_0 such that the diffeomorphism $\Phi = (g|_{U_0})^{-1}$ exists. Let $W_0 = g(U_0)$ and write $\Phi = (\varphi_1, \varphi_2)$. For all $(x, y) \in W_0$, we have

$$(x, y) = g \circ \Phi(x, y) = g(\varphi_1(x, y), \varphi_2(x, y)) = (f_1(\varphi_1(x, y), \varphi_2(x, y)), \varphi_2(x, y))$$

so

$$y = \varphi_2(x, y) \text{ and } x = f_1(\varphi_1(x, y), y)$$

We see that $f \circ \Phi(x, y) = (x, f_2(\varphi_1(x, y), y))$. Let $h(x, y) = f_2(\varphi_1(x, y), y)$ and notice that

$$J(f \circ \Phi)(x, y) = \left(\begin{array}{c|c} I_k & O \\ \hline H_1(x, y) & H_2(x, y) \end{array} \right)$$

for some matrices H_1, H_2 . But $J(f \circ \Phi)$ has rank k , so it must be true that $H_2 = 0$. It follows that for all $(x, y) \in W_0$, $h(x, y) = h(x, p_2)$. Let $\Psi : f(U_0) \rightarrow \mathbb{R}^m$ be defined as

$$\Psi(z, w) = (z, w - h(z, p_2)), \text{ where } z \in \mathbb{R}^k, w \in \mathbb{R}^{m-k}.$$

Its Jacobian is

$$\left(\begin{array}{c|c} I_k & O \\ \hline -H_1(z, p_2) & I_{m-k} \end{array} \right).$$

which has rank m , implying that Ψ is locally a diffeomorphism onto its image.

Finally, we let $(x, y) \in W_0$ for $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^{n-k}$ and see that

$$\begin{aligned} \Psi \circ f \circ \Phi(x, y) &= \Psi(x, h(x, y)) \\ &= (x, h(x, y) - h(x, p_2)) \\ &= (x, 0) \end{aligned}$$

□

It is slightly important to note that one of the assumptions in this theorem is that Jf is of constant rank, rather than just being invertible at a single point. This is because the invertibility of a matrix is not affected under small perturbations of its entries, while in general, the rank is volatile to these changes. For example, the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has rank 1, but almost any change in its entries causes the rank to increase.

The usage of local diffeomorphisms hint at applications involving manifolds, but as of now, the usage escapes my knowledge. This is still pretty cool though, as the theorem is related to the result in linear algebra that states that any matrix with rank r can be row and column reduced to the identity matrix in the top left, and if we take the Jacobian of the composition $\Psi \circ f \circ \Phi$, we get

$$J(\Psi \circ f \circ \Phi)(x) = \left(\begin{array}{c|c} I_k & O \\ \hline O & O \end{array} \right)$$

so the intuition behind this generalisation of the Inverse Function Theorem is that locally, f looks like a linear transformation with rank k , so we can reduce it to a nice looking matrix via row and column operations.

4 Major Theorems

4.1 Sequential Compactness and Compactness

4.1 Definition. A metric space X is compact if every open cover permits a finite subcover.

4.2 Definition. A metric space X is sequentially compact if every sequence has a convergent subsequence.

4.3 Theorem. Let X be a metric space. Then X is compact if and only if X is sequentially compact.

Proof. Suppose for contradiction that X is compact but not sequentially compact. Let (a_n) be a sequence with no convergent subsequence. For each $x \in X$, there exists $r_x > 0$ such that $B(x, r_x)$ contains finitely many elements of (a_n) for if not, then x is a cluster point which is a contradiction. Consider the open cover $\{B(x, r_x)\}_{x \in X}$. We can find a finite subcover by compactness. Since each ball should contain a finite number of elements in (a_n) , that implies that the sequence reaches a finite number of values, meaning that we can find a convergent subsequence, which is a contradiction. Thus X is sequentially compact.

Conversely, suppose that X is sequentially compact, and suppose again seeking contradiction that X is not compact. Let $\square = \{U_i\}_{i \in I}$ be an open cover that does not have a finite subcover. Recall the Lebesgue number lemma: for any open cover of a sequentially compact metric space, there exists $\delta > 0$ such that any subset $A \subseteq X$ with $\text{diam} A < \delta$ is a subset of some U_i in the open cover. We construct a sequence as follows. Let $x_0 \in X$ be an arbitrary choice. Consider $B(x_0, \delta)$, which is contained in some $U_0 \in \mathcal{U}$. By assumption, U_0 does not cover X so we pick $x_1 \in X \setminus U_0$. Similarly, $B(x_1, \delta)$ is contained in some U_1 , but $U_0 \cup U_1$ does not cover X , so we can find $x_3 \in X \setminus (U_0 \cup U_1)$. We continue picking x_n in this fashion. Notice by construction, each ball $B(x_n, \delta)$ necessarily contains exactly one element in the sequence. But now consider a convergent subsequence of (x_n) , and let x be the cluster point of this subsequence. Then there exists an infinite number of sequence terms in $B(x, \frac{\delta}{2})$, so for any two distinct $x_i, x_j \in B(x, \frac{\delta}{2})$, $x_j \in B(x_i, \delta)$, which is a contradiction.

Thus we have proven that sequential compactness is equivalent to compactness in metric spaces. \square

4.2 Equivalence of Norms in Finite Dimensional Vector Spaces

4.4 Theorem. Let $(V, \|\cdot\|)$ be a normed vector space with dimension n . There exists a linear homeomorphism between $(V, \|\cdot\|)$ and $(\mathbb{R}^n, \|\cdot\|_2)$, where $\|\cdot\|_2$ represents the standard Euclidean norm.

Proof. We prove that all norms are equivalent in finite dimensional vector spaces, from which the theorem will follow quickly. First, note that a vector space over \mathbb{R} with dimension n is isomorphic to \mathbb{R}^n , so there exists a linear isomorphism $\Phi : V \rightarrow \mathbb{R}^n$. Let $\|\cdot\|$ be a norm on V . We define the norm $\|\cdot\|_2$ to be $\|x\|_2 = \|\Phi(x)\|_E$, where the norm on the right represents the Euclidean norm in \mathbb{R}^n . It is easy to show that this is indeed a norm. We claim that $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent. To do this, we need to show that there exist $m, M > 0$ such that for all $x \in V$,

$$m\|x\| \leq \|x\|_2 \leq M\|x\|.$$

Recall that the unit ball (with respect to $\|\cdot\|_2$) is compact in all finite dimensional vector spaces. It follows that the closed set $C = \{x \in V : \|x\|_2 = 1\} \subseteq B_2(x, 1)$ is compact. Define a continuous function $f : C \rightarrow \mathbb{R}$ by

$$f(x) = \frac{\|x\|_2}{\|x\|}.$$

By the Extreme Value Theorem, f attains a maximum M and minimum m . That is, for all $x \in C$,

$$m \leq \frac{\|x\|_2}{\|x\|} \leq M \implies m\|x\| \leq \|x\|_2 \leq M\|x\|.$$

We claim that these values of M, m are the ones we are looking for. Let $x \in V$. If $x = 0$, the inequality trivially holds true. Otherwise, we can take $\frac{x}{\|x\|_2} \in C$, from which it follows that

$$m \frac{\|x\|}{\|x\|_2} \leq 1 \leq M \frac{\|x\|}{\|x\|_2} \implies m\|x\| \leq \|x\|_2 \leq M\|x\|.$$

as desired.

Now we prove the main result. We claim that the linear isomorphism defined above is the homeomorphism we are looking for. We show that Φ is continuous. Since Φ is linear, it suffices to show that this is the case at $x = 0$. Let $\varepsilon > 0$. Take $\delta = \frac{\varepsilon}{M}$. Then for all x with $\|x\| < \delta$,

$$\|\Phi(x)\|_E = \|x\|_2 \leq M\|x\| < \varepsilon$$

We apply the same argument to Φ^{-1} to conclude that Φ is indeed a homeomorphism and we are done. \square

4.3 Differentiability Theorem

4.5 Theorem. *Let $f : U \rightarrow \mathbb{R}^m$ be a function, where U is open in \mathbb{R}^n . Then f is continuously differentiable if and only if f is partially continuously differentiable.*

Proof. First, let us consider the case when $m = 1$. As well, for convenience, we equip our vector spaces with the 1-norm. Suppose that f is continuously differentiable. We want to show that the partial derivative with respect to each x_i is continuous. Let $p \in U$. For each i , there exists $\delta > 0$ such that for $t \in (-\delta, \delta)$, $f(p + te_i) \in U$. Define $f_i : (-\delta, \delta) \rightarrow \mathbb{R}$ by $f_i(t) = f(p + te_i)$. Using the chain rule, we get that

$$f'_i(t) = f'(p + te_i)(e_i).$$

but notice that

$$\frac{\partial f}{\partial x_i}(p) = \lim_{t \rightarrow 0} \frac{1}{t}(f(p + te_i) - f(p)) = \lim_{t \rightarrow 0} \frac{1}{t}(f_i(t) - f_i(0)) = f'_i(0) = f'(p)(e_i).$$

Since $f'(\cdot)(e_i)$ is continuous, it follows that the partial derivatives are also continuous.

Conversely, suppose that every partial derivative is continuous. We claim that

$$f'(p)(h) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(p)$$

To do this, we want to show

$$\lim_{h \rightarrow 0} \frac{1}{\|h\|} \left(f(p + h) - f(p) - \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(p) \right) = 0.$$

For $p \in U, h \in \mathbb{R}^n$, let p_k be defined by

$$p_k = p + \sum_{i=1}^k h_i e_i$$

where k is an integer that ranges from 0 to n . Using the Mean Value Theorem, we can say that there exists $t_i \in (0, h_i)$ such that

$$h_i \frac{\partial f}{\partial x_i}(p_{i-1} + t_i e_i) = f(p_i) - f(p_{i-1}).$$

Write $q_i = p_{i-1} + t_i e_i$.

Now let $\varepsilon > 0$. Then choose δ so that $B(p, \delta) \subseteq U$ and $\left| \frac{\partial f}{\partial x_i}(p + h) - \frac{\partial f}{\partial x_i}(p) \right| < \varepsilon$ for all i with $\|h\| < \delta$. Then for $h \in \mathbb{R}^n$ with $\|h\| < \delta$,

$$\begin{aligned} \left| f(p + h) - f(p) - \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(p) \right| &= \left| \sum_{i=1}^n \left(f(p_i) - f(p_{i-1}) - h_i \frac{\partial f}{\partial x_i}(p) \right) \right| \\ &\leq \sum_{i=1}^n |h_i| \left| \frac{\partial f}{\partial x_i}(q_i) - \frac{\partial f}{\partial x_i}(p) \right| \end{aligned}$$

Notice that $\|q_i - p\| < \|h\| < \delta$, so

$$\begin{aligned} \sum_{i=1}^n |h_i| \left| \frac{\partial f}{\partial x_i}(q_i) - \frac{\partial f}{\partial x_i}(p) \right| &< \sum_{i=1}^n |h_i| \varepsilon \\ &= \varepsilon \|h\|. \end{aligned}$$

Thus showing that our limit indeed converges to 0, so our derivative is valid. It remains to show that it is continuous. Let $\varepsilon > 0$. We use the same δ as above. Let $v \in \mathbb{R}^n$ with $\|v\| < \delta$. Notice that for all $h \in \mathbb{R}^n$ with $\|h\| = 1$,

$$|f'(p+v)(h) - f'(p)(h)| = \left| \sum_{i=1}^n h_i \left(\frac{\partial f}{\partial x_i}(p+v) - \frac{\partial f}{\partial x_i}(p) \right) \right| < \|h\| \varepsilon = \varepsilon$$

which means that

$$\|f'(p+v) - f'(p)\|_{\text{op}} < \varepsilon$$

as desired. \square

4.4 Chain Rule

Let X, Y, Z be normed vector spaces, and $g : X \rightarrow Y$, $f : Y \rightarrow Z$ be functions. For some $p \in X$, suppose that g is totally differentiable at p and f is totally differentiable at $g(p)$. Then the function $f \circ g$ is totally differentiable at p , and $(f \circ g)' = f'(g(p)) \circ g'(p)$.

Proof. Let $L_f = f'(g(p))$ and $L_g = g'(p)$. First, let's prove that $L_f \circ L_g$ is a linear bounded operator. For $a, b \in X$, $c \in \mathbb{R}$,

$$L_f(L_g(ca + b)) = L_f(cL_g(a) + L_g(b)) = cL_f(L_g(a)) + L_f(L_g(b))$$

Thus $L_f \circ L_g$ is linear. Now, since L_f, L_g are bounded linear operators, there exist constants $M_f, M_g > 0$ such that for all $x \in X$, $y \in Y$,

$$\|L_f(y)\|_Z \leq M_f \|y\|_Y \text{ and } \|L_g(x)\|_Y \leq M_g \|x\|_X$$

Then we have that for all $x \in X$,

$$\|L_f(L_g(x))\|_Z \leq M_f \|L_g(x)\|_Y \leq M_f \cdot M_g \|x\|_X$$

so it is bounded as well. Now we can move on to the long part of the proof.

Let $\varepsilon > 0$. By our assumption, there exists positive numbers δ_f, δ_g such that for h_f, h_g with $\|h_f\|_X < \delta_f$, $\|h_g\|_X < \delta_g$,

$$\|f(g(p) + h_f) - f(g(p)) - L_f(h_f)\|_Z \leq \min \left\{ \sqrt{\varepsilon}, \frac{\varepsilon}{4M_g} \right\} \|h_f\|_X \quad (1)$$

and

$$\|g(p + h_g) - g(p) - L_g(h_g)\|_Y \leq \min \left\{ \frac{\sqrt{\varepsilon}}{4}, \frac{\varepsilon}{2M_f} \right\} \|h_g\|_X \quad (2)$$

As well, since g being totally differentiable at p implies continuity at p , then for all $\|h_c\|_X < \delta_c$,

$$\|g(p + h_c) - g(p)\| < \delta_f \quad (3)$$

Define $\varepsilon_f = \min \left\{ \sqrt{\varepsilon}, \frac{\varepsilon}{4M_g} \right\}$ and $\varepsilon_g = \min \left\{ \frac{\sqrt{\varepsilon}}{4}, \frac{\varepsilon}{2M_f} \right\}$, and let $\delta = \min\{\delta_g, \delta_c, r\}$. For $h \in X$ with $\|h\|_X < \delta$,

$$\begin{aligned} &\frac{\|f(g(p+h)) - f(g(p)) - L_f(L_g(h))\|_Z}{\|h\|_X} \\ &\leq \frac{\|f(g(p+h)) - f(g(p)) - L_f(g(p+h) - g(p))\|_Z + \|L_f(g(p+h) - g(p)) - L_f(L_g(h))\|_Z}{\|h\|_X} \end{aligned}$$

Consider the first term. Because of (3), we can apply (1) with $h_f = g(p+h) - g(p)$ and get that

$$\frac{\|f(g(p+h)) - f(g(p)) - L_f(g(p+h) - g(p))\|_Z}{\|h\|_X} < \frac{\|g(p+h) - g(p)\|}{\|h\|_X} \varepsilon_f$$

Then applying (2) with $h_g = h$,

$$\begin{aligned} &\leq \frac{\|g(p+h) - g(p) - L_g(h)\|_Y + \|L_g(h)\|_Y}{\|h\|_X} \varepsilon_f \leq \frac{\sqrt{\varepsilon}}{4} \varepsilon_f + M_g \varepsilon_f \leq \frac{\sqrt{\varepsilon}}{4} \sqrt{\varepsilon} + M_g \frac{\varepsilon}{4M_g} \\ &= \frac{\varepsilon}{2} \end{aligned}$$

Now, we examine the second term. Using the fact that L_f is a bounded linear operator and (2),

$$\frac{\|L_f(g(p+h) - g(p) - L_g(h))\|_Z}{\|h\|_X} \leq \frac{M_f \|g(p+h) - g(p) - L_g(h)\|_Y}{\|h\|_X} < \frac{M_f \varepsilon}{2M_f} = \frac{\varepsilon}{2}$$

Adding the two together we conclude that

$$\begin{aligned} &\frac{\|f(g(p+h)) - f(g(p)) - L_f(g(p+h) - g(p))\|_Z + \|L_f(g(p+h) - g(p) - L_g(h))\|_Z}{\|h\|_X} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

and we are done. □

4.5 Inverse Function Theorem and Implicit Function Theorem

4.6 Theorem. Inverse Function Theorem. *Let $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a C^1 function. If $p \in U$ such that $Jf(p)$ is invertible, then*

- *There exists an open set $U_0 \subseteq U$ such that $\Phi = f|_{U_0}$ is injective.*
- *$V_0 = f(U_0)$ is open.*
- *$\Phi^{-1}(q) = (\Phi' \circ \Phi^{-1}(q))^{-1}$*
- *Φ^{-1} is C^1 .*

Proof. Let $L = f'(p)$. For $y \in \mathbb{R}^n$, define a function $\alpha_y(q) = q + L^{-1}(y - f(q))$, where $q \in U$. We claim that q is a fixed point of α_y if and only if $f(q) = y$. Suppose that q is a fixed point. Then

$$q = \alpha_y(q) = q + L^{-1}(y - f(q)) \implies L^{-1}(y) = L^{-1}(f(q))$$

Since L^{-1} is invertible, it must be that $y = f(q)$.

Conversely, suppose that $y = f(q)$. Then

$$\alpha_y(q) = q + L^{-1}(0) = q$$

so q is a fixed point.

Next, we see by inspection that the derivative of α_y is given by

$$\alpha'_y(q) = I - L^{-1} \circ f'(q).$$

Then

$$\|\alpha'_y(q)\|_{\text{op}} = \|I - L^{-1} \circ f'(q)\| \leq \|L^{-1}\|_{\text{op}} \|L - f'(q)\|_{\text{op}}$$

Recall that $L = f'(p)$. Since f' is continuous, by letting $\varepsilon = \frac{1}{2}\|L^{-1}\|_{\text{op}}^{-1}$, for all q in some open ball U_0 of p ,

$$\|\alpha'_y(q)\|_{\text{op}} \leq \frac{1}{2}$$

Here, we also require U_0 to be chosen such that $f'(q)$ is invertible, which will be useful later.

From this, we can say that

$$\|\alpha_y(a) - \alpha_y(b)\| \leq \frac{1}{2}\|a - b\|, \text{ where } a, b \in U_0$$

Now, we have the tools to show that f is injective on U_0 . Let $a, b \in U_0$ such that $f(a) = f(b)$. Let $y = f(a) = f(b)$. Then we immediately have that a and b are fixed points of α_y . Then

$$\|\alpha_y(a) - \alpha_y(b)\| \leq \frac{1}{2}\|a - b\| \implies \|a - b\| \leq \frac{1}{2}\|a - b\|$$

which is only possible if $a = b$. Thus $f(p)$ is locally injective, and we will call this restriction $\Phi : U_0 \rightarrow \mathbb{R}$. That concludes the proof of the first statement.

Next, to show that $V_0 = \Phi(U_0)$ is open, first select any $v \in V_0$ and write $v = \Phi(q)$, for some $q \in U_0$. Since U_0 is open, we have $C = \overline{B}(q, r) \subseteq U_0$ for some $r > 0$. Let $s = \frac{1}{2}\|L^{-1}\|_{\text{op}}^{-1}$. We claim that $B(v, s) \subseteq V_0$. Let $w \in B(v, s)$. Consider α_y when $y = w$. We claim that $\alpha_y(C) \subseteq C$. Let $\alpha_y(u) \in \alpha_y(C)$. Then

$$\begin{aligned} \|q - \alpha_y(u)\| &\leq \|q - \alpha_y(q)\| + \|\alpha_y(q) - \alpha_y(u)\| \\ &\leq \|L^{-1}(w - \Phi(q))\| + \frac{1}{2}\|q - u\| \\ &\leq \|L^{-1}\|_{\text{op}}\|w - v\| + \frac{1}{2}\|q - u\| \\ &< \frac{1}{2}r + \frac{1}{2}r \\ &= r \end{aligned}$$

so $\alpha_y(u) \in C$. Since α_y is a contraction map, it must be that it has a unique fixed point $z \in C$. But z is a fixed point only if $\Phi(z) = w$, so $w \in V_0$ which satisfies the second condition.

Let $x \in V_0$. Since V_0 is open, for small enough $k \in \mathbb{R}^n$, $x + k \in V_0$, so $x + k = \Phi(u)$ for some $u \in U_0$. We write $u = \Phi^{-1}(x) + h$, where $h = u - \Phi^{-1}(x)$. Moreover, by setting $y = x$ and once again considering α_y , we have

$$0 = k + x - \Phi(\Phi^{-1}(x) + h) \implies h = L^{-1}(k) + \alpha_y(\Phi^{-1}(x) + h) - \Phi^{-1}(x)$$

Note that $\Phi^{-1}(x)$ is a fixed point of α_y , so

$$\begin{aligned} h &= L^{-1}(k) + \alpha_y(\Phi^{-1}(x) + h) - \alpha_y(\Phi^{-1}(x)) \\ \implies \|h\| &\leq \|L^{-1}\|_{\text{op}}\|k\| + \frac{1}{2}\|h\| \implies \|h\| \leq \frac{\|k\|}{s} \end{aligned}$$

Now, we show that $(\Phi^{-1})'(x) = (\Phi' \circ \Phi^{-1})^{-1}(x)$. Let $\varepsilon > 0$. Let $k \in \mathbb{R}^n$ small enough so that $x + k \in V_0$ and $x + k = \Phi(\Phi^{-1}(x) + h)$, where $\|h\| \leq \frac{\|k\|}{s}$. Then

$$\begin{aligned} \|\Phi^{-1}(x + k) - \Phi^{-1}(x) - (\Phi' \circ \Phi^{-1})^{-1}(x)(k)\| &= \|\Phi^{-1}(\Phi(\Phi^{-1}(x) + h)) - \Phi^{-1}(x) - (\Phi' \circ \Phi^{-1})^{-1}(x)(k)\| \\ &= \|h - (\Phi' \circ \Phi^{-1})^{-1}(x)(\Phi(\Phi^{-1}(x) + h) - x)\| \\ &\leq \|(\Phi' \circ \Phi^{-1})^{-1}(x)\|_{\text{op}}\|\Phi(\Phi^{-1}(x) + h) - \Phi(\Phi^{-1}(x)) - (\Phi' \circ \Phi^{-1})^{-1}(x)(h)\| \end{aligned}$$

Since Φ is differentiable, for $\|h\| < \delta$, for some δ ,

$$\|\Phi(\Phi^{-1}(x) + h) - \Phi(\Phi^{-1}(x)) - (\Phi' \circ \Phi^{-1})^{-1}(x)(h)\| < \frac{\|h\|s\varepsilon}{\|(\Phi' \circ \Phi^{-1})^{-1}(x)\|_{\text{op}}}$$

Thus if we take $\|k\| < s\delta$, $\|h\| < \delta$, so

$$\begin{aligned}\|\Phi^{-1}(x+k) - \Phi^{-1}(x) - (\Phi' \circ \Phi^{-1})^{-1}(x)(k)\| &< \|(\Phi' \circ \Phi^{-1})^{-1}(x)\|_{\text{op}} \cdot \frac{\|h\|s\varepsilon}{\|(\Phi' \circ \Phi^{-1})^{-1}(x)\|_{\text{op}}} \\ &= \|h\|s\varepsilon \\ &\leq \|k\|\varepsilon\end{aligned}$$

So Φ^{-1} is indeed differentiable. To show that $(\Phi^{-1})'$ is continuous, we note that it is the composition of Φ^{-1} , Φ' , and the inversion map $A \mapsto A^{-1}$, all of which are continuous. To see that the inversion map is continuous, refer to the proof that the inversion map is differentiable in Big List #39, from which continuity follows immediately.

We have proven all four statements so the proof is done. \square

4.7 Theorem. Implicit Function Theorem. Let $f : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$ be a C^1 function. Let $(a, b) \in \mathbb{R}^{n+k}$, where $a \in \mathbb{R}^n$, $b \in \mathbb{R}^k$ such that $f(a, b) = 0$. Write

$$Jf(a, b) = (A \mid B)$$

If B is invertible, then there is an open neighborhood $U_0 \subseteq \mathbb{R}^k$ containing b and a C^1 function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ such that for all $x \in U_0$,

$$f(x, \Phi(x)) = 0 \text{ and } \Phi(a) = b.$$

Proof. Define $G : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^{n+k}$ by

$$G(x, y) = (x, f(x, y)).$$

Notice that

$$JG(a, b) = \left(\begin{array}{c|c} I & O \\ \hline A & B \end{array} \right)$$

Since I, B are invertible, $JG(a, b)$ is invertible, so by the Inverse Function Theorem, we have that G^{-1} exists in an open neighborhood V_0 around (a, b) . Write $G^{-1} = (G_1^{-1}, G_2^{-1})$, where $G_1^{-1} : V_0 \rightarrow \mathbb{R}^n$, $G_2^{-1} : V_0 \rightarrow \mathbb{R}^k$ are C^1 functions. Let U_0 be defined by

$$U_0 = \{x \in \pi_{\mathbb{R}^n}(V_0), (x, 0) \in V_0\} \subseteq V_0$$

where $\pi_{\mathbb{R}^n}$ is the projection of \mathbb{R}^{n+k} onto the first n components. We want to show that U_0 is non-empty and open. Note that $G(a, b) = (a, 0) \in U_0$, so U_0 is non-empty, and additionally contains a . Let $x \in U_0$. Since V_0 is open, we can find $B((x, 0), r) \subseteq V_0$. If we take the same radius, we have for $v \in B(x, r)$,

$$\|x - v\| < r \implies \|(x, 0) - (v, 0)\| < r \implies (v, 0) \in V_0 \text{ and } v \in U_0.$$

We claim that our desired function Φ is defined by

$$\Phi = G_2^{-1}(\cdot, 0).$$

We already have that Φ is C^1 . As well, for $x \in U_0$, we have

$$(x, 0) = G(G^{-1}(x, 0)) = G(G_1^{-1}(x, 0), G_2^{-1}(x, 0)) = (G_1^{-1}(x, 0), f(G_1^{-1}(x, 0), G_2^{-1}(x, 0)))$$

so $x = G_1^{-1}(x, 0)$ and $f(G_1^{-1}(x, 0), G_2^{-1}(x, 0)) = f(x, \Phi(x)) = 0$ as needed. Finally, notice that $G(a, b) = (a, 0)$, and G is injective on V_0 , so it must be true that $\Phi(a) = b$ and we are done. \square

4.6 Condition for Local Minimum

4.8 Theorem. Let $U \subseteq \mathbb{R}^n$ be an open set. Let $f : U \rightarrow \mathbb{R}$ be a C^2 function. Let $p_0 \in U$ be a point such that $\nabla f(p_0) = 0$. If the Hessian matrix $Hf(p_0)$ is positive definite, then f achieves a local minimum at p_0 .

Proof. First, we show that for a quadratic form $Q(x) = x^t A x$, $Q(x) > 0$ for all $x \neq 0$ if and only if all the eigenvalues of A are positive.

Suppose that $Q(\vec{x}) > 0$ for all $\vec{x} \neq 0$. Since A is symmetric, it is orthogonally diagonalizable, so there exists an orthogonal matrix $P \in M_n(\mathbb{R})$ such that $B = P^T A P$ is diagonal. Letting $Q_B(\vec{x}) = \vec{x}^T B \vec{x}$, we have that

$$Q(\vec{x}) = Q_B \circ P(\vec{x})$$

We will show that $Q_B(\vec{x}) > 0$ for all $\vec{x} \neq 0$.

Let $\vec{x} \in \mathbb{R}^n$ so that $\vec{x} \neq 0$. Since $Q(\vec{x}) > 0$, it follows that $Q_B(\vec{x}) > 0$ as well, which implies that its eigenvalues are strictly positive. Since A and B share the same eigenvalues, we are done.

Conversely, suppose that all the eigenvalues of A are strictly positive. By a similar argument as before, the eigenvalues of B are positive, so $Q_B(\vec{x}) > 0$ for all $\vec{x} \neq 0$. The same applies to Q and we are done.

Next, we will show that there exists a constant $\eta > 0$ so that $Q(x) \geq \eta \|x\|^2$ for all $x \in \mathbb{R}^n$.

First, equip \mathbb{R}^n with the 2-norm.

Let Q be a positive definite quadratic form, so $Q(\vec{x}) = \vec{x}^T A \vec{x}$, for some diagonalizable matrix A , where all its eigenvalues λ_i are strictly positive.

Let $\eta = \min_{1 \leq i \leq n} \{\lambda_i\}$. Let B be the diagonal matrix such that $A = P^T B P$ for an orthogonal matrix P . Denote $P(\vec{x}) = (p_1, \dots, p_n)$. We have

$$Q(\vec{x}) = Q_B \circ P(\vec{x}) = (P(\vec{x}))^T B (P(\vec{x})) = \sum_{i=1}^n \lambda_i \cdot p_i^2 \geq \eta \sum_{i=1}^n p_i^2 = \eta \|P(\vec{x})\|_2^2$$

Since P preserves distance, we have that

$$Q(\vec{x}) \geq \eta \|P(\vec{x})\|_2^2 = \eta \|\vec{x}\|_2^2$$

as desired.

Now, we prove the main result at hand.

Let $U \subseteq \mathbb{R}^n$ be an open set, let $f : U \rightarrow \mathbb{R}$ be a C^2 function. Let $p_0 \in U$ be a point such that $\nabla f(p_0) = 0$. Suppose that $Hf(p_0)$ is positive definite. We will show that $Q(x) = \frac{1}{2} x^T Hf(p_0) x$ is a quadratic approximation to f at p_0 . That is,

$$\lim_{h \rightarrow 0} \frac{f(p_0 + h) - f(p_0) - Q(h)}{\|h\|^2} = 0$$

Let $\varepsilon > 0$. Since U is open, there is a δ_o so that $p_0 + h \in U$, where $\|h\| < \delta_o$. Additionally, by the continuity of Hf , there exists δ_c so that for $\|h\| < \delta_c$, $|Hf(p_0 + h) - Hf(p_0)| < 2\varepsilon$. Let $\delta = \min\{\delta_o, \delta_c\}$. Fix $h \in \mathbb{R}^n$ such that $0 < \|h\| < \delta$.

Since U is open, there exists $r > 0$ so that $p_0 + th \in U$ for $t \in (-r, 1 + r)$. Define $g(t) = f(p_0 + th)$. Applying Taylor's quadratic remainder formula to g , there exists $\theta \in (0, 1)$ so that

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(\theta) \implies f(p_0 + h) = f(p_0) + f'(p_0)h + \frac{1}{2} f''(p_0 + \theta h)(h)h$$

Since $\nabla f(p_0) = 0$, we have

$$f(p_0 + h) = f(p_0) + \frac{1}{2} h^T Hf(p_0 + \theta h) h$$

Thus we have

$$\begin{aligned} |f(p_0 + h) - f(p_0) - Q(h)| &= \left| \frac{1}{2} h^T Hf(p_0 + \theta h) h - \frac{1}{2} h^T Hf(p_0) h \right| \\ &\leq \frac{1}{2} \|h^T\|_c \|Hf(p_0 + \theta h) - Hf(p_0)\| \|h\| \end{aligned}$$

where $\|\cdot\|_c$ is a norm on row vectors defined by $\|x\|_c = \|x^T\|$. Automatically we have that

$$|f(p_0 + h) - f(p_0) - Q(h)| \leq \frac{1}{2} \|Hf(p_0 + \theta h) - Hf(p_0)\| \|h\|^2 < \varepsilon \|h\|^2$$

$$\implies \frac{f(p_0 + h) - f(p_0) - Q(h)}{\|h\|^2} < \varepsilon$$

as desired. Thus Q is indeed a quadratic approximation for f at p_0 . It remains to show that f attains a local minimum at p_0 . Suppose that this is not the case. In particular, for every open neighborhood around p_0 there is a point q where $f(q) < f(p_0)$. We write $q = p_0 + s$, where $s = q - p_0$. Since $Hf(p_0)$ is positive definite, Q is positive definite. Using the quadratic approximation, there is a $\delta > 0$ so that for all $0 < \|h\| < \delta$,

$$|f(p_0 + h) - f(p_0) - Q(h)| < \eta \|h\|^2,$$

where η comes from part (b) applied to Q . Find a suitable s such that $f(p_0 + s) < f(p_0)$. By part (a), $Q(s) > 0$. This implies that $f(p_0 + s) - f(p_0) - Q(s) < 0$, so

$$f(p_0) - f(p_0 + s) + Q(s) < \eta \|h\|^2 \implies f(p_0) + Q(s) - \eta \|h\|^2 < f(p_0 + s).$$

From part (b), we know that $Q(s) - \eta \|h\|^2 \geq 0$, so

$$f(p_0) < f(p_0 + s)$$

which is a contradiction.

Therefore we have shown that f attains a local minimum at p_0 and we are done. \square

4.7 Change of Variables

4.9 Theorem. Let E be a measurable set in \mathbb{R}^n and let $\Phi : U \rightarrow V$ be a diffeomorphism, where U contains E . Let f be integrable on $\Phi(E)$. Then

$$\int_{\Phi(E)} f = \int_E (f \circ \Phi) \cdot |\det J\Phi|$$

Proof. We begin with a series of reductions. First, it suffices to prove the case when E is a box, for if not, then take a box B that contains E . Then

$$\int_E (f \circ \Phi) \cdot |\det J\Phi| = \int_B \chi_E \cdot (f \circ \Phi) \cdot |\det J\Phi| = \int_{\Phi(B)} \chi_{\Phi(E)} \cdot f = \int_{\Phi(E)} f.$$

Secondly, we can assume without loss of generality that $f = 1$. For an integrable function g , if we take an arbitrary partition P of $\Phi(E)$, then the upper sum

$$\begin{aligned} U(g, P) &= \sum_{i=1}^k \sup\{g(x) : x \in P_i\} \cdot \text{vol}(P_i) = \sum_{i=1}^k \sup\{g(x) : x \in P_i\} \int_{P_i} 1 = \sum_{i=1}^k \sup\{g(x) : x \in P_i\} \int_{\Phi^{-1}(P_i)} |\det J\Phi| \\ &= \sum_{i=1}^k \sup\{g \circ \Phi(y) : y \in \Phi^{-1}(P_i)\} \int_{\Phi^{-1}(P_i)} |\det J\Phi| \geq \sum_{i=1}^k \int_{\Phi^{-1}(P_i)} (g \circ \Phi) \cdot |\det J\Phi| = \int_E (g \circ \Phi) \cdot |\det J\Phi| \end{aligned}$$

A similar argument will show that the same applies for the lower sum so

$$L(g, P) \leq \int_E (g \circ \Phi) \cdot |\det J\Phi| \leq U(g, P) \implies L(g) \leq \int_E (g \circ \Phi) \cdot |\det J\Phi| \leq U(g)$$

which means that $\int_{\Phi(E)} g = \int_E (g \circ \Phi) \cdot |\det J\Phi|$.

With the reductions out of the way, we are going to show that diffeomorphisms can be locally represented as a composition of one linear mapping and multiple simple mappings. However, we first show that this is the case for when $\Phi(0) = 0$. To state the proposition more clearly, for any $p \in U$, we get a neighborhood U_0 such that

$$\Phi(q) = L \circ S_1 \circ \cdots \circ S_n(q)$$

where $q \in U_0$, L is linear, and S_i are simple mappings that are non-singular, and $S_i(0) = 0$.

We proceed with induction on the number of consecutive simple entries of Φ , starting from the first component. If all but one entry is simple, the result is immediate, as Φ is already simple. Now suppose that we can decompose such Φ if its first k entries were simple. Take a Φ with its first $k-1$ entries simple. Then

$$\Phi(x_1, \dots, x_n) = (x_1, \dots, x_{k-1}, \varphi_k(x), \dots, \varphi_n(x)).$$

We have that

$$J\Phi(0) = \left(\begin{array}{c|c} I_{k-1} & O \\ \hline A & B \end{array} \right)$$

where A, B are matrices. Since $J\Phi(0)$ is invertible, for some $i, k \leq i \leq n$, $\frac{\partial \varphi_i}{\partial x_k}(0) \neq 0$. Define

$$\Psi(x_1, \dots, x_n) = (x_1, \dots, x_{k-1}, \varphi_i, x_{k+1}, \dots, x_n).$$

Notice that

$$J\Psi(0) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & \dots & \vdots \\ \vdots & \vdots & \ddots & \dots & \dots & \vdots \\ \dots & \dots & \dots & \nabla \varphi_i(0) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & 1 \end{pmatrix}$$

so $J\Psi(0)$ is invertible. By the Inverse Function Theorem, there is an open neighborhood U_0 containing 0 such that $\Psi|_{U_0}$ is invertible. We write

$$\Psi^{-1} = (x_1, \dots, x_{k-1}, \varphi_i^{-1}, x_{k+1}, \dots, x_n), \text{ where } \varphi_i^{-1} : \Psi(U_0) \rightarrow \mathbb{R}.$$

Notice that $\varphi_i \circ \Psi^{-1} = x_k$. Let L_{ik} be the linear transformation that swaps the i and k th component. Then consider the function $\alpha : \Psi(U_0) \rightarrow V$

$$\begin{aligned} \alpha(x) &= L_{ik} \circ \Phi \circ \Psi^{-1}(x) = L_{ik}(x_1, \dots, x_{k-1}, \varphi_k(\Psi^{-1}(x)), \dots, \varphi_{i-1}(\Psi^{-1}(x)), x_k, \varphi_{i+1}(\Psi^{-1}(x)), \dots, \varphi_n(\Psi^{-1}(x))) \\ &= (x_1, \dots, x_{k-1}, x_k, \varphi_{k+1}(\Psi^{-1}(x)), \varphi_{i-1}(\Psi^{-1}(x)), \varphi_k(\Psi^{-1}(x)), \varphi_{i+1}(\Psi^{-1}(x)), \dots, \varphi_n(\Psi^{-1}(x))) \end{aligned}$$

$J\alpha(0)$ is invertible because the Jacobian of its composite functions are invertible. As well, since $\Psi(0) = 0$ and it is locally injective, we have that

$$\alpha(0) = L_{ik} \circ \Phi \circ \Psi^{-1}(0) = 0$$

Moreover, by rearranging we have

$$\Phi(x) = L_{ik} \circ \alpha \circ \Psi(x)$$

for $x \in U_0$. Notice that α is a function with its first k entries simple. By the induction hypothesis, we can shrink U_0 to locally decompose α into several non-singular simple maps β_i with $\beta_i(0) = 0$ and one linear map so that

$$\alpha(x) = T \circ \beta_1 \circ \dots \circ \beta_{n-k} \text{ for all } x \in U_0$$

Thus

$$\Phi(x) = L_{ik} \circ T \circ \beta_1 \circ \dots \circ \beta_{n-k} \circ \Psi(x)$$

$L_{ik} \circ T$ is linear and each map β_i, Ψ is simple, non-singular, and vanishes at 0. Thus we have shown the statement holds for functions with the first $k-1$ components simple.

Through induction, we have proven that the statement holds for any diffeomorphism Φ .

Now, to show that we can decompose any Φ without the restriction that $\Phi(0) = 0$, we apply the statement to $\Phi(x+p) - \Phi(p)$ to get that

$$\Phi(x+p) - \Phi(p) = L \circ S_1 \circ \dots \circ S_n(x).$$

We perform the substitution $u = x + p$ to get

$$\Phi(u) - \Phi(p) = L \circ S_1 \circ \cdots \circ S_n(u - p).$$

Let $\alpha_{p_i}(x) = (x_1, \dots, x_{i-1}, x_i - p_i, \dots, x_n)$. We can write $u - p = \alpha_1 \circ \cdots \circ \alpha_n(u)$. As well, for any function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, we have

$$f(x) - p = \alpha_{p_1} \circ \cdots \circ \alpha_{p_n} \circ f(x),$$

Recall that L is a series of component swaps, so L^{-1} exists for any $x \in \mathbb{R}^n$. Thus

$$\begin{aligned} \Phi(u) &= L \circ S_1 \circ \cdots \circ S_n(u - p) + \Phi(p) = L(S_1 \circ \cdots \circ S_n(u - p) + L^{-1} \circ \Phi(p)) \\ &= L \circ \alpha_{(L^{-1} \circ \Phi(p))_1} \circ \cdots \circ \alpha_{(L^{-1} \circ \Phi(p))_n} \circ S_1 \circ \cdots \circ S_n \circ \alpha_{p_1} \circ \cdots \circ \alpha_{p_n}(u) \end{aligned}$$

Every function in the expression is simple, save for L . Thus we can still decompose Φ regardless of whether $\Phi(0) = 0$.

Next, we show that if the main theorem holds for two diffeomorphisms $\Phi : U \rightarrow V, \Psi : V \rightarrow W$, it holds for $\Psi \circ \Phi$. This can be shown quickly through a series of algebraic manipulations:

$$\int_{\Psi \circ \Phi(E)} 1 = \int_{\Phi(E)} |\det J\Psi| = \int_E |\det J\Phi \cdot \det J\Psi| = \int_E |\det(J\Psi J\Phi)| = \int_E |\det J(\Psi \circ \Phi)|$$

Now, we show that the statement is true for linear diffeomorphisms, that is, for $\Phi \in L(\mathbb{R}^n, \mathbb{R}^n)$. But recall that for linear maps,

$$\int_{\Phi(E)} 1 = \mu(\Phi(E)) = |\det \Phi| \mu(E) = \int_E |\det \Phi|$$

Next, we prove that the theorem is valid for simple diffeomorphisms Φ . Without loss of generality, let the only non-simple component of Φ be φ_n , and let $E = \prod_{i=1}^n [a_i, b_i] \subseteq U$ be a closed box. Since Φ is a diffeomorphism, this implies that $\Phi(E)$ is measurable, and as well, because Φ is simple,

$$\Phi(E) = \left\{ (x_1, \dots, x_{n-1}, \varphi(x)) \in V : x = (x_1, \dots, x_n) \in \prod_{i=1}^n [a_i, b_i] \right\}$$

Let $E' = \prod_{i=1}^{n-1} [a_i, b_i]$, and define $\Psi : E' \rightarrow \mathbb{R}$ by

$$\Psi(x') = \int_{\varphi(x', [a_n, b_n])} 1 \, dx_n.$$

Φ is a diffeomorphism, so the x' -slice $\varphi(x', \cdot)$ is C^1 and injective. It follows that the image $\varphi(x', [a_n, b_n])$ is an interval. Now, we want to apply single-variable change of variables with $\varphi(x', u) = x_n$. Notice that

$$dx_n = \frac{\partial \varphi}{\partial u}(x', u) du$$

and

$$\det J\Phi = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi}{\partial x_1}(x', u) & \frac{\partial \varphi}{\partial x_2}(x', u) & \cdots & \frac{\partial \varphi}{\partial u}(x', u) \end{vmatrix} = \frac{\partial \varphi}{\partial u}(x', u)$$

so it follows that

$$\Psi(x') = \int_{\varphi(x', [a_n, b_n])} 1 \, dx_n = \int_{a_n}^{b_n} |\det J\Phi|$$

Finally, using Fubini's Theorem, we get that

$$\mu(\Phi(E)) = \int_{\Phi(E)} 1 = \int_{E'} \Psi = \int_{E'} \int_{a_n}^{b_n} |\det J\Phi| = \int_E |\det J\Phi|$$

as needed.

We combine the results above to prove it for an arbitrary diffeomorphism Φ . For each $p_i \in E$, we can decompose Φ into a composition of linear maps and simple mappings, and we know that the result holds for linear maps, simple mappings, and compositions of them, so the result holds for Φ in some neighborhood U_i containing p_i , so

$$\int_{\Phi(U_i)} 1 = \int_{U_i} |\det \Phi|$$

Consider the open cover $\mathcal{U} = \{U_i\}_{i \in I}$ of E . Let $\{\phi_i\}_{i \in I}$ be a partition of unity subordinate to \mathcal{U} . Then we have that

$$\begin{aligned} \int_{\Phi(E)} 1 &= \sum_{i \in I} \phi_i \int_{\Phi(E)} 1 \\ &= \sum_{i \in I} \phi_i \int_{\Phi(U_i)} 1 \\ &= \sum_{i \in I} \phi_i \int_{U_i} |\det J\Phi| \\ &= \sum_{i \in I} \phi_i \int_E |\det J\Phi| \\ &= \int_E |\det J\Phi| \end{aligned}$$

Thus we have proven the Change of Variables formula. □