

Class Notes

Math 395 - Honors Analysis
University of Michigan



Jinyan Miao
and his friends from Math 395

Fall 2021

Contents

In this text, Munkres refers to the text *Analysis on Manifolds*, 1st Edition (1990), by James R. Munkres, published by CRC Press in 2018; Spivak refers to the text *Calculus*, 4th Edition (2008), by Michael Spivak, published by Publish-or-Perish Press in 2008.

The class number for 2020 Math 295/296 and 2021 Math 395 is 57, but in this text we will use 19 as our prime number, notice that we have $19 \cdot 3 = 57$. What a coincidence!

Special thanks to Jinyan's friends in 2021 Math 395 for helping Jinyan to write this text and correcting typos in this text.

1 | The Algebra and Topology

Functions and Sets

In the setting of the Extended Reals, every subset of \mathbb{R} has a supremum and infimum. For $S \subseteq \mathbb{R}$, if $\exists x \in S$, then we have $\inf S \leq x \leq \sup S$. Notice that we might have $\sup S < \inf S$ in the case where $S = \emptyset$, because $\sup \emptyset = -\infty < +\infty = \inf \emptyset$.

Let A be an index set, let X be a set. For $\alpha \in A$, we let $S_\alpha \subseteq X$.

$$\bigcup_{\alpha \in A} S_\alpha = \{x \in X \mid x \in S_\alpha \text{ for at least one } \alpha \in A\}$$

$$\bigcap_{\alpha \in A} S_\alpha = \{x \in X \mid x \in S_\alpha \text{ for all } \alpha \in A\}$$

Here we note that, if $A = \emptyset$, we have $\bigcup_{\alpha \in \emptyset} S_\alpha = \emptyset$, and $\bigcap_{\alpha \in \emptyset} S_\alpha = X$.

Definition 1.0.0.0.1

Let A be a set, let S be a subset of A . The indicator function \mathbb{I}_S is defined by the following:

$$\mathbb{I}_S : A \rightarrow \{0, 1\} \quad a \mapsto \begin{cases} 1 & a \in S \\ 0 & a \notin S \end{cases}$$

Definition 1.0.0.0.2

Let A be a set, $\mathcal{P}^{finite}(A)$ denote the collection of all finite subsets of A .

Definition 1.0.0.0.3

Let A be a set. $\# : \mathcal{P}(A) \rightarrow [0, \infty]$ $A' \mapsto \begin{cases} \text{cardinality of } A' & \text{if } A' \in \mathcal{P}^{finite}(A) \\ +\infty & \text{if } A' \notin \mathcal{P}^{finite}(A) \end{cases}$

Let S be a set, and let $P^{finite}(S)$ be a collection of all finite subsets of S . For $A, B \in P^{finite}(S)$, let $d_1 : P^{finite}(S) \times P^{finite}(S) \rightarrow \mathbb{R}$ and $d_2 : P^{finite}(S) \times P^{finite}(S) \rightarrow \mathbb{R}$ be functions that satisfy $d_1(A, B) = 0$ if $A = B$, $d_1(A, B) = 1$ if $A \neq B$, and $d_2(A, B) = \#(A \Delta B)$. Here d_1 and d_2 are metrics for $P^{finite}(S)$, and they induce the same topology on $P^{finite}(S)$, the discrete topology on $P^{finite}(S)$.

Lemma 1.0.1

Given nonempty $S, T \subseteq \mathbb{R}$ such that $s \leq t$ for all $s \in S$ and $t \in T$. Then we have the following holds:

1. $\sup(S) \leq \inf(T)$
2. $\sup(S) = \inf(T) \iff \forall \epsilon > 0, \exists s \in S \text{ and } t \in T \text{ such that } t \leq s + \epsilon.$

Proposition 1.0.2

Let S_1, S_2, \dots, S_k be sets, then $\#(S_1 \cup S_2 \cup \dots \cup S_k) \leq \#(S_1) + \#(S_2) + \dots + \#(S_k)$

Proof. One might proceed by using induction. Or, we might want to write the following:

$$S_1 \cup S_2 \cup \dots \cup S_k = S_1 \cup (S_2 \setminus S_1) \cup (S_3 \setminus (S_1 \cup S_2)) \cup \dots \cup \left(S_k \setminus \left(\bigcup_{i=1}^{k-1} S_i \right) \right)$$

Hence we have:

$$\begin{aligned} \#(S_1 \cup S_2 \cup \dots \cup S_k) &= \#(S_1) + \#(S_2 \setminus S_1) + \#(S_3 \setminus (S_1 \cup S_2)) + \dots + \# \left(S_k \setminus \left(\bigcup_{i=1}^{k-1} S_i \right) \right) \\ &\leq \#(S_1) + \#(S_2) + \dots + \#(S_k) \end{aligned}$$

The result follows. \square

Affine Sets

Definition 2.0.0.0.1

Let V be a vector space, and let $\vec{p}, \vec{q}, \vec{x} \in V$, where $\vec{p} \neq \vec{q}$. $\vec{p}, \vec{q}, \vec{x}$ are said to be **collinear** provided that $\vec{x} - \vec{p} = t(\vec{q} - \vec{p})$ for some scalar t .

Definition 2.0.0.0.2

Let V be a vector space over a field F , and let $\vec{p}, \vec{q} \in V$, where $\vec{p} \neq \vec{q}$. A line through \vec{p}, \vec{q} is a set L defined by $L := \{(1-t)\vec{p} + t\vec{q} \mid t \in F\}$.

Definition 2.0.0.0.3

Let V be a vector space over a field F , let S be a subset of V . S is called an affine set provided that S contains all lines through any two vectors in S . That is, for $\vec{x}, \vec{y} \in S$, and for $t \in F$, S is said to be affine provided that $(1-t)\vec{x} + t\vec{y} \in S$.

Example: Let F be a field.

$S_1 = \{(t+1, t, 2t), t \in F\}$ is a line through $(1, 0, 0), (2, 1, 2)$ and S_1 is an affine set.

$S_2 = \{(x, y, 3), x, y \in F\}$ is an affine set.

Theorem 2.1

Let V be a vector space over a field F , let $\vec{0}$ be the zero vector in V , and let $S \subseteq V$. If we have $\vec{0} \in S$, and $1_F + 1_F \neq 0_F$ for $1_F, 0_F \in F$, then S is affine if and only if S is a vector subspace of V .

Proof. For the \Leftarrow direction, let $\vec{x}, \vec{y} \in S$, and let t be a scalar, then $(1-t)\vec{x} + t\vec{y} \in S$ because S is a vector subspace of V , hence S is affine. For the \Rightarrow direction, let $\vec{0}, \vec{y} \in S$, let t be a scalar, since S is affine, we have $t\vec{y} = t\vec{y} + (1-t)\vec{0} \in S$. Moreover, let $\vec{x}, \vec{y} \in S$, since S is affine, then we have $\frac{1}{2}(2\vec{x}) + \frac{1}{2}(2\vec{y}) \in S$, and here we conclude that S is a vector subspace of V . \square

Note: Let $F = \{0_F, 1_F\} = \mathbb{Z}/2\mathbb{Z}$, where we have $1_F + 1_F = 0_F$. Here, let V be a vector space over the field F , let $\vec{p}, \vec{q} \in V$, then any line through \vec{p}, \vec{q} is the set $\{\vec{p}, \vec{q}\}$, hence all two-point sets are lines, and all sets are affine.

Remark: Not all sets containing $\vec{0}$ will be vector subspaces.

Definition 2.1.0.0.1

Let V be a vector space over a field F , let A, B be subsets of V , and let $\vec{x} \in V$.

$$V - \vec{x} := \{\vec{y} - \vec{x} \mid \vec{y} \in V\} \quad A - B := \{\vec{a} - \vec{b} \mid \vec{a} \in A, \vec{b} \in B\}$$

Lemma 2.1.1

Let S be a subset of a vector space V over a field F , and let $\vec{x} \in V$.

S is affine if and only if $S - \vec{x}$ is affine.

Proof. For the \Rightarrow direction, let $\vec{a}, \vec{b} \in S - \vec{x}$, we have $\vec{a} = \vec{m} - \vec{x}$ and $\vec{b} = \vec{n} - \vec{x}$ for some $\vec{m}, \vec{n} \in S$. $\forall t \in F$, we have $(1-t)\vec{m} + t\vec{n} \in S$, hence we have $\forall t \in F$, $(1-t)\vec{m} + t\vec{n} - \vec{x} = (1-t)(\vec{m} - \vec{x}) + t(\vec{n} - \vec{x}) = \vec{m} - \vec{x} - t\vec{m} + t\vec{n} - t\vec{x} = (1-t)\vec{a} + t\vec{b} \in S - \vec{x}$, this shows that $S - \vec{x}$ is affine when S is affine. For the \Leftarrow direction, we assume $S - \vec{x}$ is affine, let $t \in F$, let $\vec{m}, \vec{n} \in S$, then $\vec{m} - \vec{x}, \vec{n} - \vec{x} \in S - \vec{x}$, hence we know that $(1-t)(\vec{m} - \vec{x}) + t(\vec{n} - \vec{x}) \in S - \vec{x}$, which implies $\exists \vec{z} \in S$ such that $\vec{z} - \vec{x} = (1-t)(\vec{m} - \vec{x}) + t(\vec{n} - \vec{x}) = \vec{m} - \vec{x} - t\vec{m} + t\vec{n} - t\vec{x} = (1-t)\vec{m} + t\vec{n} - \vec{x}$, it follows that $\vec{z} = (1-t)\vec{m} + t\vec{n} \in S$, hence we know that S is also affine. \square

Note: Let S be a subset of a vector space V over a field F , and let $\vec{x} \in S$, by Theorem 2.1 and Lemma 2.1.1, we can conclude that S is affine if and only if $S - \vec{x}$ is affine if and only if $S - \vec{x}$ is a vector subspace of V .

Definition 2.1.1.0.1

Let S be a subset of V , where V is a vector space over a field F . $\tilde{S} := \{\vec{a} - \vec{b} \mid \vec{a}, \vec{b} \in S\}$

Theorem 2.2

Let V be a vector space over a field F , let S be a subset of V , and let $\vec{x} \in S$.

If S is affine, then we have $S - \vec{x} = \tilde{S}$.

Proof. Let $\vec{y} \in S - \vec{x}$, then we have $\vec{y} = \vec{a} - \vec{x}$ for some $\vec{a}, \vec{x} \in S$, which implies $\vec{y} \in \tilde{S}$, so we have $S - \vec{x} \subseteq \tilde{S}$. Now let $\vec{k} \in \tilde{S}$, then we have $\vec{k} = \vec{a} - \vec{b}$ for some $\vec{a}, \vec{b} \in S$, which implies $\vec{k} = (\vec{a} - \vec{x}) - (\vec{b} - \vec{x}) \in S - \vec{x}$. \square

Corollary 2.2.1

Let V be a vector space over a field F , let S be a subset of V .

If S is affine, for $\vec{x}_1, \vec{x}_2 \in S$, we have $S - \vec{x}_1 = S - \vec{x}_2$.

Remark: \tilde{S} is called the vector subspace associated to S .

Example:

$\tilde{S}_1 = \{(t, t, 2t) \mid t \in F\}$ is an affine set.

$\tilde{S}_2 = \{(x, y, 0) \mid x, y \in F\}$ is an affine set.

Remark: Let $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k$ be a basis of \tilde{S} , where $\tilde{S} \subseteq V$, and V is a vector space over a field F . Write $\tilde{S} = S - \vec{x}$ for some $\vec{x} \in S$. Then we have $\tilde{S} = \{t_1\vec{a}_1 + t_2\vec{a}_2 + \dots + t_k\vec{a}_k \mid t_j \in F\}$, and $S = \{\vec{x} + t_1\vec{a}_1 + t_2\vec{a}_2 + \dots + t_k\vec{a}_k \mid t_j \in F\}$.

Definition 2.2.1.0.1

Let S be a subset of V , where V is a vector space over a field F . S is said to be convex provided that S contains all line segments joining two of its points. That is, S is said to be convex provided that $\forall \vec{x}, \vec{y} \in S$, and $\forall t \in [0, 1]$, we have $(1-t)\vec{x} + t\vec{y} \in S$.

Example:

$S_3 = \{(x, y, 3) \mid x^2 + y^2 \leq 1\}$ is convex but not affine.

Definition 2.2.1.0.2

An affine line is an affine set of dimension 1.

Exercise:

- (1) For $S \subseteq V$, where $\dim V = 1$ and V is a vector space over \mathbb{R} , S is convex $\iff S$ is connected.
- (2) Let V be a vector space over \mathbb{R} , let $S \subseteq V$, then S is convex if and only if its intersection with each affine line is connected.

In the rest of this section, when we define a field F , we assume that for $1_F, 0_F \in F$, $1_F + 1_F \neq 0_F$.

Definition 2.2.1.0.3

Let F be a field, let V, W be vector spaces over F , and let $T : V \rightarrow W$ be a function.

$\text{Graph}(T) = \{(\vec{v}, \vec{w}) \in V \times W \mid \vec{w} = T\vec{v}\} = \{(\vec{v}, T\vec{v}) \in V \times W \mid \vec{v} \in V\}$ is called the graph of T .

Definition 2.2.1.0.4

A function T is said to be affine provided that the graph of T is affine.

Note: Let F be a field, let V, W be vector spaces over F , and let $T : V \rightarrow W$ be a function. If we have $T(\vec{0}) = \vec{0}$, which is equivalent to say $(\vec{0}, \vec{0}) \in \text{Graph}(T)$, then the followings are equivalent:

1. T is affine
2. $\text{Graph}(T)$ is a vector subspace of $V \times W$
3. For $v_1, v_2 \in V$ and $t \in F$, we have:

$$(\vec{v}_1 + \vec{v}_2, T(\vec{v}_1) + T(\vec{v}_2)) = (\vec{v}_1, T(\vec{v}_1)) + (\vec{v}_2, T(\vec{v}_2)) \in \text{Graph}(T)$$

$$(t\vec{v}_1, tT(\vec{v}_1)) = t(\vec{v}_1, T(\vec{v}_1)) \in \text{Graph}(T)$$
4. For $\vec{v}_1, \vec{v}_2 \in V, t \in F$, we have:

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2), \text{ and } tT(\vec{v}_1) = T(t\vec{v}_1)$$
5. T is linear

Note: Let F be a field, let V, W be vector spaces over F , and let $T : V \rightarrow W$ be a function, in general we have the followings being equivalent:

1. T is affine
2. $\text{Graph}(T)$ is affine
3. $\text{Graph}(T) - (\vec{0}, T(\vec{0})) = \{(\vec{v}, T(\vec{v}) - T(\vec{0})) \mid \vec{v} \in V\}$ is a vector subspace of $V \times W$
4. $L : V \rightarrow V \times W \quad \vec{v} \mapsto T(\vec{v}) - T(\vec{0})$ is a linear function
5. T is of the form $T(\vec{v}) = \tilde{T}(\vec{v}) + \vec{b}$ with \tilde{T} being linear and $\vec{b} \in W$

Exercise: From (5), we have \tilde{T} being uniquely determined, and $\vec{b} = T(\vec{0})$

Metric Spaces

Definition 3.0.0.0.1

A metric on a set X is a function $d : X \times X \rightarrow \mathbb{R}$ such that, for $x, y, z \in X$, the followings hold:

1. $d(x, y) = d(y, x)$
2. $d(x, y) \geq 0$, where $d(x, y) = 0$ if and only if $x = y$
3. $d(x, z) \leq d(x, y) + d(y, z)$

Definition 3.0.0.0.2

A metric space is a set equipped with a metric.

Definition 3.0.0.0.3

Let (X, d) be a metric space, and let $Y \subseteq X$. The restriction of d on $Y \times Y$, denoted as $d|_{Y \times Y}$, is the induced metric on Y , and $(Y, d|_{Y \times Y})$ forms a metric space.

Example: $d_{euc} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ $(\vec{x}, \vec{y}) \mapsto \sqrt{\sum_{j=1}^n (y_j - x_j)^2}$ is called the Euclidean metric.

Definition 3.0.0.0.4

Let (X, d) be a metric space, and let $\epsilon > 0$. For $x_0 \in X$, the set $B_\epsilon(x_0) = \{x \in X \mid d(x_0, x) < \epsilon\}$ is called the ϵ -ball around x_0 .

Definition 3.0.0.0.5

Let (X, d) be a metric space, let $x_0 \in X$, and let $A \subseteq X$.

x_0 is interior to A provided that $\exists \epsilon > 0$ such that $B_\epsilon(x_0) \subseteq A$.

x_0 is exterior to A provided that $\exists \epsilon > 0$ such that $B_\epsilon(x_0) \cap A = \emptyset$.

x_0 is a boundary point of A provided that $\forall \epsilon > 0$, we have $B_\epsilon(x_0) \cap A \neq \emptyset \neq B_\epsilon(x_0) \cap (X \setminus A)$.

Remark: If $x_0 \in X$ is neither interior nor exterior to $A \subseteq X$, then x_0 is a boundary point of A .

Definition 3.0.0.0.6

$\text{Int}(A) :=$ The set of interior points of A

$\text{Ext}(A) :=$ The set of exterior points of A

$\text{Bd}(A) :=$ The set of boundary points of A

Remark: $X = \text{Int}(A) \cup \text{Ext}(A) \cup \text{Bd}(A)$

Definition 3.0.0.0.7

Let (X, d) be a metric space, and let $A \subseteq X$, A is said to be open provided that $A = \text{Int}(A)$.

Remark: Definition 3.0.0.0.7 defines a topology on X .

Note: Let (X, d) be a metric space, let $A \subseteq X$, and let $x_0 \in X$.

1. $\forall \epsilon > 0$, $B_\epsilon(x_0)$ is open.
2. $\text{Int}(A)$ is an open subset of A , and it contains all open subsets of A .
3. $\text{Bd}(X \setminus A) = \text{Bd}(A)$
4. $\text{Bd}(A)$ is closed

Definition 3.0.0.0.8

Let (X, d) be a metric space and let $S \subseteq X$, S is said to be closed provided that $X \setminus S$ is open.

Note: A set A is closed if and only if $A = \text{Int}(A) \cup \text{Bd}(A)$

Definition 3.0.0.0.9

The closure of a set S is the closed superset set of S that is contained in all closed superset of S .

Definition 3.0.0.0.10

Let (X, d) be a metric space, and let (x_n) be a sequence of points in X . (x_n) converges to some $x \in X$, denoted as $x_n \rightarrow x$, provided that $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \in \mathbb{N}$ with $n > N$. (x_n) is said to be convergent provided that $\exists x \in X$ such that $x_n \rightarrow x$.

Note: All metric spaces are Hausdorff, the convergence of a sequence in a Hausdorff space is unique. Hence, if (x_n) is a sequence in a metric space (X, d) with $x_n \rightarrow x$ and $x_n \rightarrow y$, then we have $x = y$.

Definition 3.0.0.0.11

Let (X, d_x) and (Y, d_y) be metric spaces, and let $f : X \rightarrow Y$ be a function. f is said to be sequentially continuous at $x \in X$ provided that for all sequence (x_n) in X , $x_n \rightarrow x$ implies $f(x_n) \rightarrow f(x)$.

Theorem 3.1

Let (X, d_x) and (Y, d_y) be metric spaces, and let $f : X \rightarrow Y$ be a function. f is sequentially continuous if and only if f is continuous.

Definition 3.1.0.0.1

Let (x_n) be a sequence of points in a metric space (X, d) . (x_n) is said to be Cauchy provided that $\forall \epsilon > 0, \exists N \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$ for all $n, m \in \mathbb{N}$ with $n > N$ and $m > N$.

Definition 3.1.0.0.2

A metric space (X, d) is said to be complete provided that all Cauchy sequences of points in X converge in X .

Definition 3.1.0.0.3

A set X is said to be sequentially closed provided that all convergent sequence of points in X converges to some $x \in X$.

Note: Here we use d_{euc} to denote the Euclidean metric.

1. The Euclidean metric space (\mathbb{R}, d_{euc}) is complete
2. For $Y \subseteq \mathbb{R}^n$, the induced metric space (Y, d_{euc}) is complete if and only if Y is closed.
3. $Z \subseteq X$ is closed if and only if Z is sequentially closed

Definition 3.1.0.0.4

A topological space (X, \mathcal{T}) is said to be compact provided that every open cover of X admits a finite subcover. That is, (X, \mathcal{T}) is said to compact provided that, if we have $X = \bigcup_{\alpha \in A} X_\alpha$ where X_α is open in X , then $\exists \alpha_1, \alpha_2, \dots, \alpha_k \in A$ such that $X = X_{\alpha_1} \cup X_{\alpha_2} \cup \dots \cup X_{\alpha_k}$.

Lemma 3.1.1

(X, d) is compact if and only if every sequence in X admits a convergent subsequence.

Definition 3.1.1.0.1

A metric space (X, d) is said to be sequentially compact provided that every sequence of points in X has a convergent subsequence converging to a point in X .

Lemma 3.1.2

Let (X, d) be a sequentially compact metric space, $\forall \epsilon > 0$, we can cover X by finitely many ϵ -ball, in other words, X is totally bounded.

Proof. We proceed by contradiction, suppose there exists $\epsilon > 0$ that X cannot be covered by finitely many ϵ -balls. let $x_1 \in X$, and let $x_2 \notin B_\epsilon(x_1)$. In general, let for $i \in \mathbb{N}$, $x_{i+1} \notin B_\epsilon(x_1) \cup \dots \cup B_\epsilon(x_i)$, we collect x_i for all i and form a sequence (x_n) . It is immediate that (x_n) has no Cauchy subsequence, and hence (x_n) has no convergent subsequence, we reach a contradiction, which completes the proof. \square

Theorem 3.2 (Bolzano-Weierstrass Theorem for metric spaces)

A metric space (X, d) is compact if and only if (X, d) is sequentially compact.

Proof. For the \Rightarrow direction, we proceed by contradiction, let (X, d) be a compact metric space, let (x_n) be a sequence of points in X that does not have convergent subsequence. We claim that for all $y \in X$, there exists an open neighborhood B_y for y such that B_y contains only finitely many x_n . Here we have $X \subseteq \bigcup_{y \in X} B_y$. Since X is compact, there exists some $k \in \mathbb{N}$ such that $X \subseteq \bigcup_{y_k \in X} B_{y_k}$. Note that each B_{y_k} contains finitely many x_n , but the finite union of finite sets is finite, and $\{x_n \mid n \in \mathbb{N}\}$ is not finite, here we reach a contradiction, which completes the proof of the \Rightarrow direction. For the \Leftarrow direction, we let $X \subseteq \bigcup_{a \in A} X_a$, where $X_a \subseteq X$ is open, we will show that $X \subseteq X_{a_1} \cup \dots \cup X_{a_k}$ for some $k \in \mathbb{N}$, and $a_k \in A$. We call $S \subseteq X$ to be Good if S is contained in a finite union of X_a , and we call $S \subseteq X$ Bad otherwise. Note that, the finite union of Good sets is Good. We will show that X is Good. We proceed by contradiction, suppose X is bad, then by Lemma 3.1.2, for all $m \in \mathbb{N}$, $\exists v_m \in X$ such that $B_{\frac{1}{m}}(v_m)$ is Bad. We collect the v_m , for $m \in \mathbb{N}$, to be a sequence (v_m) . By hypothesis, we know that a subsequence (v_{m_k}) of (v_m) converges to some v in X . Pick α such that $v \in X_\alpha$, then $\exists \epsilon > 0$ such that $B_\epsilon(v) \subseteq X_\alpha$. Now pick k such that $m_k > \frac{2}{\epsilon}$, and $d(v, v_{m_k}) < \frac{\epsilon}{2}$. Now we can write $B_{\frac{1}{m_k}}(v_{m_k}) \subseteq B_{\epsilon/2}(v_{m_k}) \subseteq B_\epsilon(v) \subseteq X_\alpha$ by Triangle inequality, so here we know that $B_{\frac{1}{m_k}}(v_{m_k})$ is Good, a contradiction. This completes the proof. \square

Theorem 3.3 (Bolzano-Weierstrass Theorem for \mathbb{R}^n space)

For $X \subseteq \mathbb{R}^n$ and given Euclidean metric, X is sequentially compact if and only if X is closed and bounded.

Theorem 3.4 (Heine-Borel Theorem)

For $X \subseteq \mathbb{R}^n$ with Euclidean metric, X is compact if and only if X is closed and bounded.

Notice here, a metric space (X, d) is compact if and only if (X, d) is complete and totally bounded.

Exercise: Let (X, d) be a metric space, let $\tilde{d} : X \rightarrow \mathbb{R}$ $x \mapsto \min\{d(x), 1\}$. Here \tilde{d} is a metric on X . Moreover, d , \tilde{d} induce the same topology on X .

Connectivity and Continuity

Definition 4.0.0.0.1

Let (X, d_X) and (Y, d_Y) be metric spaces, let $f : X \rightarrow Y$ be a function. f is said to be Lipschitz provided that $\exists C \in [0, +\infty)$ such that $d_Y(f(x_1), f(x_2)) \leq C \cdot d_X(x_1, x_2)$ for all $x_1, x_2 \in X$. Here C is called a Lipschitz constant for f .

Note: Lipschitz functions are continuous, but continuous functions are not necessarily Lipschitz.

Example: $f : \mathbb{R} \rightarrow \mathbb{R}$ $x \mapsto x^2$ with Euclidean metric, f is continuous but not Lipschitz.

Exercise: If a function f is Lipschitz, then $\inf(\{x \mid x \text{ is a Lipschitz constant for } f\})$ is also a Lipschitz constant for f , that is, the set $\{x \mid x \text{ is a Lipschitz constant for } f\}$ has a minimum value.

Definition 4.0.0.0.2

A function f is said to be contraction provided that f is Lipschitz with a Lipschitz constant $C < 1$.

Definition 4.0.0.0.3

A function f is said to be bi-Lipschitz provided that f is a bijection and f and f^{-1} are both Lipschitz. That is, f is said to be bi-Lipschitz provided that $\exists C \in [0, +\infty), \tilde{C} \in (0, \infty)$ such that $\tilde{C} \cdot d_X(x_1, x_2) \leq d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2)$ for all $x_1, x_2 \in X$.

Example: The function $f : \mathbb{R} \rightarrow \mathbb{R}$ $x \mapsto x + \frac{\sin(x)}{19}$ is bi-Lipschitz.

Definition 4.0.0.0.4

A function f is said to be isometry provided that $d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2)$ for all $x_1, x_2 \in X$.

Definition 4.0.0.0.5

Let $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a function, h is called an affine isometry provided that $h(\vec{x}) = A\vec{x} + \vec{b}$ for some matrix A and vector $\vec{b} \in \mathbb{R}^n$, with $A^T A = I$ where I is the identity matrix.

Theorem 4.1 (Theorem 20.6 on Munkres)

All isometries from \mathbb{R}^n to \mathbb{R}^n are affine isometry.

Definition 4.1.0.0.1

Let X be a vector space over a field F . A function $f : X \times X \rightarrow \mathbb{R}$ is a translation-invariant function provided that $f(\vec{x}_1 + \vec{y}, \vec{x}_2 + \vec{y}) = f(\vec{x}_1, \vec{x}_2)$ for all $\vec{x}_1, \vec{x}_2, \vec{y} \in X$.

Exercise: For vector space X , if a function d is translation-invariant, then $d(\vec{x}_1, \vec{x}_2) = d(\vec{0}, \vec{x}_2 - \vec{x}_1)$.

Definition 4.1.0.0.2

Let X be a vector space over a field F . A function $f : X \times X \rightarrow \mathbb{R}$ is a said to be homogeneous provided that $f(t\vec{x}_1, t\vec{x}_2) = |t| \cdot f(\vec{x}_1, \vec{x}_2)$ for all $\vec{x}_1, \vec{x}_2 \in X$ and $t \in F$.

Exercise: For vector space X , if a function d is homogeneous, then $d(\vec{0}, t\vec{x}) = |t| \cdot d(\vec{0}, \vec{x})$.

Note: Let X be a vector space over \mathbb{R} , or \mathbb{C} , let d be a metric on X . d is both translation-invariant and homogeneous, and $d(\vec{0}, \vec{x})$ is called a norm on X .

Example: In the vector space \mathbb{R}^n , $\|\vec{x}\|_p := (|x_1|^p + \dots + |x_n|^p)^{\frac{1}{p}}$ is a norm when $1 \leq p < \infty$.

Example: In the vector space \mathbb{R}^n , $\|\vec{x}\| := \max\{|x_1|, \dots, |x_n|\}$ is a norm on \mathbb{R}^n .

Note: For Euclidean norm $\|\vec{x}\|$, $|\vec{x}| \leq \|\vec{x}\| \leq \sqrt{n}|\vec{x}|$. Also, $\|\vec{x}\|_p$ tends to $|\vec{x}|$ as p tends to infinity.

Lemma 4.1.1

Let (X, d) and (Y, d) be metric spaces, and let $f : X \rightarrow Y$ be a function. Then f is continuous if and only if $f^{-1}(U)$ is open in X when U is open in Y , if and only if f is sequentially continuous.

Proposition 4.1.2

Let f be a Lipschitz function, f is uniformly continuous.

Proposition 4.1.3

Let (X, \mathcal{T}) be a topological space, the followings are equivalent:

1. There exists $f : X \rightarrow \{0, 1\}$ that is a continuous surjective function.
2. There exists nonempty $A \subsetneq X$ that is open and closed in X

Proof. First we note that, the topology on $\{0, 1\}$ is the subspace topology inherited from the Euclidean topology on \mathbb{R} , hence we know that $\{1\}$ and $\{0\}$ are both open and closed. To show that (1) implies (2), let $A = f^{-1}(\{1\})$, the result follows. To show that (2) implies (1), we can make use of the indicator function $\mathbb{I}_A : X \rightarrow \{0, 1\}$ $x \mapsto 1$ if $x \in A$, 0 if $x \notin A$, then the result follows. \square

Proposition 4.1.4

In Euclidean topological space $(\mathbb{R}, \mathcal{T}_{EUC})$, the interval $[0, 1]$ is connected.

Proof. For any continuous function $f : [0, 1] \rightarrow \{0, 1\}$, f cannot be surjective by the Intermediate Value Theorem. Then the result follows. \square

Definition 4.1.4.0.1

Let $(X, \mathcal{T}), (\{0, 1\}, \mathcal{T}_E)$ be a topological spaces, where \mathcal{T}_E is the subspace topology on $\{0, 1\}$ inherited from the Euclidean topology on \mathbb{R} . X is said to be connected provided that for all continuous function f from X to the set $\{0, 1\}$, f is not surjective.

Definition 4.1.4.0.2

A set X is said to be path connected provided that there exists a continuous function $f : [0, 1] \rightarrow X$ with $f(0) = \alpha$ and $f(1) = \beta$ for all $\alpha, \beta \in X$.

Proposition 4.1.5

A path connected topological space is connected.

Proof. Let (X, \mathcal{T}) be a path connected topological space. We will show that X is connected, we proceed by contradiction, suppose X is not connected, then there exists $f : X \rightarrow \{0, 1\}$ that is a continuous surjective function. Pick $\alpha, \beta \in X$ with $f(\alpha) = 0$ and $f(\beta) = 1$. Let $\phi : [0, 1] \rightarrow X$ be a continuous function with $\phi(0) = \alpha$ and $\phi(1) = \beta$, then we have $f \circ \phi : [0, 1] \rightarrow \{0, 1\}$ being a continuous surjective function, which implies $[0, 1]$ is not connected, a contradiction. Hence X must be connected, this completes the proof. \square

Note:

1. If X is a convex subset of \mathbb{R}^n , then X is path connected.
2. If X is a convex subset of \mathbb{R}^n , then X is connected.
3. If X is an open connected subset of \mathbb{R}^n , then X is path connected.
4. Let $X = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y = \sin(\frac{1}{x})\}$, let $Y = \{(0, y) \mid -1 \leq y \leq 1\}$, then the set $X \cup Y$ is connected, but not path connected.
5. $GL(n, \mathbb{R}) := \{\text{Invertible } n \times n \text{ real matrices}\}$ is not connected.
6. $GL_+(n, \mathbb{R}) := \{M \in GL(n, \mathbb{R}) \mid \det(M) > 0\}$ is path connected.
7. If X is a connected subset of \mathbb{R} , then X is an interval or a singleton.

Here (5) can be proved by using the function $f : GL(n, \mathbb{R}) \rightarrow \{0, 1\} \quad M \mapsto \frac{\det(M)}{|\det(M)|} + 1$, which is continuous and surjective.

Remark: The function $p_j : \mathbb{R}^n \rightarrow \mathbb{R} \quad (x_1, \dots, x_j, \dots, x_n) \mapsto x_j$ is Lipschitz with $C = 1$, and hence the function p_j is continuous.

Proposition 4.1.6

Let (X, \mathcal{T}) be a topological space. The function $f : X \rightarrow \mathbb{R}^n \quad x \mapsto (f_1(x), f_2(x), \dots, f_n(x))$ is continuous if and only if each function $f_j : X \rightarrow \mathbb{R}$ is continuous.

Proof. For the \Rightarrow direction, note that $f_j = p_j \circ f$, and the composition of continuous functions is continuous. For the \Leftarrow direction, fix $x_0 \in X$, let $\epsilon > 0$, there exists $\delta > 0$ such that $|p_j(x) - p_j(x_0)| < \frac{\epsilon}{\sqrt{n}}$ for $1 \leq j \leq n$. For $x \in X$, if we have $d(x, x_0) < \delta$, then we can write $d(f(x), f(x_0)) = \|f(x) - f(x_0)\| \leq \sqrt{n}|f(x) - f(x_0)| < \sqrt{n} \cdot \frac{\epsilon}{\sqrt{n}} = \epsilon$, hence f is continuous. \square

Example: Consider a function f of the form given by the following:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad (x_1, x_2) \mapsto \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & (x_1, x_2) \neq (0, 0) \\ 0 & (x_1, x_2) = (0, 0) \end{cases}$$

The function f is continuous when $x_1 \in \mathbb{R}$ is fixed, and continuous when $x_2 \in \mathbb{R}$ is fixed, but the function f is not continuous on \mathbb{R}^2 . This is because the sequence (a_n) , where $a_n = (\frac{1}{n}, \frac{1}{n})$ converges to $(0, 0)$, but $f(\frac{1}{n}, \frac{1}{n}) = \frac{1}{2}$ converges to $(\frac{1}{2}, \frac{1}{2})$, and $f(0, 0) = (0, 0)$.

2 | Multivariate Differentiation

Differentiability

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function. f is continuous at $a \in \mathbb{R}$ as a if and only if the graph of f is "almost horizontal" when magnified horizontally at a . f is differentiable at a if and only if the graph of f at a is "almost affine," and non-vertical when magnified the same scale in both vertical and horizontal direction. Here we note that the continuity of f at a does not imply its differentiability at a .

Definition 5.0.0.0.1

Let V be a vector space over a field F . A function $f : V \rightarrow V$ $\vec{x} \mapsto \lambda \vec{x}$, for some scalar $\lambda \in F$, is called a dilation centered at $\vec{0}$. A function $d : V \rightarrow V$ $\vec{x} \mapsto \lambda \vec{x} + (1 - \lambda) \vec{p}$, for $\lambda \in F$ and $\vec{p} \in V$, is called a dilation centered at \vec{p} .

In the following discussion, we will try to find a way to define the differentiability of functions in vector spaces. Let V and W be vector spaces over the field \mathbb{R} or the field \mathbb{C} . Given $f : V \rightarrow W$ where $\text{Graph}(f) = \{(\vec{x}, f(\vec{x})) \mid \vec{x} \in V\}$. Dilate $\text{Graph}(f)$ about the point $(\vec{a}, f(\vec{a}))$, we get $(\vec{x}, f(\vec{x})) \mapsto (\lambda(\vec{x} - \vec{a}) + \vec{a}, \lambda(f(\vec{x}) - f(\vec{a})) + f(\vec{a}))$. Set $t = \frac{1}{\lambda}$, and $\vec{u} = \frac{\vec{x} - \vec{a}}{t} \Rightarrow \vec{x} = \vec{a} + t\vec{u}$. The dilated graph is given by $\{(\vec{a} + t\vec{u}, f(\vec{a}) + \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}) \mid \vec{u} = \frac{\vec{x} - \vec{a}}{t}, \vec{x} \in V\}$. Consider the case where t tends to 0. We are hoping that $f(\vec{a}) + \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}$ is an affine function of \vec{u} . Note that \vec{a} is fixed, so we hope to find some $T(\vec{u}) + \vec{b}$ where $\vec{b} = f(\vec{a})$, and $T(\vec{u}) = \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}$ is linear. Now the problem reduces to finding a linear function T that satisfies $\lim_{t \rightarrow 0} \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t} = T(\vec{u})$. If such T exists, then we can write the directional derivative of f at \vec{a} in the direction of \vec{u} as $f'(\vec{a}; \vec{u}) := \lim_{t \rightarrow 0} \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}$. We could try to use this as a core definition for multivariate differential calculus. However, we need to be careful with the following cases:

An example in Munkres Section 5 shows that there exists a function f such that all $f'(\vec{0}; \vec{u})$ exists but not linear at \vec{u} . Also, we might consider the function:

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

here we have $f'(\vec{0}; \vec{u}) = 0$ for all \vec{u} , and $f(\frac{1}{n}, \frac{1}{n^4}) = n$ which does not tends to 0 when n tends to ∞ , so f is not sequentially continuous, and hence not continuous at $(0, 0)$. Moreover, the Chain Rule for multivariate calculus will fail without stronger assumptions for the differentiability of a function.

Note: Let I be an open interval in \mathbb{R} , let W be a vector space over a field F . Suppose W is equipped with a norm which induces a metric d . If $\dim W < \infty$, then all norms on W induce the same topology on W , and for metric d_1 and d_2 for W , the identity function $Id : (W, d_1) \rightarrow (W, d_2)$ is a bi-Lipschitz function. On another note, for function $f : I \rightarrow W$, we write $f'(x) = \lim_{t \rightarrow 0} \frac{f(x+t) - f(x)}{t}$.

Example: Let V and W be normed vector spaces, let $f : V \rightarrow W$ be a function. For $\vec{a}, \vec{u} \in V$, we define $g_{\vec{a}, \vec{u}} : \mathbb{R} \rightarrow V$ $t \mapsto \vec{a} + t\vec{u}$, then $g'_{\vec{a}, \vec{u}}(t) = \vec{u}$ for all $t \in \mathbb{R}$. We write:

$$(f \circ g_{\vec{a}, \vec{u}})'(0) = \lim_{t \rightarrow 0} \frac{(f \circ g_{\vec{a}, \vec{u}})(t) - (f \circ g_{\vec{a}, \vec{u}})(0)}{t} = \lim_{t \rightarrow 0} \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t} = f'(\vec{a}; \vec{u})$$

Now the directional derivative $f'(\vec{a}; \vec{u})$ reduces to derivative of vector-valued function of a scalar. We do not really need f to be defined on all of V , we just need $f \circ g_{\vec{a}, \vec{u}}$ defined in neighborhood of 0 in \mathbb{R} , in particular, it is good enough to have $\vec{a} \in \text{Int}(\text{domain of } f)$.

Note: Let V and W be normed vector spaces over field \mathbb{R} or \mathbb{C} . For $T \in \text{hom}(V, W)$, TFAE:

1. T is Lipschitz
2. T is continuous
3. T is continuous at $\vec{0} \in V$
4. $\exists M \in [0, \infty)$ such that $\|T\vec{v}\| \leq M\|\vec{v}\| \forall \vec{v} \in V$, in which we say T is bounded.

Definition 5.0.0.0.2

For normed vector spaces V and W :

$$B(V, W) := \{T \in \text{hom}(V, W) \mid T \text{ is continuous, bounded, and Lipschitz}\}$$

Note: For normed vector spaces V and W , if $\dim(V) < \infty$, then we have $\text{hom}(V, W) = B(V, W)$. That is, the linearity of a function $f \in \text{hom}(V, W)$ implies the continuity of f

Let V and W be vector spaces over the field \mathbb{R} or the field \mathbb{C} . Given $f : V \rightarrow W$, we dilate $\text{Graph}(f)$ about the point $(\vec{a}, f(\vec{a}))$, take the limit, the dilated graphs converge point-wise to $\{(\vec{a} + \vec{u}, f(\vec{a}) + f'(\vec{a}; \vec{u})) \mid \vec{u} \in V\} = \{(\vec{y}, f(\vec{a}) + f'(\vec{a}, \vec{y} - \vec{a})) \mid \vec{y} \in V\} = \text{Graph}(\vec{y} \mapsto f(\vec{a}) + f'(\vec{a}, \vec{y} - \vec{a}))$. To get a good theory, we need the followings hold:

1. $f'(\vec{a}; \vec{u}) = T(\vec{u})$ is linear in \vec{u} .
2. T need to be continuous.
3. We need $f(\vec{a}) + T(\vec{y} - \vec{a}) \approx f(\vec{y})$. That is, let $\vec{h} = \vec{y} - \vec{a}$, we need $f(\vec{a}) + T(\vec{h}) \approx f(\vec{a} + \vec{h})$. Here, the error, $\frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - T(\vec{h})}{\|\vec{h}\|}$, should be small compared to \vec{h} , or it should tends to $\vec{0}$.

Definition 5.0.0.0.3

Let V, W be normed vector spaces over the field \mathbb{R} or the field \mathbb{C} , let $\vec{a} \in A$, where A is an open subset of V . For function $f : V \rightarrow W$, we write $Df(\vec{a}) = T$ provided that there exists $T \in B(V, W)$ that satisfies the following:

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

If such T exists, then we write $Df(\vec{a})(\vec{h}) = T(\vec{h})$ and f is said to be differentiable at \vec{a} .

Proposition 5.0.1 (Munkres Theorem 5.1)

Let V, W be normed vector spaces over the field \mathbb{R} or the field \mathbb{C} , let $\vec{a} \in A$, where A is an open subset of V , and let $f : V \rightarrow W$ be a function that is differentiable at \vec{a} , for $\vec{u} \in V$, we have $f'(\vec{a}; \vec{u}) = Df(\vec{a})(\vec{u})$ where $Df(\vec{a})$ is a linear map from V to W .

Corollary 5.0.2

Let V, W be normed vector spaces over the field \mathbb{R} or the field \mathbb{C} , let $\vec{a} \in A$, where A is an open subset of V , and let $f : V \rightarrow W$ be a function that is differentiable at \vec{a} . If we have $Df(\vec{a}) = T_1$ and $Df(\vec{a}) = T_2$, then $T_1 = T_2$.

Exercise: Let V and W be vector spaces over the field \mathbb{R} or the field \mathbb{C} , let $f : V \rightarrow W$ $\vec{v} \mapsto T(\vec{v}) + \vec{b}$ be a function where $T : V \rightarrow W$ is a linear transformation, and $\vec{b} \in W$. Then the function f is differentiable and $Df(\vec{a}) = T$ for all $\vec{a} \in V$.

Proposition 5.0.3

Let V, W be normed vector spaces over the field \mathbb{R} or the field \mathbb{C} , let $\vec{a} \in A$, where A is an open subset of V , and let $f : V \rightarrow W$ be a function that is differentiable at \vec{a} . f is continuous at \vec{a} .

Proof. It suffices to show that $f(\vec{x})$ approaches $f(\vec{a})$ as \vec{x} approaches \vec{a} . Write $\vec{x} = \vec{a} + \vec{h} \Rightarrow \vec{h} = \vec{x} - \vec{a}$. We need to show that $f(\vec{a} + \vec{h}) - f(\vec{a})$ approaches $\vec{0}$ as \vec{h} approaches $\vec{0}$. Let $T = Df(\vec{a})$. Notice that:

$$f(\vec{a} + \vec{h}) - f(\vec{a}) = \|\vec{h}\| \cdot \frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - T(\vec{h})}{\|\vec{h}\|} + T(\vec{h}) \quad (*) := \frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - T(\vec{h})}{\|\vec{h}\|}$$

Here we have $\|f(\vec{a} + \vec{h}) - f(\vec{a})\| \leq \|\vec{h}\| \cdot \|(*)\| + \|T(\vec{h})\|$. Since T is linear, then T is continuous and hence it is bounded, and tends to $\vec{0}$ as \vec{h} tends to $\vec{0}$. Since f is differentiable, then we know that $(*)$ tends to $\vec{0}$ as \vec{h} tends to $\vec{0}$, hence we conclude that $\|\vec{h}\| \cdot \|(*)\| + \|T(\vec{h})\|$ tends to $\vec{0}$ as \vec{h} tends to $\vec{0}$. This completes the proof. \square

Remark: Check Munkres Theorem 5.4 for details about the following:

Let V be a vector space, and let $f : V \rightarrow \mathbb{R}^n$ be a function. For $\vec{v} \in V$, the function f can be written as $f(\vec{v}) = (f_1(\vec{v}), f_2(\vec{v}), \dots, f_n(\vec{v}))$. We can also write f as the following column matrix:

$$f(\vec{v}) = \begin{bmatrix} f_1(\vec{v}) \\ f_2(\vec{v}) \\ \vdots \\ f_n(\vec{v}) \end{bmatrix}$$

The function f is differentiable at $\vec{a} \in V$ if and only if each f_j is differentiable at \vec{a} , and if each f_j is differentiable at \vec{a} , then for $\vec{u} \in V$, we can write the following:

$$Df(\vec{a})(\vec{u}) = \begin{bmatrix} Df_1(\vec{a})(\vec{u}) \\ Df_2(\vec{a})(\vec{u}) \\ \vdots \\ Df_n(\vec{a})(\vec{u}) \end{bmatrix}$$

Remark:

Let W be a vector space, and let $f : \mathbb{R}^m \rightarrow W$ be a function. For $\vec{u} = (u_1, u_2, \dots, u_m) \in \mathbb{R}^m$,

$$Df(\vec{a}) \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \sum u_j Df(\vec{a})(\vec{e}_j) = \sum u_j f'(\vec{a}; \vec{e}_j) := \sum u_j D_j f(\vec{a})$$

here $D_j f(\vec{a})$ is called a partial derivative of f , and we have the following holds:

$$Df(\vec{a})(\vec{u}) = [D_1 f(\vec{a}) \ \cdots \ D_m f(\vec{a})] [u_1 \ \cdots \ u_m]^T$$

here we note that $[D_1 f(\vec{a}) \cdots D_m f(\vec{a})]$ is a list of vectors in W .

Remark:

For $V = \mathbb{R}^m$, $W = \mathbb{R}^n$, and a differentiable function $f : V \rightarrow W$. Each $D_j f(\vec{a})$ is a column vector in \mathbb{R}^n . Here we have $[D_1 f(\vec{a}) \ \cdots \ D_m f(\vec{a})]$ being an $n \times m$ matrix representing $Df(\vec{a})$. In such matrix, the (j, k) entry is $D_k f_j(\vec{a})$. If we have $n = 1$, then $Df(\vec{a}) = [D_1 f(\vec{a}) \cdots D_m f(\vec{a})] \in \mathbb{R}_{row}^m$ is called the gradient of f .

Remark:

Generally, or loosely, if f is differentiable, then $f(\vec{a} + \vec{h}) \approx f(\vec{a}) + Df(\vec{a})(\vec{h})$ for $\vec{h} \approx \vec{0}$, as we result, we can write $f(\vec{x}) \approx f(\vec{a}) + Df(\vec{a})(\vec{x} - \vec{a})$ for $\vec{x} \approx \vec{a}$.

In the following, we will show the Chain Rule holds to multivariate differentiation. Let V, W, Z be normed vector spaces over the field \mathbb{R} or the field \mathbb{C} , let $f : V \rightarrow W$ be a differentiable function. Suppose we have $f(\vec{a}) = \vec{b} \in B$ for some $\vec{a} \in V$ and open subset B of V , and suppose we have another differentiable function $g : B \rightarrow Z$, where $g(f(\vec{x})) \approx g(\vec{b}) + Dg(\vec{b})(f(\vec{x}) - f(\vec{a}))$ for $f(\vec{x}) \approx f(\vec{a})$, here $f(\vec{x}) \approx f(\vec{a})$ implies $\vec{x} \approx \vec{a}$ suffices. Note that we have the following holds:

$$g(f(\vec{x})) \approx g(\vec{b}) + Dg(\vec{b})(f(\vec{x}) - f(\vec{a})) \approx g(f(\vec{a})) + Dg(\vec{b})(Df(\vec{a})(\vec{x} - \vec{a}))$$

which suggests that $D(g \circ f)(\vec{a}) = Dg(\vec{b}) \circ Df(\vec{a})$ here $Dg(\vec{b})$ and $Df(\vec{a})$ are both matrices, hence \circ should be replaced by a \cdot , and this suggests the Chain Rule in multivariate calculus.

Theorem 5.1 (Chain Rule for Multivariate Differentiation)

Let V, W, Z be normed vector spaces over the field \mathbb{R} or the field \mathbb{C} , let $f : V \rightarrow W$ be a differentiable function, denote $\text{im}(f) = B$, where B is an open subset of W , and let $g : B \rightarrow Z$ be a differentiable function. For $\vec{a} \in V$, $D(g \circ f)(\vec{a}) = Dg(\vec{b}) \circ Df(\vec{a})$, where $f(\vec{a}) = \vec{b}$.

Proof. Write $f(\vec{a}) = \vec{b} \in B$. We want to show that $(**)$ approaches $\vec{0}$ as $\vec{h} \in V$ approaches $\vec{0}$, where

$$(**) := \frac{g(f(\vec{a} + \vec{h})) - g(f(\vec{a})) - Dg(\vec{b})(Df(\vec{a})(\vec{h}))}{\|\vec{h}\|}$$

Define

$$F(\vec{h}) := \begin{cases} \frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - Df(\vec{a})(\vec{h})}{\|\vec{h}\|} & \vec{h} \neq \vec{0} \\ \vec{0} & \vec{h} = \vec{0} \end{cases} \quad G(\vec{t}) := \begin{cases} \frac{g(\vec{b} + \vec{t}) - g(\vec{b}) - Dg(\vec{b})(\vec{t})}{\|\vec{t}\|} & \vec{t} \neq \vec{0} \\ \vec{0} & \vec{t} = \vec{0} \end{cases}$$

F is continuous at $\vec{0}$ with $F(\vec{0}) = \vec{0}$. Note that here we have $f(\vec{a} + \vec{h}) - f(\vec{a}) = Df(\vec{a})(\vec{h}) + \|\vec{h}\|F(\vec{h})$. Similarly, we can write $g(\vec{b} + \vec{k}) - g(\vec{b}) = Dg(\vec{b})(\vec{k}) + \|\vec{k}\|G(\vec{k})$, with $G(\vec{0}) = \vec{0}$ and G is continuous at $\vec{0}$, and here we let $\vec{k} := f(\vec{a} + \vec{h}) - f(\vec{a})$, then we can write the following:

$$\begin{aligned} (**)&= \frac{g(\vec{b} + \vec{k}) - g(\vec{b}) - Dg(\vec{b})(Df(\vec{a})(\vec{h}))}{\|\vec{h}\|} \\ &= \frac{Dg(\vec{b})(\vec{k}) + \|\vec{k}\|G(\vec{k}) - Dg(\vec{b})(Df(\vec{a})(\vec{h}))}{\|\vec{h}\|} \\ &= \frac{Dg(\vec{b})(Df(\vec{a})(\vec{h})) + \|\vec{h}\|Dg(\vec{b})F(\vec{h}) + \|\vec{k}\|G(\vec{k}) - Dg(\vec{b})(Df(\vec{a})(\vec{h}))}{\|\vec{h}\|} \\ &= Dg(\vec{b})(F(\vec{h})) + \frac{\|\vec{k}\|}{\|\vec{h}\|}G(\vec{k}) \end{aligned}$$

As \vec{h} goes to $\vec{0}$, we have $F(\vec{h})$ tends to $\vec{0}$ by continuity, and hence $Dg(\vec{b})(F(\vec{h}))$ tends to $\vec{0}$. Also, as \vec{h} tends to $\vec{0}$, \vec{k} also tends to $\vec{0}$ by definition of \vec{k} , and hence $G(\vec{k})$ tends to $\vec{0}$ by continuity.

$$\frac{\|\vec{k}\|}{\|\vec{h}\|} = \left\| Df(\vec{a}) \left(\frac{\vec{h}}{\|\vec{h}\|} \right) + F(\vec{h}) \right\| \leq M \left\| \frac{\vec{h}}{\|\vec{h}\|} \right\| = M$$

for some $M \in [0, \infty)$ as $Df(\vec{a})$ and F being continuous, then we know that $(**)$ tends to $\vec{0}$ as \vec{h} tends to $\vec{0}$, and hence the result follows. \square

Note: Let f, g be functions over vector spaces, if both f and g have all directional derivatives, it is not necessary that $g \circ f$ has all directional derivatives.

Remark: Let W be a normed vector space, let $\alpha : W \times W \rightarrow W$ $(\vec{w}_1, \vec{w}_2) \mapsto \vec{w}_1 + \vec{w}_2$, then we know that α is linear and continuous. $D\alpha(\vec{w}_1, \vec{w}_2) = \alpha$.

Remark: Let V, W be normed vector spaces, let A be an open subset of V , let $f_1 : A \rightarrow W$ and $f_2 : A \rightarrow W$ be functions, let $g : A \rightarrow W \times W$ $\vec{a} \mapsto (f_1(\vec{a}), f_2(\vec{a}))$. If f_1 and f_2 are both differentiable at \vec{a} , then we have $Dg(\vec{a}) = (Df_1(\vec{a}), Df_2(\vec{a}))$. Let $\alpha : W \times W \rightarrow W$ $(\vec{w}_1, \vec{w}_2) \mapsto \vec{w}_1 + \vec{w}_2$, then if f_1 and f_2 are differentiable at \vec{a} , then $\alpha \circ g : A \rightarrow W$ is differentiable at \vec{a} by the Chain Rule, and we have $D(\alpha \circ g)(\vec{a}) = D\alpha(g(\vec{a}))(Df_1(\vec{a}), Df_2(\vec{a})) = Df_1(\vec{a}) + Df_2(\vec{a})$

Exercise: Let $\mu : \mathbb{R}^2 \rightarrow \mathbb{R}$ $(x_1, x_2) \mapsto x_1 x_2$. $D\mu \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} D_1 \mu \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & D_2 \mu \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{bmatrix} = [x_2 \ x_1]$

Example: Let V be a normed vector space. Let A be an open subset of V .

Let $f_1 : A \rightarrow \mathbb{R}$ and $f_2 : A \rightarrow \mathbb{R}$, both are differentiable at $\vec{a} \in A$.

$$g : A \rightarrow \mathbb{R}^2 \quad \vec{a} \mapsto \begin{bmatrix} f_1(\vec{a}) \\ f_2(\vec{a}) \end{bmatrix} \quad \mu : \mathbb{R}^2 \rightarrow \mathbb{R} \quad (x_1, x_2) \mapsto x_1 x_2$$

By Chain Rule, we write the following:

$$D(\mu \circ g)(\vec{a}) = D\mu \begin{bmatrix} f_1(\vec{a}) \\ f_2(\vec{a}) \end{bmatrix} \begin{bmatrix} Df_1(\vec{a}) \\ Df_2(\vec{a}) \end{bmatrix} = [f_2(\vec{a}), f_1(\vec{a})] \begin{bmatrix} Df_1(\vec{a}) \\ Df_2(\vec{a}) \end{bmatrix} = f_2(\vec{a})Df_1(\vec{a}) + f_1(\vec{a})Df_2(\vec{a})$$

Example: Consider $f_1 : A \rightarrow \text{Mat}(n, m, \mathbb{R})$, $f_2 : A \rightarrow \text{Mat}(m, p, \mathbb{R})$, $f_1 f_2 : A \rightarrow \text{Mat}(n, p, \mathbb{R})$

$$D(f_1 f_2)(\vec{a})(\vec{h}) = Df_1(\vec{a})(\vec{h}) \cdot f_2(\vec{a}) + f_1(\vec{a}) \cdot Df_2(\vec{a})(\vec{h})$$

We can also let f_1 and f_2 be \mathbb{C} -valued, and identify $a + ib$ with $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ to get expected results.

Continuous Differentiable Functions

Definition 6.0.0.0.1

Let V, W be vector spaces, let A be an open subset of V , let $f : A \rightarrow W$ be a function. f is said to be continuously differentiable on A provided that $Df : A \rightarrow B(V, W)$ exists and Df is continuous on A . In such case, we say f belongs to $C^1(A, W)$.

Note: The continuity of Df is not automatic. (Example on Munkres Section 6 Exercise 2)

Let V and W be vector spaces, and let A be an open subset of V , let f be a function from V to W . Suppose f is differentiable at each $\vec{a} \in A$. For $\vec{h} \in V$, we have $Df(\vec{a})(\vec{h}) \in W$. Here we can write $Df : A \rightarrow B(V, W)$, where $B(V, W)$ is a normed space of linear continuous functions. On another note, if we have $V = \mathbb{R}^m$ and $W = \mathbb{R}^n$, then $B(V, W)$ is the matrix space $\text{Mat}_{n \times m}(\mathbb{R})$. In such case, A is an open subset of \mathbb{R}^m , $f : A \rightarrow \mathbb{R}^n \in C^1$ if and only if f is differentiable at each $\vec{a} \in A$ and $Df : A \rightarrow \text{Mat}_{n \times n}(\mathbb{R})$ is continuous. In the following, we will show that $f \in C^1$ if and only if each $D_k f_j = f'_j(\vec{a}; \vec{e}_k)$ exists and is continuous on A . Here f_j denote the component function of f .

Theorem 6.1

Let A be an open subset of \mathbb{R}^m , and let $f : A \rightarrow \mathbb{R}^n \quad \vec{a} \mapsto (f_1(\vec{a}), f_2(\vec{a}), \dots, f_n(\vec{a}))$ be a function. If $D_k f_j$ exists and is continuous, then $f \in C^1(A, \mathbb{R}^n)$.

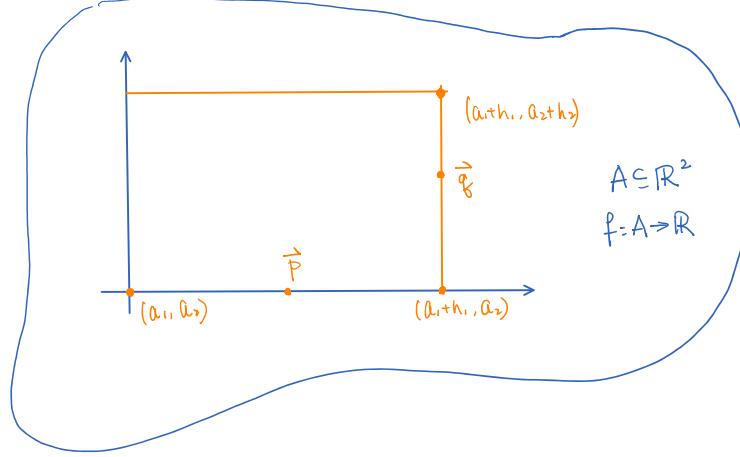
Proof. It suffices to show that f is differentiable at each $\vec{a} \in A$, and it suffices to consider $n = 1$ by Theorem 5.4 from Munkres Section 5. For $m = 2$, let $\vec{a} = (a_1, a_2)$ and $\vec{h} = (h_1, h_2)$, then we can write $f(a_1 + h_1, a_2) - f(a_1, a_2) = D_1 f(\vec{p}) \cdot h_1$ for some $\vec{p} \in \mathbb{R}^2$ by the Mean Value Theorem (MVT), and $f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) = D_2 f(\vec{q}) \cdot h_2$ for some $\vec{q} \in \mathbb{R}^2$ by the MVT, which implies $f(\vec{a} + \vec{h}) - f(\vec{a}) = D_1 f(\vec{p})h_1 + D_2 f(\vec{q})h_2$. Consider the following:

$$\frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - (D_1 f(\vec{a})h_1 + D_2 f(\vec{a})h_2)}{\|\vec{h}\|} \quad (1)$$

Since $f(\vec{a} + \vec{h}) - f(\vec{a}) = D_1 f(\vec{p})h_1 + D_2 f(\vec{q})h_2$, we can rewrite equation (1) as the following:

$$(D_1 f(\vec{p}) - D_1 f(\vec{a})) \frac{h_1}{\|\vec{h}\|} - (D_2 f(\vec{q}) - D_2 f(\vec{a})) \frac{h_2}{\|\vec{h}\|}$$

Since $D_1 f$ and $D_2 f$ are continuous, then we know that, when \vec{h} tends to $\vec{0}$, $D_1 f(\vec{p}) - D_1 f(\vec{a})$ and $D_2 f(\vec{q}) - D_2 f(\vec{a})$ tends to zero, and we have $\frac{h_1}{\|\vec{h}\|}$ and $\frac{h_2}{\|\vec{h}\|}$ both bounded above by 1. This shows that the limit for equation (1) as \vec{h} tends to $\vec{0}$ is 0, and the result follows.



The proof for $m > 2$ is left for reader, and can be found in Munkres Theorem 6.2. \square

Suppose we have a continuous differentiable map from A to W , where A is a open subset of a vector space V and W is a vector space. Then we know that $Df : A \rightarrow B(V, W)$ is continuous, and Df might be differentiable. If Df is indeed differentiable, then we can define the following:

Definition 6.1.0.0.1

Let V, W be vector spaces, let A be an open subset of V , let $f : A \rightarrow W$ be a function that belongs to $C^1(A, W)$. We write $f \in C^2(A, W)$ provided that $Df \in C^1(A, B(V, W))$, that is, we write $f \in C^2(A, W)$ provided that Df is differentiable at each $\vec{a} \in A$, and $D^2 f$ is continuous on A , where $D^2 f := D(Df) : A \rightarrow B(V, B(V, W))$

Remark: For vector spaces V, W

The set $B(V, B(V, W))$ can be viewed as a set of functions from $V \times V$ to W .

Note: For vector spaces $\mathbb{R}^m, \mathbb{R}^n$, let A be an open subset of \mathbb{R}^m , if $f : A \rightarrow \mathbb{R}^n \in C^2(A, \mathbb{R}^n)$, then $f \in C^1(A, \mathbb{R}^n)$ and each $D_{j_1}D_{j_2}f$ exists and is continuous on A , Theorem 6.1 provides the converse.

Definition 6.1.0.0.2

Let V, W be vector spaces, let A be an open subset of V ,
 $C^0(A, W) = C(A, W)$ is the set of continuous function from A to W .

Definition 6.1.0.0.3

Let V, W be vector spaces, let A be an open subset of V ,

$$C^\infty(A, W) := \bigcap_{r \in \mathbb{N}} C^r(A, W)$$

Note: For $r \in \mathbb{N}$, a function f is of C^r type if and only if Df is of C^{r-1} type.

Theorem 6.2

Given $f \in C^2(A, \mathbb{R})$, where A is an open subset of \mathbb{R}^2 . Then we can write the following:

$$D_2D_1f(a, b) = \lim_{(h,k) \rightarrow (0,0)} \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)}{hk}$$

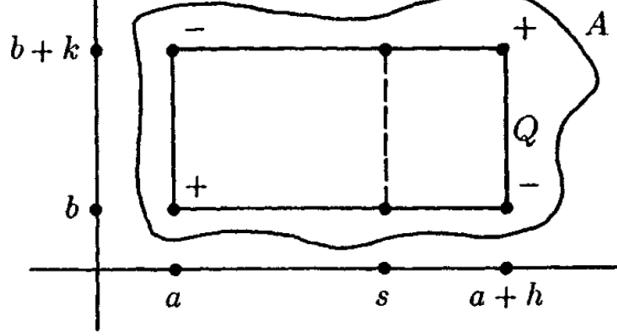
Proof. Let $\phi : M \rightarrow \mathbb{R}$ $s \mapsto f(s, b+k) - f(s, b)$ where M is an open interval contained in $[a, a+h]$. Differentiate ϕ using the Chain Rule, then we can write $\phi'(s) = D_1f(s, b+k) - D_1f(s, b)$.

$$\text{Let } \text{num} := f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)$$

Using $\phi'(s) = D_1f(s, b+k) - D_1f(s, b)$, we can write:

$$\text{num} = \phi(a+h) - \phi(a) = \phi'(s_0) \cdot h = (D_1f(s_0, b+k) - D_1f(s_0, b)) \cdot h = D_2D_1f(s_0, t_0) \cdot k \cdot h$$

for some $s_0 \in [a, a+h]$ and $t_0 \in [b, b+k]$ by the Mean Value Theorem. For $\frac{\text{num}}{hk} = D_2D_1f(s_0, t_0)$, as (h, k) tends to $(0, 0)$, by continuity, we have $D_2D_1f(s_0, t_0)$ tends to $D_2D_1f(a, b)$, the result follows.



(Figure 6.2 from Munkres Section 6)

□

Note: The existence of D^2f is not enough for this result.

Corollary 6.2.1 (Clairaut's Theorem)

Given $f \in C^2(A, \mathbb{R})$, where A is an open subset of normed vector space V .

We have $D_1D_2f = D_2D_1f$

Definition 6.2.1.0.1

Let V, W be normed vector spaces, let A be an open subset of V , let $f : A \rightarrow W$ be a function that is differentiable at $\vec{u} \in A$, we write $D_{\vec{u}}f : A \rightarrow W$ $\vec{a} \mapsto f'(\vec{a}; \vec{u}) = (Df(\vec{a}))(\vec{u})$

Corollary 6.2.2

Given $f \in C^2(A, \mathbb{R})$, where A is an open subset of normed vector space V .

We have $D_{\vec{u}_1}D_{\vec{u}_2}f = D_{\vec{u}_2}D_{\vec{u}_1}f$

Proof plan for Corollary 6.2.2. Apply the Chain Rule to $(x_1, x_2) \mapsto f(\vec{a} + x_1\vec{u}_1 + x_2\vec{u}_2)$ □

Another proof plan for Corollary 6.2.2. Adapt from the proof of Theorem 6.2, with the limit:

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(\vec{a} + h\vec{u}_1 + k\vec{u}_2) - f(\vec{a} + h\vec{u}_1) - f(\vec{a} + k\vec{u}_2) + f(\vec{a})}{hk} = D_{\vec{u}_1}D_{\vec{u}_2}f(\vec{a}) = D_{\vec{u}_2}D_{\vec{u}_1}f(\vec{a})$$

□

Corollary 6.2.3

Given $f \in C^2(A, \mathbb{R}^m)$, where A is an open subset of normed vector space V .

We have $D_{\vec{u}_1} D_{\vec{u}_2} f = D_{\vec{u}_2} D_{\vec{u}_1} f$

Proof. Apply the proof of Theorem 6.2 to each component function of f and reassemble. \square

Remark: Theorem 6.2 and its corollaries also work for $f \in C^2(A, W)$ where W is an arbitrary real normed vector space and A is an open subset of \mathbb{R}^2 . If $\dim W < \infty$, we can use the equivalence of norms to get the result similar to $W = \mathbb{R}^n$. If $\dim W = \infty$, the statements in the theorem and the corollaries are still true, but we need addition tools to prove.

Suppose $f \in C^r$, consider $D_{\vec{u}_1} \cdots D_{\vec{u}_r} f$. We can interchange $D_{\vec{u}_j}$ with $D_{\vec{u}_{j+1}}$ by Theorem 6.2 and its corollaries above, and hence any arbitrary permutation of $D_{\vec{u}_i}$ can be achieved.

Exercise: For $f \in C^r(\mathbb{R}^m, \mathbb{R})$, find the number of distinct $D_{j_1} \cdots D_{j_r} f$.

Note: For open set $A \subseteq \mathbb{R}^m$ and the function defined by the following:

$$f =: A \rightarrow \mathbb{R}^n \quad \vec{a} \mapsto \begin{bmatrix} f_1(\vec{a}) \\ f_2(\vec{a}) \\ \vdots \\ f_n(\vec{a}) \end{bmatrix}$$

Here $f \in C^r(A, \mathbb{R}^n)$ if and only if for all $p \leq r$, $D_{j_1} D_{j_2} \cdots D_{j_p} f_r$ exists and is continuous. Also, for each of the $D_{j_1} D_{j_2} \cdots D_{j_p} f_r$, we can interchange the order of D_{j_t} by the discussion above, and rewrite as $D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_m^{\alpha_m} f_r$, where $\alpha_j \in \mathbb{N} \cup \{0\}$, $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_m \leq r$.

Note: Let V, W be normed vector spaces, and let A be an open subset of V , B be an open subset of W . Suppose the function $f : A \rightarrow B$ is differentiable at $\vec{a} \in A$, and the function $g : B \rightarrow A$ is differentiable at $\vec{b} = f(\vec{a}) \in B$, with the property that $g \circ f = Id_A : A \rightarrow A$ $\vec{x} \mapsto \vec{x}$. Then by the Chain Rule, we can write $Dg(\vec{b}) \cdot Df(\vec{a}) = I$, where I is the identity transformation, which implies $Dg(\vec{b})$ is a left inverse for $Df(\vec{a})$, and we must have $\dim(V) \leq \dim(W)$. Here, if we have $\dim(V) = \dim(W) < \infty$, then it is immediate that $Dg(\vec{b})$ is a two-sided inverse for $Df(\vec{a})$. Or, if we have $f \circ g = Id_B : B \rightarrow B$ $\vec{y} \mapsto \vec{y}$. Then $Dg(\vec{b})$ will also become a two-sided inverse for $Df(\vec{a})$, and we have $\dim(V) = \dim(W)$.

Proposition 6.2.4

Let V, W be normed vector spaces with $\dim(V) < \infty$ and $\dim(W) < \infty$, and let A be an open subset of V , B be an open subset of W . Suppose $f : A \rightarrow B$ and $g : B \rightarrow A$ are continuous, not necessarily differentiable, with the property that $g \circ f = Id_A$ and $f \circ g = Id_B$, then we have $\dim(V) = \dim(W)$.

[L. E. J. Brouwer, 1912]

Exercise: There exists a continuous surjection from \mathbb{R} to \mathbb{R}^2 .

Homeomorphism and Diffeomorphism

Definition 7.0.0.0.1

A homeomorphism is a continuous bijective function f whose inverse is also continuous.

Example: Let $S^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$.

Consider the function $f : [0, 2\pi) \rightarrow S^1 \quad t \mapsto (\cos(t), \sin(t))$. Notice here f is a bijection, but not a homeomorphism from $[0, 2\pi)$ to S^1 . This can be shown by considering the sequence (a_n) , where $a_i = (\cos(1/i), -\sin(1/i))$. (a_n) converges to $(1, 0)$. The sequence $(f^{-1}(a_n))$ converges to 2π , but we have $f^{-1}(-1, 0) = 0$, hence f^{-1} is not continuous.

Definition 7.0.0.0.2

Let A be an open subset of normed vector space V and let B be an open subset of normed vector space W . A function $f : A \rightarrow B$ is called a C^r -diffeomorphism provided that f is a bijection, $f \in C^r(A, B)$, and $f^{-1} \in C^r(B, A)$.

Example: $f : \mathbb{R} \rightarrow \mathbb{R} \quad t \mapsto t^3$ is a homeomorphism, but not a C^r -diffeomorphism.

Definition 7.0.0.0.3

A complete normed vector space is called a Banach space.

Example: Consider the set $A = \left\{ \begin{bmatrix} r \\ \theta \end{bmatrix} \in \mathbb{R}^2 \mid r > 0 \right\}$. Let $f : A \rightarrow \mathbb{R}^2 \quad \begin{bmatrix} r \\ \theta \end{bmatrix} \mapsto \begin{bmatrix} r \cos(\theta) \\ r \sin(\theta) \end{bmatrix}$.

Here we note that f is of C^∞ type, and we can write the following:

$$Df \left(\begin{bmatrix} r \\ \theta \end{bmatrix} \right) = \begin{bmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{bmatrix}$$

where we have:

$$\det \left(Df \left(\begin{bmatrix} r \\ \theta \end{bmatrix} \right) \right) = r$$

here we know that $Df \left(\begin{bmatrix} r \\ \theta \end{bmatrix} \right)$ is invertible for all $\begin{bmatrix} r \\ \theta \end{bmatrix} \in A$, but f is not injective on A .

Note that $f(A) = \mathbb{R}^2 \setminus \{0\}$, and there exists local C^∞ inverse for f , but not global inverse.

Definition 7.0.0.0.4

Let A, B be topological spaces, a function $f : A \rightarrow B$ is said to be open proved that $f(U)$ is open in B when U is open in A .

Note: If a function f is a bijection.

1. f is continuous if and only if f^{-1} is open.
2. f is open if and only if f^{-1} is continuous.
3. f is a homeomorphism if and only if f is continuous and open.

Theorem 7.1 (Inverse Function Theorem)

Let A be an open subset of \mathbb{R}^n , let $\vec{a} \in A$, let $f \in C^r(A, \mathbb{R}^n)$ with $r \geq 1$, and given $Df(\vec{a})$ is invertible. There exists an open neighborhood U of \vec{a} such that $f|_U$ is a C^r -diffeomorphism, that is, f maps U bijectively to some open set in \mathbb{R}^n , and $f^{-1} : f(U) \rightarrow U$ is of C^r type.

Note: $E = \{\vec{x} \in A \mid Df(\vec{x}) \text{ is invertible}\} = \{\vec{x} \in A \mid \det(Df(\vec{x})) \neq 0\}$ is an open set containing \vec{a} .

Note: The theorem does not assume that $E = A$, but we could assume such by replacing A by E .

Proof of the Inverse Function Theorem.

First we need some preliminaries:

1. Let $T_{\vec{a}}$ be a function such that $\vec{x} \mapsto \vec{x} + \vec{a}$
2. Let g denote the function $Df(\vec{a})^{-1} \circ T_{-f(\vec{a})} \circ f \circ T_{\vec{a}}$

From (2), we note that the following holds:

1. $g(\vec{0}) = \vec{0}$
2. $Dg(\vec{0}) = I$ where I is the identity transformation
3. $f = T_{f(\vec{a})} \circ Df(\vec{a}) \circ g \circ T_{\vec{a}}$

In the following, diffeomorphism refers to C^r -diffeomorphism. It suffices to show that g is a diffeomorphism on some open set containing $\vec{0}$. First we let $h = g - Id$, then we know that $Dh = Dg - I$, and $Dh(\vec{0}) = 0$, which implies that each $D_j h_k(\vec{0}) = 0$, $\forall 1 \leq j, k \leq n$. Fix $\mu > 0$, then there exists $\delta > 0$ such that for $\vec{x} \in B_\delta(\vec{0})$, we have each $|D_j h_k(\vec{x})| < \mu$. Let $\vec{u} \in B_\delta(\vec{0})$, consider $Dh(\vec{x})(\vec{u})$.

$$Dh(\vec{x})(\vec{u}) = \begin{bmatrix} \sum_j D_j h_1(\vec{x}) u_j \\ \sum_j D_j h_2(\vec{x}) u_j \\ \vdots \\ \sum_j D_j h_n(\vec{x}) u_j \end{bmatrix}$$

$$\|Dh(\vec{x})(\vec{u})\| = \left\| \begin{bmatrix} \sum_j D_j h_1(\vec{x}) u_j \\ \vdots \\ \sum_j D_j h_n(\vec{x}) u_j \end{bmatrix} \right\| \leq \sqrt{n} \left\| \begin{bmatrix} \sum_j D_j h_1(\vec{x}) u_j \\ \vdots \\ \sum_j D_j h_n(\vec{x}) u_j \end{bmatrix} \right\| \leq \sqrt{n} n \cdot \mu |\vec{u}| \leq n^{\frac{3}{2}} \mu \|\vec{u}\|$$

which implies the following holds:

$$\|Dh(\vec{x})\| \leq n^{\frac{3}{2}} \mu \quad \text{for } \vec{x} \in B_\epsilon(\vec{0})$$

Fix $0 < \epsilon < 1$, let $\mu = \frac{\epsilon}{n^{3/2}}$, then by previous argument, there exists $\delta > 0$ such that $\forall \vec{x} \in B_\delta(\vec{0})$, we have $\|Dh(\vec{x})\| < n^{3/2} \mu = \epsilon$. By using Lipschitz Estimate from Math 395 HW4, we get that $\|h(\vec{x}) - h(\vec{y})\| \leq \epsilon \|\vec{x} - \vec{y}\|$ for $\vec{x}, \vec{y} \in B_\delta(\vec{0})$. Then for $\vec{x}, \vec{y} \in B_\delta(\vec{0})$, we have:

$$\|g(\vec{x}) - g(\vec{y})\| = \|\vec{x} - \vec{y} + h(\vec{x}) - h(\vec{y})\| \leq \|\vec{x} - \vec{y}\| + \|h(\vec{x}) - h(\vec{y})\| \leq (1 + \epsilon) \|\vec{x} - \vec{y}\|$$

Also, we have the following by the reverse triangle inequality:

$$\|g(\vec{x}) - g(\vec{y})\| = \|\vec{x} - \vec{y} + h(\vec{x}) - h(\vec{y})\| \geq \|\vec{x} - \vec{y}\| - \|h(\vec{x}) - h(\vec{y})\| \geq (1 - \epsilon) \|\vec{x} - \vec{y}\|$$

because we have $0 < \epsilon < 1$, here we see that g is injective on $B_\delta(\vec{0})$.

Here $B_\delta(\vec{0})$ is a convex set. Let $\psi_y : \vec{x} \mapsto \vec{y} - h(\vec{x})$.

Pick $0 < \tilde{\delta} < \delta$ such that for $\vec{y} \in B_{(1-\epsilon)\tilde{\delta}}(\vec{0})$ and $\vec{x} \in \overline{B_{\tilde{\delta}}(\vec{0})}$, we have:

$$\|\psi_y(\vec{x})\| \leq (1 - \epsilon) \tilde{\delta} + \epsilon \tilde{\delta} = \tilde{\delta}$$

so the function $\psi_y : \overline{B_{\tilde{\delta}}(\vec{0})} \rightarrow \overline{B_{\tilde{\delta}}(\vec{0})}$ is well defined, with the following holds:

$$\|\psi_y(\vec{x}_1) - \psi_y(\vec{x}_2)\| \leq \|h(\vec{x}_2) - h(\vec{x}_1)\| \leq \epsilon \|\vec{x}_2 - \vec{x}_1\|$$

So ψ_y is a contraction on $\overline{B_{\tilde{\delta}}(\vec{0})}$ where $\overline{B_{\tilde{\delta}}(\vec{0})}$ is a complete metric space. By the Contraction Mapping Theorem, there exists a unique $\vec{x} \in \overline{B_{\tilde{\delta}}(\vec{0})}$ such that $\psi_y(\vec{x}) = \vec{y} - h(\vec{x}) = \vec{x}$, that is, $g(\vec{x}) = \vec{y}$. Take unions $\bigcup_{\tilde{\delta} < \delta} \overline{B_{\tilde{\delta}}(\vec{0})} = B_\delta(\vec{0})$ and $\bigcup_{\tilde{\delta} < \delta} B_{(1-\epsilon)\tilde{\delta}}(\vec{0}) = B_{(1-\epsilon)\delta}(\vec{0})$ to get $B_{(1-\epsilon)\delta}(\vec{0}) \subseteq g(B_\delta(\vec{0})) \subseteq g(A)$. Notice we get $\vec{0} \in \text{Int}(g(A))$.

We conclude the results that we shown above as a lemma:

Lemma 7.1.1

Let A be an open subset of \mathbb{R}^n , let $\vec{a} \in A$, let $f \in C^r(A, \mathbb{R}^n)$ with $r \geq 1$, and given $Df(\vec{a})$ is invertible. Let $T_{\vec{a}}$ be a function that maps \vec{x} in its domain to $\vec{x} + \vec{a}$ in its codomian, with some \vec{a} in its domain. Let g denote the function $Df(\vec{a})^{-1} \circ T_{-f(\vec{a})} \circ f \circ T_{\vec{a}}$, and let $h = g - Id$ where Id is the identity function. We have the followings hold:

1. For $\mu > 0$, there exists $\delta > 0$ such that for all $\vec{x} \in B_\delta(\vec{0})$, we have $\|Dh(\vec{x})\| \leq n^{3/2} \mu$.
2. For $0 < \epsilon < 1$, $\exists \delta > 0$ s.t. $\forall \vec{x}, \vec{y} \in B_\delta(\vec{0})$, $\|h(\vec{x}) - h(\vec{y})\| \leq \epsilon \|\vec{x} - \vec{y}\|$
3. For $0 < \epsilon < 1$, $\exists \delta > 0$ s.t. $\forall \vec{x}, \vec{y} \in B_\delta(\vec{0})$, $(1 - \epsilon) \|\vec{x} - \vec{y}\| \leq \|g(\vec{x}) - g(\vec{y})\| \leq (1 + \epsilon) \|\vec{x} - \vec{y}\|$
4. For $0 < \epsilon < 1$, $\exists \delta > 0$ s.t. g is injective on $B_\delta(\vec{0})$.
5. For $0 < \epsilon < 1$, $\exists \delta > 0$ s.t. $B_{(1-\epsilon)\delta}(\vec{0}) \subseteq g(B_\delta(\vec{0})) \subseteq g(A)$.
6. $\vec{0} \in \text{Int}(g(A))$.

To proceed proving the Inverse Function Theorem, we will make use of the following Theorem:

Theorem 7.2

Given E as a open subset of \mathbb{R}^n , $f \in C^1(E, \mathbb{R}^n)$, and $\det(Df(\vec{x})) \neq 0$ for all $\vec{x} \in E$.

Then we have the followings hold:

1. $\vec{a} \in E$ implies $f(\vec{a}) \in \text{Int}(f(E))$.
2. $f(E)$ is open in \mathbb{R}^n .
3. $f : E \rightarrow f(E)$ is an open map.

Proof of Theorem 7.2. Fix $\vec{a} \in E$, apply Lemma 7.1.1 to $g = Df(\vec{a})^{-1} \circ T_{-f(\vec{a})} \circ f \circ T_{\vec{a}}$ defined on the set $T_{-\vec{a}}(E)$. Notice here $g(\vec{0}) = \vec{0}$, and $Dg(\vec{0}) = I$, and we have $\vec{0} \in \text{Int}(\text{Im}(g))$. Consider $Df(\vec{a}) \circ g$, we see that $\vec{0} \in \text{Int}(\text{Im}(Df(\vec{a}) \circ g)) = \text{Int}(\text{Im}(T_{-f(\vec{a})} \circ f \circ T_{\vec{a}}))$. Then we consider $T_{f(\vec{a})} \circ Df(\vec{a}) \circ g$, we have $f(\vec{a}) \in \text{Int}(\text{Im}(f \circ T_{\vec{a}})) = \text{Int}(f(E))$, which implies (1) holds, and (2) follows immediately as $f(E)$ contains all its interior points. For (3), let U be an open subset of E , apply (2) to $f|_U$ and the result follows. \square

Corollary 7.2.1

Given E as a open subset of \mathbb{R}^n , $f \in C^1(E, \mathbb{R}^n)$, and $\det(Df(\vec{x})) \neq 0$ for all $\vec{x} \in E$.

If f is injective, then $f : E \rightarrow f(E)$ is a homeomorphism.

Using Lemma 7.1.1, we know that for $0 < \epsilon < 1$, $\exists \delta > 0$ such that g is injective on $B_\delta(\vec{0})$. Since $Df(\vec{a})$ is invertible, then $Df(\vec{a})$ is injective. So we know that the composition of injective functions $f|_{B_\delta(\vec{a})} = T_{f(\vec{a})} \circ Df(\vec{a}) \circ g|_{B_\delta(\vec{0})} \circ T_{-\vec{a}}$ is injective, then by Corollary 7.2.1, let $U = E \cap B_\delta(\vec{a})$, where E is the open neighborhood of \vec{a} in which $\det(Df(\vec{t})) \neq 0$ for all $\vec{t} \in E$, then we know that $f|_U : U \rightarrow f(U)$ defines a homeomorphism. On the other hand, we know that for all $\vec{x}, \vec{y} \in U = B_\delta(\vec{0})$, we have $(1 - \epsilon)\|\vec{x} - \vec{y}\| \leq \|g(\vec{x}) - g(\vec{y})\| \leq (1 + \epsilon)\|\vec{x} - \vec{y}\|$, so we know that g is bi-Lipschitz from $B_\delta(\vec{0})$ to $g(B_\delta(\vec{0}))$. It follows that $f|_U$ is bi-Lipschitz as the composition of bi-Lipschitz functions is bi-Lipschitz. To show that $f|_U$ is a diffeomorphism, we will make use of the following Proposition:

Proposition 7.2.2

Given E as a open subset of \mathbb{R}^n , $f \in C^1(E, \mathbb{R}^n)$ is an injective function, with $\det(Df(\vec{x})) \neq 0$ for all $\vec{x} \in E$. Then we have $f \in C^r$ implies $f^{-1} \in C^r$.

Proof. It suffices to show that $\psi := f^{-1}$ is C^1 . Fix $\vec{a} \in E$. Let $\vec{b} \in f(\vec{a})$, and let $M := Df(\vec{a})$. To show that ψ is differentiable, we want to find some T such that the following holds:

$$\lim_{\vec{k} \rightarrow \vec{0}} \frac{\psi(\vec{b} + \vec{k}) - \psi(\vec{b}) - T\vec{k}}{\|\vec{k}\|} = 0 \quad (1)$$

Let $T = M^{-1}$, $\vec{h} = \psi(\vec{b} - \vec{a}) - \psi(\vec{b})$, then we can write the following:

$$\begin{aligned} \frac{\psi(\vec{b} + \vec{k}) - \psi(\vec{b}) - M^{-1}\vec{k}}{\|\vec{k}\|} &= \frac{\vec{h} - M^{-1}\vec{k}}{\|\vec{k}\|} \\ &= -M^{-1} \left(\frac{\vec{h} - M\vec{k}}{\|\vec{k}\|} \right) \\ &= -M^{-1} \left(\frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - M\vec{h}}{\|\vec{h}\|} \right) \left(\frac{\|\vec{h}\|}{\|\vec{k}\|} \right) \\ &:= -M^{-1}(*)(**) \end{aligned}$$

where $(**)$ is bounded for small $\|\vec{k}\|$ by Lemma 7.1.1 (3), and we have $\vec{h} \rightarrow \vec{0}$ as $\vec{k} \rightarrow \vec{0}$. While $(*)$ tends to 0 as f is differentiable, we know that equation (1) must hold. Now we have shown that ψ is differentiable and we have $D\psi(\vec{b}) = [Df(\psi(\vec{b}))]^{-1}$. Lastly, we will show that $D\psi : f(E) \rightarrow \text{Mat}(n, n, \mathbb{R})$ is continuous. Here we have the following holds:

$$f(E) \xrightarrow[\text{continuous}]{\psi} E \xrightarrow[\text{continuous}]{Df} GL_n(\mathbb{R}) \xrightarrow[\text{continuous}]{\text{Inverse}} GL_n(\mathbb{R})$$

with $D\psi = \text{Inverse} \circ Df \circ \phi$, here *Inverse* is the function that sends an invertible matrix $M \in GL_n(\mathbb{R})$ to its inverse $M^{-1} \in GL_n(\mathbb{R})$. *Inverse* is a continuous function by Cramer's Rule and Theorem 2.14 on Munkres. We know that the composition of continuous functions is continuous, hence $D\psi$ is continuous. This finishes the proof for $r = 1$. For $r > 1$, we can proceed by induction. Assume the result is known for C^{r-1} , where ψ and Df are of C^{r-1} type, and we know that C^{r-1} is closed under composition, then the result follows immediately. \square

Here we see that the results for the Inverse Function Theorem follows from Proposition 7.2.2 and Corollary 7.2.1. This completes the proof for the Inverse Function Theorem. \square

For V, W, Z be normed vector spaces, and let A be a open subset of V , B be a subset of W . Consider function $f : A \rightarrow B$ and $g : B \rightarrow Z$, then we can write the chain rule $D(g \circ f) = ((Dg) \circ f) \circ Df$. If we have stronger condition where $f, g \in C^1$, then Dg, f, Df are all continuous, hence we know that the composition of continuous functions $D(g \circ f) \in C^1$. If we have $f \in C^2$, then $D^2f : A \rightarrow B(V, B(V, W))$. Now suppose $f, g \in C^2$, then $Df, Dg \in C^1$, and hence $D(g \circ f) \in C^1$ by the Chain Rule and composition of C^1 function, we have $g \circ f \in C^2$.

Note: Consider an affine function f which sends \vec{x} to $T\vec{x} + b$, where T is a linear transformation. We have $Df(\vec{x}) = T$, and $D^2f(\vec{x}) = 0$, because T is a "constant" in the space $B(V, W)$.

Let V be a normed vector space, and let A be an open subset of V .

Exercise: Let $f, g \in C^r(A, \mathbb{R})$, then we have $g \circ f \in C^r$.

Exercise: Let $f \in C^r(A, \mathbb{R} \setminus \{0\})$, then we have $\frac{1}{f} \in C^r$.

Exercise: Let $f, g \in C^r(A, \mathbb{R})$, then we have $f + g \in C^r$, and $fg \in C^r$.

Leibniz's Notation in Several Variables

Suppose we have five quantities, x, y, u, v, t , such that any two of these variables determine the other three. Consider the following:

$$f : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} u \\ v \end{bmatrix}$$

where we define

$$\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} := Df = \begin{bmatrix} D_1 f_1 & D_2 f_2 \\ D_1 f_2 & D_2 f_2 \end{bmatrix}$$

Consider another function defined by the following:

$$g : \begin{bmatrix} u \\ v \end{bmatrix} \mapsto t$$

and we define:

$$\begin{bmatrix} \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix} := Dg = \begin{bmatrix} D_1 g & D_2 g \end{bmatrix}$$

Then we write the following:

$$\begin{bmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{bmatrix} = D(g \circ f) = (Dg \circ f) \circ Df = \begin{bmatrix} \frac{\partial t}{\partial u} & \frac{\partial t}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial t}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial t}{\partial v} \frac{\partial v}{\partial x} & \frac{\partial t}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial t}{\partial v} \frac{\partial v}{\partial y} \end{bmatrix}$$

Notice here, one disadvantage of using the Leibniz's Notation is that the notation " $\frac{\partial t}{\partial u}$ " does not specify what quantity is being held "fixed" when taking the derivative.

Implicit Function Theorem

Consider $M \in \text{Mat}(n, k+n, \mathbb{R})$, $\vec{c} \in \mathbb{R}^n$, let $E = \{\vec{v} \in \mathbb{R}^{k+n} \mid M\vec{v} = \vec{c}\} = \{\vec{v} \in \mathbb{R}^{k+n} \mid M\vec{v} - \vec{c} = \vec{0}\}$. Now define $T : \mathbb{R}^{k+n} \rightarrow \mathbb{R}^n$ $\vec{v} \mapsto M\vec{v} - \vec{c}$. Here T is an affine. Let $\hat{M} : \mathbb{R}^{k+n} \rightarrow \mathbb{R}^n$ $\vec{v} \mapsto M\vec{v}$. Note here E could be empty, if E is non-empty, then E is an affine set because $E = \ker(\hat{M}) + \vec{v}_p$ where \vec{v}_p is any particular solution of $M\vec{v}_p = \vec{c}$. We also have $\dim(E) = \dim(\ker(\hat{M})) = k+n - \text{rank}(M)$ by the Rank-Nullity Theorem. Here we have $k \leq \dim(E) \leq k+n$.

In the following discussion, we will investigate if we can write E as $\text{Graph}(g)$ for some $g : \mathbb{R}^k \rightarrow \mathbb{R}^n$. If such is true, then g has to be affine. Consider the following:

$$M = [X \ Y] \text{ where } X \text{ is an } n \times k \text{ matrix, and } Y \text{ is an } n \times n \text{ matrix}$$

For $\vec{x} \in \mathbb{R}^k$ and $\vec{y} \in \mathbb{R}^n$, we can write the following:

$$M \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} = X\vec{x} + Y\vec{y}$$

Case (1): Y is non-invertible, then there exists $\vec{w} \in \mathbb{R}^n$ such that $Y\vec{w} = \vec{0}$ with $\vec{w} \neq \vec{0}$. Then for $(\vec{x}, \vec{y}) \in E$, we have the following holds:

$$\begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \in E \Rightarrow \begin{bmatrix} \vec{x} \\ \vec{y} + \vec{w} \end{bmatrix} \in E$$

Case (2): Y is invertible, then we have the following holds:

$$M \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} = \vec{c} \iff Y^{-1}X\vec{x} + \vec{y} = Y^{-1}\vec{c} \iff \vec{y} = -Y^{-1}X\vec{x} + Y^{-1}\vec{c} := T(\vec{x})$$

Here we see that T is an affine function.

Consider $\text{Proj}_{k+1, \dots, n} = [Z \ I]$ where Z is an $n \times k$ zero matrix and I is an $n \times n$ identity matrix.

$$\text{Proj}_{k+1, \dots, n} : \mathbb{R}^{k+n} \rightarrow \mathbb{R}^n \quad \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \mapsto \vec{y}$$

Note here we have $Y = \text{Proj}_{k+1, \dots, n} \cdot M$.

Here, one might want a different set of k independent vectors on the graph.

Let $\hat{P} : \mathbb{R}^{k+n} \rightarrow \mathbb{R}^{k+n}$ $\vec{x} \mapsto P\vec{x}$ be a coordinate permutation, where we have:

$$P = \begin{bmatrix} \vec{e}_{j_1} \\ \vec{e}_{j_2} \\ \vdots \\ \vec{e}_{j_{k+n}} \end{bmatrix}$$

here we view each standard basis vector \vec{e}_{j_i} as a row vector.

The image $\hat{P}(\text{Graph}(g))$ is a coordinate permuted graph with k independent variables.

$$\begin{aligned} \dim(E) = k &\iff \text{rank}(M) = n \iff \text{some } n \times n \text{ submatrix of } M \text{ is invertible} \\ &\iff E \text{ is permuted graph of an affine function with } k \text{ independent variables} \end{aligned}$$

Consider a general vector in \mathbb{R}^{k+n} defined by the following:

$$\vec{G} = \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \quad \vec{x} \in \mathbb{R}^k, \vec{y} \in \mathbb{R}^n$$

and a special vector in \mathbb{R}^{k+n} defined by the following:

$$\vec{S} = \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} \quad \vec{a} \in \mathbb{R}^k, \vec{b} \in \mathbb{R}^n$$

The Implicit Function Theorem states that, for $f \in C^r(A, \mathbb{R}^n)$, where A is an open subset of \mathbb{R}^{k+n} , and given that the following holds:

$$\begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} \in f^{-1}(\vec{0}) := E \quad \frac{\partial f}{\partial \vec{y}} \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} \text{ is invertible}$$

Then there exists an open neighborhood U of \vec{S} such that $E \cap U = \text{Graph}(g)$ for some $g \in C^r(B, \mathbb{R}^n)$ where we have $\vec{a} \in B$ and B is an open subset of \mathbb{R}^k . That is, we can solve for \vec{y} in terms of \vec{x} near \vec{S} .

Theorem 8.1 (Implicit Function Theorem)

Let \vec{G} in \mathbb{R}^{k+n} defined by the following:

$$\vec{G} = \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \quad \vec{x} \in \mathbb{R}^k, \vec{y} \in \mathbb{R}^n$$

and fix \vec{S} in \mathbb{R}^{k+n} by the following:

$$\vec{S} = \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} \quad \vec{a} \in \mathbb{R}^k, \vec{b} \in \mathbb{R}^n$$

For $f \in C^r(A, \mathbb{R}^n)$, where A is an open subset of \mathbb{R}^{k+n} . If we have the followings hold:

$$\vec{S} \in f^{-1}(\vec{0}) := E \quad \left[\frac{\partial f}{\partial \vec{y}} \vec{S} \right] \text{ is invertible}$$

Then there exists a neighborhood U of \vec{S} such that $E \cap U = \text{Graph}(g)$ for a unique function $g \in C^r(B, \mathbb{R}^n)$, where $\vec{a} \in B$, and B is an open subset of \mathbb{R}^k . In other words, \exists an open neighborhood B of \vec{x} such that $\vec{y} = g(\vec{x})$ and $f(\vec{x}, g(\vec{x})) = \vec{0}$ for all $\vec{x} \in B$, with a unique function $g \in C^r(B, \mathbb{R}^n)$.

Proof. We will construct a function F to which we can apply the Inverse Function Theorem:

$$F : A \rightarrow \mathbb{R}^{k+n} \quad \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \mapsto \begin{bmatrix} \vec{x} \\ f \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \end{bmatrix} \quad \text{for } \vec{x} \in \mathbb{R}^k, \vec{y} \in \mathbb{R}^n$$

Here we note that $f \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \in \mathbb{R}^n$, and $F \in C^r(A, \mathbb{R}^{k+n})$ with:

$$DF = \begin{bmatrix} I & Z \\ \frac{\partial f}{\partial \vec{x}} & \frac{\partial f}{\partial \vec{y}} \end{bmatrix}$$

Here I is the $k \times k$ identity matrix, Z is the $k \times n$ zero matrix. $\frac{\partial f}{\partial \vec{x}}$ is an $n \times k$ matrix, and $\frac{\partial f}{\partial \vec{y}}$ is an $n \times n$ matrix. Note that $DF = \det I \cdot \det \frac{\partial f}{\partial \vec{y}} \neq 0$ by Munkres Lemma 2.12. Hence we know that the DF is invertible when evaluated at \vec{S} . Here the inverse of DF is given by the following:

$$(DF)^{-1} = \begin{bmatrix} I & Z \\ -\left(\frac{\partial f}{\partial \vec{x}}\right)^{-1} \frac{\partial f}{\partial \vec{x}} & \left(\frac{\partial f}{\partial \vec{y}}\right)^{-1} \end{bmatrix}$$

By the Inverse Function Theorem, we know that F has local inverse of C^1 type, then we can write the following with a function h of C^r mapping an open set $U \subseteq \mathbb{R}^{k+n}$ to \mathbb{R}^n :

$$F^{-1} : W \rightarrow \mathbb{R}^n \quad \begin{bmatrix} \vec{x} \\ \vec{z} \end{bmatrix} \mapsto \begin{bmatrix} \vec{x} \\ h \begin{bmatrix} \vec{x} \\ \vec{z} \end{bmatrix} \end{bmatrix} \quad \text{for } \vec{x} \in \mathbb{R}^k, \vec{z} \in \mathbb{R}^n$$

Near \vec{S} , we have the following holds:

$$\begin{aligned} \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \in E &\iff f \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} = \vec{0} \iff F \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} = \begin{bmatrix} \vec{x} \\ \vec{0} \end{bmatrix} \iff \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} = F^{-1} \begin{bmatrix} \vec{x} \\ \vec{0} \end{bmatrix} = \begin{bmatrix} \vec{x} \\ h \begin{bmatrix} \vec{x} \\ \vec{0} \end{bmatrix} \end{bmatrix} \\ &\iff \vec{y} = h \begin{bmatrix} \vec{x} \\ \vec{0} \end{bmatrix} := g(\vec{x}) \end{aligned}$$

One can show that g is unique and the result follows. \square

More details of this proof can be found on Munkres Theorem 9.2.

In the proof of the Implicit Function Theorem:

Remark: F gives "change of coordinates" that locally straightens all levels set of f .

Remark: E is locally a coordinate permuted graph if $\text{rank}(Df(\vec{S})) = n$.

Optimization

Theorem 9.1 (First Derivative Test for Higher Dimensions)

Let Ω be an open subset of \mathbb{R}^n , and let $h : \Omega \rightarrow \mathbb{R}$ be a function that achieves a local maximum, or minimum, at $\vec{p} \in \Omega$, then $Dh(\vec{p}) = 0$.

Proof. Suppose we have $h \in C^1(\Omega, \mathbb{R})$ where Ω is an open subset of \mathbb{R}^n , with the property that h achieves a maximum, or a minimum, at $\vec{p} \in \Omega$. Consider arbitrary $\gamma \in C^1(A, \Omega)$ with $\gamma(0) = \vec{p}$, where A is an open subset of \mathbb{R} . Then $h \circ \gamma$ has a maximum, or a minimum, at $t = 0$. So we have $0 = (h \circ \gamma)'(0) = Dh(\gamma(0))\gamma'(0)$. Since this is true for arbitrary γ that satisfies $\gamma(0) = \vec{p}$, here we can conclude that $Dh(\gamma(0)) = Dh(\vec{p}) = \vec{0}$, and hence we can write:

$$D_1 h(\vec{p}) = D_2 h(\vec{p}) = \cdots = D_n h(\vec{p}) = 0$$

Now assume h has global maximum, or global minimum, at \vec{p} . Then for an open subset U that contains \vec{p} , $h|_U$ has a local maximum, or a local minimum, and we can reach the same conclusion that $Dh(\vec{p}) = \vec{0}$ with the same argument. That is, $Dh(\vec{p}) = 0$ holds true regardless whether the maximum, or minimum, is global or local. \square

Theorem 9.2 (Lagrange Multiplier Theorem)

Let U be an open subset of \mathbb{R}^{k+n} , let the constraint function be $f \in C^1(U, \mathbb{R}^n)$ with $E = f^{-1}(\vec{0})$, let the objective function be $h \in C^1(U, \mathbb{R})$, with the property that $h|_E$ has a local maximum, or a local minimum, at $\vec{p} \in E$. Given $\text{rank}(Df(\vec{p})) = n$, we have $Dh(\vec{p})$ belongs to the row space of $Df(\vec{p})$, that is, we can write $Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p})$ for $\lambda_j \in \mathbb{R}$.

Note: The goal of Theorem 9.2 is to maximize, or minimize, the value of the objective function h restricted on the kernel of the constraint function f . That is, we want to find $\vec{t} \in \mathbb{R}^{n+k}$ such that $h(\vec{t})$ is maximized, or minimized, whenever $f(\vec{t}) = \vec{0}$ holds. We call such maximum, or minimum, point as $\vec{p} \in \mathbb{R}^{n+k}$. If we have $\text{rank}(Df(\vec{p})) = n$, then we can find such \vec{p} with the equation $Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p})$ for $\lambda_j \in \mathbb{R}$.

Proof of Theorem 9.2. Here we will show the general idea of proving Theorem 9.2, a more general version of this proof is given on Theorem E.2 from Supplement to Munkres. Here we want to maximize, or minimize, h while applying the constraints given by f and its kernel E . Consider $\gamma \in C^1(A, E)$, where A is an open subset of \mathbb{R} , with $\gamma(0) = \vec{p}$. By Theorem 9.1, we have the following:

$$Dh(\vec{p}) \cdot \gamma'(0) = 0 \tag{1}$$

Notice here, we have $f \circ \gamma = \vec{0}$ by construction. Then $Df(\gamma(t)) \cdot \gamma'(t) = 0$, and hence we have:

$$\gamma'(0) \in \ker(Df(\vec{p})) \tag{2}$$

where $Df(\vec{p}) \in \text{hom}(\mathbb{R}^{k+n}, \mathbb{R}^n)$. From Math 395 HW6, we will show that, if $Df(\vec{p})$ has maximal rank n , then there are no other constraints applied to $\gamma'(0)$. Then we can combine equation (1) and (2) to get the following:

$$\begin{aligned} Dh(\vec{p}) &\in (\ker(Df(\vec{p})))^\perp = ((\text{row space of } Df(\vec{p}))^\perp)^\perp \\ &= (\text{row space of } Df(\vec{p})) \\ &= \text{span}\{Df_1(\vec{p}), \dots, Df_n(\vec{p})\} \end{aligned}$$

Here we get:

$$Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \lambda_2 Df_2(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p}) \tag{3}$$

where λ_i are scalars called the Lagrange multipliers. \square

Remark: From the settings of Theorem 9.2, let f_i denote the component functions of f . When maximizing, or minimizing, h with the constraints from f , we require $f_i(\vec{p}) = 0$ for $1 \leq i \leq n$. By equation (3), in order to have to columns of $Dh(\vec{p})$ to match the RHS of the equation (3), we get a set of $n+k$ equations, as we require $D_i h(\vec{p}) = \sum_{j=1}^n D_i f_j(\vec{p})$ for $1 \leq i \leq n+k$. In total, we will need to solve $k+2n$ unknowns, $k+n$ coordinate components in \mathbb{R}^{k+n} and the n Lagrange multipliers, using the $k+2n$ equations. Here Theorem 9.2 provides us a possible way to solve such optimization problem.

Note: Suppose K is an compact subset of \mathbb{R}^n , let $h : K \rightarrow \mathbb{R}$ be a continuous function, then by the Extreme Value Theorem, we know that h has a maximum and a minimum on K . By Theorem 9.1, we might want to check the point $\vec{p} \in K$ that satisfies one of the followings to find the local maximum, or local minimum, of h :

1. $\vec{p} \in \text{Int}(K)$ and $Dh(\vec{p}) = \vec{0}$
2. $\vec{p} \in \text{Int}(K)$ and h is not differentiable at \vec{p}
3. $\vec{p} \in \text{Bd}(K)$

Example: Let the objective function $h(x, y) = x^4 + y^6$ be defined on $K = \{(x, y) \mid x^2 + y^2 \leq 1\}$. We observe that $Dh = [4x^3 \ 6y^5]$, so $Dh(\vec{x}) = \vec{0}$ only if $\vec{x} = \vec{0}$, and we deduce that h has a minimum at $\vec{0}$, and a maximum somewhere on $\text{Bd}(K) := E$. Now we want to find such maximum on E . Let the constraint function $f(x, y) = x^2 + y^2 - 1$ be defined on K , and we have $Df = [2x \ 2y]$. Notice here, for all $x, y \in \mathbb{R}$ that satisfy $f(x, y) = \vec{0}$, $(x, y) \in \text{Bd}(K)$. Now we get the following equations by Theorem 9.2:

$$\begin{cases} 4x^3 = \lambda 2x & \text{if } x = 0 \Rightarrow y = \pm 1 \Rightarrow h = 1 \\ 6y^5 = \lambda 2y & \text{if } y = 0 \Rightarrow x = \pm 1 \Rightarrow h = 1 \\ x^2 + y^2 = 1 & \Rightarrow \lambda/2 + \sqrt{\lambda/3} = 1 \end{cases}$$

Here λ is the Lagrange multiplier. From such system of equations, we solve that $\lambda = \frac{2}{3}(4 - \sqrt{7})$. Check all the critical points, if we have $x = \pm\sqrt{\lambda/2}$, and $y = \pm\sqrt[4]{\lambda/3}$, we see that $h(x, y) = 0.368$ gives the minimum of h along the boundary $\text{Bd}(K)$. On the other hand, when $x = 0, y = \pm 1$, or $y = 0, x = \pm 1$, we have $h(x, y) = 1$, hence $\max\{\text{im}(h|_E)\} = 1$.

Lemma 9.2.1

Let X be a closed subset of \mathbb{R}^n , let $\vec{p} \in \mathbb{R}^n$. Then $f : X \rightarrow \mathbb{R}$ $\|\vec{x} - \vec{p}\|$ has a minimum.

The proof of Lemma 9.2.1 is given on Math 395 HW5 Q5.

Example: We want to investigate which points in the set $\{(x, y, z) \mid xyz = 1\}$ lie closest to $\vec{0}$. Let constraint function be $f(x, y, z) = xyz - 1$, and let objective function be $h(x, y, z) = x^2 + y^2 + z^2$. By Theorem 9.2, we can look for solutions of $f(x, y, z) = 0$, with the following holds:

$$[2x \ 2y \ 2z] = Dh = \lambda Df = \lambda[yz \ xz \ xy] = [\lambda yz \ \lambda xz \ \lambda xy]$$

Here we get the following set of equations:

$$\begin{cases} 2x = \lambda yz \Rightarrow 2x^2 = \lambda \\ 2y = \lambda xz \Rightarrow 2y^2 = \lambda \\ 2z = \lambda xy \Rightarrow 2z^2 = \lambda \\ xyz = 1 \Rightarrow 3\sqrt{\lambda/2} = 1 \end{cases}$$

We solve that $\lambda = 2$, and $(x, y, z) = (\pm 1, \pm 1, \pm 1)$ with even numbers of minus sign, so there is a total of four (x, y, z) solutions. All these solutions are the points closest to $\vec{0}$.

Example: Let $B \in \text{Mat}(n, n, \mathbb{R})$ be a symmetric matrix, that is, we have $B = B^T$, let $h(\vec{x}) = \vec{x}^T B \vec{x}$. We want to maximize h on the unit sphere $\{\vec{x} \mid \|\vec{x}\| = 1\}$. Let $f(\vec{x}) = \|\vec{x}\|^2 - 1$. One might check that $Df(\vec{x})$ has rank 1 when we have $\|\vec{x}\|^2 = 1$, and $Df(\vec{x}) = 2\vec{x}^T$. By Theorem 9.2, there exists a solution of this maximization problem when we have $Dh(\vec{x}) = \lambda Df(\vec{x})$. Here we have:

$$h(\vec{x}) = \sum_{j,k} b_{j,k} x_j x_k \Rightarrow Dh(\vec{x}) = \sum_k b_{m,k} x_k + \sum_j b_{j,m} x_j = 2 \sum_j b_{j,m} x_j = (2\vec{x}^T B)_m$$

It follows that $Dh(\vec{x}) = 2\vec{x}^T B$. Note that one might also write the following to show $Dh(\vec{x}) = 2\vec{x}^T B$:

$$Dh(\vec{x})(\vec{u}) = h'(\vec{x}; \vec{u}) = \vec{u}^T B \vec{x} + \vec{x}^T B \vec{u} = 2\vec{x}^T B \vec{u}$$

Then by Theorem 9.1, we can write the following:

$$\begin{aligned} Dh(\vec{x}) &= \lambda Df(\vec{x}) \\ 2\vec{x}^T B &= 2\lambda \vec{x}^T \\ B \vec{x} &= \lambda \vec{x} \end{aligned}$$

Here we get a system of equations:

$$\begin{cases} B \vec{x} = \lambda \vec{x} \\ \|\vec{x}\| = \vec{x}^T \vec{x} = 1 \end{cases} \tag{S1}$$

Note here, we are trying to find the eigenvalues and the eigenvectors of unit length for the matrix B . Moreover, since the unit sphere $S^{n-1} = \{\vec{x} \mid \|\vec{x}\| = 1\}$ is compact, and we see that h is a continuous function, then we know that $h(S^{n-1})$ is compact and hence $h(S^{n-1})$ admits a maximum value. Combining with Theorem 9.2, we know that there exists at least one solution to the system (S1). That is, there exists at least one eigenvalue μ_1 for B with its eigenvector \vec{x}_1 , such that $h(\vec{x}_1) = \vec{x}_1^T B \vec{x}_1 = \vec{x}_1^T \mu \vec{x}_1 = \mu \vec{x}_1^T \vec{x} = \mu$ is a maximum of h over the unit sphere S^{n-1} . Now we would like to maximize h over $S^{n-1} \cap (\text{span}\{\vec{x}_1\})^\perp$, we define $f_{\vec{x}_1}$ by the following:

$$f_{\vec{x}_1}(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{bmatrix} = \begin{bmatrix} \|\vec{x}\|^2 - 1 \\ \vec{x}_1^T \vec{x} \end{bmatrix} = \begin{bmatrix} \vec{x}^T \vec{x} - 1 \\ \vec{x}_1^T \vec{x} \end{bmatrix}$$

Notice here we have: $f_{\vec{x}_1}^{-1}(\vec{0}) = S^{n-1} \cap (\text{span}\{\vec{x}_1\})^\perp$. By Theorem 9.2, if such local maximum of h under the constraint from f exists, then the following system of equations must holds:

$$\begin{cases} \vec{x}^T \cdot \vec{x} = 1 \\ \vec{x}_1^T \cdot \vec{x} = 0 \\ 2\vec{x}^T B = 2\lambda_1 \vec{x}^T + \lambda_2 \vec{x}_1^T \end{cases} \quad (\text{S2})$$

Here we see that:

$$2\vec{x}^T B = 2\lambda_1 \vec{x}^T + \lambda_2 \vec{x}_1^T \Rightarrow 2\vec{x}^T B \vec{x}_1 = \lambda_2 \Rightarrow 2\mu \vec{x}^T \vec{x}_1 = 0 = \lambda_2$$

So $\lambda_2 = 0$, $\vec{x}^T B = \lambda_1 \vec{x}^T \Rightarrow B \vec{x} = \lambda_1 \vec{x}$. Rewriting system (S2) we get the following:

$$\begin{cases} \vec{x}^T \cdot \vec{x} = 1 \\ \vec{x}_1^T \cdot \vec{x} = 0 \\ B \vec{x} = \lambda_1 \vec{x} \end{cases} \quad (\text{S3})$$

Since the orthogonal complement of a subset of a vector space is always closed, then we know that $S^{n-1} \cap (\text{span}\{\vec{x}_1\})^\perp$ is closed and bounded, and hence it is compact, then $h(S^{n-1} \cap (\text{span}\{\vec{x}_1\})^\perp)$ is also compact and admits a maximum value. So we know that the system S3 has at least one solution for λ_1 and \vec{x} , here we get the second eigenvalue $\mu_2 = \lambda_1$, with eigenvector $\vec{x}_2 \in S^{n-1} \cap (\text{span}\{\vec{x}_1\})^\perp$, for the symmetric matrix B . One might proceed using induction to show that B admits an orthonormal basis of eigenvectors $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$, with $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$. This example proposes a proof for the Spectral Theorem, we conclude our results as the followings:

Lemma 9.2.2

Every symmetric matrix has at least one real eigenvalue.

Theorem 9.3 (Spectral Theorem)

Every symmetric square matrix admits an orthonormal basis of eigenvectors. The corresponding eigenvalues are all real.

Exercise:

For a symmetric $n \times n$ matrix B , from the previous exercise, B admits a set of maximized real eigenvalues $\{\mu_1, \mu_2, \dots, \mu_n\}$, with their corresponding eigenvectors $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$. Let $h(\vec{x}) = \vec{x}^T B \vec{x}$, we have $h(c_1 \vec{x}_1 + \dots + c_n \vec{x}_n) = c_1^2 \mu_1 + \dots + c_n^2 \mu_n$, with the following statements hold:

- (1) All $\mu_j \geq 0$ if and only if $\vec{x}^T B \vec{x} \geq 0 \forall \vec{x} \in \mathbb{R}^n$.
- (2) All $\mu_j > 0$ if and only if $\vec{x}^T B \vec{x} > 0 \forall \vec{x} \neq \vec{0} \in \mathbb{R}^n$.
- (3) All $\mu_j \leq 0$ if and only if $\vec{x}^T B \vec{x} \leq 0 \forall \vec{x} \in \mathbb{R}^n$.
- (4) All $\mu_j < 0$ if and only if $\vec{x}^T B \vec{x} < 0 \forall \vec{x} \neq \vec{0} \in \mathbb{R}^n$.

Definition 9.3.0.0.1

For a symmetric $n \times n$ matrix B , B admits a set of maximized real eigenvalues $\{\mu_1, \mu_2, \dots, \mu_n\}$.

1. If $\mu_j \geq 0$ for all $1 \leq j \leq n$, then B is said to be positive semi-definite.
2. If $\mu_j > 0$ for all $1 \leq j \leq n$, then B is said to be positive definite.
3. If $\mu_j \leq 0$ for all $1 \leq j \leq n$, then B is said to be negative semi-definite.
4. If $\mu_j < 0$ for all $1 \leq j \leq n$, then B is said to be negative definite.

Definition 9.3.0.0.2

Let f be a C^2 type function defined in a neighborhood of $\vec{x} \in \mathbb{R}^n$. The Hessian $Hf(\vec{x})$ of f at \vec{x} is the $n \times n$ matrix whose entry at i -th row, j -th column is given by $D_k D_j f(\vec{x})$.

Definition 9.3.0.0.3

$$\text{Symm}(n) := \{M \in \text{Mat}(n, n) \mid M^T = M\}$$

Theorem 9.4 (Second Derivative Test for Higher Dimensions)

Let Ω be an open subset of \mathbb{R}^n , let $f \in C^2(\Omega, \mathbb{R})$, with $Hf(\vec{x}) \geq 0$ for all $\vec{x} \in \Omega$.

If $Df(\vec{x}_0) = \vec{0}$ for some $\vec{x}_0 \in \Omega$, then $f(\vec{x}) \geq f(\vec{x}_0)$ for all $\vec{x} \in \Omega$.

Proof. We first note that $Hf(\vec{x}) \in \text{Symm}(n)$ by Corollary 6.2.1. Let $\phi(t) = (1-t)\vec{x}_0 + t\vec{x}$. Let $(a_1, a_2, \dots, a_n) = \vec{a} := \phi'(t) = \vec{x} - \vec{x}_0$. We see that $(f \circ \phi)'(t) = Df(\phi(t))\vec{a} = \sum_j D_j f(\phi(t))a_j$. Notice here we have $(f \circ \phi)'(0) = 0$, and the following:

$$(f \circ \phi)''(t) = \sum_j DD_j f(\phi(t)) \cdot \phi'(t)a_j = \sum_{j,k} D_k D_j f(\phi(t))a_k a_j = \vec{a}^T Hf(\phi(t))\vec{a} = Hf(\phi(t))\vec{a}^T \vec{a} \geq 0$$

We conclude that $f(\vec{x}) = (f \circ \phi)(1) \geq (f \circ \phi)(0) = f(\vec{x}_0)$ □

Corollary 9.4.1

Let Ω be an open subset of \mathbb{R}^n , let $f \in C^2(\Omega, \mathbb{R})$, with $Hf(\vec{x}) > 0$ for all $\vec{x} \in \Omega$.

If $Df(\vec{x}_0) = \vec{0}$ for some $\vec{x}_0 \in \Omega$, then $f(\vec{x}) > f(\vec{x}_0)$ for all $\vec{x} \in \Omega \setminus \{\vec{x}_0\}$.

Definition 9.4.1.0.1

$\text{Pos}(n)$ is defined to be the set of positive definite symmetric matrices.

Note: $\text{Pos}(n)$ is an open subset of $\text{Symm}(n)$.

Theorem 9.5 (Local Second Derivative Test)

Let A be an open subset of \mathbb{R}^n , let $f \in C^2(A, \mathbb{R})$, let $\vec{x}_0 \in A$ with $Df(\vec{x}_0) = \vec{0}$. Then we have the followings hold:

1. If we have $Hf(\vec{x}_0) > 0$, then $Hf(\vec{x}) > 0$ for all $\vec{x} \in B_\delta(\vec{x}_0)$ with some $\delta > 0$, and the function f achieves a local minimum at \vec{x}_0
2. If we have $Hf(\vec{x}_0) \not\geq 0$, then the function f has a strict local maximum at the point \vec{x}_0 along some line in A , and hence f does not have a local minimum at the point \vec{x}_0 .
3. If we have $Hf(\vec{x}_0) < 0$, then f has strict local maximum at \vec{x}_0
4. If we have $Hf(\vec{x}_0) \not\leq 0$, then f does not have a local maximum at \vec{x}_0
5. If we have $Hf(\vec{x}_0)$ is not definite nor semi-definite, then the function f does not have a local max, nor local min, at the point \vec{x}_0 .

Corollary 9.5.1

Let Ω be an open subset of \mathbb{R}^n , let $f \in C^2(\Omega, \mathbb{R})$ with the property that $Hf(\vec{x}) \geq 0$ for all $\vec{x} \in \Omega$. Then $f(\vec{x}) \geq f(\vec{x}_0) + Df(\vec{x}_0)(\vec{x} - \vec{x}_0)$ for all $\vec{x}_0, \vec{x} \in \Omega$.

Proof. Let $g(\vec{x}) = f(\vec{x}) - Df(\vec{x}_0)(\vec{x} - \vec{x}_0)$, then we have $Dg(\vec{x}) = Df(\vec{x}) - Df(\vec{x}_0)$, here we note that $Dg(\vec{x}_0) = 0$. Also note that we have $Hg(\vec{x}) = Hf(\vec{x})$. Then by Theorem 9.4, we know that $g(\vec{x}) \geq g(\vec{x}_0) = f(\vec{x}_0)$. The result follows. □

Corollary 9.5.2

Let Ω be an open subset of \mathbb{R}^n , let $f \in C^2(\Omega, \mathbb{R})$ with the property that $Hf(\vec{x}) > 0$ for all $\vec{x} \in \Omega$. Then $f(\vec{x}) > f(\vec{x}_0) + Df(\vec{x}_0)(\vec{x} - \vec{x}_0)$ for all $\vec{x}_0, \vec{x} \in \Omega$ with $\vec{x}_0 \neq \vec{x}$.

Definition 9.5.2.0.1

Let Ω be an open subset of \mathbb{R}^n , let $\psi : \Omega \rightarrow \mathbb{R}$ be a function. The set $\{(\vec{x}, y) \in \Omega \times \mathbb{R} \mid y \geq \psi(\vec{x})\}$ is called the epigraph of ψ . The set $\{(\vec{x}, y) \in \Omega \times \mathbb{R} \mid y \leq \psi(\vec{x})\}$ is called the hypergraph of ψ .

Corollary 9.5.3

Let Ω be an open subset of \mathbb{R}^n , let $\psi \in C^2(\Omega, \mathbb{R})$ with the property that $H\psi(\vec{x}) > 0$ for all $\vec{x} \in \Omega$.

$$\text{epigraph}(\psi) = \bigcap_{\vec{x}_0 \in \Omega} \{(\vec{x}, y) \in \Omega \times \mathbb{R} \mid y \geq \psi(\vec{x}_0) + D\psi(\vec{x}_0)(\vec{x} - \vec{x}_0)\}$$

Proof. Here we see that $(\vec{x}, y) \in \text{epigraph}(\psi) \Rightarrow (\vec{x}, y) \in \text{RHS}$ by Corollary 9.5.2. On the other hand, we have $(\vec{x}, y) \in \text{RHS} \Rightarrow y \in \psi(\vec{x}) \Rightarrow (\vec{x}, y) \in \text{epigraph}(\psi)$ □

Corollary 9.5.4

Let Ω be an open subset of \mathbb{R}^n , let $\psi \in C^2(\Omega, \mathbb{R})$ with the property that $H\psi(\vec{x}) > 0$ for all $\vec{x} \in \Omega$. Then the set $\text{epigraph}(\psi)$ is convex.

Definition 9.5.4.0.1

Let Ω be a convex subset of \mathbb{R}^n , and let $f : \Omega \rightarrow \mathbb{R}^n$. The function f is said to be convex provided that the set $\text{epigraph}(f)$ is a convex set.

Remark: Corollary 9.5.4 provides us a way to determine if a function f is convex.

Lemma 9.5.5

Let Ω be a convex subset of \mathbb{R}^n , let $f : \Omega \rightarrow \mathbb{R}$ be a function.

The function f is convex if and only if the following holds:

$$f((1-t)\vec{x}_0 + t\vec{x}_1) \leq (1-t)f(\vec{x}_0) + tf(\vec{x}_1), \quad \forall \vec{x}_1, \vec{x}_0 \in \Omega, 0 \leq t \leq 1$$

3 | Multivariate Integration

Integrating Over a Box

Definition 10.0.0.0.1

Let $a_1, a_2, \dots, a_n \in \mathbb{R}$, and $b_1, b_2, \dots, b_n \in \mathbb{R}$, with $a_i \leq b_i$ for all $1 \leq i \leq n$, $I_i = [a_i, b_i] \subseteq \mathbb{R}$. $Q := [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n] := I_1 \times I_2 \times \dots \times I_n$ is called a rectangle, or a box, in \mathbb{R}^n , and $V(Q) := (b_1 - a_1)(b_2 - a_2) \dots (b_n - a_n)$ gives the volume of box Q .

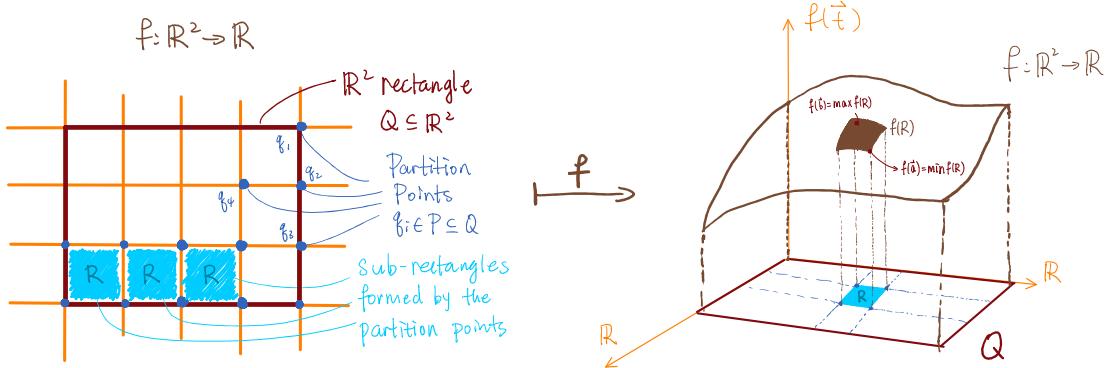
Definition 10.0.0.0.2

For $[a, b] \subseteq \mathbb{R}$ with $a < b$, a partition of $[a, b]$ is a finite collection P of points in $[a, b]$ that includes a, b . That is, $P \in \mathcal{P}^{\text{finite}}([a, b])$ with $a, b \in P$. We index the elements of P in increasing order, as $a = t_0 < t_1 < \dots < t_k = b$. Each $[t_{i-1}, t_i]$ for $1 \leq i \leq k$ is called a subinterval determined by P .

Definition 10.0.0.0.3

For a box $Q = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n] \in \mathbb{R}^n$, a partition P of Q is an n -tuple (P_1, P_2, \dots, P_n) such that P_j is a partition of $[a_j, b_j]$ for $1 \leq j \leq n$. Let I_j be one of the subintervals determined by P_j of the interval $[a_j, b_j]$, the box $R := I_1 \times I_2 \times \dots \times I_n$ is called a subbox, or subrectangle, determined by P , of the box Q .

Let Q be a box in \mathbb{R}^n , and let P be a partition of Q . Consider a bounded function $f : Q \rightarrow \mathbb{R}$. In the following figure, we assume $n = 2$ for simplicity.



We want to find a way to define the integral $\int_Q f$ for bounded functions over Q . For f being defined above, bounded function $g : Q \rightarrow \mathbb{R}$, and a constant $C \in \mathbb{R}$, $\int_Q g$ should satisfy the followings:

$$\int_Q C = C \cdot V(Q) \quad f \leq g \implies \int_Q f \leq \int_Q g \quad \int_Q f = \sum_R \int_R f$$

Definition 10.0.0.0.4

For $f : Q \rightarrow \mathbb{R}$ with P being a partition of a box $Q \subseteq \mathbb{R}^n$, we define the followings:

$$L(f, P) := \sum_R (\inf(f(R)))V(R) \quad U(f, P) := \sum_R (\sup(f(R)))V(R)$$

Note: From Definition 10.0.0.0.4, it is immediate that $L(f, P) \leq U(f, P)$.

Definition 10.0.0.0.5

Let P' and P be partitions for a box $Q \subseteq \mathbb{R}^n$, we say P' refines P provided that P' is obtained by adding more partition points to P , that is, P' refines P provided that $P \subseteq P'$.

Lemma 10.0.1

For bounded function $f : Q \rightarrow \mathbb{R}$ with P being a partition of a box $Q \subseteq \mathbb{R}^n$, we have the following:

$$L(f, P) \leq L(f, P') \leq U(f, P') \leq U(f, P)$$

Proof. The proof for this Lemma for $n \geq 1$ is on Munkres Lemma 10.1.

One might also want to check Spivak pg.255 for $n = 1$. \square

Lemma 10.0.2

Let P and P' be arbitrary partitions of a box $Q \subseteq \mathbb{R}^n$, and let f be a bounded function from Q to \mathbb{R} . Then we have $L(f, P) \leq U(f, P')$.

Proof. First we form a partition P'' using all partition points from P and P' . Here P'' refines P and P' . By Lemma 10.0.1, we see that $L(f, P) \leq L(f, P'') \leq U(f, P'') \leq U(f, P')$. \square

Definition 10.0.2.0.1

Let A be a subset of \mathbb{R} , and let $f : \mathcal{P}(A) \rightarrow \mathbb{R}$ be a bounded function.

$$\inf_{A_i} f(A_i) := \inf\{f(A_i) \mid A_i \subseteq A\} \quad \sup_{A_i} f(A_i) := \sup\{f(A_i) \mid A_i \subseteq A\}$$

Definition 10.0.2.0.2

Let P be a partition of a box $Q \subseteq \mathbb{R}^n$, and let $f : Q \rightarrow \mathbb{R}$ be a bounded function.

The Darboux Upper Integral is defined to be:

$$\bar{\int}_Q f := \inf_P U(f, P)$$

The Darboux Lower Integral is defined to be:

$$\underline{\int}_Q f := \sup_P L(f, P)$$

Proposition 10.0.3

For a bounded function $f : Q \rightarrow \mathbb{R}$ with Q being a box of \mathbb{R}^n , we have the following holds:

$$\underline{\int}_Q f \leq \bar{\int}_Q f$$

Proof. Combining Lemma 1.0.1 with Lemma 10.0.2, the result follows. \square

Definition 10.0.3.0.1

For a bounded function $f : Q \rightarrow \mathbb{R}$ with Q being a box of \mathbb{R}^n , f is said to be Riemann integrable over Q provided that $\bar{\int}_Q f = \underline{\int}_Q f$, in which case we write $\int_Q f := \underline{\int}_Q f$.

Theorem 10.1 (Cauchy Criterion for Integrability, or the Riemann Condition for Integrability)
For a bounded function $f : Q \rightarrow \mathbb{R}$ with Q being a box of \mathbb{R}^n , f is Riemann integrable on Q if and only if for all $\epsilon > 0$, there exists a partition P such that $U(f, P) - L(f, P) < \epsilon$.

Proof. For the \Leftarrow direction, for all $\epsilon > 0$, there exists some partition P of Q such that we have:

$$0 \leq \bar{\int}_Q f - \underline{\int}_Q f \leq U(f, P) - L(f, P) < \epsilon$$

so we must have $\bar{\int}_Q f = \underline{\int}_Q f$ by Lemma 1.0.1, this completes the proof for the \Leftarrow direction. For the \Rightarrow direction, for a given $\epsilon > 0$, we see that there exists partitions P, P' of Q with $U(f, P') < L(f, P) + \epsilon$, take refinement P'' of P and P' , we have $U(f, P'') < L(f, P'') + \epsilon$, the result follows. \square

Definition 10.1.0.0.1

Let Q be a box of \mathbb{R}^n , the osculation of a bounded function $f : Q \rightarrow \mathbb{R}$ at $\vec{a} \in \mathbb{R}^n$ is defined to be:

$$\text{osc}(f, \vec{a}) := \inf \{ \sup\{f(\vec{t}) \mid \vec{t} \in B_\delta(\vec{a}) \cap Q\} - \inf\{f(\vec{m}) \mid \vec{m} \in B_\delta(\vec{a}) \cap Q\} \mid \delta > 0 \}$$

Let $f : Q \rightarrow \mathbb{R}$ be a bounded function with Q being a box of \mathbb{R}^n .

Note: $\text{osc}(f, \vec{a}) < \epsilon \iff \exists$ open neighborhood U of \vec{a} such that $\sup_{U \cap Q} f - \inf_{U \cap Q} f < \epsilon$.

Note: $\{\vec{a} \in Q, \text{osc}(f, \vec{a}) < \epsilon\}$ is relatively open. $g : Q \rightarrow \mathbb{R}$ $\vec{a} \mapsto \text{osc}(f, \vec{a})$ is upper semi-continuous.

Example:

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad x \mapsto \begin{cases} \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases} \quad \text{osc}(f, a) = \begin{cases} 0 & a \neq 0 \\ 2 & a = 0 \end{cases}$$

Definition 10.1.0.0.2

Let $f : Q \rightarrow \mathbb{R}$ be a bounded function with Q being a box of \mathbb{R}^n . $\mathcal{D}_k(f) := \{\vec{a} \in Q \mid \text{osc}(f, \vec{a}) \geq \frac{1}{k}\}$

Note: The set $\mathcal{D}_k(f)$ defined in Definition 10.1.0.0.2 is closed.

Definition 10.1.0.0.3

Let $f : Q \rightarrow \mathbb{R}$ be a bounded function with Q being a box of \mathbb{R}^n .

$$\mathcal{D}(f) := \bigcup_{k=1}^{\infty} \mathcal{D}_k(f) = \{\vec{a} \in Q \mid f \text{ is not continuous at } \vec{a}\}$$

Note: The set $\mathcal{D}(f)$ defined in Definition 10.1.0.0.3 is an F_σ set.

Definition 10.1.0.0.4

Let (X, \mathcal{T}) be a topological space, let A be a subset of X equipped with the subspace topology inherited from X , let M be a subset of A , the relative interior of M , denoted as $rInt(M)$, is defined to be:

$$rInt(M) := \{x \in M \mid \exists \epsilon > 0, \text{ s.t. } B_\epsilon(x) \cap A \subseteq M\}$$

Definition 10.1.0.0.5

Let A be a subset of \mathbb{R}^n . A is said to have measure zero in \mathbb{R}^n provided that for every $\epsilon > 0$, there exists countably many boxes Q_1, Q_2, \dots of \mathbb{R}^n such that $A \subseteq \bigcup_{i=1}^{\infty} Q_i$ with $\sum_{i=1}^{\infty} V(Q_i) < \epsilon$.

Theorem 10.2

Let $f : Q \rightarrow \mathbb{R}$ be a bounded function with Q being a box of \mathbb{R}^n .

Given $\epsilon > 0$ and some $k \in \mathbb{N}$, the followings are equivalent:

1. The function f is Riemann integrable on Q
2. There exists a partition P such that $U(f, P) < L(f, P) + \epsilon$
3. \exists some subboxes R_1, R_2, \dots, R_j of Q s.t. $\mathcal{D}_k(f) \subseteq R_1 \cup R_2 \cup \dots \cup R_j$ and $\sum_{i=1}^j V(R_i) < \epsilon$
4. There exists some subboxes R_p of Q such that $\mathcal{D}(f) \subseteq \bigcup_{p=1}^{\infty} R_p$, with $\sum_{p=1}^{\infty} V(R_p) < \epsilon$
5. There exists some subboxes R_p of Q such that $\mathcal{D}(f) \subseteq \bigcup_{p=1}^{\infty} rInt(R_p)$ with $\sum_{p=1}^{\infty} V(R_p) < \epsilon$

In the context of this theorem, a box R is called a subbox of Q provided that R is a box in \mathbb{R}^n contained in Q , and here R is allowed to have zero volume, or in other words, measure zero.

Proof. The equivalency between (1) and (2) is immediate from Theorem 10.1. Now we will show that (2) implies (3). Given $k \in \mathbb{N}$, pick a partition P of Q such that for subboxes R_i determined by P , we have the following holds:

$$U(f, P) - L(f, P) = \sum_{R_i} \left(\sup_{R_i} f - \inf_{R_i} f \right) V(R_i) < \epsilon / (19k)$$

Let $R_{i,1}, R_{i,2}, \dots, R_{i,l}$ denote the subboxes determined by P that satisfies $Int(R_{i,j}) \cap \mathcal{D}_k(f) \neq \emptyset$, then we can write the following:

$$\frac{1}{k} \sum_{p=1}^l V(R_p) = \sum_{p=1}^l \frac{1}{k} V(R_{i,p}) \leq \sum_{p=1}^l \left(\sup_{R_{i,p}} f - \inf_{R_{i,p}} f \right) V(R_{i,p}) \leq \frac{\epsilon}{19k} \quad \Rightarrow \quad \sum_{p=1}^l V(R_{i,p}) < \epsilon$$

Here we note that, $\mathcal{D}_k(f)$ might not necessarily be a subset of $R_{i,1} \cup R_{i,2} \cup \dots \cup R_{i,l}$ because $\mathcal{D}_k(f)$ might have nonempty intersection with the boundaries of some other subboxes $R_{b,1}, R_{b,2}, \dots, R_{b,t}$ determined by P , so we can write $\mathcal{D}_k(f) \subseteq R_{i,1} \cup R_{i,2} \cup \dots \cup R_{i,l} \cup Bd(R_{b,1}) \cup Bd(R_{b,2}) \cup \dots \cup Bd(R_{b,t})$, where each $Bd(R_{b,j})$ is a finite union of subboxes of Q that have measure zero, hence we have:

$$\mathcal{D}_k(f) \subseteq R_{i,1} \cup R_{i,2} \cup \dots \cup R_{i,l} \cup R_{z,1} \cup R_{z,2} \cup \dots \cup R_{z,m}$$

for some zero volume subboxes $R_{z,j}$ of Q . With the definition of a box in \mathbb{R}^n , it follows that we have $\sum_{p=1}^l V(R_{i,p}) + \sum_{j=1}^m V(R_{z,j}) = \sum_{p=1}^l V(R_{i,p}) < \epsilon$. This shows that (2) implies (3). To show that (3) implies (4), for each $k \in \mathbb{N}$, pick finitely many subboxes of Q such that $\mathcal{D}_k(f)$ is contained in the union of the subboxes and the volume sum of those subboxes is less than $\frac{\epsilon}{2^k}$, the result follows. Now we will show (4) implies (5), let $\epsilon > 0$ be given, pick subboxes R_p such that $\mathcal{D}(f) \subseteq \bigcup_{p=1}^{\infty} R_p$ with $\sum_{p=1}^{\infty} V(R_p) < \frac{\epsilon}{19}$. For each p , pick subboxes \hat{R}_p of Q such that the following condition holds:

- (1) If $V(R_p) > 0$, then $R_p \subseteq rInt(\hat{R}_p)$ with $V(\hat{R}_p) < 2V(R_p)$
- (2) If $V(R_p) = 0$, then $R_p \subseteq rInt(\hat{R}_p)$ with $V(\hat{R}_p) < \frac{\epsilon}{2^{p+1}}$

Here we get $\sum_{p=1}^{\infty} V(\hat{R}_p) < 2 \cdot \frac{\epsilon}{19} + \sum \frac{\epsilon}{2^{p+1}} < \epsilon$. This shows that (4) implies (5). Lastly, we will show (5) implies (2). Given $\epsilon > 0$, let $\hat{\epsilon} = \frac{\epsilon}{3(M+V(Q))}$ where $M = \max(|f||_Q)$. By assumption, there exists some subboxes R_p such that $\mathcal{D}(f) \subseteq \bigcup_{p=1}^{\infty} rInt(R_p)$ with $\sum V(R_p) < \hat{\epsilon}$. For each $\vec{a} \in Q \setminus \mathcal{D}(f)$, f is continuous at \vec{a} , so there exists a subbox $Q_{\vec{a}}$ of Q with $\vec{a} \in rInt(Q_{\vec{a}})$ and $|f(\vec{x}) - f(\vec{a})| < \hat{\epsilon}$ for $\vec{x} \in Q_{\vec{a}}$. Note here $Q \subseteq \bigcup_{p=1}^{\infty} rInt(R_p) \cup (\bigcup_{\vec{a} \in Q \setminus \mathcal{D}(f)} rInt(Q_{\vec{a}}))$. Here we note that Q is compact, and each $rInt(Q_{\vec{a}})$, or $rInt(R_p)$, is open, so there exists a finite subcover such that we have:

$$Q \subseteq \left(\bigcup_{p=1}^s rInt(R_p) \right) \cup \left(\bigcup_{\text{finitely many } \vec{a} \in Q \setminus \mathcal{D}(f)} rInt(Q_{\vec{a}}) \right)$$

Here we denote $rInt(Q_{\vec{a}_1}) \cup rInt(Q_{\vec{a}_2}) \cup rInt(Q_{\vec{a}_t}) = \bigcup_{\text{finitely many } \vec{a} \in Q \setminus \mathcal{D}(f)} \text{, and } rInt(R_{p_1}) \cup rInt(R_{p_2}) \cup \dots \cup rInt(R_{p_s}) = \bigcup_{p=1}^s rInt(R_p)$. Pick partition P' such that each subbox R determined by P' satisfies at least one of the two conditions: (1) $R \subseteq Q_{\vec{a}_j}$, or (2) $R \subseteq R_{p_k}$. If R satisfies condition (1), we say R is good, and if R does not satisfies (1), we say R is bad. Then we can write the following:

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_R (\sup_R f - \inf_R f) V(R) \\ &= \sum_{\text{good } R} (\sup_R f - \inf_R f) V(R) + \sum_{\text{bad } R} (\sup_R f - \inf_R f) V(R) \\ &\leq \sum_{\text{good } R} 2\hat{\epsilon}(R) + \sum_{\text{bad } R} 2MV(R) \\ &\leq 2\hat{\epsilon}V(Q) + 2M\hat{\epsilon} \\ &< \epsilon \end{aligned}$$

This shows that (5) implies (2). Now we have shown that (2), (3), (4), (5), are equivalent, and we know that (1) and (2) are equivalent by Theorem 10.1, the result of this theorem follows. \square

Lemma 10.2.1

Let $f : Q \rightarrow \mathbb{R}$ and $g : Q \rightarrow \mathbb{R}$ be function defined on a box $Q \subseteq \mathbb{R}^n$. If we have $f(\vec{x}) \leq g(\vec{x})$ for all \vec{x} in Q , then we have $\underline{\int}_Q f \leq \underline{\int}_Q g$ and $\bar{\int}_Q f \leq \bar{\int}_Q g$.

Proof. For partition P of the box Q , we have $L(f, P) \leq L(g, P) \leq \underline{\int}_Q g$, it follows that $\underline{\int}_Q f \leq \underline{\int}_Q g$. On the other hand, we have $\bar{\int}_Q f \leq U(f, P) \leq U(g, P)$, it follows that $\bar{\int}_Q f \leq \bar{\int}_Q g$. A more detailed proof can be found on Section 10 on Munkres. \square

Theorem 10.3 (Fubini's Theorem)

Let A be a box in \mathbb{R}^k , B be a box in \mathbb{R}^n , let $Q = A \times B$, and let $f : Q \rightarrow \mathbb{R}$ be a bounded function. Then we can write:

$$\int_Q f \leq \int_{\vec{x} \in A} \int_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \left\{ \frac{\int_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})}{\int_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})} \right\} \leq \bar{\int}_{\vec{x} \in A} \int_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \bar{\int}_Q f$$

Proof. By Proposition 10.0.3, we know that $\underline{\int}_Q f \leq \bar{\int}_Q f$, we will refer this result as (*). By Lemma 10.2.1, we know that $\underline{\int}_Q f \leq \underline{\int}_Q g$ when $f(\vec{x}) \leq g(\vec{x})$ for all $\vec{x} \in Q$, we refer this result as (a). Also by Lemma 10.2.1, we know that $\bar{\int}_Q f \leq \bar{\int}_Q g$ when $f(\vec{x}) \leq g(\vec{x})$ for all $\vec{x} \in Q$, we refer this result as (b). Here we see that, Using result from (*), we get $\underline{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \bar{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$, and $\underline{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \bar{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$. Combining the result from (*) and (a), we get $\underline{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \underline{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$. Combining the result from (*) and (b), we get $\bar{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \bar{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$. To show $\underline{\int}_Q f \leq \bar{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$, we denote $\vec{x}_0 \in R_A$ where R_A is a subrectangle of A , denote R_b as subrectangle of B , we can write:

$$\int_{\vec{y} \in B} f(\vec{x}_0, \vec{y}) \geq \sum_{R_B} \inf_{\vec{y} \in R_B} (f(\vec{x}_0, \vec{y})) \cdot V(R_B) \geq \sum_{R_B} \inf_{R_A \times R_B} (f) \cdot V(R_B)$$

On the other hand, we have:

$$\begin{aligned} \int_{\vec{x} \in A} \left(\int_{\vec{y} \in B} f(\vec{x}, \vec{y}) \right) &\geq \sum_{R_A} \inf_{\vec{x} \in R_A} \left(\int_{\vec{y} \in B} f(\vec{x}, \vec{y}) \right) V(R_A) \\ &\geq \sum_{R_A} \sum_{R_B} \inf_{R_A \times R_B} (f) V(R_A \times R_B) \\ &= L(f, P) \end{aligned}$$

Hence we see that $\underline{\int}_Q f \leq \bar{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$. Lastly, one can show that $\bar{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \bar{\int}_Q f$ by combining $\underline{\int}_Q f \leq \bar{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$ with the fact that $\bar{\int}_Q f = -\underline{\int}_Q (-f)$. This completes the proof of this theorem. \square

Corollary 10.3.1

Let A be a box in \mathbb{R}^k , B be a box in \mathbb{R}^n , let $Q = A \times B$, and let $f : Q \rightarrow \mathbb{R}$ be a bounded function. If f is integrable on Q , then we have:

$$\int_Q f = \int_{\vec{x} \in A} \int_{\vec{y} \in B} f(\vec{x}, \vec{y}) = \left\{ \begin{array}{l} \underline{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \\ \bar{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \end{array} \right\} = \bar{\int}_{\vec{x} \in A} \int_{\vec{y} \in B} f(\vec{x}, \vec{y}) = \bar{\int}_Q f$$

and we can write:

$$\int_Q f = \int_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) = \int_{\vec{x} \in A} \int_{\vec{y} \in B} f(\vec{x}, \vec{y})$$

Note: Given the settings in Corollary 10.3.1, $\int_{\vec{y} \in B} f(\vec{x}, \vec{y})$ may not necessarily exist.

Corollary 10.3.2

A function f is continuous on $Q = I_1 \times \cdots \times I_n$ implies $\int_Q f = \int_{x_1 \in I_1} \int_{x_{n-1} \in I_{n-1}} \int_{x_n \in I_n} f(x_1, \dots, x_n)$

Lemma 10.3.3

Let R be a subbox of a box $Q \in \mathbb{R}^n$.

For $y > V(R)$, there exists a subbox \hat{R} of Q with $R \subseteq rInt(\hat{R}) \subseteq Q$ such that $V(\hat{R}) < y$.

Note: Consider a bounded function defined on a box Q .

1. If $f^{-1}(0)$ is dense in Q , then all $L(f, P) \leq 0$, all $U(f, P) \geq 0$, which implies $\underline{\int}_Q f \leq 0 \leq \bar{\int}_Q f$.
2. If $f^{-1}(0)$ is dense in Q and f is integrable, then $\int_Q f = 0$.
3. If $f \geq 0$, $f(\vec{a}) > 0$, f is continuous at \vec{a} , then $\int_Q f > 0$.
4. If $f \geq 0$, $f(\vec{a}) > 0$, and $\int_Q f = 0$, then f is discontinuous.
5. If $f \geq 0$ and f is integrable on Q , with $\int_Q f = 0$, then $Q \setminus f^{-1}(0)$ has measure zero.

Proof of (3). Choose partition P such that $f > \frac{f(\vec{a})}{2}$ on some boxes R^* determined by P . Then we can write $\int_Q f \geq L(f, P) = \inf_{R^*} (f) \cdot V(R^*) + \sum_R \inf_R f \cdot V(R) \geq \frac{f(\vec{a})}{2} V(R^*) + 0$. \square

Example: Let $Q = [-1, 1] \times [-1, 1]$, and $f = \mathbb{I}_{\{0\} \times \mathbb{Q}}$, where \mathbb{I} is the indicator function, we see that $\mathcal{D}(f) = \{0\} \times [-1, 1]$, which is a set that has zero measure, so we know that f is integrable on Q . Notice that, integrating over y with fixed x , we have the followings:

$$\int_{y \in [-1, 1]} f(x, y) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases} \quad \int_{y \in [-1, 1]} f(x, y) = \begin{cases} 0 & x \neq 0 \\ 0 & x = 0 \end{cases}$$

One might also integrate over x with fixed y and get $\int_{x \in [-1, 1]} f(x, y) = 0$. From Fubini's Theorem, we can write $\int_Q f = 0$.

Integrating Over a Bounded Set

Definition 11.0.0.0.1

Let S be a bounded subset of \mathbb{R}^n , and let $f : S \rightarrow \mathbb{R}$ be a bounded function.

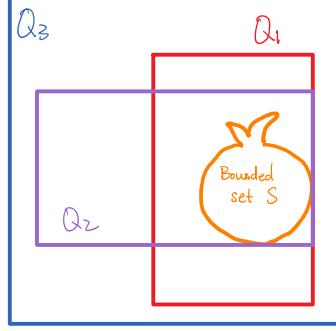
$$f_S : \mathbb{R}^n \rightarrow \mathbb{R} \quad \vec{x} \mapsto \begin{cases} f(\vec{x}) & \vec{x} \in S \\ 0 & \vec{x} \notin S \end{cases} \quad \int_S f := \int_Q f_S \quad \text{for box } Q \text{ that contains } S$$

Proposition 11.0.1

Let S be a bounded subset of \mathbb{R}^n , and let $f : S \rightarrow \mathbb{R}$ be a bounded function.

The existence and the value of $\int_Q f_S$ does not depend on the choice of the box Q which contains S .

Proof. For the existence of $\int_Q f_S$, let Q_1, Q_2 be boxes that contain S , and let Q_3 be a box of S such that Q_1 and Q_2 are contained in the interior of Q_3 .



We will investigate $\mathcal{D}(f_S)$ on Q_3 , denoted as $\mathcal{D}(f_S)_3$, and $\mathcal{D}(f_S)$ on Q_1 , denoted as $\mathcal{D}(f_S)_1$. We see that $\mathcal{D}(f_S)_3 = \mathcal{D}(f_S)_1 \cup \{\text{subset of } Bd(Q_1)\}$, here $\{\text{subset of } Bd(Q_1)\}$ can be covered by finitely many zero-volume boxes. Here we conclude that f_S is integrable on Q_1 if and only if f_S is integrable on Q_3 by Theorem 10.2, and similarly, f_S is integrable on Q_2 if and only if f_S is integrable on Q_3 . Now we can conclude that f_S is integrable on Q_1 if and only if f_S is integrable on Q_2 . That is, the integrability of f_S does not depend on the choice of Q . Now we will show that $\int_{Q_1} f_S = \int_{Q_3} f_S$. Let P be a partition of Q_3 , and refine P to P' such that Q_1 is contained in a finite union of boxes determined by P' . We denote the boxes determined by P' as P' -boxes. Then we can write:

$$L(f_S, P) \leq L(f_S, P') = \sum_{\substack{P'-\text{boxes } R \\ \text{contained in } Q_1}} \inf_R(f_S) V(R) + \sum_{\substack{\text{other } P'-\text{boxes } R' \\ \text{contained in } Q \setminus rIntQ_1}} \inf_{R'}(f_S) V(R') \leq \int_{Q_1} f_S$$

Here we can take the supremum over choice of P to get $\int_{Q_3} f_S \leq \int_{Q_1} f_S$. With similar argument, we get $\int_{Q_3} f_S \geq \int_{Q_1} f_S$ using $U(f_S, P)$, and the result follows. \square

Note: The proof of the followings can be found on Section 13 in Munkres, and Chapter 13 in Spivak. Let $S \subseteq \mathbb{R}^n$ be bounded, let $a, b \in \mathbb{R}$, and let $f : S \rightarrow \mathbb{R}$ and $g : S \rightarrow \mathbb{R}$ be integrable functions.

1. $\int_S (af + bg) = a \int_S f + b \int_S g$
2. If $f(\vec{x}) \leq g(\vec{x})$ for all $\vec{x} \in S$, then we have $\int_S f \leq \int_S g$
3. $|f|$ is integrable on S
4. If f_S is continuous, then $|f|_S$ is continuous.
5. $|\int_S f| = \max\{(\int_S f), (-\int_S f)\} \leq \int_S |f|$
6. Let $T \subseteq S$, if $f(\vec{x}) \geq 0$ for all $\vec{x} \in T$, and f is integrable on T , then $\int_T f \leq \int_S f$.
7. If f is integrable on bounded sets S_1 and S_2 , then f is integrable on $S_1 \cup S_2$ and $S_1 \cap S_2$, with

$$\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f - \int_{S_1 \cap S_2} f$$

Proof of (7). Let $A = \{(x, y) \in \mathbb{R}^2 \mid x = 0 \text{ or } y = 0 \text{ or } x = y\}$. Let $\phi : A \rightarrow \mathbb{R}$ be defined by:

$$\phi : A \rightarrow \mathbb{Q} \quad \begin{cases} (x, 0) \mapsto x \\ (0, y) \mapsto y \\ (x, x) \mapsto x \end{cases}$$

It is easy to check that ϕ is continuous, and we have $\phi(f_{S_1}, f_{S_2}) = f_{S_1 \cup S_2}$, so we know that $f_{S_1 \cup S_2}$ is continuous at the points where f_{S_1} and f_{S_2} are continuous. The discontinuous sets for f_{S_1} and f_{S_2} satisfies the condition of Theorem 10.2 because f is integrable on S_1 and S_2 , then it follows that the discontinuous set for $f_{S_1 \cup S_2}$ also satisfies the condition in Theorem 10.2, and hence $f_{S_1 \cup S_2}$ is integrable on $S_1 \cup S_2$, where one can easily show that $f_{S_1 \cup S_2} + f_{S_1 \cap S_2} = f_{S_1} + f_{S_2}$. \square

Exercise: Let $S = \{(x_1, x_2, x_3) \mid x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1\}$, here S is a tetrahedron, or a 3-simplex, with $S = ch(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{0})$. Let $Q = [0, 1] \times [0, 1] \times [0, 1] \supset S$, we have:

$$\begin{aligned} \int_S 1 = \int_Q \mathbb{I}_S &= \int_{0 \leq x_1 \leq 1} \int_{0 \leq x_2 \leq 1} \int_{0 \leq x_3 \leq 1} \mathbb{I}_S \\ &= \int_{0 \leq x_1 \leq 1} \int_{0 \leq x_2 \leq 1} \max\{1 - x_1 - x_2, 0\} \\ &= \int_{0 \leq x_1 \leq 1} \int_{0 \leq x_2 \leq 1-x_1} 1 - x_1 - x_2 \\ &= \int_{0 \leq x_1 \leq 1} \left[(1 - x_1)x_2 - \frac{x_2^2}{2} \right]_{x_2=0}^{x_2=1-x_1} \\ &= \int_{0 \leq x_1 \leq 1} \frac{(1 - x_1)^2}{2} = \left[-\frac{(1 - x_1)^3}{2} \right]_{x_1=0}^{x_1=1} = \frac{1}{6} \end{aligned}$$

Definition 11.0.1.0.1

Let E be a subset of \mathbb{R}^n .

$$m^*(E) := \inf \left\{ \sum_{j=1}^{\infty} V(Q_j) \mid E \subseteq \bigcup_{j=1}^{\infty} Q_j, \text{ where } Q_j \text{ are boxes in } \mathbb{R}^n \right\}$$

Here $m^*(E)$ is called the outer Lebesgue measure of E .

Note: A set E has outer Lebesgue measure 0 provided that $m^*(E) = 0$, if and only if E has measure 0 as described in Munkres Section 11 and Math 395 IBL.

Remark: Here we do not allow arbitrary coverage by boxes to avoid having $m^*(E) \equiv 0$.

Remark: Here we allow infinite coverings. We might consider building the theory using finite coverings, called the outer Jardon measure, defined by the following when $E \subseteq \mathbb{R}^n$ is bounded:

$$m^{*,J}(E) := \inf \left\{ \sum_{j=1}^k V(Q_j) \mid E \subseteq \bigcup_{j=1}^k Q_j, \text{ where } Q_j \text{ are boxes in } \mathbb{R}^n \right\}$$

and setting $m^{*,J}(E) := \infty$ when E is unbounded.

Note: For a set $E \subseteq \mathbb{R}^n$, we have $m^*(E) \leq m^{*,J}(E)$.

Proposition 11.0.2

For sets $E_j \subseteq \mathbb{R}^n$.

If $m^*(E_j) = 0$ for $j = 1, 2, \dots$, then we have $m^*(\bigcup_{j=1}^{\infty} E_j) = 0$.

If $m^{*,J}(E_j) = 0$ for $j = 1, 2, \dots$, then we have $m^{*,J}(\bigcup_{j=1}^{\infty} E_j) = 0$.

Proof. The proof of this proposition is given on Munkres Theorem 11.1. \square

Lemma 11.0.3

For $E \subseteq \mathbb{R}^n$, we have:

$$\begin{aligned} m^*(E) &= \inf \left\{ \sum_{j=1}^{\infty} V(Q_j) \mid E \subseteq \bigcup_{j=1}^{\infty} Int(Q_j), \text{ where } Q_j \text{ are boxes in } \mathbb{R}^n \right\} \\ m^{*,J}(E) &= \inf \left\{ \sum_{j=1}^k V(Q_j) \mid E \subseteq \bigcup_{j=1}^k Int(Q_j), \text{ where } Q_j \text{ are boxes in } \mathbb{R}^n \right\} \end{aligned}$$

here we use relative interior when appropriate.

Proof. Comparing to the definition of the Lebesgue and Jardon measure, such restrictions in this Lemma can only increase the infimum. Hence we need to show no increase in the infimum occurs. For the case $m^*(E) = \infty$, the result follows trivially. Let $\epsilon > 0$ be given, for nonempty E , pick $\tilde{Q}_j \supseteq Int(\tilde{Q}_j) \supseteq Q_j$ with $V(\tilde{Q}_j) < V(Q_j) + \frac{\epsilon}{2^j}$, by Lemma 10.3.3, here we get $Int(\tilde{Q}_j)$ covering E with $\sum V(\tilde{Q}_j) < \sum V(Q_j) + \epsilon$. Since ϵ is arbitrary, we see that there is no change in infimum. \square

Proposition 11.0.4

Let K be a compact subset of \mathbb{R}^n , then we have $m^{*,J}(K) = m^*(K)$

Proof. It suffices to show that $m^*(K) \geq m^{*,J}(K)$ holds. Pick Q_j with $\bigcup_{j=1}^{\infty} \text{Int}(Q_j) \supseteq K$. By compactness of K , there exists $M \in \mathbb{N}$ such that $\bigcup_{j=1}^M \text{Int}(Q_j) \supseteq K$, then we can write:

$$\sum_{j=1}^{\infty} V(Q_j) \geq \sum_{j=1}^M V(Q_j) \geq m^{*,J}(K)$$

Now take infimum over choice of Q_j 's to get the desired result. \square

Note: In general, $m^{*,J}(E) \neq m^*(E)$ for $E \subseteq \mathbb{R}^n$.

Example: Let $E = Q \cap [0, 1]$, we have $m^*(E) = 0 \neq m^{*,J}(E) = 1$.

Definition 11.0.4.0.1

Let S be a bounded subset of \mathbb{R}^n , let $f : S \rightarrow \mathbb{R}$ be a bounded function.

$$f_S : \mathbb{R}^n \rightarrow \mathbb{R} \quad \vec{x} \mapsto \begin{cases} f(\vec{x}) & \vec{x} \in S \\ 0 & \vec{x} \notin S \end{cases}$$

The function f is said to be Riemann integrable on S provided that the function f_S is Riemann integrable on a box Q that contains S .

Theorem 11.1

Let S be a bounded subset of \mathbb{R}^n , let $f : S \rightarrow \mathbb{R}$ be a bounded continuous function, let $E = \{\vec{x}_0 \in \text{Bd}(S) \mid \lim_{\vec{x} \in S, \vec{x} \rightarrow \vec{x}_0} f(\vec{x}) \neq 0\}$. If we have $m^*(E) = 0$, then f is Riemann integrable on S .

Proof. Note that E also contains isolated points of S , and for such isolated point \vec{p} , if exists, $\exists \epsilon > 0$ such that $\#(B_\epsilon(\vec{p}) \cap S) = 1$, and $\vec{p} \notin \mathcal{D}(f_S)$. Here we have $\mathcal{D}(f_S) \subseteq E$, the result follows immediately from Theorem 10.2. \square

Corollary 11.1.1

Let S be a bounded subset of \mathbb{R}^n , let $f : S \rightarrow \mathbb{R}$ be a bounded continuous function.

If $m^*(\text{Bd}(S)) = 0$, then f is Riemann integrable on S .

Definition 11.1.1.0.1

A bounded set $S \subseteq \mathbb{R}^n$ is said to be rectifiable provided that any one of the following holds:

1. The function $\mathbb{I} : S \rightarrow \mathbb{R} \quad \vec{x} \mapsto 1$ is Riemann integrable on S
2. The indicator function \mathbb{I}_S is integrable on some box $Q \subseteq \mathbb{R}^n$ that contains S .
3. $m^*(\text{Bd}(S)) = 0$
4. $m^{*,J}(\text{Bd}(S)) = 0$

Note: By Corollary 10.4.1, if a bounded set $S \subseteq \mathbb{R}^n$ is rectifiable, then all bounded continuous functions defined on the set S are integrable.

Definition 11.1.1.0.2

Let $S \subseteq \mathbb{R}^n$ be a rectifiable set, let $\mathcal{J} : S \rightarrow \mathbb{R} \quad \vec{x} \mapsto 1$. $V(S) := \int_S \mathcal{J}$.

Note: Let $S \subseteq \mathbb{R}^n$ be a rectifiable set, let $A = \text{Int}(S)$. Then we have $\text{Bd}(A) \subseteq \text{Bd}(S) \Rightarrow m^*(\text{Bd}(A)) \leq m^*(\text{Bd}(S)) = 0 \Rightarrow A$ is rectifiable $\Rightarrow \mathbb{I}_S$ and \mathbb{I}_A are integrable, where $\mathbb{I}_S, \mathbb{I}_A$ are indicator functions on S and on A . Moreover, $\mathbb{I}_{S \setminus A} = \mathbb{I}_S - \mathbb{I}_A$ is integrable on a box Q that contains $\bar{S} \Rightarrow S \setminus A$ is rectifiable, $\text{Int}(S \setminus A) = \emptyset$, for all partition P , $L(\mathbb{I}_{S \setminus A}, P) = 0 \Rightarrow \int_Q \mathbb{I}_{S \setminus A} = 0 \Rightarrow \int \mathbb{I}_A = \int \mathbb{I}_S \Rightarrow V(A) = V(S)$

Extended Riemann Integrals

In this section we will discuss the integrability of a function f defined on a set $S \subseteq \mathbb{R}^n$, not necessarily bounded, with the property that f is not necessarily bounded. We will focus on discussing the integrability of a continuous real-valued function f defined on an open set $A \subseteq \mathbb{R}^n$, with A and f being not necessarily bounded. **We will start by assuming that $f(\vec{a}) \geq 0$ for all $\vec{a} \in A$.**

Definition 12.0.0.0.1

Let A be an open subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ be a continuous function with $f(\vec{a}) \geq 0$, $\forall \vec{a} \in A$.

$$\text{ext } \int_A f := \sup \left\{ \int_E f \mid E \text{ is a compact rectifiable set that is contained in } A \right\}$$

Here $\text{ext } \int_A f$ is called the Extended Riemann Integral of f on A .

Lemma 12.0.1

Let A be an open subset of \mathbb{R}^n and let B be an open subset of A , then $\text{ext } \int_B f \leq \text{ext } \int_A f$ for non-negative function f defined on A and on B .

Definition 12.0.1.0.1

For bounded open set $A \subseteq \mathbb{R}^n$ and bounded non-negative continuous function f defined on \mathbb{R}^n .

$$\underline{\int}_A f := \int_Q f_A$$

for a box Q that contains A .

Note: For bounded open $A \subseteq \mathbb{R}^n$ and bounded non-negative continuous function f defined on \mathbb{R}^n .

1. When A is rectifiable, then $\underline{\int}_A f$ exists.
2. For compact rectifiable subset E of A , we have $\int_E f = \underline{\int}_E f \leq \underline{\int}_A f$, so $\text{ext } \int_A f \leq \underline{\int}_A f$.
3. Let P be a partition of a box Q that contains A , let M be the union of all the boxes determined by P that are contained in A , we have $L(f_A, P) \leq \int_M f \leq \text{ext } \int_A f$, so $\underline{\int}_A f \leq \text{ext } \int_A f$.
4. Combining (2) and (3) we have $\underline{\int}_A f = \text{ext } \int_A f$.
5. From (4), we conclude that if f is integrable on A , then $\int_A f = \text{ext } \int_A f$

Lemma 12.0.2 (Computational Methods for Extended Riemann Integral)

Let $E_1 \subseteq E_2 \subseteq E_3 \subseteq \dots \subseteq A \subseteq \mathbb{R}^n$ with each E_j being compact rectifiable, $\bigcup_j \text{Int}(E_j) = A$, and A being open. We have $\text{ext } \int_A f = \lim_{j \rightarrow \infty} \int_{E_j} f$ for non-negative continuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Proof. We have each of $\int_{E_j} f \leq \text{ext } \int_A f$. So $\lim_{j \rightarrow \infty} \int_{E_j} f \leq \text{ext } \int_A f$. If $E \subseteq A$ is compact rectifiable, then by the compactness of E , we have $E \subseteq E_M$ for some M , because $E \subseteq \text{Int}(E_1) \cup \text{Int}(E_2) \cup \dots \cup \text{Int}(E_M) = \text{Int}(E_M) \subseteq E_M$. Here we can write $\int_E f \leq \int_{E_M} f \leq \lim_{j \rightarrow \infty} \int_{E_j} f$, so we have $\text{ext } \int_A f \leq \lim_{j \rightarrow \infty} \int_{E_j} f$. The result follows. \square

Note: Given the setting in Lemma 12.0.2. It is not automatic to have $\bigcup E_j = A$. As an example, consider $E_j = [-j, 0] \cup [\frac{1}{j}, j]$, we have $\bigcup E_j = \mathbb{R}$, and $\bigcup \text{Int}(E_j) = \mathbb{R} \setminus \{0\}$.

Proposition 12.0.3

For open subset A of \mathbb{R}^n . One can always find E_j that satisfies $E_1 \subseteq E_2 \subseteq E_3 \subseteq \dots \subseteq A$ with each E_j being compact rectifiable, and $\bigcup_j \text{Int}(E_j) = A$.

Proof. One might propose a proof that is similar to the proof of Lemma 15.1 on Munkres. Alternatively, one can construct each E_j by letting E_j be the union of all closed hyper cubes, where each cube has side length $\frac{1}{2^j}$, each cube is a subset of A , and the vertex of each cube has coordinates in $\frac{\mathbb{Z}}{2^j} \cap [-j, j]$. \square

Example:

$$\int_{\mathbb{R}} \frac{1}{1+x^2} = \lim_{j \rightarrow \infty} \int_{\{x|-j \leq x \leq j\}} \frac{1}{1+x^2} = \lim_{j \rightarrow \infty} \arctan(x)|_{x=-j}^{x=j} = \lim_{j \rightarrow \infty} 2 \arctan(j) = \pi$$

Example:

$$\begin{aligned} \int_{\mathbb{R}^2} \frac{1}{1+x^2+y^2} &= \lim_{j \rightarrow \infty} \int_{\{x|-j \leq x \leq j\} \times \{y|-j \leq y \leq j\}} \frac{1}{1+x^2+y^2} \\ &= \lim_{j \rightarrow \infty} \int_{\{y|-j \leq y \leq j\}} \int_{\{x|-j \leq x \leq j\}} \frac{1}{1+x^2+y^2} \\ &= \lim_{j \rightarrow \infty} \int_{\{y|-j \leq y \leq j\}} \frac{2 \arctan((1+y^2)^{-\frac{1}{2}})}{\sqrt{1+y^2}} \end{aligned}$$

To evaluate such integral, one might want to use the Risch Algorithm to determine whether such integral is an elementary function. However, there is another easier approach, by Spivak Chapter 24 Theorem 1, we can rewrite the following:

$$\begin{aligned} \int_{\mathbb{R}^2} \frac{1}{1+x^2+y^2} &= \lim_{j \rightarrow \infty} \lim_{k \rightarrow \infty} \int_{\{x|-k \leq x \leq k\} \times \{y|-j \leq y \leq j\}} \frac{1}{1+x^2+y^2} \\ &= \lim_{j \rightarrow \infty} \int_{-j}^j \left(\lim_{k \rightarrow \infty} \int_{-k}^k \frac{1}{1+x^2+y^2} dx \right) dy \\ &= \lim_{j \rightarrow \infty} \lim_{k \rightarrow \infty} \int_{-j}^j \frac{2 \arctan(k(1+y^2)^{-\frac{1}{2}})}{\sqrt{1+y^2}} dy \\ &= \lim_{j \rightarrow \infty} \int_{-j}^j \frac{\pi}{\sqrt{1+y^2}} dy \end{aligned}$$

Let $y = \sinh(u)$, we get: $\lim_{j \rightarrow \infty} \int_{-j}^j \frac{\pi}{\sqrt{1+y^2}} dy = \lim_{j \rightarrow \infty} \pi u \Big|_{u=-\operatorname{arcsinh}(j)}^{u=\operatorname{arcsinh}(j)} = \pi(\infty) - \pi(-\infty) = \infty$.

Note: We conclude the results from above: Let A be an open subset of \mathbb{R}^n , let $f \in C(A, [0, \infty))$. $\operatorname{ext} \int_A f = \int_A f$ if $\int_A f$ exists. $\operatorname{ext} \int_A f = \lim_{j \rightarrow \infty} \int_{E_j} f$ provided that each E_j is a compact rectifiable subset of A , $E_1 \subseteq E_2 \subseteq \dots \subseteq A$, and $\bigcup_j \operatorname{Int}(E_j) = A$.

Lemma 12.0.4

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be non-negative continuous function, let A be open subset of \mathbb{R}^n , let U_1, U_2, \dots be open subsets of A with $U_1 \subseteq U_2 \subseteq \dots \subseteq A$, and $\bigcup_j U_j = A$. We have $\operatorname{ext} \int_A f = \lim_{j \rightarrow \infty} \operatorname{ext} \int_{U_j} f$.

Proof. Note here $\operatorname{ext} \int_A f$ exists as the sequence $(\operatorname{ext} \int_{U_j} f)$ is monotonic. Now we write each $\operatorname{ext} \int_{U_j} f \leq \operatorname{ext} \int_A f$, hence we have $\lim_{j \rightarrow \infty} \operatorname{ext} \int_{U_j} f \leq \operatorname{ext} \int_A f$. For each compact rectifiable E subset of A , one can find finite subcover U_j such that $E \subseteq U_1 \cup \dots \cup U_j$, so we can write $\int_E f \leq \operatorname{ext} \int_{U_j} f \leq \lim_{j \rightarrow \infty} \operatorname{ext} \int_{U_j} f$. Take supremum over the choice of E , then we have $\operatorname{ext} \int_A f \leq \lim_{j \rightarrow \infty} \operatorname{ext} \int_{U_j} f$. This completes the proof. \square

Definition 12.0.4.0.1

For $x \in [-\infty, \infty]$:

$$x_+ := \max\{x, 0\} \quad x_- := \max\{-x, 0\}$$

Note: For $x \in [-\infty, \infty]$, both x_+ and x_- are non-negative. $x_+ \cdot x_- = 0$, $x = x_+ - x_-$, $|x| = x_+ + x_-$.

Remark: Alternatively, for $x \in [-\infty, \infty]$, one might define $x_+ = \frac{|x|+x}{2}$ and $x_- = \frac{|x|-x}{2}$ instead.

In the discussion below, we will consider continuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, not necessarily non-negative.

Definition 12.0.4.0.2

Let X be a set, for $f : X \rightarrow [-\infty, \infty]$:

$$f_+ : x \mapsto (f(x))_+ \quad f_- : x \mapsto (f(x))_-$$

Definition 12.0.4.0.3

Let X be a set, let $f : X \rightarrow \mathbb{R}$ be a function.

If f is non-negative, we denote $f \geq 0$. If f is non-positive, we denote $f \leq 0$.

If f is negative, we denote $f < 0$. If f is positive, we denote $f > 0$.

Note: For $f : X \rightarrow [-\infty, \infty]$, $f_+ \geq 0$, $f_- \geq 0$, $f_+ \cdot f_- = 0$, $f = f_+ - f_-$, and $|f| = f_+ + f_-$.

Definition 12.0.4.0.4

For $f \in C(A, \mathbb{R})$ where A is an open subset of \mathbb{R}^n . f is said to be extended-integrable on A , or integrable on A in the extended sense, provided that both $\operatorname{ext} \int_A f_+ < \infty$ and $\operatorname{ext} \int_A f_- < \infty$ holds. In other words, f is extended-integrable on A provided that $\operatorname{ext} \int_A f_+ + \operatorname{ext} \int_A f_- = \operatorname{ext} \int_A |f| < \infty$. If f is extended-integrable on A , then¹ we write $\operatorname{ext} \int_A f := \operatorname{ext} \int_A f_+ - \operatorname{ext} \int_A f_-$.

¹By Munkres, $\operatorname{ext} \int_A f$ exists provided that f is extended-integrable on A . For the purpose of this course, Professor Barrett propose another definition for the existence of $\operatorname{ext} \int_A f$, given in Definition 12.0.4.0.5

Definition 12.0.4.0.5

For $f \in C(A, \mathbb{R})$ where A is an open subset of \mathbb{R}^n , $\text{ext } \int_A f$ exists¹ provided that at least one of $\text{ext } \int_A f_+$ and $\text{ext } \int_A f_-$ is finite.

Exercise: Let f be a real-valued continuous bounded non-positive function defined on an open subset A of \mathbb{R}^n . Then we have $f_- = -f$, and here $\text{ext } \int_A f = -\text{ext } \int_A (-f) = -\int_A (-f) = \bar{\int}_A f$.

Lemma 12.0.5

Let $a, b \in \mathbb{R}$, let $f : A \rightarrow \mathbb{R}$ and $g : A \rightarrow \mathbb{R}$ be continuous functions that are extended-integrable on an open subset A of \mathbb{R}^n , we have $\text{ext } \int_A (a \cdot f + b \cdot g) = a \cdot \text{ext } \int_A f + b \cdot \text{ext } \int_A g$.

Lemma 12.0.6

Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{R}$ and $g : A \rightarrow \mathbb{R}$ be continuous functions with $f(\vec{a}) \leq g(\vec{a})$ for all $\vec{a} \in A$, then $\text{ext } \int_A f \leq \text{ext } \int_A g$ whenever f and g are extended-integrable on A .

Theorem 12.1 (Theorem 15.2 on Munkres)

Let A be an open subset of \mathbb{R}^n , let $E_1 \subseteq E_2 \subseteq \dots \subseteq A$ be compact rectifiable sets with $\bigcup_j \text{Int}(E_j) = A$, then we have $\text{ext } \int_A f = \lim_{j \rightarrow \infty} \int_{E_j} f$ when $\text{ext } \int_A f$ exists for continuous function $f : A \rightarrow \mathbb{R}$.

Theorem 12.2 (Theorem 15.6 on Munkres)

Let A be an open subset of \mathbb{R}^n , let $U_1 \subseteq U_2 \subseteq \dots \subseteq A$ be open subsets of A with $\bigcup_j U_j = A$, then we have $\text{ext } \int_A f = \lim_{j \rightarrow \infty} \text{ext } \int_{U_j} f$ whenever $\text{ext } \int_A f$ exists for the continuous function $f : A \rightarrow \mathbb{R}$.

Lemma 12.2.1

Let $f \in C(\mathbb{R}^2, \mathbb{R})$. We have the following holds:

$$\begin{aligned} \text{ext } \int_{\mathbb{R}^2} f_{\pm} &= \lim_{j \rightarrow \infty} \int_{|y| < j} \left(\text{ext } \int_{x \in \mathbb{R}} f_{\pm}(x, y) \right) \\ &= \text{ext } \int_{y \in \mathbb{R}} \left(\text{ext } \int_{x \in \mathbb{R}} f_{\pm}(x, y) \right) = \text{ext } \int_{x \in \mathbb{R}} \left(\text{ext } \int_{y \in \mathbb{R}} f_{\pm}(x, y) \right) \end{aligned}$$

Corollary 12.2.1.1

Let $f \in C(\mathbb{R}^2, \mathbb{R})$. We have the following holds:

$$\begin{aligned} \text{ext } \int_{\mathbb{R}^2} f &= \left(\text{ext } \int_{y \in \mathbb{R}} \text{ext } \int_{x \in \mathbb{R}} f_+(x, y) \right) - \left(\text{ext } \int_{y \in \mathbb{R}} \text{ext } \int_{x \in \mathbb{R}} f_-(x, y) \right) \\ &= \left(\text{ext } \int_{x \in \mathbb{R}} \text{ext } \int_{y \in \mathbb{R}} f_+(x, y) \right) - \left(\text{ext } \int_{x \in \mathbb{R}} \text{ext } \int_{y \in \mathbb{R}} f_-(x, y) \right) \\ &= \text{ext } \int_{y \in \mathbb{R}} \text{ext } \int_{x \in \mathbb{R}} f(x, y) = \text{ext } \int_{x \in \mathbb{R}} \text{ext } \int_{y \in \mathbb{R}} f(x, y) \end{aligned}$$

whenever $\text{ext } \int_{\mathbb{R}^2} f$ exists.

Proposition 12.2.2

Let Q be a box in \mathbb{R}^n , then we have $m^{*,J}(Q) = m^*(Q) = V(Q)$

Proof. Note here we have $m^*(Q) = m^{*,J}(Q)$ since Q is compact. By the definition of m^* , one must show that, if we have $Q \subseteq Q_1 \cup Q_2 \cup \dots \cup Q_k$ for some boxes Q_1, Q_2, \dots, Q_k , then we have $V(Q) \leq V(Q_1) + V(Q_2) + \dots + V(Q_k)$. We proceed by defining m_{pixel} as the following:

$$m_{pixel}(E) := \lim_{N \rightarrow \infty} \frac{\#(E \cap \frac{\mathbb{Z}^n}{2^N})}{2^{nN}}$$

Here we leave it to the reader to prove that $m_{pixel}(R) = V(R)$ for any box R . By Proposition 1.0.2, we can write the following:

$$\frac{\#(Q \cap \frac{\mathbb{Z}^n}{2^N})}{2^{nN}} \leq \sum_{j=1}^k \frac{\#(Q_j \cap \frac{\mathbb{Z}^n}{2^N})}{2^{nN}}$$

Hence we have:

$$V(Q) = \lim_{N \rightarrow \infty} \frac{\#(Q \cap \frac{\mathbb{Z}^n}{2^N})}{2^{nN}} \leq \lim_{N \rightarrow \infty} \sum_{j=1}^k \frac{\#(Q_j \cap \frac{\mathbb{Z}^n}{2^N})}{2^{nN}} = \sum_{j=1}^k V(Q_j)$$

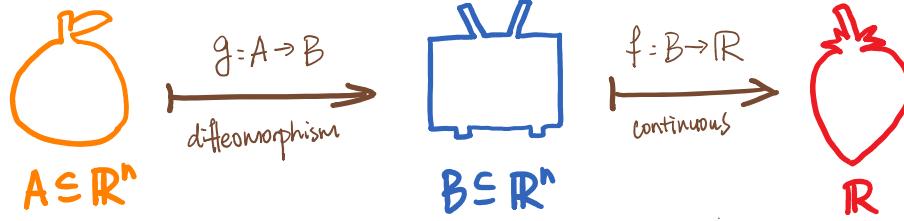
The result follows. \square

From previous result, for bounded function $f : Q \rightarrow [0, \infty)$ where Q is a box in \mathbb{R}^n , if f is continuous at $\vec{a} \in Q$, and if we have $f(\vec{a}) > 0$, then $\int_Q f > 0$. So if a bounded function $g : Q \rightarrow [0, \infty)$ is integrable, and if $\int_Q g = 0$ with $g(\vec{a}) > 0$, then g is discontinuous at each point in $Q \setminus f^{-1}(0)$, then we know that $m^*(Q \setminus f^{-1}(0)) = 0$ by Theorem 11.3 on Munkres.

¹In the rest of this text, the existence of $\text{ext } \int_A f$ refers to Definition 12.0.4.0.5

Change of Variables

Let A be an open subset of \mathbb{R}^n , and let B be an open subset of \mathbb{R}^n . Let g be a diffeomorphism from A to B , and let f be a continuous function from B to \mathbb{R} . We want to find some function $h : A \rightarrow \mathbb{R}$ such that $\text{ext } \int_B f = \text{ext } \int_A h$. In this section, we will show that $h = f \circ g \cdot |\det Dg|$.



Definition 13.0.0.0.1

Let Q be a box in \mathbb{R}^n , and consider an affine map $T : \mathbb{R}^n \rightarrow \mathbb{R}^k$ $\vec{x} \mapsto M\vec{x} + \vec{b}$ that maps Q to P . P is called a parallelepiped.

Note: Given the settings in Definition 13.0.0.0.1, P is rectifiable, with $V(P) = |\det(M)| \cdot V(Q)$.

Definition 13.0.0.0.2

Let C be a compact rectifiable set in \mathbb{R}^{n-1} for $n \geq 2$, let $\phi : C \rightarrow \mathbb{R}$ and $\psi : C \rightarrow \mathbb{R}$ be continuous functions such that $\phi(x) \leq \psi(x)$ for all $x \in C$. The set $S := \{(x, t) \mid x \in C, \phi(x) \leq t \leq \psi(x)\}$ is called a simple region in \mathbb{R}^n .

Lemma 13.0.1

(Lemma 14.3 on Munkres)

Simple regions are compact and rectifiable.

Theorem 13.1

(Fubini's Theorem for Simple Regions)

Let $S := \{(x, t) \mid x \in C, \phi(x) \leq t \leq \psi(x)\}$ be a simple region in \mathbb{R}^n , where C is a compact rectifiable set in \mathbb{R}^{n-1} for $n \geq 2$, $\phi : C \rightarrow \mathbb{R}$ and $\psi : C \rightarrow \mathbb{R}$ are continuous functions with the property $\phi(x) \leq \psi(x)$ for all $x \in C$, let $f : S \rightarrow \mathbb{R}$ be a continuous function. Then f is integrable over S and we have the following holds:

$$\int_S f = \int_{x \in C} \int_{t=\phi(x)}^{t=\psi(x)} f(x, t) dt dx$$

Proof. The proof of this theorem is given by Theorem 14.4 on Munkres. \square

Theorem 13.2

(Substitution Rule)

Let $I = [a, b] \subseteq \mathbb{R}$, let $g : I \rightarrow \mathbb{R}$ be a function of C^1 type, with $g'(x) \neq 0$ for all $x \in (a, b)$. Then the set $g(I)$ is a closed interval J with end points $g(a)$ and $g(b)$. For continuous function $f : J \rightarrow \mathbb{R}$, we have the following holds:

$$\int_{g(a)}^{g(b)} f = \int_a^b (f \circ g) g' \quad \int_J f = \int_I (f \circ g) |g'|$$

Proof. The proof of this theorem is given by Theorem 17.1 on Munkres. \square

Theorem 13.3

(Change of Variable Theorem)

Let A be an open subset of \mathbb{R}^n , let B be an open subset of \mathbb{R}^n , and let g be a diffeomorphism from A to B . For continuous function $f : B \rightarrow \mathbb{R}$, f is integrable over B if and only if the function $(f \circ g) \cdot |\det Dg|$ is integrable over A . Moreover, if f is integrable over B , we have

$$\text{ext } \int_B f = \text{ext } \int_A (f \circ g) \cdot |\det Dg|$$

That is, for continuous function $f : B \rightarrow \mathbb{R}$, we have either $\text{ext } \int_B f = \text{ext } \int_A (f \circ g) \cdot |\det Dg|$, or neither $\text{ext } \int_B f$ nor $\text{ext } \int_A (f \circ g) \cdot |\det Dg|$ exists.

Note: (The restriction of the Change of Variable Theorem on \mathbb{R})

First we generalize the Substitution Rule as a consequence of the Change of Variable Theorem: Given the settings in Theorem 13.3, suppose we have $n = 1$, and suppose further $A \subseteq \mathbb{R}$ is an open connected set, that is, we can write $A = (\alpha, \beta)$. In such case, the diffeomorphism g defined on A must be strictly monotonic, and we have either (1) $B = (g(\alpha), g(\beta))$, or (2) $B = (g(\beta), g(\alpha))$. For case (1), we can write:

$$\text{ext } \int_{(g(\alpha), g(\beta))} f = \text{ext } \int_{(\alpha, \beta)} (f \circ g) \cdot g'$$

For case (2), we have

$$\text{ext} \int_{(g(\beta), g(\alpha))} f = -\text{ext} \int_{g(\alpha)}^{g(\beta)} f = -\text{ext} \int_{(\alpha, \beta)} (f \circ g) \cdot g'$$

These results follows immediately from the Substitution Rule, one can use Chain Rule and Fundamental Theorem of Calculus to prove these results easily.

Proof of Theorem 13.3.

We start with some temporary added assumptions:

1. A, B are bounded and rectifiable.
2. g is a diffeomorphism from a neighborhood of \bar{A} to neighborhood of \bar{B}
3. $f \in C(\bar{B}, \mathbb{R})$

First we consider Special Case 1:

Suppose the function g defined in the settings of Theorem 13.3 is a coordinate transposition that maps $(x_1, \dots, x_l, \dots, x_k, \dots, x_n)$ to $(x_1, \dots, x_k, \dots, x_l, \dots, x_n)$. Then we know that $\det(Dg(\vec{x})) = -1 \Rightarrow |\det(Dg(\vec{x}))| = 1$ for all $\vec{x} \in A$. Note that reorienting boxes does not alter their volume, then we can write:

$$\int_B f = \int_A f \circ g = \int_A (f \circ g) |\det(Dg)|$$

use extended integral when appropriate. This completes the proof of Theorem 13.3 for this Special Case 1. In this rest of the proof of Theorem 13.3, we will call such coordinate transposition function g as Type-1 diffeomorphism.

Now consider Special Case 2:

Suppose the set B defined in the settings of Theorem 13.3 is given by the following:

$$B = \{\vec{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid (x_1, \dots, x_{n-1}) \in E, \phi(x_1, \dots, x_{n-1}) < x_n < \psi(x_1, \dots, x_{n-1})\}$$

where E is a bounded open rectifiable subset of \mathbb{R}^{n-1} , and $\phi, \psi \in C(\bar{E}, \mathbb{R})$ with $\phi(\vec{x}) < \psi(\vec{x})$ for all $\vec{x} \in E$. Suppose further that the function g defined in the settings of Theorem 13.3 maps $(x_1, x_2, \dots, x_{n-1}, x_n)$ to $(x_1, \dots, x_{n-1}, \alpha(\vec{x}))$, for some function $\alpha : A \rightarrow \mathbb{R}$. Here we can write:

$$A = \{\vec{x} \mid (x_1, \dots, x_{n-1}) \in E, \phi(x_1, \dots, x_{n-1}) < \alpha(\vec{x}) < \psi(x_1, \dots, x_{n-1})\}$$

And we see that we get the followings:

1. A, B are rectifiable by Lemma 14.3 on Munkres.
2. Given $\vec{x}' = (x_1, x_2, \dots, x_{n-1}) \in E$, we see that:

$$\vec{x} = (\vec{x}', x_n) \in A \iff x_n \text{ belongs to an interval determined by } \vec{x}'$$

3. Given $\vec{x} \in A$, we have $\det(Dg(\vec{x})) = D_n \alpha(\vec{x})$ because we have:

$$Dg(\vec{x}) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ D_1 \alpha(\vec{x}) & D_2 \alpha(\vec{x}) & \cdots & D_n \alpha(\vec{x}) \end{bmatrix}$$

Let $\vec{x}' = (x_1, x_2, \dots, x_{n-1})$ denote a vector in E , let $\vec{x} = (\vec{x}', x_n)$ denote a vector in A determined by \vec{x}' , and let $I_{\vec{x}'} \subseteq \mathbb{R}$ denote the interval determined by \vec{x}' . By Fubini's Theorem on Simple Regions and Substitution Rule, one can write the following:

$$\int_B f = \int_{\vec{x}' \in E} \int_{x_n \in (\phi(\vec{x}'), \psi(\vec{x}'))} f = \int_{\vec{x}' \in E} \int_{\alpha(\vec{x}') \in I_{\vec{x}'}} (f \circ g) |D_n \alpha| = \int_A (f \circ g) |D_n \alpha| = \int_A (f \circ g) |\det Dg|$$

This completes the proof of Theorem 13.3 for Special Case 2. In the rest of the proof of Theorem 13.3, we will call the function g defined in Special Case 2 as Type-2 diffeomorphism.

To proceed, we will make use of the following proposition:

Proposition 13.3.1

Let A, B, C be open subsets of \mathbb{R}^n , let $g : A \rightarrow B$ and $h : B \rightarrow C$ be diffeomorphisms, and let $f : C \rightarrow \mathbb{R}$ be a continuous function. If f is integrable over C , and $f \circ h$ is integrable over B , then we have $\int_C f = \int_A (f \circ (h \circ g)) |\det(D(h \circ g))|$.

Proof. By the Change of Variable Theorem, one can write the following:

$$\int_C f = \int_B (f \circ h) |\det(Dh)| = \int_A ((f \circ h) \circ g) |\det(Dh)| \cdot |g| |\det(Dg)| = \int_A f \circ h \circ g |\det(D(h \circ g))|$$

Use extended integral here when appropriate. \square

Now suppose the function g defined in the settings of Theorem 13.3 can be characterized by:

$$g : A \rightarrow B \quad \vec{x} = (x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n) \mapsto (x_1, \dots, x_{j-1}, \eta(\vec{x}), x_{j+1}, \dots, x_n)$$

for some function $\eta : A \rightarrow \mathbb{R}$. Such diffeomorphism is called a generalized shear. In the rest of the proof of Theorem 13.3, we will call generalized shear as Type-3 diffeomorphism. Notice that, Type-3 diffeomorphisms can be obtained by compositions of Type-1 and Type-2 diffeomorphism. Combing Proposition 13.3.1 with Special Case 1 and Special Case 2, it follows that the results of Theorem 13.3 holds for Type-3 diffeomorphism. Note that Type-1 diffeomorphism is a composition of two Type-3 diffeomorphism. Check Page 157 on Munkres for details.

The following proposition and theorem ensure that one can locally factorize any diffeomorphism into compositions of Type-1 and Type-2 diffeomorphism.

Proposition 13.3.2

Any invertible affine map can be factored into composition of affine maps of Type-1 and Type-2, or equivalently, composition of Type-1 and Type-3.

Proof. For linear maps, this follows from element matrix factorization from Theorem 2.4 on Munkres. For translation maps, one can move one coordinate at a time by using Type-3 diffeomorphism. All invertible affine maps can be obtained from composition of linear maps and translation maps. This completes the argument of this proposition. \square

Theorem 13.4 (Theorem 18.3 on Munkres)

Let A, B be open subset of \mathbb{R}^n , let $g : A \rightarrow B$ be a diffeomorphism, let $\vec{p} \in A$ with $g(\vec{p}) = \vec{q} \in B$. In some neighborhood of \vec{p} , the diffeomorphism g can be factored into a composition of Type-1 diffeomorphisms and Type-2 diffeomorphisms.

Proof. Let C denote the matrix $Dg(\vec{p})$. Since g is a diffeomorphism, then by the Inverse Function Theorem, we know that C is invertible. Define the following invertible affine maps:

$$t_1 : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \vec{x} \mapsto \vec{x} + \vec{p} \quad t_2 : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \vec{x} - \vec{q} \quad t_3 : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \vec{x} \mapsto C^{-1}\vec{x}$$

Now let $T_1 = t_2^{-1} \circ t_3^{-1}$ and $T_2 = t_1^{-1}$, we get the followings:

1. $\hat{g} := T_1^{-1} \circ g \circ T_2^{-1}$ is a diffeomorphism, $g = T_1 \circ \hat{g} \circ T_2$.
2. $T_2(\vec{p}) = \vec{0}$, and $T_1(\vec{0}) = \vec{q}$
3. $D\hat{g}(\vec{0}) = I$, where I is the $n \times n$ identity matrix.

From (2), we get $\hat{g}(\vec{0}) = \vec{0}$.

From (3), we get $D\hat{g}_k(\vec{0}) = \vec{e}_k^T$.

From Chain Rule, we get $DT_1 = Dg(\vec{p})$ and $DT_2 = I$.

Now we define the following maps:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} \hat{g}_1(\vec{x}) \\ x_2 \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} \hat{g}_1(\vec{x}) \\ \hat{g}_2(\vec{x}) \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \quad \dots \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \mapsto \hat{g}(\vec{x})$$

By the Inverse Function Theorem, each of these maps is a diffeomorphism in some neighborhood of $\vec{0}$ because the derivative of each of them at $\vec{0}$ is I . We observe that each of these maps is a Type-3 diffeomorphism. Then we get an induced map given by the following:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} \hat{g}_1(\vec{x}) \\ x_2 \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} \hat{g}_1(\vec{x}) \\ \hat{g}_2(\vec{x}) \\ x_3 \\ \dots, \\ x_n \end{bmatrix} \mapsto \dots \mapsto \hat{g}(\vec{x})$$

Combining with Proposition 13.3.2, the result follows. \square

Concluding the above, we have got a local version of Theorem 13.3. To finish the proof of Theorem 13.3, we will need the Partition of Unity Theorem to partition the function f on the open set B .

Definition 13.4.0.0.1

Let X be a metric space, and let V be a vector space, let $f : X \rightarrow V$.

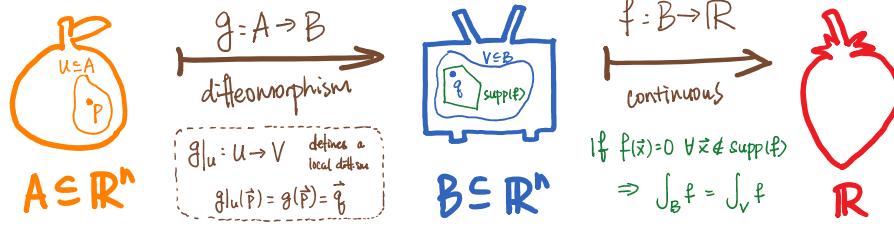
The support of f , denoted as $\text{supp}(f)$, is defined to be the closure of the set $\{\vec{x} \mid f(\vec{x}) \neq \vec{0}\}$ in X .

Note: Let X be a metric space, and let V be a vector space, let $f : X \rightarrow V$.

For $\vec{x} \in X$, $\vec{x} \notin \text{supp}(f)$ if and only if $\exists \epsilon > 0$ such that $f(\vec{t}) = \vec{0}, \forall \vec{t} \in B_\epsilon(\vec{x})$.

Corollary 13.4.1

Let A, B be open subset of \mathbb{R}^n , let $g : A \rightarrow B$ be a diffeomorphism, let $\vec{p} \in A$ with $g(\vec{p}) = \vec{q} \in B$. There exists some open neighborhood U of \vec{p} , and some open neighborhood V of \vec{q} such that $g|_U : U \rightarrow V$ defines a diffeomorphism that can be factored into Type-1 and Type-2 diffeomorphisms. For some continuous function $f : B \rightarrow \mathbb{R}$, if $\text{supp}(f) \subseteq U$, then we have either $\text{ext} \int_B f = \text{ext} \int_A (f \circ g) \cdot |\det Dg|$, or neither $\text{ext} \int_B f$ nor $\text{ext} \int_A (f \circ g) \cdot |\det Dg|$ exists.



The Partition of Unity Theorem gives us a way to partition the function f on the open set B such that the support of each partition of f is contained in some desired open subset of B , in which case we can apply Corollary 13.4.1 to complete the proof of the Change of Variable Theorem.

Lemma 13.4.2

The function f given by the following belongs to $C^\infty(\mathbb{R}, [0, \infty))$:

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Proof. The proof of this lemma is given in Spivak Chapter 18. \square

Corollary 13.4.2.1

Consider the function:

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

For $Q = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$, the function $\Phi_Q(\vec{x}) = f(x_1 - a_1)f(b_1 - x_1) \cdots f(x_n - a_n)$ is in $C^\infty(\mathbb{R}^n, [0, \infty))$, with $\text{supp}(\Phi_Q) = Q$ and $\Phi_Q(\vec{x}) > 0$ for all $\vec{x} \in Q$.

Lemma 13.4.3

One can find positive volume boxes Q_1, Q_2, \dots such that the following holds:

1. $\Omega = \bigcup_j \text{Int}(Q_j)$
2. Each Q_j is contained in some U_{α_j}
3. Each $y \in \Omega$ has neighborhood meeting only finitely many Q_j , such property is called the local finiteness

Proof. Write $\Omega = K_1 \cup K_2 \cup \dots$ with each K_j being compact, and $K_j \subseteq \text{Int}(K_{j+1})$. Check Lemma 15.1, or a previous theorem in this text. Letting $E_j = K_j \setminus \text{Int}(K_{j-1})$, each E_j is compact, and disjoint from K_{j-2} for $j \geq 3$. $E_j \subseteq \text{Int}(\hat{Q}_{j,1}) \cup \dots \cup \text{Int}(\hat{Q}_{j,M_j})$, with each $\hat{Q}_{j,k}$ contained in some U_α , and each $\hat{Q}_{j,k}$ is disjoint from K_{j-2} . Then we can write $\Omega = \bigcup_{j,k} \text{Int}(\hat{Q}_{j,k})$. K_j meets at most $M_i + \dots + M_j$ of the $\hat{Q}_{j,k}$. Relabel the $\hat{Q}_{j,k}$ as Q_1, Q_2, Q_3, \dots This completes the proof of the lemma. \square

Now we have got our setup for the Partition of Unity Theorem.

Theorem 13.5 (Partition of Unity Theorem)

Let Ω be an open subset of \mathbb{R}^n . If $\Omega = \bigcup_{\alpha \in A} U_\alpha$ for some open subsets U_α of \mathbb{R}^n , then there exist some functions $\phi_1, \phi_2, \dots \in C^\infty(\Omega, [0, \infty))$ such that the followings hold:

1. Each $\text{supp}(\phi_j)$ is compact
2. Each $\text{supp}(\phi_j)$ is contained in some U_α
3. Each $\vec{x} \in \Omega$ has an open neighborhood that intersects only finitely many $\text{supp}(\phi_j)$
4. $\sum_{j=1}^{\infty} \phi_j(\vec{x}) = 1$ for all $\vec{x} \in \Omega$, such sum is called the locally finite sum.

Proof. The proof of this theorem follows from the Fundamental Theorem of Laziness.

Section 16 on Munkres gives a proof for this theorem. Alternatively, consider the following: Let $\Phi(Q_i)$ be functions defined appropriately similar to those given in Corollary 13.4.2.1. Let λ be a function defined by $\lambda(\vec{x}) = \sum_{j=1}^{\infty} \Phi_{Q_j}(\vec{x})$. Here we have $\lambda \in C^\infty(\Omega, (0, \infty))$, and the following holds:

$$1 = \sum_{j=1}^{\infty} \frac{\Phi_{Q_j}(\vec{x})}{\lambda(\vec{x})}$$

We let $\phi_j(\vec{x}) := \frac{\Phi_{Q_j}(\vec{x})}{\lambda(\vec{x})}$, which is a C^∞ function, and we have $\text{supp}(\phi_j) = Q_j$. Then the results of Theorem 13.5 follow immediately. \square

Lemma 13.5.1

Let B be an open subset of \mathbb{R}^n , let $f \in C(B, \mathbb{R})$ such that $\text{ext} \int_B f$ exists, let (ϕ_j) be a sequence of functions in $C^\infty(B, [0, \infty])$ that satisfies conditions (1), (3), (4) in the Partition of Unity Theorem. Then we can write $\text{ext} \int_B f = \sum_{j=1}^{\infty} \int_B \phi_j \cdot f$.

Proof. First we will tackle the case where $f \geq 0$, then apply to f_+ and f_- and combine the results. Assume that $f \geq 0$, suppose E is a compact rectifiable subset of B . Then there exists M such that $\phi_j = 0$ on E for $j \geq M$. Since $f \geq 0$, here we have:

$$\int_E f = \int_E \sum_{j=1}^{\infty} \phi_j \cdot f = \int_E \sum_{j=1}^M \phi_j \cdot f = \sum_{j=1}^M \int_E \phi_j \cdot f \leq \sum_{j=1}^M \int_B \phi_j \cdot f \leq \sum_{j=1}^{\infty} \int_B \phi_j \cdot f$$

Take the supremum over E , and get the following:

$$\text{ext} \int_B f \leq \sum_{j=1}^{\infty} \int_B \phi_j \cdot f$$

On the other hand, we have the following:

$$\sum_{j=1}^{\infty} \int_B \phi_j \cdot f = \lim_{M \rightarrow \infty} \sum_{j=1}^M \int_B \phi_j \cdot f = \lim_{M \rightarrow \infty} \int_B \sum_{j=1}^M \phi_j \cdot f \leq \text{ext} \int_B f$$

Combine the two inequalities and apply to f_+ and f_- to get the result for this Lemma. \square

Combining the results from above to complete the proof of Theorem 13.3:

For $\vec{y} \in B$, choose an open subset $U_{\vec{y}}$ of B such that g can be factorized into Type-1 and Type-2 diffeomorphisms on $U_{\vec{y}}$, then choose a partition of unity (ϕ_j) dominated by $\{U_{\vec{y}} \mid \vec{y} \in B\}$. Now we can write the following:

$$\begin{aligned} \text{ext} \int_B f_+ &= \sum_{j=1}^{\infty} \int_A (\phi_j \circ g)(f_+ \circ g) |\det Dg| = \text{ext} \int_A (f_+ \circ g) |\det Dg| \\ \text{ext} \int_B f_- &= \sum_{j=1}^{\infty} \int_A (\phi_j \circ g)(f_- \circ g) |\det Dg| = \text{ext} \int_A (f_- \circ g) |\det Dg| \end{aligned}$$

Combining the results for f_+ and f_- to get the desired results for Theorem 13.3. \square

Note: In the settings of Theorem 13.3, suppose we have $g : A \rightarrow B \quad \vec{x} \mapsto r\vec{x}$ for some $r > 0$.

$$\text{ext} \int_{\vec{x} \in B} f(\vec{x}) = \text{ext} \int_{\vec{x} \in A} f(r\vec{x}) \cdot r^n$$

Definition 13.5.1.0.1

The number π is defined to be the following:

$$\pi := \lambda_2 := V(B_1^2(\vec{0}))$$

Example: Let $B_r^n(\vec{0}) := \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\| < r\}$, we have $V(B_r^n(\vec{0})) = r^n V(B_1^n(\vec{0})) := r^n \lambda_n$.

Example: We have $\frac{V(B_1^n(\vec{0}))}{V([-1,1]^n)} \leq 1$.

$$\lambda_n = \int_{B_1^n(\vec{0})} 1 = \int_{\vec{x} \in B_1^n(\vec{0})} \int_{\substack{\vec{y} \in B_1^{n-k} \\ \sqrt{1-\|\vec{x}\|^2}}} 1 = \int_{\vec{x} \in B_1^n(\vec{0})} (1 - \|\vec{x}\|^2)^{\frac{n-k}{2}} \lambda_{n-k}$$

Here we choose $k = 2$, then we can rewrite:

$$\begin{aligned} \lambda_n &= \lambda_{n-2} \int_{B_1^2(\vec{0})} (1 - \|\vec{x}\|^2)^{\frac{n}{2}-1} \\ &= \lambda_{n-2} \int_{B_1^2(\vec{0}) \setminus ((-1,0] \times \{0\})} (1 - \|\vec{x}\|^2)^{\frac{n}{2}-1} + \lambda_{n-2} \int_{(-1,0] \times \{0\}} (1 - \|\vec{x}\|^2)^{\frac{n}{2}-1} \end{aligned}$$

Here $\lambda_{n-2} \int_{(-1,0] \times \{0\}} (1 - \|\vec{x}\|^2)^{\frac{n}{2}-1} = 0$ by Theorem 11.13 from Munkres. Let $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ $(r, \theta) \mapsto (r \cos(\theta), r \sin(\theta))$. Notice that g is a local diffeomorphism and $\det Dg(r, \theta) = r$.

$$\begin{aligned} \lambda_n &= \lambda_{n-2} \int_{B_1^2(\vec{0}) \setminus ((-1,0] \times \{0\})} (1 - \|\vec{x}\|^2)^{\frac{n}{2}-1} \\ &= \lambda_{n-2} \int_{\{r|0 < r < 1\} \times \{\theta|-\pi < \theta < \pi\}} (1 - r^2)^{\frac{n}{2}-1} r \\ &= \lambda_{n-2} \int_0^1 \int_{-\pi}^{\pi} (1 - r^2)^{\frac{n}{2}-1} r d\theta dr \\ &= 2\pi \lambda_{n-2} \int_0^1 (1 - r^2)^{\frac{n}{2}-1} r dr \\ &= \pi \lambda_{n-2} \int_0^1 t^{\frac{n}{2}-1} dt \\ &= \frac{2\pi \lambda_{n-2}}{n} \end{aligned}$$

with $t = 1 - r^2$ and $dt = 2r dr$. It follows that, for $u \in \mathbb{N}$, we have:

$$\lambda_{2u} = \frac{\pi^u}{u!} \quad \lambda_{2u+1} = \frac{2^{u+1} \pi^u}{(2u+1)(2u-1)(2u-3)\cdots(3)(1)}$$

Exercise: For $n \in \mathbb{N}$, let $\mu_n = V([-1, 1]^n) = 2^n$, and let κ_n be defined as:

$$\kappa_n = V \left(\left\{ \vec{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n |x_i| < 1 \right\} \right) = \frac{2^n}{n!}$$

We get $\kappa_n \leq \lambda_n \leq \mu_n$, $\lim_{n \rightarrow \infty} \frac{\lambda_n}{\mu_n} = 0$, and $\lim_{n \rightarrow \infty} \frac{\kappa_n}{\lambda_n} = 0$.

Exercise: Let Ω be a rectifiable subset of \mathbb{R}^n , let h be an isometry that maps Ω to $h(\Omega)$, we have:

$$V(h(\Omega)) = \int_{h(\Omega)} 1 = \int_{\Omega} 1 \cdot |\det Dh| = V(\Omega)$$

Definition 13.5.1.0.2

Let A be a rectifiable subset of \mathbb{R}^n with positive volume, let $f : A \rightarrow \mathbb{R}$ be an integrable function.

$$\text{avg}_A f = \frac{\int_A f}{V(A)}$$

Note: In the setting of Definition 13.5.1.0.2, we have $\inf_A f \leq \text{avg}_A f \leq \sup_A f$.

Definition 13.5.1.0.3

Let A be a rectifiable subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}^n \quad \vec{x} \mapsto (f_1(\vec{x}), f_2(\vec{x}), \dots, f_n(\vec{x}))$ be an integrable function.

$$\int_A f := \begin{bmatrix} \int_A f_1 \\ \int_A f_2 \\ \vdots \\ \int_A f_n \end{bmatrix}$$

Note: In the settings of Definition 13.5.1.0.3, we have the following holds:

$$\text{avg}_A f := \frac{\int_A f}{V(A)} \in \mathbb{R}^n$$

and for a special case where $f : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \vec{x} \mapsto \vec{x}$, $\text{avg}_A f$ is called the centroid of A .

Here we can write:

$$\text{avg}_A f = \begin{bmatrix} \text{avg}_A f_1 \\ \text{avg}_A f_2 \\ \vdots \\ \text{avg}_A f_n \end{bmatrix}$$

Remark: The height of the centroid of the upper half of a unit sphere is greater than the height of the centroid of the upper half of a unit ball.

Volume of Parallelipeds

Definition 14.0.0.1

Let Q be a box in \mathbb{R}^k , and let T be an affine injection from \mathbb{R}^k to \mathbb{R}^n of the form $T(\vec{x}) = A\vec{x} + \vec{b}$ for some matrix A and $\vec{b} \in \mathbb{R}^n$. $T(Q)$ is called a k -paralleliped.

For $k < n$, let Q be a box in \mathbb{R}^k , and let T be an affine injection from \mathbb{R}^k to \mathbb{R}^n of the form $\vec{x} \mapsto A\vec{x} + \vec{b}$ for some matrix A and vector $\vec{b} \in \mathbb{R}^n$. $T(Q)$ is a k -paralleliped. Here we note that the n -volume of the k -paralleliped $T(Q)$ is always zero because $T(Q)$ lies in a k -dimensional subspace of \mathbb{R}^n which has measure zero. We want to find a way to define the k -volume of $T(Q)$, denoted as $V_k(T(Q))$. Here are some expectations for such definition:

1. In the case where $A = \begin{bmatrix} M \\ Z \end{bmatrix}$ for some $k \times k$ matrix M and zero matrix Z , we should have

$$V_k(T(Q)) = |\det M| \cdot V(Q)$$

2. For isometry $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$, we should have $V_k(h(T(Q))) = V_k(T(Q))$

Theorem 14.1 (Theorem 21.2 from Munkres)

For $k \leq n$, let $T : \mathbb{R}^k \rightarrow \mathbb{R}^n$ $\vec{x} \mapsto A\vec{x} + \vec{b}$ be an affine injection for some matrix A and $\vec{b} \in \mathbb{R}^n$. One can pick an orthogonal $n \times n$ matrix B such that we have:

$$B \cdot A = \begin{bmatrix} M \\ Z \end{bmatrix}$$

for some $k \times k$ matrix M and zero matrix Z .

Proof. First k rows of B forms an orthonormal basis for column space of A . Extend such orthonormal basis to an orthonormal basis of \mathbb{R}^n to obtain the remaining rows of B by applying the Gram-Schmidt Process. \square

Given the settings in Theorem 14.1, we have $(BA)^T = [M \ Z]$, and $(BA)^T(BA) = M^T M$. Now let $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ $\vec{y} \mapsto B\vec{y}$, note that h is an isometry, and we have the following being satisfied:

$$\begin{aligned} V_k(T(Q)) &= V_k(h(T(Q))) = |\det(M)| \cdot V(Q) \\ &= \sqrt{\det(M^T M)} \cdot V(Q) \\ &= \sqrt{\det((BA)^T(BA))} V(Q) \\ &= \sqrt{\det(A^T A)} \cdot V(Q) \end{aligned}$$

Lemma 14.1.1

Let A be an $n \times k$ matrix, we have $\ker(A^T A) = \ker(A)$

Proof. It is trivial that we have $\ker(A) \subseteq \ker(A^T A)$. On the other hand, if $A^T A\vec{x} = 0$ for some \vec{x} , then $\|\vec{x}\|^2 = \langle A\vec{x}, A\vec{x} \rangle = \vec{x}^T A^T A\vec{x} = 0 \Rightarrow A\vec{x} = \vec{0}$. This shows that $\ker(A^T A) \subseteq \ker(A)$. \square

Corollary 14.1.1.1

Let A be an $n \times k$ matrix, then $\text{rank}(A^T A) = k - \dim(\ker(A^T A)) = k - \dim(\ker(A)) = \text{rank}(A)$.

Corollary 14.1.1.2

Let A be an $n \times k$ matrix. If $k > n$, then $\det(A^T A) = 0$.

Proof. $k > n$ implies $\text{rank}(A) \leq n < k$, which implies $\text{rank}(A^T A) < k$, hence $\det(A^T A) = 0$. \square

Corollary 14.1.1.3

Let A be an $n \times k$ matrix. If $k < n$, then $\det(A A^T) = 0$.

Lemma 14.1.2

Let A be an $n \times k$ matrix. We have $A^T A \geq 0$.

Proof. For \vec{x} , we have $\vec{x}^T A^T A\vec{x} = \langle A\vec{x}, A\vec{x} \rangle \geq 0$. \square

Corollary 14.1.2.1

Let A be an $n \times k$ matrix. All eigenvalues of $A^T A$ are greater than or equal to 0.

Corollary 14.1.2.2

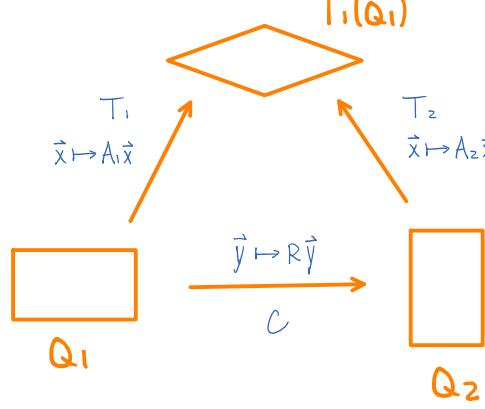
Let A be an $n \times k$ matrix, $\det(A^T A) \geq 0$.

Definition 14.1.2.2.1

Let Q be a box in \mathbb{R}^k , and let T be an affine injection from \mathbb{R}^k to \mathbb{R}^n of the form $\vec{x} \mapsto A\vec{x} + \vec{b}$ for some matrix A and vector $\vec{b} \in \mathbb{R}^n$. We define $V_k(T(Q)) = \sqrt{\det(A^T A)} \cdot V(Q)$.

In the following discussion, we will show that the k -volume of a parallelopiped defined in Definition 14.1.2.2.1 is legitimate. Let Q_1 be a box in \mathbb{R}^k , and let T_1 be an affine injection from \mathbb{R}^k to \mathbb{R}^n of the form $\vec{x} \mapsto A_1\vec{x}$ for some matrix A_1 .

First check independent of choices. Suppose $T_1(Q_1)$ can also be obtained by a composition of affine maps T_2 and C , $T_1(Q_1) = T_2(C(Q_1))$. Here C is a diffeomorphism.



Let A_1 denote the matrix for T_1 , and let A_2 denote the matrix for T_2 . Here we have $A_1 = A_2 R$:

$$\begin{aligned} \sqrt{\det(A_1^T A_1)} \cdot V(Q_1) &= \sqrt{\det((A_2 R)^T (A_2 R))} \cdot V(Q_1) \\ &= \sqrt{\det(R^T A_2^T A_2 R)} \cdot V(Q_1) \\ &= \sqrt{\det(R^T) \det(A_2^T A_2) \det(R)} \cdot V(Q_1) \\ &= \sqrt{\det(A_2^T A_2)} \cdot |\det(R)| \cdot V(Q_1) \\ &= \sqrt{\det(A_2^T A_2)} \cdot V(Q_2) \end{aligned}$$

This shows that the k -volume of $T_1(Q_1)$ is well-defined. Now we will check the expectations for the definition of k -volume of a parallelopiped:

1. If $A_1 = \begin{bmatrix} M \\ Z \end{bmatrix}$ for some $k \times k$ matrix M and zero matrix Z , $V_k(T_1(Q_1)) = |\det M| \cdot V(Q_1)$.
2. For isometry $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$, we should have $V_k(h(T(Q))) = V_k(T(Q))$.

For expectation (1). Suppose we have $A_1 = [M \ Z]^T$, then we see that $\sqrt{\det(A_1^T A_1)} = \sqrt{\det(M^T M)} = |\det(M)|$, which implies $V_k(T_1(Q_1)) = |\det(M)|V(Q_1)$. For expectation (2). Consider an isometry $h : \vec{x} \mapsto B\vec{x} + \vec{p}$ for some orthogonal matrix B , we see that $\det((BA_1)^T (BA_1)) = \det(A_1^T A_1)$, hence it follows that $V_k(h(T_1(Q_1))) = V_k(T_1(Q_1))$.

So we see that there exists a map $V_k : \{k - \text{parallelopipeds}\} \mapsto (0, \infty)$ that satisfies expectation (1) and expectation (2) that can define the k -volume of a parallelopiped, given by the following:

$$V_k(T(Q)) = \sqrt{\det(A^T A)} \cdot V(Q)$$

Definition 14.1.2.2.2

Let A be an $n \times k$ matrix. $\mathcal{V}(A) := \sqrt{\det(A^T A)}$

Theorem 14.2 (Pythagorean Theorem)

For $n \geq k$, let A be an $n \times k$ matrix.

$(\mathcal{V}(A))^2 = \text{the sum of the squares of the determinant of } k \times k \text{ sub-matrices of } A$.

Proof. The proof of this theorem is given in Theorem 21.4 on Munkres. \square

Corollary 14.2.1

Let $\vec{a} = (a_1, a_2, \dots, a_n), \vec{b} = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n$.

$$\left[\text{the one dimensional volume of the line segment from } \vec{a} \text{ to } \vec{b} \right]^2 = \sum_{j=1}^n (b_j - a_j)^2$$

Proof. Proceed by using the affine map $f : [0, 1] \rightarrow \mathbb{R}^n \quad t \mapsto t(\vec{b} - \vec{a}) + \vec{a}$. \square

4 | Manifolds

There are three kinds of non-affine k -dimensional objects.

1. Parametrized k -manifold, discussed in section 22 on Munkres
2. k -manifold in \mathbb{R}^n , discussed in section 23 on Munkres
3. Abstract k -manifold, discussed in section 41 on Munkres

Parametrized Manifolds

Definition 15.0.0.0.1

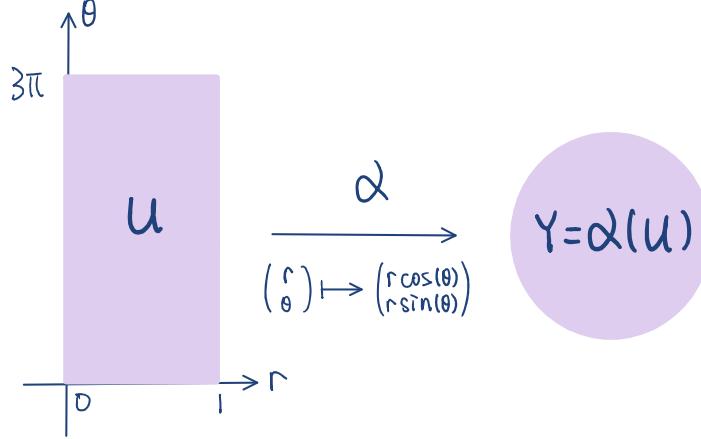
Let $k, n, r \in \mathbb{N}$ with $k \leq n$, let $\alpha \in C^r(U, \mathbb{R}^n)$, where U is an open subset of \mathbb{R}^k . The set $Y := \alpha(U)$, equipped with the map α , constitute a parametrized k -manifold of class C^r , denoted as Y_α . The k -volume of Y_α is defined by the following:

$$V_k(Y_\alpha) := \text{ext} \int_U \mathcal{V}(D\alpha)$$

Remark: In the settings of Definition 15.0.0.0.1, for a special case where we have $k = n$, we will have $\mathcal{V}(D\alpha) = |\det(D\alpha)|$. Suppose further Y is an open rectifiable subset of \mathbb{R}^n , and $\alpha : U \rightarrow Y$ is a diffeomorphism, then we have:

$$V_n(Y_\alpha) = \int_U |\det(D\alpha)| = \int_Y 1 = V(Y)$$

In such case, we consider an example where $U = (0, 1) \times (0, 3\pi)$, and α defined by the following:



Here we have: $V_2(Y_\alpha) = \int_U r = \int_0^1 \int_0^{3\pi} r d\theta dr = \frac{3}{2}\pi$.

Note: Let U and V be open sets, $\alpha \in C^r(U, \mathbb{R}^n)$, $\beta \in C^r(V, \mathbb{R}^n)$ with $\alpha(U) = \beta(V)$. If there is a diffeomorphism g from U to V , then we can write:

$$\begin{aligned} V_k(Y_\alpha) &= \int_U \sqrt{\det((D(\beta \circ g))^T D(\beta \circ g))} \\ &= \int_U \sqrt{\det((Dg)^T ((D\beta)^T \circ g) (D\beta) \circ g) Dg} \\ &= \int_U \sqrt{\det(Dg^T) \cdot \det(((D\beta)^T \circ g)(D\beta \circ g)) \cdot \det(Dg)} \\ &= \int_U \sqrt{\det(((D\beta)^T \circ g)(D\beta \circ g))} |\det(Dg)| \\ &= \int_V \sqrt{\det((D\beta)^T (D\beta))} \\ &= \int_V \mathcal{V}(D\beta) \\ &= V_k(Y_\beta) \end{aligned}$$

Definition 15.0.0.2

Let Y_α be a parametrized k -manifold of class C^r with $Y = \alpha(U)$. Given $f \in C(Y, \mathbb{R})$, we define:

$$\int_{Y_\alpha} f dV := \text{ext} \int_U (f \circ \alpha) \psi(D\alpha)$$

Exercise: Let U and V be open sets, $\alpha \in C^r(U, \mathbb{R}^n)$, $\beta \in C^r(V, \mathbb{R}^n)$ with $\alpha(U) = \beta(V)$. If there is a diffeomorphism g from U to V , we have $\int_{Y_\alpha} f dV = \int_{Y_\beta} f dV$ for $f \in C(Y, \mathbb{R})$.

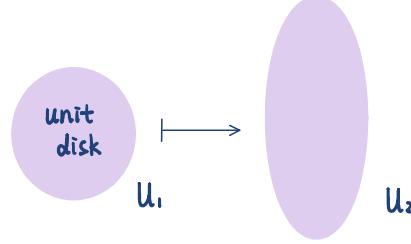
Note: Let Y_α be a parametrized k -manifold of class C^r , let $h : Y \rightarrow Z$ $\vec{z} \mapsto B\vec{z} + \vec{h}$ be an isometry with orthogonal matrix B and $Z = h(Y)$, and let $f : Z \rightarrow \mathbb{R}$ be a continuous function. Here $Z_{h \circ \alpha}$ defines a parametrized k -manifold, we can write the following:

$$\int_{Z_{h \circ \alpha}} f dV = \int_U (f \circ h \circ \alpha) V(B \cdot D\alpha) = \int_U (f \circ h \circ \alpha) V(D\alpha) = \int_{Y_\alpha} (f \circ h) dV$$

Consider $\delta : Y \rightarrow M$ $\vec{x} \mapsto r\vec{x}$ for some $r \in \mathbb{R}$ with $M = \delta(Y)$. $M_{\delta \circ \alpha}$ defines a parametrized k -manifold, let $f : M \rightarrow \mathbb{R}$ be a continuous function, we can write the following:

$$\int_{M_{\delta \circ \alpha}} f dV = r^k \int_{Y_\alpha} f \circ \delta dV$$

Example:



U_2 is the image of the function $\alpha : U_1 \rightarrow U_2$ $(x, y) \mapsto (x, 2y)$. Here we can write:

$$V_2(U_2) = 2V_2(U_1) \quad V_1(Bd(U_1)) = 2\pi$$

Manifolds without Boundaries

Definition 16.0.0.0.1

Given $r, k, n \in \mathbb{N}$, a set $M \subseteq \mathbb{R}^n$ is called a k -manifold without boundary of class C^r provided that for all $\vec{p} \in M$, there exist a set $V \subseteq M$ that contains \vec{p} , a set $U \subseteq \mathbb{R}^k$, with V being open in M and U being open in \mathbb{R}^k , and a homeomorphism $\alpha \in C^r(U, V)$, with $\text{rank}(D\alpha(\vec{x})) = k$ for all $\vec{x} \in U$. The map α is called a coordinate patch on M about \vec{p} .

Example: Consider the set $M = \{(x, y) \mid x^2 = y^3\}$, and the map $\alpha : \mathbb{R} \rightarrow M \quad t \mapsto (t^3, t^2)$, here α is a homeomorphism, but $\alpha'(0) = 0$, so we have $\text{rank}(\alpha'(0)) = 0$. With such α , M is not a 1-manifold, but $M \setminus \{(0, 0)\}$ is a 1-manifold.

Example: Consider the set $M = \{(x, y) \mid y = |x|\}$, and the map $\alpha : \mathbb{R} \rightarrow M \quad t \mapsto (t^3, |t|^3)$, here α is a homeomorphism, but $\alpha'(0) = 0$, so we have $\text{rank}(\alpha'(0)) = 0$. With such α , M is not a 1-manifold.

Example: Consider the set $M = \{(x, y) \mid y = x^2\}$. The map $\alpha : \mathbb{R} \rightarrow M \quad t \mapsto (t^3, t^6)$ defines a homeomorphism, but $\alpha'(0) = 0$. The map $\tilde{\alpha} : \mathbb{R} \rightarrow M \quad t \mapsto (t, t^2)$ also defines a homeomorphism, and we have $\tilde{\alpha}'(x) \neq 0$ for all $x \in \mathbb{R}$, here we found $\tilde{\alpha}$ that satisfies the conditions that makes M a 1-manifold, and $\tilde{\alpha}$ is called a coordinate patch on M

Example: Check Section 23 Example 3 on Munkres Page 198.

Example: Consider the set $M = \{(x, y) \mid x^2 + y^2 = 1\}$. Note that continuous functions map compact set to compact set, here M is compact, hence we cannot handle M with a single coordinate patch, but we can make M a 1-manifold without boundary using the following four coordinate patches to cover all points on M :

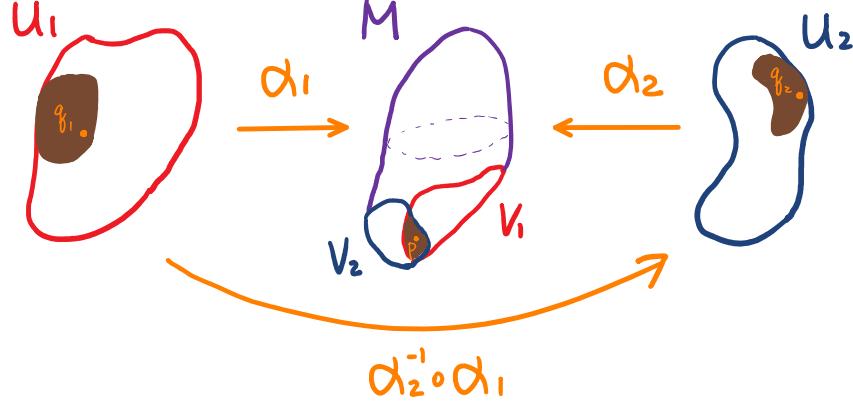
$$\alpha_1 : (-1, 1) \rightarrow M \quad x \mapsto \left(x, \sqrt{1 - x^2} \right) \qquad \alpha_2 : (-1, 1) \rightarrow M \quad y \mapsto \left(\sqrt{1 - y^2}, y \right)$$

$$\alpha_3 : (-1, 1) \rightarrow M \quad y \mapsto \left(-\sqrt{1 - y^2}, y \right) \qquad \alpha_4 : (-1, 1) \rightarrow M \quad x \mapsto \left(x, -\sqrt{1 - x^2} \right)$$

Example: Consider the set $M = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\| = 1\}$. One might use the following $2n$ coordinate patches to make M a $n - 1$ -manifold:

$$\begin{aligned} \alpha_1 : \mathbb{R}^{n-1} &\rightarrow \mathbb{R}^n \quad (x_1, x_2, \dots, x_{n-1}) \mapsto \left(x_1, x_2, \dots, x_{n-1}, \sqrt{1 - \sum_{i=1}^{n-1} x_i^2} \right) \\ \alpha_2 : \mathbb{R}^{n-1} &\rightarrow \mathbb{R}^n \quad (x_1, x_2, \dots, x_{n-1}) \mapsto \left(x_1, x_2, \dots, x_{n-1}, -\sqrt{1 - \sum_{i=1}^{n-1} x_i^2} \right) \\ \alpha_3 : \mathbb{R}^{n-1} &\rightarrow \mathbb{R}^n \quad (x_1, x_2, \dots, x_{n-1}) \mapsto \left(x_1, x_2, \dots, x_{n-2}, \sqrt{1 - \sum_{i=1}^{n-1} x_i^2}, x_{n-1} \right) \\ \alpha_4 : \mathbb{R}^{n-1} &\rightarrow \mathbb{R}^n \quad (x_1, x_2, \dots, x_{n-1}) \mapsto \left(x_1, x_2, \dots, x_{n-2}, -\sqrt{1 - \sum_{i=1}^{n-1} x_i^2}, x_{n-1} \right) \\ &\vdots \\ \alpha_{2n-1} : \mathbb{R}^{n-1} &\rightarrow \mathbb{R}^n \quad (x_1, x_2, \dots, x_{n-1}) \mapsto \left(\sqrt{1 - \sum_{i=1}^{n-1} x_i^2}, x_1, x_2, \dots, x_{n-1} \right) \\ \alpha_{2n} : \mathbb{R}^{n-1} &\rightarrow \mathbb{R}^n \quad (x_1, x_2, \dots, x_{n-1}) \mapsto \left(-\sqrt{1 - \sum_{i=1}^{n-1} x_i^2}, x_1, x_2, \dots, x_{n-1} \right) \end{aligned}$$

Let M be a manifold. For $p \in M$, there might be two coordinate patches for p , $\alpha_1 : U_1 \rightarrow V_1$ and $\alpha_2 : U_2 \rightarrow V_2$, such that $\alpha_1(q_1) = p$ and $\alpha_2(q_2) = p$.



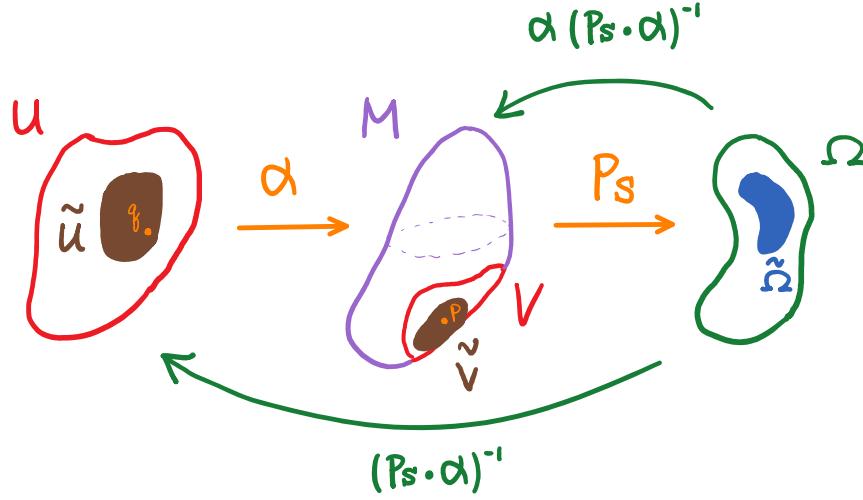
Here we note that $\alpha_2^{-1} \circ \alpha_1$ is of C^r type when defined, and $(\alpha_2^{-1} \circ \alpha_1)^{-1} = \alpha_1^{-1} \circ \alpha_2$ is also of C^r type. Hence we know that $\alpha_2^{-1} \circ \alpha_1$ is a C^r -diffeomorphism.

Theorem 16.1

For $M \subseteq \mathbb{R}^n$, the followings are equivalent:

1. For all $\vec{p} \in M$, there exist a set $U \subseteq \mathbb{R}^k$ open in \mathbb{R}^k , a set $V \subseteq M$ open in M that contains \vec{p} , and a homeomorphism $\alpha \in C^r(U, V)$ with rank $D\alpha(\vec{x}) = k$ for all $\vec{x} \in U$.
2. For all $\vec{p} \in M$, there exist a set $A \subseteq \mathbb{R}^k$ open in \mathbb{R}^k , a set $V \subseteq M$ open in M that contains \vec{p} , a function $g \in C^r(A, \mathbb{R}^{n-k})$, and a coordinate permutation $\rho : \mathbb{R}^n \rightarrow \mathbb{R}^n$, such that we have $\rho(V) = \text{Graph}(g)$.

Proof. First we will show that (2) implies (1). Take $\alpha : A \rightarrow V$ $\vec{x} \mapsto \rho^{-1}(\vec{x}, g(\vec{x}))$. We see that $\alpha^{-1} = (\text{projection onto } \mathbb{R}^k) \circ \rho$ is a continuous function. This completes the proof for this direction. Now we will show that (1) implies (2). For $1 \leq s_1 < s_2 < \dots < s_k \leq n$ where $s_i \in \mathbb{N}$ for all $1 \leq i \leq k$, let $\mathbb{S} = \{s_1, \dots, s_k\}$. Define $P_{\mathbb{S}} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ $\vec{x} = (x_1, x_2, \dots, x_n) \mapsto (x_{s_1}, x_{s_2}, \dots, x_{s_k})$. Let $\vec{q} := \alpha^{-1}(\vec{p})$. Here we can write $D(P_{\mathbb{S}} \circ \alpha)(\vec{q}) = (DP_{\mathbb{S}}(\vec{p})) \cdot D\alpha(\vec{q}) = P_{\mathbb{S}} D\alpha(\vec{q}) = (k \text{ selected row of } D\alpha(\vec{q}))$. Hence we can choose $P_{\mathbb{S}}$ such that $D(P_{\mathbb{S}} \circ \alpha)$ is invertible. Let Ω denote the image of $P_{\mathbb{S}}$. By the Inverse Function Theorem, we can shrink U, V, Ω to $\tilde{U}, \tilde{V}, \tilde{\Omega}$ such that $P_{\mathbb{S}} \circ \alpha$ is a diffeomorphism from \tilde{U} to $\tilde{\Omega}$. Let $h = \alpha \circ (P_{\mathbb{S}} \circ \alpha)^{-1}$, note here we have $P_{\mathbb{S}} \circ h = P_{\mathbb{S}} \circ \alpha \circ (P_{\mathbb{S}} \circ \alpha)^{-1} = \text{identity transformation}$. Let ρ takes \mathbb{S} to $\{1, \dots, k\}$, one can check $\rho(\tilde{V}) = \text{Graph}(h)$. This completes the proof.



Remark: $(P_{\mathbb{S}} \circ \alpha)^{-1} \circ P_{\mathbb{S}} \in C^r(\rho^{-1}(\tilde{\Omega} \times \mathbb{R}^{n-k}), \tilde{U})$. □

In the settings of Theorem 16.1, a region of M can be viewed as an affine set after a dilation $\mathcal{T}_{\vec{p}}$ of the region of M centered at \vec{p} , we denote this affine set by $\vec{p} + \mathcal{T}_{\vec{p}}(M)$, which is called the tangent space of M at \vec{p} . For \vec{x} near \vec{q} , we have:

$$\alpha(\vec{x}) \approx \alpha(\vec{q}) + D\alpha(\vec{q}) \cdot (\vec{x} - \vec{q})$$

We observe that $\mathcal{T}_{\vec{p}}(M)$ is approximately $D\alpha(\vec{q})(\mathbb{R}^k)$. Now we suppose there is another coordinate patch $\tilde{\alpha}$ for $\vec{p} \in M$, with $\tilde{\alpha}(\vec{q}') = \vec{p}$. By Chain Rule, we can write:

$$D\tilde{\alpha}(\vec{q}')(\mathbb{R}^k) = (D\tilde{\alpha}(\vec{q}'))(D(\tilde{\alpha}^{-1} \circ \alpha)(\vec{q}))(\mathbb{R}^k) = D\alpha(\vec{q})(\mathbb{R}^k)$$

Theorem 16.2

Let U be an open subset of \mathbb{R}^n , let $F \in C^r(U, \mathbb{R}^{n-k})$, let $M = F^{-1}(\vec{0})$. If $\text{rank}(DF(\vec{x})) = n - k$ for all $\vec{x} \in M$. Then M is a k -manifolds without boundary of class C^r .

Proof. One can apply the Implicit Function Theorem, and show that M is locally a coordinate-permuted C^r graph. Theorem 24.4 on Munkres provides a very similar proof. \square

Note: The converse of Theorem 16.2 might fail.

Note: In the settings of Theorem 16.2 and Theorem 16.1, we have $F(\alpha(\vec{q})) = 0$ for all \vec{q} in the domain of α , hence by Chain Rule, we can write the following:

$$D(F \circ \alpha)(\vec{q}) = DF(\vec{p}) \cdot D\alpha(\vec{q}) = 0 \quad \Rightarrow \quad \mathcal{T}_{\vec{p}}(M) = D\alpha(\vec{q})(\mathbb{R}^k) \subseteq \ker(DF(\vec{p}))$$

Here $D\alpha(\vec{q})(\mathbb{R}^k)$ has dimension k , and $\ker(DF(\vec{p}))$ has dimension $n - (n - k) = k$, so we have:

$$\mathcal{T}_{\vec{p}}(M) = \ker(DF(\vec{p}))$$

Exercise: $F : \mathbb{R}^n \rightarrow \mathbb{R}$ $\vec{x} \mapsto \|\vec{x}\|^2 - 1$. $F^{-1}(\vec{0}) := S^{n-1}$.

Here we have $DF(\vec{p}) = 2 \cdot \vec{p}^T \Rightarrow \ker(DF(\vec{p})) = \{\vec{p}\}^\perp = \mathcal{T}_{\vec{p}}(S^{n-1})$.

Remark: Let U be an open subset of \mathbb{R}^{n+k} , let $f \in C^1(U, \mathbb{R}^n)$ and $h \in C^1(U, \mathbb{R})$, with $Df = n$. Each $f^{-1}(\vec{a})$ is a k -manifold without boundary by Theorem 16.2. Say we want to maximize, or minimize, h under the constraint $f^{-1}(\vec{a})$. Lagrange Multiplier Theorem states that, if h is maximized, or minimized at \vec{x} , we must have:

$$f(\vec{x}) = \vec{a} \quad \text{with} \quad Dh(\vec{x}) = \vec{\lambda} \cdot Df(\vec{x})$$

with some row vector $\vec{\lambda}$. If we vary \vec{a} , one might view the pair $(\vec{x}, \vec{\lambda})$ as a function of \vec{a} locally via Implicit Function Theorem. In the case where $n = 1$, $\vec{a} \in \mathbb{R}$, here we denote \vec{a} as scalar a , by argument above, \vec{x} and λ are functions of a . Here we have $f(\vec{x}(a)) = a$, with the following holds:

$$\frac{d}{da} h(\vec{x}(a)) = Dh(\vec{x}(a)) \cdot \vec{x}'(a) = \lambda(a) Df(\vec{x}(a)) \cdot \vec{x}'(a) = \lambda(a) \cdot \frac{d}{da} (f(\vec{x}(a))) = \lambda(a)$$

Manifold which might have Boundaries

Definition 17.0.0.0.1

For $S \subseteq \mathbb{R}^k$, $g : S \rightarrow \mathbb{R}^n$. For some $r \in \mathbb{N}$. g is of C^r type provided that there exists open subset of \mathbb{R}^k that contains S , and $G \in C^r(U, \mathbb{R}^n)$ such that $G|_S = g$. If such G exists, then G extends g , and we write $g \in C^r(S, \mathbb{R}^n)$.

Note: Let $S \subseteq \mathbb{R}^k$, for $g \in C^r(S, \mathbb{R}^n)$, the extension G of g is not unique in general. For $|\alpha| \leq r$, we have $D^\alpha g := D^\alpha G$ on $\text{Int}(S)$ and on $\overline{\text{Int}(S)} \cap S$. Suppose we have $g(S) \subseteq T \subseteq \mathbb{R}^n$, and let $h \in C^r(T, \mathbb{R}^m)$. We have $h \circ g \in C^r(S, \mathbb{R}^m)$. The extension G of g can be obtained through Whitney Extension Theorem, or, sometimes, simply by using the Taylor Polynomial.

Definition 17.0.0.0.2

$$\mathbb{H}_k := \{(x_1, x_2, \dots, x_k) \in \mathbb{R}^k \mid x_k \geq 0\} \quad \mathbb{H}_+^k := \{(x_1, x_2, \dots, x_k) \in \mathbb{R}^k \mid x_k > 0\}$$

Definition 17.0.0.0.3

Given $r, k, n \in \mathbb{N}$, a set $M \subseteq \mathbb{R}^n$ is called a k -manifold of class C^r provided that, for all $\vec{p} \in M$, there exist a subset U of \mathbb{R}^k open in either \mathbb{R}^k or \mathbb{H}^k , a subset V of M open in M , and a homeomorphism $\alpha \in C^r(U, V)$ with $\text{rank}(D\alpha(\vec{x})) = k$ for all $\vec{x} \in U$. If such α exists for $\vec{p} \in M$, then α is called the coordinate patch on M about \vec{p} , and M is also called a C^r manifold which might have boundary.

Note: Let M be a C^r manifold which might have boundary, for $\vec{p} \in M$, let α_1 and α_2 be coordinate patches on M about \vec{p} . Then α_1^{-1} is a C^r function. $\alpha_2^{-1} \circ \alpha_1$ is a C^r diffeomorphism between the domain of $\alpha_2^{-1} \circ \alpha_1$ and domain of $\alpha_1^{-1} \circ \alpha_2$.

Let U and W be relatively open in \mathbb{H}_k , let $\gamma : U \rightarrow W$ be a diffeomorphism.

Then $U \cap \mathbb{H}_+^k$ is open in \mathbb{R}^k . $D\gamma(\vec{x})$ is invertible for all $\vec{x} \in U \cap \mathbb{H}_+^k$. By the Inverse Function Theorem, $\gamma(U \cap \mathbb{H}_+^k)$ is open in \mathbb{H}^k and must be contained in \mathbb{H}_+^k . Here we have $\gamma(U \cap \mathbb{H}_+^k) \subseteq (W \cap \mathbb{H}_+^k)$. With similar argument applying to γ^{-1} , we get $\gamma(U \cap \mathbb{H}_+^k) = W \cap \mathbb{H}_+^k$, so we have $\gamma(U \cap (\mathbb{R}^{k-1} \times \{0\})) = W \cap (\mathbb{R}^{k-1} \times \{0\})$.

Definition 17.0.0.0.4

Let M be a k -manifold. For $\vec{p} \in M$, \vec{p} is called a boundary point of M provided that there exists a coordinate patch $\alpha : U \rightarrow V$ on M about \vec{p} such that U is open in \mathbb{H}^k , V is open in M , and $\vec{p} = \alpha((x_1, x_2, \dots, x_{k-1}, 0))$. The set of boundary points of M is called the manifold boundary of M . denoted as ∂M . For $\vec{q} \in M \setminus \partial M$, \vec{q} is called an interior point of M .

Lemma 17.0.1

Let M be a k -manifold in \mathbb{R}^n , and let $\alpha : U \rightarrow V$ be a coordinate patch about the point $\vec{p} \in M$.

If U is open in \mathbb{H}_k and $\vec{p} = \alpha(\vec{x})$ for some $\vec{x} \in \mathbb{H}_+^k$, then \vec{p} is an interior point of M .

If U is open in \mathbb{R}^k , then \vec{p} is an interior point of M .

Theorem 17.1 (Theorem 24.4 on Munkres)

Let M be a k -manifold, ∂M is a C^r ($k-1$)-manifold without boundary.

Proof. For coordinate patch $\alpha : U \rightarrow V$ where $V \subseteq M$, it induces coordinate patch defined by:

$$\hat{\alpha} : \{(x_1, x_2, \dots, x_{k-1}) \mid (x_1, x_2, \dots, x_{k-1}, 0) \in U\} \rightarrow \partial M \\ (x_1, x_2, \dots, x_{k-1}) \mapsto \alpha(x_1, x_2, \dots, x_{k-1}, 0)$$

□

Exercise: For k -manifold M_1 and M_2 in \mathbb{R}^n . If there exists a diffeomorphism $\alpha : M_1 \rightarrow M_2$, then $\alpha|_{\partial M_1}$ is a diffeomorphism from ∂M_1 to ∂M_2 .

Example: Let M be a k -manifold in \mathbb{R}^n . Consider the extreme case where we have $k = n$, n -manifold without boundary is an open subset of \mathbb{R}^n . n -manifold which might have boundary is an open subset of \mathbb{R}^n union with a set ∂M that is an $n-1$ -manifold without boundary, and ∂M is contained in the topological boundary of M .

Example: Consider a closed triangle in \mathbb{R}^2 , which is not a 2-manifold, because the boundary of the closed triangle fails to be a 1-manifold. If we take out the vertices of the closed triangle, that becomes a 2-manifold. If we take out all boundary points of the closed triangle, then we get a 2-manifold without boundary.

Here we extend the definition for k -manifolds in \mathbb{R}^n for the case where $k = 0$:

Definition 17.1.0.0.1

A countable collection of points in \mathbb{R}^n is called a 0-manifold in \mathbb{R}^n .

Remark: 0-manifold M has $\partial M = \emptyset$. M is a 0-manifold if and only if each singleton subset of M is relatively open, that is, if and only if each point of M is isolated.

Note: A set M is a compact 0-manifold $\iff M$ is a finite set. For such we write

$$\int_M f dV := \sum_{x \in M} f(x)$$

Note: All 0-manifolds are countable sets.

Exercise: For open subset U of a k -manifold M , U is a k -manifold.

Theorem 17.2

Every k -manifold M can be decomposed uniquely as a disjoint union of open connected k -manifolds, which are called the components of M .

Proof. The proof of this theorem can be found on Math 395 Supplement materials Section H. \square

Theorem 17.3

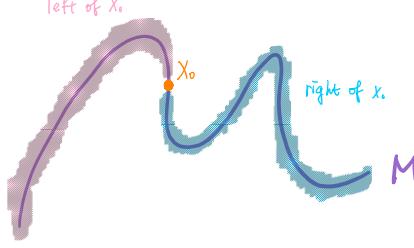
Every connected 1-manifold of class C^r is C^r -diffeomorphic to an interval in \mathbb{R} or to the circle S^1 .

Outline of the proof. Suppose M is a connected 1 manifold. Let $x_0 \in M \setminus \partial M$. One might show:

Exercise: For any $x_1 \in M \setminus \{x_0\}$, there is $I \subseteq M$ such that I is homeomorphic to closed interval with $\partial I = \{x_0, x_1\}$, this can be viewed as a path connecting x_0 and x_1 , one can mimic the proof of Math 395 HW3 Q4 to construct such path.

Now consider the following two cases:

(1) There exists only one such I connecting x_0 and x_1 for each $x_1 \in M \setminus \{x_0\}$. Denote such I as I_{x_0, x_1} . Consider a partition of $M \setminus \{x_0\}$ into two subsets, one is called the left of x_0 and the other one is called the right of x_0 . $x_1 \in M \setminus \{x_0\}$ belongs to left or right of x_0 according to whether I_{x_0, x_1} is contained in the left of x_0 union with $\{x_0\}$, or the right of x_0 union with $\{x_0\}$, as defined by the following:



Now consider a function $f : M \rightarrow \mathbb{R}$ defined by the following:

$$f : M \rightarrow \mathbb{R} \quad x_1 \mapsto \begin{cases} 0 & x_1 = x_0 \\ V_1(I_{x_0, x_1}) & x_1 \text{ lies on the right of } x_0 \\ -V_1(I_{x_0, x_1}) & x_1 \text{ lies on the left of } x_0 \end{cases}$$

One can check that $f : M \rightarrow \mathbb{R}$ is continuous, and here $f(M)$ is connected because M is connected, then we know that $f(M)$ must be an interval. Consider the following exercise:

Exercise: $\{y \in f(M) \mid \#(f^{-1}(y)) = 1\}$ is open in $f(M)$, closed in $f(M)$, and contains 0.

with this exercise, we know that $f : M \rightarrow f(M)$ is a continuous bijection. Consider coordinate patch $\alpha : U \rightarrow V \subseteq M$ with U being an interval and $[t_0, t_1] \subseteq U$, we have $f(\alpha(t_2))$ either being greater than or less than $f(\alpha(t_1))$. WLOG, we suppose $f(\alpha(t_2)) > f(\alpha(t_1))$. We can write the following:

$$\int_{[t_1, t_2]} \|D\alpha\| = \int_{[t_1, t_2]} \mathcal{V}(D\alpha) = V_1(\alpha([t_1, t_2])) = f(\alpha(t_2)) - f(\alpha(t_1))$$

Then by the Fundamental Theorem of Calculus, we see that $D(f \circ \alpha) = \|D\alpha\|$. From the definition of coordinate patch, we know that $D(f \circ \alpha) = \|D\alpha\| > 0$. Here $D\alpha$ and $\|D\alpha\|$ are both C^{r-1}

functions, note that the norm function is of C^∞ type except at $\vec{0}$, but $D\alpha$ is nonzero. Then we know that $D(f \circ \alpha)$ is of C^{r-1} type and hence $f \circ \alpha$ is of C^r type. Hence $f = (f \circ \alpha) \circ \alpha^{-1}$ is also of C^r type. Since $D(f \circ \alpha) \neq 0$, then by Inverse Function Theorem, we know that $(f \circ \alpha)^{-1}$ is of C^r type, and hence $f^{-1} = \alpha \circ (f \circ \alpha)^{-1}$ is also of C^r type. This completes the proof for case (1).

Now consider case (2), in which case there exists $I_1, I_2 \subseteq M$ that are homeomorphic to closed intervals in \mathbb{R}^2 with $\partial I_1 = \{x_0, x_1\} = I_2$ and $I_1 \neq I_2$, for some $\vec{x}_1 \in M \setminus \{x_0\}$. Without loss of generality, here we assume $I_1 \not\subseteq I_2$. We will make use of the following exercises:

Exercise: $I_1 \setminus I_2$ is relatively open in $I_1 \setminus \partial I_1$, relatively closed in $I_1 \setminus \partial I_1$, and nonempty.

From this exercise, we know that $I_1 \setminus I_2 = I_1 \setminus \partial I_1$, hence we have $I_1 \cap I_2 = \{x_0, x_1\}$.

Exercise: $I_1 \cup I_2$ is relatively open in M , relatively closed in M , and non-empty.

With these exercises, we conclude that $M = I_1 \cup I_2$. From similar constructions in case (1), one can obtain the following C^r -diffeomorphisms:

$$\begin{aligned} f_1 : I_1 &\rightarrow [0, V_1(I_1)] \quad \vec{x}_2 \mapsto V_1(I_{x_0, x_2}) \\ f_2 : I_2 &\rightarrow [-V_1(I_2), 0] \quad \vec{x}_2 \mapsto -V_1(I_{x_0, x_2}) \end{aligned}$$

Let $l_1 := V_1(I_1)$ and $l_2 := V_1(I_2)$, and consider the following map:

$$\mu : M \rightarrow \mathbb{R}^2 \quad \vec{x} \mapsto \begin{cases} \left(\cos \frac{2\pi f_1(\vec{x})}{l_1+l_2}, \sin \frac{2\pi f_1(\vec{x})}{l_1+l_2} \right) & \vec{x} \in I_1 \\ \left(\cos \frac{2\pi f_2(\vec{x})}{l_1+l_2}, \sin \frac{2\pi f_2(\vec{x})}{l_1+l_2} \right) & \vec{x} \in I_2 \end{cases}$$

One can verify that μ defines a diffeomorphism from M to S^1 . \square

Corollary 17.3.1

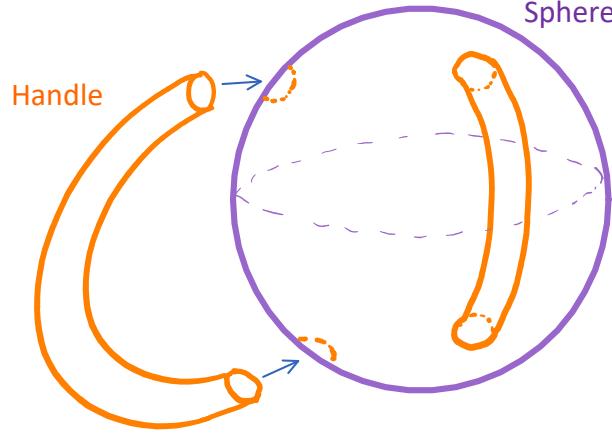
Every connected 1-manifold is diffeomorphic to one of the followings:

1. $(0, 1)$
2. $[0, 1]$
3. $[0, 1]$
4. S^1

Proof. Use the arctangent function and some affine maps. For the uniqueness, we note that diffeomorphism preserves compactness, and also note that if two 1-manifolds are diffeomorphic, then their boundary points of manifolds are also diffeomorphic to each other, hence the manifold boundaries of these two 1-manifolds have the same cardinality. \square

Theorem 17.4

Every compact connected 2-manifold without boundary in \mathbb{R}^3 is diffeomorphic to a sphere in \mathbb{R}^3 , or a sphere glued with some handles.



If a 2-manifold with boundary lives in \mathbb{R}^4 or higher dimensions, one can glue a disk along the boundary of a Möbius strip to get a projective plane.

Theorem 17.5

More generally, every compact connected 2-manifold without boundary in \mathbb{R}^n is diffeomorphic to one of these with handles: (1) a sphere with handles, (2) a projective plane, (3) a Klein bottle.

One can glue two disks to get a sphere S^2 in \mathbb{R}^3 , glue a disk to boundary of Möbius strip to get a projective plane in \mathbb{R}^4 , and glue two Möbius strip to get a Klein Bottle in \mathbb{R}^4 .

One-forms and Pullbacks

Definition 18.0.0.0.1

A 1-form on an open subset A of \mathbb{R}^n is a continuous function from A to \mathbb{R}_{row}^n

Note: Here \mathbb{R}_{row}^n is the space of vectors in \mathbb{R}^n represented by row vector.

Remark: For notation, one might write $(\mathbb{R}^n)^* := \mathbb{R}_{row}^n$.

Example: Let A be an open subset of \mathbb{R}^n . For $f \in C^1(A, \mathbb{R})$, the function Df which maps \vec{x} to $Df(\vec{x})$ is a 1-form. In multivariate differentiation, the operator D takes functions defined on some open subset of \mathbb{R}^n to matrices in $Mat(m, n)$. In the case we $m = 1$, we denote the operator D as d , which takes functions defined on some open subset of \mathbb{R}^n to 1-forms.

Definition 18.0.0.0.2

Let A be an open subset of \mathbb{R}^n , B be an open subset of \mathbb{R}^m , let $\alpha \in C^1(B, A)$, and let ω be an 1-form defined on A . $\alpha^*\omega := (\omega \circ \alpha) \cdot D\alpha$ is called the pullback of ω by α .

Remark: In the settings of definition 18.0.0.0.2, $\omega \circ \alpha$ is an $1 \times n$ matrix and $D\alpha$ is an $n \times m$ matrix, hence $\alpha^*\omega$ is an $1 \times m$ matrix, which is an 1-form on B .

Note: For \mathbb{R} -valued function f defined on an open subset A of \mathbb{R}^n , and $\alpha \in C^1(B, A)$ where B is an open subset of \mathbb{R}^m , we can write the following:

$$\alpha^*df = (df \circ \alpha) \cdot D\alpha = (Df \circ \alpha) \cdot D\alpha = D(f \circ \alpha) = d(f \circ \alpha)$$

Note: Fundamental Theorem of Calculus II for one-forms

For C^1 type function $\alpha : [a, b] = I \rightarrow A$ where A is an open subset of \mathbb{R}^n , with ω being an 1-form on A . Then $\alpha^*\omega$ is a 1-form on I , which can be viewed as \mathbb{R} -valued function. Consider the 1-manifold Y_α , we can define the following:

$$\int_{Y_\alpha} \omega := \int_I \alpha^* \omega$$

For C^1 type function f defined on A , we can write the following:

$$\int_{Y_\alpha} df = \int_I \alpha^* df = \int_I d(f \circ \alpha) = \int_I (f \circ \alpha)' = (f \circ \alpha)(b) - (f \circ \alpha)(a) = f(\alpha(b)) - f(\alpha(a)) := \Delta_{Y_\alpha} f$$

Here we get the results by the Fundamental Theorem of Calculus II.

Example: Let A be an open subset of \mathbb{R}^n . For $f \in C^1(A, \mathbb{R})$, $df : \vec{x} \rightarrow Df(\vec{x})$ is a 1-form. We write:

$$\begin{aligned} df &= \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \\ &= \frac{\partial f}{\partial x_1} [1 \ 0 \ \cdots \ 0] + \frac{\partial f}{\partial x_2} [0 \ 1 \ 0 \ \cdots \ 0] + \cdots + \frac{\partial f}{\partial x_n} [0 \ 0 \ \cdots \ 1] \\ &= \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n \end{aligned}$$

Here $dx_i := d\pi_i$, where $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ $(x_1, x_2, \dots, x_n) \mapsto x_i$ is the projection function. In general, for 1-form defined on A , we write $\omega = [\omega_1 \ \omega_2 \ \cdots \ \omega_n] = \omega_1 dx_1 + \omega_2 dx_2 + \cdots + \omega_n dx_n$.

Definition 18.0.0.0.3

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^1 function, and let $\alpha : \mathbb{R}^m \rightarrow \mathbb{R}^n$, we write $\alpha^{*,0} f = f \circ \alpha$.

Note: In the settings of Definition 18.0.0.3, $\alpha^*df = d(\alpha^{*,0} f)$

Note: Let ω be a 1-form defined on open subset A of \mathbb{R}^n , let B be an open subset of \mathbb{R}^m , let C be open subset of \mathbb{R}^k , and let $\alpha_1 \in C^1(B, A)$, and let $\alpha_2 \in C^1(C, B)$, we have $(\alpha_2 \circ \alpha_1)^*\omega = \alpha_1^*(\alpha_2^*\omega)$

Example: Consider the map $\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $(r, \theta) \mapsto (r \cos(\theta), r \sin(\theta))$. We can write the following:

$$\begin{aligned}\alpha^* \left(\frac{x_1}{x_1^2 + x_2^2} dx_1 + \frac{x_2}{x_1^2 + x_2^2} dx_2 \right) &= \frac{r \cos(\theta)}{r^2} d(r \cos(\theta)) + \frac{r \sin(\theta)}{r^2} d(r \sin(\theta)) \\ &= \frac{r \cos(\theta)}{r^2} (\cos(\theta) dr - r \sin(\theta) d\theta) + \frac{r \sin(\theta)}{r^2} (\sin(\theta) dr + r \cos(\theta) d\theta) \\ &= \frac{1}{r} dr\end{aligned}$$

Similarly, we have:

$$\alpha^* \left(\frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2 \right) = \frac{r \cos(\theta)}{r^2} d(r \cos(\theta)) + \frac{r \sin(\theta)}{r^2} d(r \sin(\theta)) = d\theta$$

Remark: Suppose $\omega : [a, b] \rightarrow \mathbb{R}$ is a 1-form on $[a, b]$. Here we can write the following:

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

Note: Let I, J be intervals of \mathbb{R} such that $\beta : I \rightarrow J$ defines a diffeomorphism.

For $\alpha \in C^1(I, \mathbb{R}^n)$, let ω be a 1-form defined on Y_α , we can write the following:

$$\int_{Y_{\alpha \circ \beta}} \omega = \int_J (\alpha \circ \beta)^* = \int_J \beta^* \alpha^* \omega = \int_J ((\alpha^* \omega) \circ \beta) \cdot D\beta$$

Here we see that $D\beta = \det(D\beta)$ because $D\beta$ is a 1×1 matrix. Also note that, since β defines a diffeomorphism, then $D\beta(x) \neq 0$ for all $x \in J$, and $D\beta$ is continuous, hence $D\beta$ is a negative, or positive function. Then by the Change of Variable Theorem, we can write the following:

$$\int_{Y_{\alpha \circ \beta}} \omega = \int_J ((\alpha^* \omega) \circ \beta) \cdot D\beta = \begin{cases} \int_I \alpha^* \omega & D\beta(x) > 0, \forall x \in J \\ -\int_I \alpha^* \omega & D\beta(x) < 0, \forall x \in J \end{cases}$$

On the other hand, we have:

$$\int_{Y_\alpha} \omega = \int_I \alpha^* \omega$$

Combining, we can write the following:

$$\int_{Y_{\alpha \circ \beta}} \omega = \begin{cases} \int_{Y_\alpha} \omega & D\beta(x) > 0, \forall x \in J \\ -\int_{Y_\alpha} \omega & D\beta(x) < 0, \forall x \in J \end{cases}$$

Note: For 1-form ω defined on open subset A of \mathbb{R}^n , and $\alpha \in C^1([a, b] \rightarrow A)$ we can write:

$$\begin{aligned}\int_{Y_\alpha} \omega &= \int_{Y_\alpha} \omega_1 dx_1 + \omega_2 dx_2 + \cdots + \omega_n dx_n \\ &= \int_I (\omega_1 \circ \alpha) \cdot D\alpha_1 + (\omega_2 \circ \alpha) \cdot D\alpha_2 + \cdots + (\omega_n \circ \alpha) \cdot D\alpha_n\end{aligned}$$

Example: In Physics, a force field can be described by a 1-form ω defined on an open subset A of \mathbb{R}^n . Let $I \subseteq \mathbb{R}$ be an interval, and let $\alpha \in C^1(I, A)$, then α can be viewed as a path of a particle in the \mathbb{R}^n space. The formula $\int_{Y_\alpha} \omega$ describes the work acted on the particle through the path Y_α .

Example: Consider $\alpha : [0, 1] \rightarrow \mathbb{R}^n$ $t \mapsto (t, t^n)$.

Let $\pi_1 : \mathbb{R}^2 \rightarrow \mathbb{R}$ $(x, y) \mapsto x$, $\pi_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ $(x, y) \mapsto y$ be the projection functions, we can write:

$$\int_{Y_\alpha} \pi_2 d\pi_1 - \pi_1 d\pi_2 = \int_0^1 t^n - nt^n = \frac{1-n}{n+1}$$

Proposition 18.0.1

Let $f : [a, b] \rightarrow \mathbb{R}^n$, the followings are equivalent:

1. There exists a function $g \in C^k(A, \mathbb{R}^n)$ such that $g(x) = f(x)$ for $x \in [a, b]$, with A being an open subset of \mathbb{R} that contains $[a, b]$.
2. $f \in C([a, b])$, with $f|_{(a,b)} \in C^k((a, b), \mathbb{R}^n)$, and both $\lim_{t \rightarrow a^+} f^{(j)}(t)$ and $\lim_{t \rightarrow b^-} f^{(j)}(t)$ exist and are finite for $j \in \mathbb{N}$ with $1 \leq j \leq k$.

Proof. It is trivial to show that (1) implies (2). For (2) implies (1), one can use the Taylor Polynomials to extend the function f . \square

Definition 18.0.1.0.1

For interval $[a, b] \subseteq \mathbb{R}$, a function $f : [a, b] \rightarrow \mathbb{R}^n$ is said to be of C^k type provided that there exists a function $g \in C^k(A, \mathbb{R}^n)$ such that $g(x) = f(x)$ for $x \in [a, b]$, with A being an open subset of \mathbb{R} that contains $[a, b]$. In such case, we write $f \in C^k([a, b], \mathbb{R}^n)$.

Definition 18.0.1.0.2

For interval $[a, b] \subseteq \mathbb{R}$, a function $f : [a, b] \rightarrow \mathbb{R}^n$ is said to be of piecewise C^k type provided that f is continuous on $[a, b]$, and $f|_{[t_{j-1}, t_j]} \in C^k([t_{j-1}, t_j], \mathbb{R}^n)$ for each j , where $\{t_1, t_2, \dots, t_m\}$ defines a partition on $[a, b]$. In such case, we write $f \in C_{pw}^k([a, b], \mathbb{R}^n)$.

Example: The function $M : [a, b] \rightarrow \mathbb{R}^2$ defined by the following is of piecewise C^k type.

**Definition 18.0.1.0.3**

Let $[a, b] \subseteq \mathbb{R}$ be an interval, let $\{t_1, t_2, \dots, t_n\}$ defines a partition of $[a, b]$. For piecewise function $\alpha : [a, b] \rightarrow \mathbb{R}^n$ that satisfies $\alpha|_{[t_{j-1}, t_j]}$ being continuous on $[t_{j-1}, t_j]$, we denote $\alpha_{(j)} := \alpha|_{[t_{j-1}, t_j]}$.

Definition 18.0.1.0.4

For $\alpha \in C_{pw}^k([a, b], \mathbb{R}^n)$, $Y := \alpha([a, b])$ is usually denoted as Y_α , and we write the following:

$$\int_{Y_\alpha} \omega = \int_{[a, b]} \alpha^* \omega$$

where ω is a 1-form defined on Y . Moreover, for function f such that df is a 1-form defined on $[a, b]$, we write the following:

$$\int_{Y_\alpha} df = \sum_j \int_{Y_{\alpha(j)}} df = \sum_j \Delta_{Y_{\alpha(j)}} f$$

Theorem 18.1

Let A be an open connected subset of \mathbb{R}^n , let $f \in C^1(A, \mathbb{R})$. $df(x) = 0$ for $x \in A$ if and only if f is a constant function.

Proof. The \Leftarrow direction is trivial. For the \Rightarrow direction, we know that open sets can be decomposed into path connected components, one can choose $\alpha \in C_{pw}^1$ such that $\Delta_{Y_\alpha} f = \int_{Y_\alpha} df = 0 \Rightarrow f(b) = f(a)$. This completes the proof of this theorem. \square

Integration of One-forms on Manifolds

We now have defined two kinds of integrals over a parametrized k -manifold Y_α in \mathbb{R}^n , with coordinate patch α defined on I , one for usual function g and one for 1-form ω :

1. For function $g : Y \rightarrow \mathbb{R}$, we have $\int_{Y_\alpha} g dV = \int_I (g \circ \alpha) \mathcal{V}(D\alpha)$
2. For 1-form ω defined on Y , $\int_{Y_\alpha} \omega = \int_I \alpha^* \omega$

If $I \subseteq \mathbb{R}$, then we can write the following:

$$\int_{Y_\alpha} g dV = \int_I (g \circ \alpha) \sqrt{\det((D\alpha)^T D\alpha)} = \int_I (g \circ \alpha) \|D\alpha\| = \int_I (g \circ \alpha) \|\alpha'\|$$

On the other hand, we have:

$$\int_{Y_\alpha} (\omega \circ \alpha) \cdot D\alpha = \int_I (\omega \circ \alpha) \cdot \alpha'$$

Here we have $\int_{Y_\alpha} g dV = \int_{Y_\alpha} \omega$ provided that we have the following holds:

$$(g \circ \alpha) \|\alpha'\| = (\omega \circ \alpha) \alpha'$$

That is, for all $t \in I$, with $\vec{p} = \alpha(t)$, we must have $g(\vec{p}) \|\alpha'(t)\| = \omega(\vec{p}) \cdot \alpha'(t)$, which is equivalent to the following:

$$g(\vec{p}) = \omega(\vec{p}) \cdot \frac{\alpha'(t)}{\|\alpha'(t)\|}$$

Here we require $\alpha'(t) \neq 0$ and α being injective. If $\alpha \in C^1$ is injective and $\alpha'(t) \neq 0$ for all $t \in I$, then one can set $T : Y \rightarrow \mathbb{R}^n \quad \alpha(t) \mapsto \frac{\alpha'(t)}{\|\alpha'(t)\|}$, and get the followings:

$$\int_{Y_\alpha} (\omega \cdot T) dV = \int_{Y_\alpha} \omega \quad \int_{Y_\alpha} g dV = \int_{Y_\alpha} (g \cdot T^T)$$

Theorem 19.1 (Fundamental Theorem of Calculus I(a) for 1-forms)

Let ω be a 1-form on A , where A is a connected open subset of \mathbb{R}^m . The followings are equivalent:

1. $\omega = df$ for some $f \in C^1(A, \mathbb{R})$, in which case ω is said to be exact on A .
2. For $\alpha \in C_{pw}^1([a, b], A)$ with $\alpha(a) = \alpha(b)$, we have $\int_{Y_\alpha} \omega = 0$.
3. For $\alpha_j \in C_{pw}^1([a_j, b_j], A)$ with $\alpha_1(a_1) = \alpha_2(a_2)$ and $\alpha_1(b_1) = \alpha_2(b_2)$, we have $\int_{Y_{\alpha_1}} \omega = \int_{Y_{\alpha_2}} \omega$, in which case ω is said to be path independent.

Proof. First we note that (1) implies (2) is immediate by previous discussion:

$$\int_{Y_\alpha} df = f(\alpha(b)) - f(\alpha(a)) = 0$$

Details of (2) implying (3) can be found on Theorem M.1 on the Math 395 Supplement materials. The idea is reversing the orientation of α_2 in (3) and form a closed loop to get $\int \omega = 0$. Now we will show (3) implies (1), for $a, b \in A$, one can define the integral $\int_a^b \omega$ with the path independent property of ω such that the following holds:

$$\int_x^y \omega + \int_y^z \omega = \int_x^z \omega$$

for $x, y, z \in A$. Fix $x_0 \in A$, define $f(x) = \int_{x_0}^x \omega$, we will show that we have $df = \omega$ for some $f \in C^1(A, \mathbb{R})$, that is, we want to evaluate the following as $h \in \mathbb{R}^m$ tends to zero:

$$\frac{f(x+h) - f(x) - \omega(x) \cdot h}{\|h\|} \tag{*}$$

Assuming that h is small enough, with $\alpha : [0, 1] \rightarrow A \quad t \mapsto x+th$ such that $\alpha'(t) = h$, we can write:

$$f(x+h) - f(x) = \int_x^{x+h} \omega = \int_0^1 \omega(x+th) \cdot h dt$$

So we can write:

$$(*) = \frac{\int_0^1 (\omega(x+th) - \omega(x)) \cdot h dt}{\|h\|}$$

Here we see that:

$$|(*)| \leq \frac{\int_0^1 \|\omega(x+th) - \omega(x)\| \cdot \|h\| dt}{\|h\|} \leq \max_{0 \leq t \leq 1} \|\omega(x+th) - \omega(x)\|$$

It follows that $\lim_{h \rightarrow 0} (*) = 0$. The results of this theorem follows immediately. \square

Let ω be a one form defined on an open subset A of \mathbb{R}^m . By Corollary 6.2.1, the Clairaut's Theorem, suppose there exists $f \in C^2$ such that $D_j f = \omega_j$ and $D_k f = \omega_k$, here we have $D_k D_j f = D_j D_k f$, so if we have $\omega \in C^1$ and is exact, then $D_k \omega_j = D_j \omega_k$, in which case ω is said to be closed on A .

Definition 19.1.0.0.1

Let A be an open subset of \mathbb{R}^m , let ω be a 1-form defined on A .

ω is said to be closed on A provided that $D_k \omega_j = D_j \omega_k$. ω is said to be exact on A provided that $\omega = df$ for some $f \in C^1(A, \mathbb{R})$. ω is said to be path independent provided that $\int_{Y_{\alpha_1}} \omega = \int_{Y_{\alpha_2}} \omega$ for $\alpha_j \in C_{pw}^1([a_j, b_j], A)$ with $\alpha_1(a_1) = \alpha_2(a_2)$ and $\alpha_1(b_1) = \alpha_2(b_2)$.

Example: Consider the 1-form ω defined on \mathbb{R}^2 that maps (x_1, x_2) to $(-\frac{x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2})$.

One can verify that ω is closed on $\mathbb{R}^2 \setminus \{\vec{0}\}$ with $\omega = d(\arctan(\frac{x_2}{x_1}))$ on $(0, \infty) \times \mathbb{R}$.

For $\alpha : [0, 2\pi] \rightarrow \mathbb{R}^2$ $t \mapsto (\cos(t), \sin(t))$, we can write the following:

$$\begin{aligned} \int_{Y_\alpha} \omega &= \int_0^{2\pi} \alpha^* \omega = \int_0^{2\pi} (-\sin(t))d(\cos(t)) + \cos(t)d(\sin(t)) \\ &= \int_0^{2\pi} (-\sin(t))(-\sin(t))dt + \cos(t)\cos(t)dt \\ &= \int_0^{2\pi} dt \\ &= 2\pi \neq 0 \end{aligned}$$

Here we conclude that ω is not exact on $\mathbb{R}^2 \setminus \{\vec{0}\}$.

Lemma 19.1.1

Let ω be a closed 1-form of C^1 type, and let α be a C^2 type function. Then $\alpha^ \omega$ is closed.*

Proof. Check Theorem 32.3 on Munkres and Lemma M.6 on Math 395 Supplement material \square

Lemma 19.1.2 (Green's Theorem for Two-dimensional Boxes)

Let ω be a 1-form defined on an open set $A \subseteq \mathbb{R}^2$ which contains a box R of \mathbb{R}^2 , then we have:

$$\int_{\substack{Bd(R) \\ \text{counter-clockwise orientation}}} \omega = \int_R (D_1 \omega_2 - D_2 \omega_1)$$

where ω_1, ω_2 are component functions of ω . The counter-clockwise orientation of $Bd(R)$ refers to a path which maps an interval in \mathbb{R} to $Bd(R)$ that goes in counter-clockwise direction on $Bd(R)$.

Proof. Denote $a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2$.

$$\begin{aligned} \int_{Bd(R)} \omega &= \int_{a_1}^{b_1} \omega_1(x_1, a_2) dx_1 + \int_{a_2}^{b_2} \omega_2(b_1, x_2) dx_2 - \int_{a_1}^{b_1} \omega_1(x_1, b_2) dx_1 - \int_{a_2}^{b_2} \omega_2(a_1, x_2) dx_2 \\ &= \int_{a_2}^{b_2} \int_{a_1}^{b_1} D_1 \omega_2(x_1, x_2) dx_1 dx_2 - \int_{a_1}^{b_1} \int_{a_2}^{b_2} D_2 \omega_1(x_1, x_2) dx_2 dx_1 \\ &= \int_R (D_1 \omega_2 - D_2 \omega_1) \end{aligned}$$

The result of this lemma follows. \square

Corollary 19.1.2.1

Let ω be a 1-form on an open set containing a box R of \mathbb{R}^2 , if ω is closed, then we have $\int_{Bd(R)} \omega = 0$.

Theorem 19.2 (Fundamental Theorem of Calculus I(b) for 1-forms)

Let ω be a closed 1-form on $A \subseteq \mathbb{R}^m$. If A is a convex open subset of \mathbb{R}^m , then ω is exact on A .

Proof. Theorem M.5 on the Math 395 Supplement material gives a proof to this theorem. \square

Remark: In the settings of Theorem 19.2, the result also holds when A is C^2 -diffeomorphic to a convex set. Suppose a set A is C^2 -diffeomorphic to a convex set B , let ω be a closed 1-form defined on A . Suppose further that γ defines a C^2 -diffeomorphism from B to A . Then by Lemma 19.1.1, we know that $\gamma^* \omega$ defined on B is closed, and hence $\gamma^* \omega$ is exact on B . That is, there exists some $f \in C^1$ such that $\gamma^* \omega = df$. Then we can write $\omega = (\gamma \circ \gamma^{-1})^* \omega = \gamma^{-1*}(\gamma^* \omega) = \gamma^{-1*}(df) = d(\gamma^{-1*} f)$.

Example: Consider the 1-form ω defined on \mathbb{R}^2 that maps (x_1, x_2) to $(-\frac{x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2})$.

We see that such 1-form is closed but not exact on $\mathbb{R}^2 \setminus \{\vec{0}\}$, hence it follows that $\mathbb{R}^2 \setminus \{\vec{0}\}$ is not diffeomorphic to any convex set.

Theorem 19.3

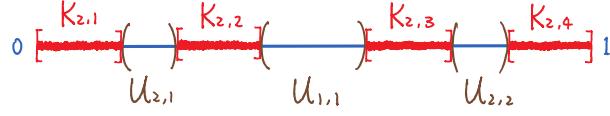
Let $I = [0, 1]$, there exists a continuous surjection $f : I \rightarrow I^2$

Proof. Let $C_1 = [0, 1] \setminus (1/3, 2/3)$, $C_2 = C_1 \setminus ((1/9, 2/9) \cup (7/9, 8/9))$, \dots . We let $C = \bigcap_{j=1}^{\infty} C_j$, which is called the middle-third Cantor set. Now let $U_1 = (1/3, 2/3)$, $U_2 = (1/9, 2/9) \cup (7/9, 8/9)$, \dots . Here each U_j is disjoint, let $U = \bigcup_j U_j = [0, 1] \setminus C$. We know that $m_{*,J}(U) = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots = \frac{1}{3} \frac{1}{1-2/3} = 1$. On the other hand $m^{*,J}(C) = 0$. By previous result from this text, we have $0 \leq m^*(C) \leq m^{*,J}(C)$, so we have $m^*(C) = 0$. For $0 < x < 1$, we have $x \in C$ if and only if x admits unique ternary expansion with no 1's. First we build $C \rightarrow I \times I$ by splitting of alternate digits in the ternary expansion of an element in C , convert all 2's to 1's, and interpret base 2. That is, $g(0.20020022\dots) = (0.1001\dots, 0.0101\dots)$, here g is surjective. To show that g is continuous, suppose $x, y \in C$, we have $|x - y| < 9^{-n}$, first $2n$ ternary digits match, then the first $ndigits$ of g_1, g_2 match, hence we have $\|g(x) - g(y)\| \leq 2^{-n}$. Now construct continuous $f : I \rightarrow I^2$ with $f|_C = g$ by insisting that f be affine on the closure of each component of U .

Exercise: $\|f(x) - f(y)\|_{sup} \leq 18|x - y|^{\frac{\log 2}{2\log 3}}$. That is, f is Holder-continuous with exponent $\frac{\log 2}{2\log 3}$. Theorem from the Supplement, if $h : I \rightarrow I^2$, and if h satisfies $\|h(x) - h(y)\|_{sup} \leq M|x - y|^\alpha$ with $\alpha > \frac{1}{2}$, then $f(I)$ has measure zero. \square

Definition 19.3.0.0.1

Let $\alpha_1, \alpha_2, \alpha_3, \dots \in (0, 1)$. $I = [0, 1] := K_0$.



$l(K_{j,p})$ is independent of the choice of p , with $l(U_{j,p}) = \alpha_j \cdot l(K_{j-1,p})$.

Let $K_j = \bigcup_{i=1}^{2^j} K_{j,i}$, and let $K = \bigcap_j K_j$. K is called the generalized cantor set.

In the settings of the Definition. The generalized cantor set K is compact, with empty interior, $U = [0, 1] \setminus K$, which is the union of all $U_{j,p}$ defined in definition, $U := \bigcup_{i,j} U_{i,j}$ is a dense open set. Pick $1 > n_1 > n_2 > \dots > 0$, one can pick α_j such that $m_{\text{elem}}(K_j) = n_j$, with the following holds:

$$m_{\text{elem}}(I \setminus K_j) = m_{\text{elem}} \left(\bigcup_{s \leq j, 1 \leq p \leq 2^{j-1}} U_{s,p} \right) = 1 - n_j$$

With results from Math 395 IBL6 Q3, we can write the following:

$$m_{*,J}(U) = \lim_{j \rightarrow \infty} 1 - n_j = 1 - \lim_{j \rightarrow \infty} n_j \in (0, 1]$$

With results from Math 395 IBL6 Q1, we can write the following:

$$m^{*,J}(U) = m^{*,J}(\bar{U}) = 1$$

Here U is Jordan measurable if and only if $\lim_{j \rightarrow \infty} n_j = 0$. Note that, between any two components of U , there infinitely many other components of U . Also note that, $Bd(U_{s,p}) \subseteq K$, and we see that the set $\bigcup_{s,p} (Bd(U_{s,p}))$ is countable, but K is not countable because there is a surjection from K to $[0, 1] \times [0, 1]$, so we know that most points of K are not boundary points of some $U_{s,p}$, but we still have $K = Bd(K) = Bd(U)$. Moreover, with the settings given above, we can write the following:

$$n_j = (1 - \alpha_j) = n_{j-1}$$

with induction argument, we conclude:

$$n_j = (1 - \alpha_1)(1 - \alpha_2) \cdots (1 - \alpha_j)$$

Hence we have the following holds:

$$\lim_{j \rightarrow \infty} n_j = \lim_{N \rightarrow \infty} \prod_{j=1}^N (1 - \alpha_j) := \prod_{j=1}^{\infty} (1 - \alpha_j)$$

hence we conclude that we have:

$$m^*(K) = m^{*,J}(K) = 1 - m_{*,J}(U) = 1 - (1 - \lim_{j \rightarrow \infty} n_j) = \lim_{j \rightarrow \infty} n_j$$

$$m^*(K) > 0 \iff \sum_{j=1}^{\infty} \alpha_j < \infty$$

Here we see that if α_j is independent of j , then $m^*(K) = 0$, and in the case where we have $\alpha_j = 2^{-j}$, we also have $m^*(K) = 0$. In fact, each K is compact, each $x \in K$ is the limit of a sequence in $K \setminus \{x\}$. All connected subsets of K are singletons.

Theorem 19.4

Every metric spaces X with the following properties is homeomorphic to the general cantor set K :

1. X is compact
2. Each $x \in X$ is the limit of some sequence in $K \setminus \{x\}$
3. All connected subsets of K are singletons

Given a vector space V , let S be a vector subspace of V , let $\vec{a} \in V$, then $S + \vec{a}$ is an affine subset of V , S is the vector subspace associated to $S + \vec{a}$. Here we define the followings:

Definition 19.4.0.0.1

Given a vector space V , let S be a vector subspace of V , let $\vec{a} \in V$, $V \setminus S :=$ the collection of all $S + \vec{a}$ is called the quotient space.

Exercise: Given a vector space V , let S be a vector subspace of V , $V \setminus S$ is a vector space. One can write the followings:

$$(S + \vec{a}) + (S + \vec{b}) = S + (\vec{a} + \vec{b}) \quad \lambda(S + \vec{a}) = S + \lambda\vec{a}$$

with $\vec{a}, \vec{b} \in V$, and $\lambda \in F$ where F is the field of V .

Definition 19.4.0.0.2

$$Q_s : V \rightarrow V \setminus S \quad \vec{a} \mapsto S + \vec{a}$$

Exercise: Q_s is linear and surjective. $\ker(Q_s) = S$. $T_V \rightarrow U$ is linear with $T|_S 0$, which induces $\tilde{T} : V \setminus S \rightarrow W$, $T = \tilde{T} \circ Q_s$.

k -form on manifold M is a map ω with $\omega(\vec{p}) \in \text{Alt}^k(\mathcal{T}_p M)$ for all $\vec{p} \in M$.

In the following discussion, we assume that all maps are C^∞ type.

When $k > \dim(M)$, then $\text{Alt}^k(\mathcal{T}_p M) = 0$.

There exists operators $d_k : \{k\text{-forms on } M\} \rightarrow \{(k+1)\text{-forms on } M\}$ such that we have (1) $d_{k+1} \circ d_k = 0$, and (2) Stokes' Theorem holds.

Here (1) implies that the image of d_{k+1} is contained in the kernel of d_k . Here the dimension of the quotient $\frac{\ker(d_k)}{\text{im}(d_{k-1})}$ is called the k -th Betti Number $b_k(M)$. Notice that $d_0 f = 0$ if and only if f is a constant function on each component, and here $b_0(M) = \text{number of components}$.

Focus now on compact connected 1-manifold. If we have $d_1 = 0$, then $\ker(d_1)$ is all 1-manifold on M . M is diffeomorphic to $I = [0, 1]$ or S^1 . On I , every 1-form is dg for some 0-form g , so we have $b_1(I) = 0$. On S^1 , one might consider $\beta : \mathbb{R} \rightarrow S^1$ $x \mapsto (\cos(x), \sin(x))$, with ω being 1-form on S^1 , we have $\beta^* S$ being 1-form on \mathbb{R} . Here we can write $\beta^* \omega = f(x) dx$ with $f(x + 2\pi) = f(x)$. $\omega = dg$ on S^1 implies $f(x) dx = \beta^* \omega = \beta^* dg = d\beta^* g = d(g \circ \beta) = (g \circ \beta)'(x)$ on \mathbb{R} . Here we need antiderivative for f of the form $g \circ \beta$, need $h(x + 2\pi) = h(x)$ **Exercise:** $h(x + 2\pi) - h(x) = \int_{S^1} \omega$. **Exercise:** $\{1\text{-forms on } S^1\} \rightarrow \mathbb{R}$ $\omega \mapsto \int_{S^1} \omega$ induces isomorphism from the quotient $\frac{\ker(d_1)}{\text{im}(d_0)} \rightarrow \mathbb{R}$.

For $k = 2$, here we will focus on M compact without boundary. There are the following kinds of such:

- (1) Sphere with g handles, $b_1 = 2g$, $b_2 = 1$.
- (2) Projection plane with g handles, $b_1 = 2g$, $b_2 = 0$.
- (3) Klein bottle with g handles, $b_1 = 2g + 1$, $b_2 = 0$.

Corollary 19.4.0.1

Connected compact 2-manifolds without boundaries classified to diffeomorphism by b_1 and b_2 .

Exercise: (Poincare)

M being the Poincare. M is a topological space. M is a topological manifold. M is a differentiable manifold diffeomorphic to 3-manifold without boundaries in \mathbb{R}^5 .

$$b_k(M) = \begin{cases} 1 & k = 0, 3 \\ 0 & k = 1, 2 \end{cases}$$

Note that this is the same for $b_k(S^3)$, but M is not homeomorphic to S^3 .

Theorem 19.5 (Perelman's Theorem)

Given M being a compact connected 3-manifold without boundary, and given that every continuous map from S^1 to M extends to continuous map B^2 (ball) to M . Then M is diffeomorphic to S^3 .

For $k = 4$, there exists compact connected topological 4-manifold without boundary that is not homeomorphic to any differentiable manifold. (Freedman 1982) There exists open subset A of \mathbb{R}^4 with A being homeomorphic to \mathbb{R}^4 but not diffeomorphic to \mathbb{R}^4 .

Definition 19.5.0.0.1

An \mathbb{R} -action on a set S is a map $G : S \times \mathbb{R} \rightarrow S$ that satisfies $G(x, 0) = x$, and $G(G(x, t), s) = G(x, t+s)$.

Exercise: For S^1 ,

$$G(\vec{x}, t) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} \vec{x}$$

defines an \mathbb{R} -action on S^1 .

Exercise: For $S^2 \subseteq \mathbb{R}^3$.

$$G(\vec{x}, t) = \begin{bmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \vec{x}$$

defines an \mathbb{R} -action on S^2 .

Exercise: For $S^3 \in \mathbb{R}^4$,

$$G(\vec{x}, t) = \begin{bmatrix} \cos(t) & -\sin(t) & 0 & 0 \\ \sin(t) & \cos(t) & 0 & 0 \\ 0 & 0 & \cos(t) & -\sin(t) \\ 0 & 0 & \sin(t) & \cos(t) \end{bmatrix} \vec{x}$$

defines an \mathbb{R} -action on S^3 .

Definition 19.5.0.0.2

An equilibrium point for the group action $G : S \times \mathbb{R} \rightarrow S$ is a point $x \in S$ such that $G(x, t) = x$ for all $t \in \mathbb{R}$.

Exercise:

$$G(\vec{x}, t) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} \vec{x}$$

has no equilibrium point.

$$G(\vec{x}, t) = \begin{bmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \vec{x}$$

has equilibrium points $(0, 0, \pm 1)$.

$$G(\vec{x}, t) = \begin{bmatrix} \cos(t) & -\sin(t) & 0 & 0 \\ \sin(t) & \cos(t) & 0 & 0 \\ 0 & 0 & \cos(t) & -\sin(t) \\ 0 & 0 & \sin(t) & \cos(t) \end{bmatrix} \vec{x}$$

has no equilibrium point.

Similarly to the exercises above, one can easily write down \mathbb{R} -actions on S^n without equilibrium point if n is odd.

Theorem 19.6 (See You All in Math 396 Theorem)

When n is even, any C^∞ type \mathbb{R} -action on S^n has at least one equilibrium point.

Given \mathbb{R} -action on M , for $x \in M$, we can write the following:

$$\frac{\partial}{\partial t} G(x, t) \Big|_{t=0} \in \mathcal{T}_x(M) = x^\perp$$

A map taking $x \in M$ to $f(x) \in \mathcal{T}_x(M)$ is called a vector field.

Theorem 19.7

For M compact manifold without boundary, the map $\{\mathbb{R}\text{-actions on } M\} \rightarrow \{\text{vector field on } M\}$ given above is a bijection.

Math 396

Let ω be a 1-form on an open subset A of \mathbb{R}^n , here ω defines a function from A to $(\mathbb{R}^n)^* = \text{hom}(\mathbb{R}^n, \mathbb{R})$. For $\vec{p} \in A$, we have $\omega(\vec{p}) : \mathbb{R}^n \rightarrow \mathbb{R}$, and $\dim(\ker(\omega(\vec{p}))) = n - 1$ or n .

Proposition 20.0.1

For k -manifold M being a subset of an open subset A of \mathbb{R}^n , the followings are equivalent:

1. $\mathcal{T}_{\vec{p}}(M) \subseteq \ker(\omega(\vec{p}))$ for all $\vec{p} \in M$
2. $\alpha^*\omega = 0$ for all coordinate patches α for M
3. $\int_C \omega = 0$ for all 1-manifold $C \subseteq M$.

Proof. The proof of this proposition is given in Math 396 HW1. \square

Definition 20.0.1.0.1

Let ω be a 1-form on an open subset A of \mathbb{R}^n , for k -manifold $M \subseteq A$, M is said to be an integral manifold for ω provided that $\int \omega = 0$ for all 1-manifold $C \subseteq M$.

Integral manifolds for ω are also integral manifold for $g\omega$ where g is a scalar function. Note that $\alpha^*(g\omega) = (\alpha^*g)(\alpha^*\omega)$.

Lemma 20.0.2

Let $f \in C^1(A, \mathbb{R})$ where A is an open subset of \mathbb{R}^n , with $df \neq 0$ on A .

Then, for $c \in \mathbb{R}$, the level set $f^{-1}(c)$ is $(n - 1)$ -manifold without boundary.

Proof. The proof of this lemma is given in Theorem I.6 on Math 396 Supplement Materials. \square

Corollary 20.0.2.1

Let $f \in C^1(A, \mathbb{R})$ where A is an open subset of \mathbb{R}^n , with $df \neq 0$ on A .

Each level set of f is an integral manifold for df .

Proof. $\alpha^*(df) = d(\alpha^*f) = d(f \circ \alpha) = d(c) = 0$ for some $c \in \mathbb{R}$. \square

Let $\omega = U(x, y)dx + V(x, y)dy$ be defined on an open subset A of \mathbb{R}^2 , with functions $U : \mathbb{R}^2 \rightarrow \mathbb{R}$, $V : \mathbb{R}^2 \rightarrow \mathbb{R}$. Let V be non-vanishing on A . Let $f \in C^1((a, b), \mathbb{R})$ with $M := \text{Graph}(f) \subseteq A$. Let $\alpha : (a, b) \rightarrow M$ $x \mapsto (x, f(x))$.

M is an integral manifold for ω if and only if $\alpha^*\omega = 0$, where $\alpha^*\omega = (U(x, f(x)) + V(x, f(x))f'(x))dx = 0$ if and only if $f'(x) = -\frac{U(x, f(x))}{V(x, f(x))}$.

Suppose $U, V \in C^1(\mathbb{R}^2, \mathbb{R})$, then $-\frac{U(x, y)}{V(x, y)} \in C^1$, consider the point $(0, y_0) \in A$, we claim that there exists $\psi \in C^1(\mathbb{R}^2, \mathbb{R})$ with bounded function $D\psi : \mathbb{R}^2 \rightarrow \mathbb{R}_{\text{row}}^2$ such that $\psi(x, y) = -\frac{U(x, y)}{V(x, y)}$ for (x, y) in the neighborhood of $(0, y_0)$. Pick $\eta \in C^\infty(\mathbb{R}^2, \mathbb{R})$ with $\eta(x, y) = 1$ in neighborhood of $(0, y_0)$ such that $\text{supp}(\eta)$ is a compact support of A . Define the following:

$$\psi(x, y) = \begin{cases} -\frac{U(x, y)}{V(x, y)}\eta(x, y) & (x, y) \in A \\ 0 & (x, y) \notin A \end{cases}$$

Here $D\psi$ is bounded because it has compact support, and it satisfies the conditions in the claim. By Math 395 HW4 Q6, ψ is Lipschitz on \mathbb{R}^2 , hence is partially Lipschitz on \mathbb{R}^2 . By Math 395 HW10 Q5, $\exists \epsilon > 0$ such that $f'(x) = \psi(x, f(x))$ has a unique solution on $(-\epsilon, \epsilon)$. One can shrink ϵ if necessary to $\bar{\epsilon} > 0$ such that $f'(x) = -\frac{U(x, f(x))}{V(x, f(x))}$ has unique solution on $(-\bar{\epsilon}, \bar{\epsilon})$, so there exists a unique local integral curve for ω passing through $(0, y_0)$. We claim that we get the same result for each $(x_0, y_0) \in A$, one can show by using a translation in the x -direction.

For continuous functions U and V , one can show that the solutions guaranteed to exist, but the uniqueness can fail. One example is given by Math 395 HW11 Q5.

Conversely given differential equation:

$$f'(x) = \psi(x, f(x))D \quad ((*)$$

then the graphs of solutions of $(*)$ are integral curves for $\omega := -\psi(x, y)dx + dy$ and for $\tilde{\omega} = B(x, y)\omega = -B(x, y)\psi(x, y)dx + B(x, y)dy$. If $\tilde{\omega}$ is exact, that is, if $\tilde{\omega} = dg$ for some function g , then B is an integrating factor for ω . Since level curves $g^{-1}(c)$ are integral manifolds for $dg = \tilde{\omega}$,

assuming $B \neq 0$, then $g^{-1}(c)$ are also integral manifolds for $\omega = \psi(x, y) dx + dy$, and hence $g^{-1}(c)$ are graphs of solutions of (D).

Conversely, suppose f solves the differential equation (D), then we can write the following:

$$\begin{aligned} \frac{d}{dx} g \begin{bmatrix} x \\ f(x) \end{bmatrix} &= dg \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \cdot \begin{bmatrix} 1 \\ f'(x) \end{bmatrix} \\ &= \left[-B \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \psi \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) - B \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \right] \cdot \begin{bmatrix} 1 \\ f'(x) \end{bmatrix} \\ &= -B \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \psi \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) + B \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) f'(x) = 0 \end{aligned}$$

Here we see that $g \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right)$ is locally constant.

The good news are, such B always exists, looking for B is often useful when solving differential equation (D), but the bad news is find B is not always easier than trying to solve (D).