

Class Notes

Math 395/396 - Honors Analysis

Professor David Barrett

University of Michigan



Jinyan Miao
and his friends from UMich Honors Math

Winter 2022

Contents

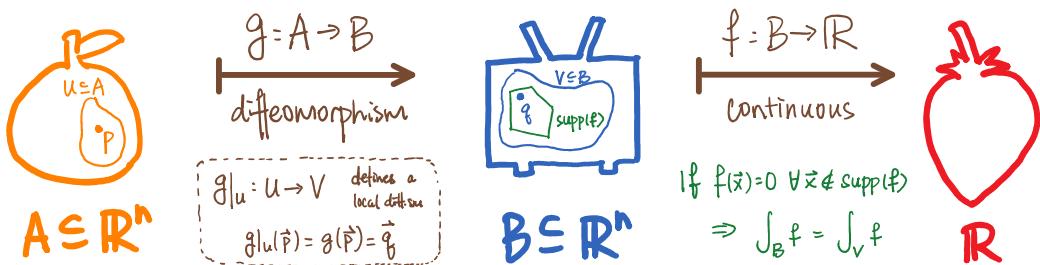


Figure. Corollary 13.4 - Change of Variables and Partition of Unity

This text is intended as a supplemental class notes for students who is taking, or has taken, Math 395/396 at the University of Michigan. We assume readers have taken one of the course sequences, Math 295-296, or Math 217-297, at the University of Michigan. We will make use of some usual notation that were introduced in those courses. The two course sequences each cover topics related to sets, groups, topology, series and sequences, differentiation and integration of functions of single variable, and linear algebra.

In this text, Munkres refers the 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 295/296/297 and 2021 Math 395/396 is 57, but in this text we will use 19 as our prime number, notice that we have $19 \cdot 3 = 57$.

This text is transcribed by Jinyan Miao. The course 2021 Math 395/396 is taught by Professor David Barrett at the University of Michigan - Ann Arbor. Here we give special thanks to Jinyan's friends in 2021 Math 395/396 for helping Jinyan to write and correcting typos in this text.

Except as permitted by both Jinyan and Professor Barrett, no part of this text is allowed to be distributed. If you have any questions or concerns regarding this text, or if you find any typos in this text, please contact Jinyan through jmiu@umich.edu.

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 1.0.0.1

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 followings:

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 $\forall t \in F$, $(1-t)\vec{m} + t\vec{n} - \vec{x} = (1-t)(\vec{m} - \vec{x}) + t(\vec{n} - \vec{x}) = \vec{m} - \vec{x} - t\vec{m} + t\vec{x} + t\vec{n} - t\vec{x} = (1-t)\vec{a} + t\vec{b} \in S - \vec{x}$, 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 $\vec{z} - \vec{x} = (1-t)(\vec{m} - \vec{x}) + t(\vec{n} - \vec{x}) = \vec{m} - \vec{x} - t\vec{m} + t\vec{x} + t\vec{n} - t\vec{x} = (1-t)\vec{m} + t\vec{n} - \vec{x}$, 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:
 $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$, 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 , $\|\vec{x}\|_p := (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$ 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 $|p_j(x) - f_j(x_0)| < \frac{\epsilon}{\sqrt{n}}$ for $1 \leq j \leq n$. For $x \in X$, if we have $d(x, x_0) < \delta$, then we can write $d(f(x), f(x_0)) = \|f(x) - f(x_0)\| \leq \sqrt{n}|f(x) - f(x_0)| < \sqrt{n} \frac{\epsilon}{\sqrt{n}} = \epsilon$, hence f is continuous. \square

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

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

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 $a \in \mathbb{R}$ 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 tend 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.

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

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

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.

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

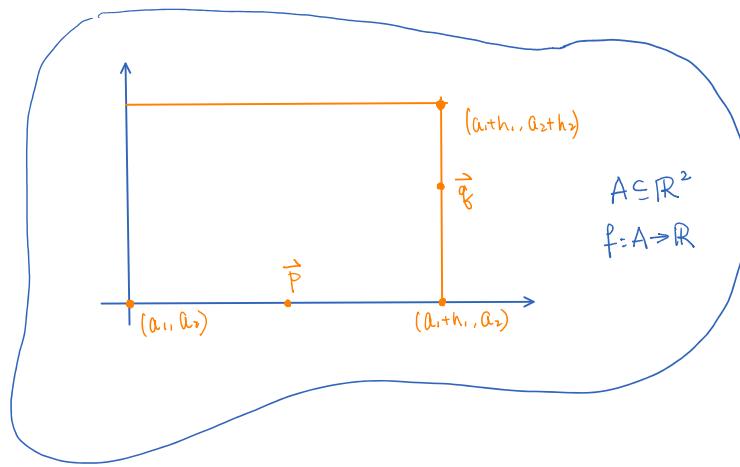
Proof. It suffices to show that f is differentiable at each $\vec{a} \in A$, and it suffices to consider $n = 1$ by Theorem 5.4 from Munkres Section 5. For $m = 2$, let $\vec{a} = (a_1, a_2)$ and $\vec{h} = (h_1, h_2)$, then we can write $f(a_1 + h_1, a_2) - f(a_1, a_2) = D_1 f(\vec{p}) \cdot h_1$ for 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^2f is continuous on A , where $D^2f := D(Df) : A \rightarrow B(V, B(V, W))$

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:

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

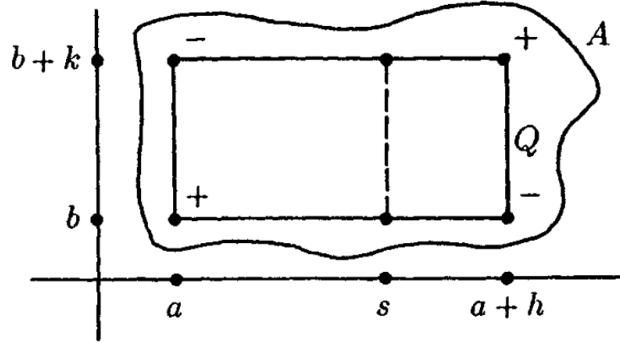
Proof. Let $\phi : M \rightarrow \mathbb{R}$ $s \mapsto f(s, b+k) - f(s, b)$ where M is an open interval contained in $[a, a+h]$. Differentiate ϕ using the Chain Rule, then we can write $\phi'(s) = D_1f(s, b+k) - D_1f(s, b)$.

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

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 .

Homeomorphisms and Diffeomorphisms

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

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

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

where we have:

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

here we know that $Df \left(\begin{bmatrix} r \\ \Phi \end{bmatrix} \right)$ is invertible for all $\begin{bmatrix} r \\ \Phi \end{bmatrix} \in A$, 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 $\vec{x} \in B_\delta(\vec{0})$, 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:

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

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.

□

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

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

Consider another function defined by the following:

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

and we define:

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

Then we write the following:

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

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.

The 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 $T : \mathbb{R}^{k+n} \rightarrow \mathbb{R}^n \quad \vec{v} \mapsto M\vec{v} - \vec{c}$. Here T is an affine. Let $\hat{M} : \mathbb{R}^{k+n} \rightarrow \mathbb{R}^n \quad \vec{v} \mapsto M\vec{v}$. 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:

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

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

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

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

Here, one might want a different set of k independent vectors on the graph. Let $\hat{p} : \mathbb{R}^{k+n} \rightarrow \mathbb{R}^{k+n} \quad \vec{x} \mapsto P\vec{x}$ 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 $Dh(\vec{p}) = \vec{0}$.

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:

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

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 $f(\vec{p}) = \vec{0}$, here we will get a set of n equations as we require $f_i(\vec{p}) = 0$ for $1 \leq i \leq n$. By equation (3), in order to have the 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 $Dh = [4x^3 \quad 6y^5]$, 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 $Df = [2x \quad 2y]$. 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:

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

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:

$$\begin{cases} B \vec{x} = \lambda \vec{x} \\ \|\vec{x}\| = \vec{x}^T \vec{x} = 1 \end{cases} \quad (\text{S1})$$

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 $h(\vec{x}_1) = \vec{x}_1^T B \vec{x}_1 = \vec{x}_1^T \mu \vec{x}_1 = \mu \vec{x}_1^T \vec{x}_1 = \mu$ 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:

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

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 $Dg(\vec{x}_0) = \vec{0}$. 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 $(\vec{x}, y) \in \text{RHS} \Rightarrow y \in \psi(\vec{x}) \Rightarrow (\vec{x}, y) \in \text{epigraph}(\psi)$ \square

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 $V(Q) := (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$ gives the volume of box Q .

Definition 10.0.0.0.2

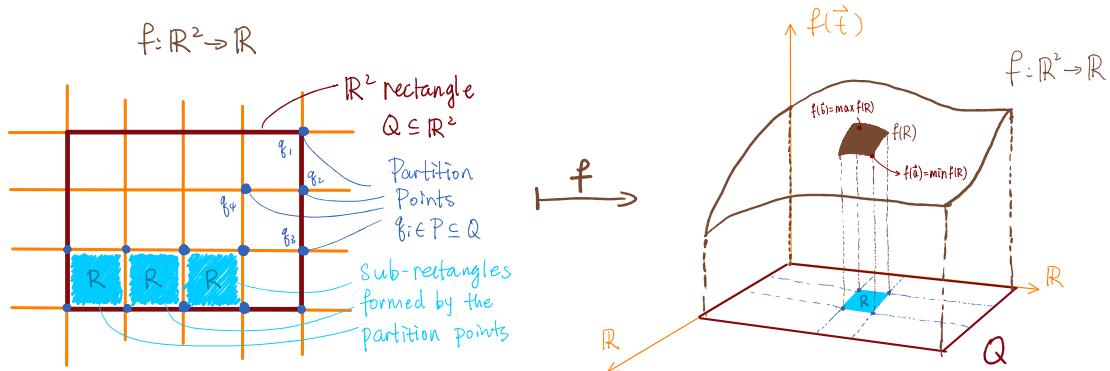
For $[a, b] \subseteq \mathbb{R}$ with $a < b$, a partition of $[a, b]$ is a finite collection P of points in $[a, b]$ that includes a, b . That is, $P \in \mathcal{P}^{finite}([a, b])$ with $a, b \in P$. We index the elements of P in increasing order, as $a = t_0 < t_1 < \dots < t_k = b$. Each $[t_{i-1},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

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

Definition 10.0.0.4

For $f : Q \rightarrow \mathbb{R}$ with P being a partition of a box $Q \subseteq \mathbb{R}^n$, let $\{R_i\}$ be the set of subboxes of Q determined by P , we define the followings:

$$L(f, P) := \sum_{R \in \{R_i\}} (\inf(f(R))V(R)) \quad U(f, P) := \sum_{R \in \{R_i\}} (\sup(f(R))V(R))$$

Note: From Definition 10.0.0.4, it is immediate that $L(f, P) \leq U(f, P)$.

Definition 10.0.0.5

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 $\underline{\int}_Q f = \bar{\int}_Q f$, 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. $g : Q \rightarrow \mathbb{R} \quad \vec{a} \mapsto \text{osc}(f, \vec{a})$ 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 $\text{Int}(R_{i,j}) \cap \mathcal{D}_k(f) \neq \emptyset$, 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 $\mathcal{D}_k(f) \subseteq R_{i,1} \cup R_{i,2} \cup \dots \cup R_{i,l} \cup \text{Bd}(R_{b,1}) \cup \text{Bd}(R_{b,2}) \cup \dots \cup \text{Bd}(R_{b,t})$, where each $\text{Bd}(R_{b,j})$ 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 $\bigcup_{p=1}^{\infty} V(R_p) < \frac{\epsilon}{19}$. For each p , pick subboxes \hat{R}_p of Q such that the following condition holds:

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

Here we get $\sum_{p=1}^{\infty} V(\hat{R}_p) < 2 \cdot \frac{\epsilon}{19} + \sum_{p=1}^{\infty} \frac{\epsilon}{2^{p+1}} < \epsilon$. 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 $\mathcal{D}(f) \subseteq \bigcup_{p=1}^{\infty} \text{rInt}(R_p)$ 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 $\vec{a} \in \text{rInt}(Q_{\vec{a}})$ and $|f(\vec{x}) - f(\vec{a})| < \hat{\epsilon}$ for $\vec{x} \in Q_{\vec{a}}$. Note here $Q \subseteq \bigcup_{p=1}^{\infty} \text{rInt}(R_p) \cup (\bigcup_{\vec{a} \in Q \setminus \mathcal{D}(f)} \text{rInt}(Q_{\vec{a}}))$. Here we note that Q is compact, and each $\text{rInt}(Q_{\vec{a}})$, or $\text{rInt}(R_p)$, is open, so there exists a finite subcover such that we have:

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

Here we denote $\text{rInt}(Q_{\vec{a}_1}) \cup \text{rInt}(Q_{\vec{a}_2}) \cup \dots \cup \text{rInt}(Q_{\vec{a}_t}) = \bigcup_{\text{finitely many } \vec{a} \in Q \setminus \mathcal{D}(f)}$, and $\text{rInt}(R_{p_1}) \cup \text{rInt}(R_{p_2}) \cup \dots \cup \text{rInt}(R_{p_s}) = \bigcup_{p=1}^s \text{rInt}(R_p)$. 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:

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

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 $\bar{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \bar{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$. Combining the result from (*) and (a), we get $\underline{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y}) \leq \underline{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$. 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 $\underline{\int}_Q f \leq \underline{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$, 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:

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

On the other hand, we have:

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

Hence we see that $\underline{\int}_Q f \leq \underline{\int}_{\vec{x} \in A} \underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$. 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 $\bar{\int}_Q f \leq \bar{\int}_{\vec{x} \in A} \bar{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$ 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:

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

and we can write:

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

Note: Given the settings in Corollary 10.3.1, $\underline{\int}_{\vec{y} \in B} f(\vec{x}, \vec{y})$ 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 $\underline{\int}_Q f > 0$.
4. If $f \geq 0$, $f(\vec{a}) > 0$, and $\underline{\int}_Q f = 0$, 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 $\underline{\int}_Q f \geq L(f, P) = \inf_{R^*} f \cdot V(R^*) + \sum_R \inf_R f \cdot V(R) \geq \frac{f(\vec{a})}{2} V(R^*) + 0$. \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:

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

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.

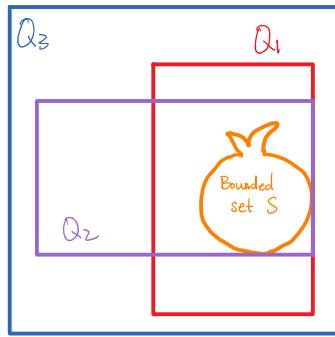
$$f_S : \mathbb{R}^n \rightarrow \mathbb{R} \quad \vec{x} \mapsto \begin{cases} f(\vec{x}) & \vec{x} \in S \\ 0 & \vec{x} \notin S \end{cases} \quad \int_S f := \int_Q f_S \quad \text{for box } Q \text{ that contains } S$$

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 $\mathcal{D}(f_S)_3 = \mathcal{D}(f_S)_1 \cup \{\text{subset of } \text{Bd}(Q_1)\}$, here $\{\text{subset of } \text{Bd}(Q_1)\}$ 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:

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

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 $f : Q \rightarrow \mathbb{R}$ and $g : Q \rightarrow \mathbb{R}$ 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 $\vec{x} \in S$, 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:

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

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

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 $\tilde{Q}_j \supseteq \text{Int}(\tilde{Q}_j) \supseteq Q_j$ with $V(\tilde{Q}_j) < V(Q_j) + \frac{\epsilon}{2^j}$, by Lemma 10.3.3, here we get $\text{Int}(\tilde{Q}_j)$ 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 $\mathcal{I} : S \rightarrow \mathbb{R} \quad \vec{x} \mapsto 1$. $V(S) := \int_S \mathcal{I}$.

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 $S \setminus A \Rightarrow S \setminus A$ 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 .

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

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 $\underline{\int}_E f = \underline{\int}_E f \leq \underline{\int}_A f$, 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 $L(f_A, P) \leq \underline{\int}_M f \leq \text{ext } \int_A f$, 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 $\underline{\int}_A f = \text{ext } \int_A f$

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 $\text{ext } \int_A f = \lim_{j \rightarrow \infty} \underline{\int}_{E_j} f$ for non-negative continuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Proof. We have each of $\underline{\int}_{E_j} f \leq \text{ext } \int_A f$. So $\lim_{j \rightarrow \infty} \underline{\int}_{E_j} f \leq \text{ext } \int_A f$. 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 $\underline{\int}_E f \leq \underline{\int}_{E_M} f \leq \lim_{j \rightarrow \infty} \underline{\int}_{E_j} f$, so we have $\text{ext } \int_A f \leq \lim_{j \rightarrow \infty} \underline{\int}_{E_j} f$. 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:

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

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: $\lim_{j \rightarrow \infty} \int_{-j}^j \frac{\pi}{\sqrt{1+y^2}} dy = \lim_{j \rightarrow \infty} \pi u|_{u=-\text{arcsinh}(j)}^{u=\text{arcsinh}(j)} = \pi(\infty) - \pi(-\infty) = \infty$.

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

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 $\text{ext } \int_A f = \lim_{j \rightarrow \infty} \text{ext } \int_{U_j} f$.

Proof. Note here $\text{ext } \int_A f$ exists as the sequence $(\text{ext } \int_{U_j} f)$ is monotonic. Now we write each $\text{ext } \int_{U_j} f \leq \text{ext } \int_A f$, hence we have $\lim_{j \rightarrow \infty} \text{ext } \int_{U_j} f \leq \text{ext } \int_A f$. 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 $\int_E f \leq \text{ext } \int_{U_j} f \leq \lim_{j \rightarrow \infty} \text{ext } \int_{U_j} f$. Take supremum over the choice of E , then we have $\text{ext } \int_A f \leq \lim_{j \rightarrow \infty} \text{ext } \int_{U_j} f$. 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 $\text{ext} \int_A f_+ < \infty$ and $\text{ext} \int_A f_- < \infty$ holds. In other words, f is extended-integrable on A provided that $\text{ext} \int_A f_+ + \text{ext} \int_A f_- = \text{ext} \int_A |f| < \infty$. If f is extended-integrable on A , then¹ we write $\text{ext} \int_A f := \text{ext} \int_A f_+ - \text{ext} \int_A f_-$.

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 $\text{ext} \int_A f = -\text{ext} \int_A (-f) = -\underline{\int}_A (-f) = \bar{\int}_A f$.

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 $f : A \rightarrow \mathbb{R}$.

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

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

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:

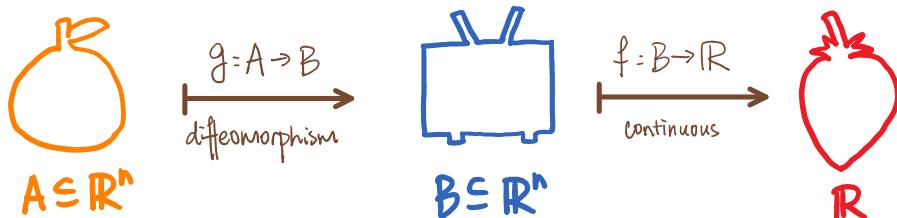
$$V(Q) = \lim_{N \rightarrow \infty} \frac{\#(Q \cap \frac{\mathbb{Z}^n}{2^N})}{2^{nN}} \leq \lim_{N \rightarrow \infty} \sum_{j=1}^k \frac{\#(Q_j \cap \frac{\mathbb{Z}^n}{2^N})}{2^{nN}} = \sum_{j=1}^k V(Q_j)$$

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.

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 $T : \mathbb{R}^n \rightarrow \mathbb{R}^k \quad \vec{x} \mapsto M\vec{x} + \vec{b}$ 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. □

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. □

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. $f \in C(\bar{B}, \mathbb{R})$
2. A, B are bounded and rectifiable
3. g is a diffeomorphism from an open set containing \bar{A} to another open set containing \bar{B}

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 $f : X \rightarrow V$.

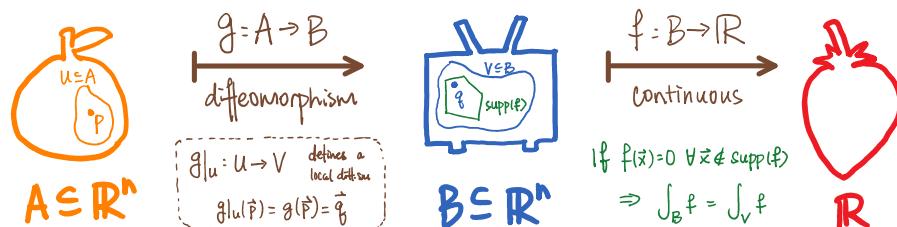
The support of f , denoted as $\text{supp}(f)$, is defined to be the closure of the set $\{\vec{x} \mid f(\vec{x}) \neq \vec{0}\}$ in X .

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 $M_1 + \dots + M_j$ 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:

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

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:

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

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

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 $g : A \rightarrow B$ $\vec{x} \mapsto r\vec{x}$ 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 $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ $(r, \Phi) \mapsto (r \cos(\Phi), r \sin(\Phi))$. Notice that g is a local diffeomorphism and $\det Dg(r, \Phi) = r$.

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

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 Parallelopipeds

Definition 14.0.0.0.1

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 -parallelopiped.

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 -parallelopiped. Here we note that the n -volume of the k -parallelopiped $T(Q)$ 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 $\|A\vec{x}\|^2 = \langle A\vec{x}, A\vec{x} \rangle = \vec{x}^T A^T A\vec{x} = 0 \Rightarrow A\vec{x} = \vec{0}$. 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

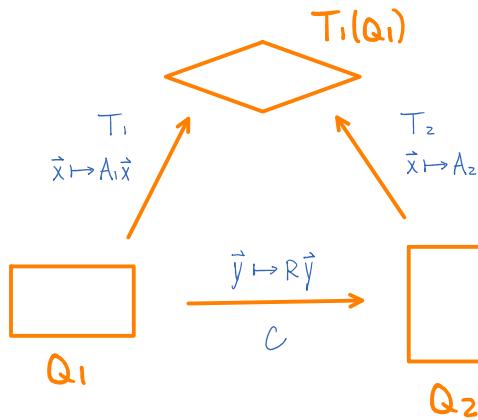
Let A be an $n \times k$ matrix, $\det(A^T A) \geq 0$.

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 | Differentiable Manifolds I

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

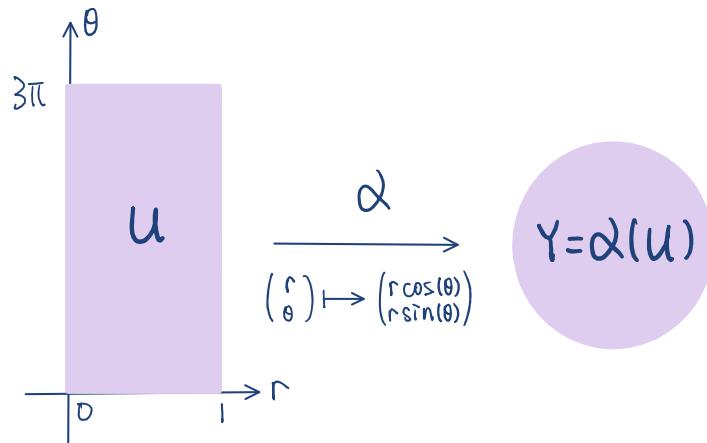
Let $k, n, r \in \mathbb{N}$ with $k \leq n$, let $\alpha \in C^r(U, \mathbb{R}^k)$, where U is an open subset of \mathbb{R}^k . The set $Y := \alpha(U)$, equipped with the map α , constitute a parametrized k -manifold of class C^r , denoted as Y_α . The k -volume of Y_α is defined by the following:

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

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

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 15.0.0.0.2

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:

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

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 $f \in C(Y, \mathbb{R})$.

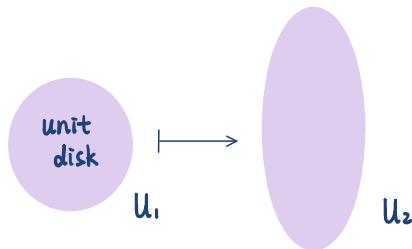
Note: Let Y_α be a parametrized k -manifold of class C^r , let $h : Y \rightarrow Z$ $\vec{z} \mapsto B\vec{z} + \vec{h}$ be an isometry with orthogonal matrix B and $Z = h(Y)$, and let $f : Z \rightarrow \mathbb{R}$ be a continuous function. Here $Z_{h \circ \alpha}$ defines a parametrized k -manifold,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

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

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 :

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

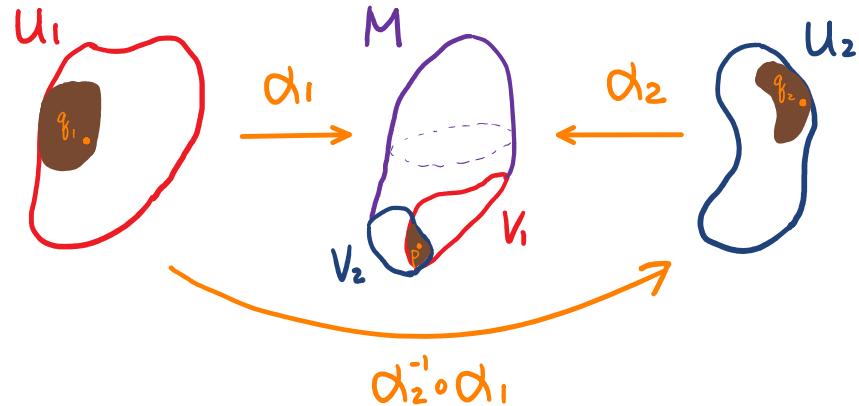
$$\alpha_3 : (-1, 1) \rightarrow M \quad y \mapsto \left(-\sqrt{1 - y^2}, y\right) \qquad \alpha_4 : (-1, 1) \rightarrow M \quad x \mapsto \left(x, -\sqrt{1 - x^2}\right)$$

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 $\alpha_2(q_2) = p$.



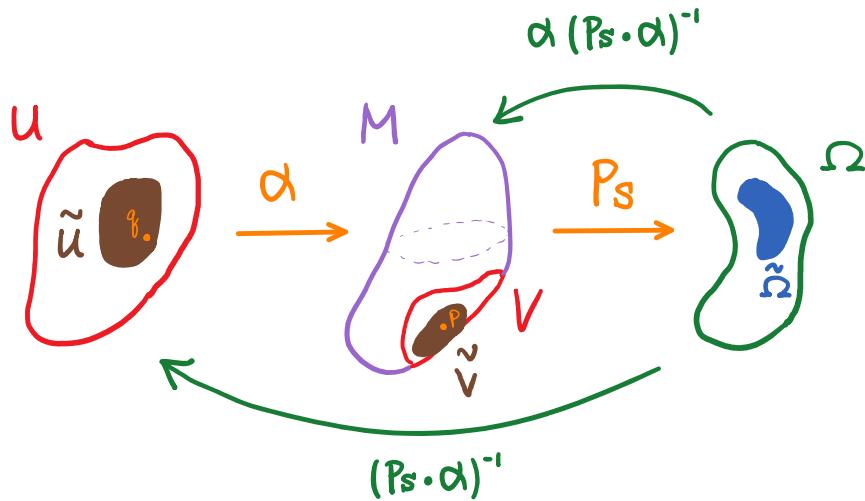
Here we note that $\alpha_2^{-1} \circ \alpha_1$ is of C^r type when defined, and $(\alpha_2^{-1} \circ \alpha_1)^{-1} = \alpha_1^{-1} \circ \alpha_2$ is also of C^r type. Hence we know that $\alpha_2^{-1} \circ \alpha_1$ is a C^r -diffeomorphism.

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 $\text{rank } D\alpha(\vec{x}) = k$ 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: $(P_S \circ \alpha)^{-1} \circ P_S \in C^r(\rho^{-1}(\tilde{\Omega}) \times \mathbb{R}^{n-k}), \tilde{U}).$ □

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. □

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: $F : \mathbb{R}^n \rightarrow \mathbb{R} \quad \vec{x} \mapsto \|\vec{x}\|^2 - 1. \quad F^{-1}(\vec{0}) := S^{n-1}.$

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

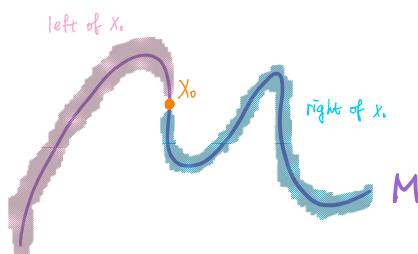
Every connected 1-manifold of class C^r is C^r -diffeomorphic to an interval in \mathbb{R} or to the circle S^1 .

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 lost 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

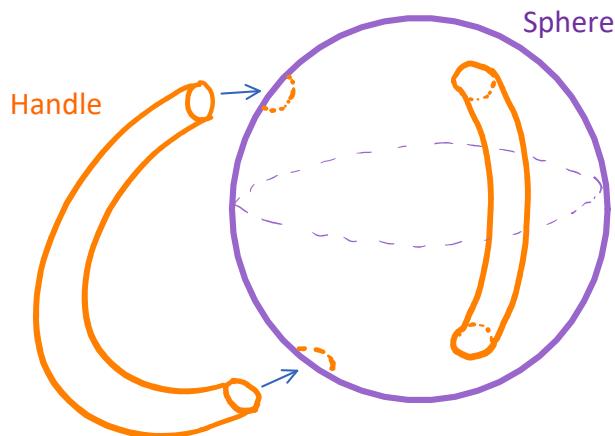
Every connected 1-manifold is diffeomorphic to one of the followings:

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:

$$\begin{aligned} df &= \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right] \\ &= \frac{\partial f}{\partial x_1} [1 \quad 0 \quad \cdots \quad 0] + \frac{\partial f}{\partial x_2} [0 \quad 1 \quad 0 \quad \cdots \quad 0] + \cdots + \frac{\partial f}{\partial x_n} [0 \quad 0 \quad \cdots \quad 1] \\ &= \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n \end{aligned}$$

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 $\omega = [\omega_1 \ \omega_2 \ \dots \ \omega_n] = \omega_1 dx_1 + \omega_2 dx_2 + \dots + \omega_n dx_n$.

Definition 18.0.0.3

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 $\alpha_1 \in C^1(A, B)$, and let $\alpha_2 \in C^1(B, C)$, 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 $(r, \Phi) \mapsto (r \cos(\Phi), r \sin(\Phi))$. We can write the following:

$$\begin{aligned}\alpha^* \left(\frac{x_1}{x_1^2 + x_2^2} dx_1 + \frac{x_2}{x_1^2 + x_2^2} dx_2 \right) &= \frac{r \cos(\Phi)}{r^2} d(r \cos(\Phi)) + \frac{r \sin(\Phi)}{r^2} d(r \sin(\Phi)) \\ &= \frac{r \cos(\Phi)}{r^2} (\cos(\Phi) dr - r \sin(\Phi) d\Phi) + \frac{r \sin(\Phi)}{r^2} (\sin(\Phi) dr + r \cos(\Phi) d\Phi) \\ &= \frac{1}{r} dr\end{aligned}$$

Similarly, we have:

$$\alpha^* \left(\frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2 \right) = \frac{r \cos(\Phi)}{r^2} d(r \cos(\Phi)) + \frac{r \sin(\Phi)}{r^2} d(r \sin(\Phi)) = d\Phi$$

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:

$$\int_{Y_{\alpha \circ \beta}} \omega = \int_J (\alpha \circ \beta)^* \omega = \int_J \beta^* \alpha^* \omega = \int_J ((\alpha^* \omega) \circ \beta) \cdot D\beta$$

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

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 $T : Y \rightarrow \mathbb{R}^n$ $\alpha(t) \mapsto \frac{\alpha'(t)}{\|\alpha'(t)\|}$, 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:

$$\frac{f(x+h) - f(x) - \omega(x) \cdot h}{\|h\|} \quad (*)$$

Assuming that h is small enough, with $\alpha : [0, 1] \rightarrow A$ $t \mapsto x + th$ 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.

Integral Manifolds

Let ω be a 1-form defined on an open subset A of \mathbb{R}^n . Here ω can be viewed as 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

Let A be an open subset of \mathbb{R}^n , and let ω be a 1-form defined on A .

For k -manifold $M \subseteq A$, 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 defined on an open subset A of \mathbb{R}^n , for k -manifold $M \subseteq A$, M is called an integral manifold for ω provided that $\int_C \omega = 0$ for all 1-manifold $C \subseteq M$.

Note: Let ω be a 1-form defined on an open subset A of \mathbb{R}^n . Integral manifolds for ω are also integral manifold for $g\omega$ where g is a scalar function, because we have $\alpha^*(g\omega) = (\alpha^*g)(\alpha^*\omega)$.

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 an $(n - 1)$ -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. For some $c \in \mathbb{R}$, the level set $M := f^{-1}(c)$ is an $(n - 1)$ -manifold without boundary. let α be a coordinate patch for M , we see that $\alpha^*(df) = d(\alpha^*f) = d(f \circ \alpha) = 0$, and here we have $f \circ \alpha = c$, hence $d(f \circ \alpha) = 0$, so M is an integral manifold for df . \square

Integral curves for solving differential equations

In this section, a function f denoted as $f(x)$ has domain being a subset of \mathbb{R} and codomain being a subset of \mathbb{R} ; A function g denoted as $g(x, y)$ has domain being a subset of \mathbb{R}^2 and codomain being a subset of \mathbb{R} . The 1-form ω has domain \mathbb{R}^2 and codomain \mathbb{R}_{row}^2 .

Let $\omega = U(x, y)dx + V(x, y)dy$ be defined on an open subset A of \mathbb{R}^2 , with V being a non-vanishing function on A . Let $f \in C^1((a, b), \mathbb{R})$. Consider a parametrized manifold $M := \text{Graph}(f) \subseteq A$ parametrized by $\alpha : (a, b) \rightarrow M$ $x \mapsto (x, f(x))$ for some continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$. Here we observe that M is an integral manifold for ω if and only if $\alpha^*\omega = 0$, where we have:

$$\alpha^*\omega = (U(x, f(x)) + V(x, f(x))f'(x)) = 0 \iff f'(x) = -\frac{U(x, f(x))}{V(x, f(x))}$$

Suppose further that $U, V \in C^1(\mathbb{R}^2, \mathbb{R})$, then $-\frac{U(x,y)}{V(x,y)} \in C^1$, consider the point $(0, y_0) \in A$, we claim that there exists $\Phi \in C^1(\mathbb{R}^2, \mathbb{R})$ with bounded function $D\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}_{row}^2$ such that $\Phi(x, y) = -\frac{U(x,y)}{V(x,y)}$ for (x, y) in the neighborhood of $(0, y_0)$. Here we will prove the claim:

Proof. 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 subset of A . Then we define the following:

$$\Phi(x, y) = \begin{cases} -\frac{U(x,y)}{V(x,y)}\eta(x, y) & (x, y) \in A \\ 0 & (x, y) \notin A \end{cases}$$

Here $D\Phi$ is bounded because Φ has compact support, and Φ satisfies the conditions in the claim. \square

By Math 395 HW4 Q6, Φ is partially Lipschitz on \mathbb{R}^2 . By Math 395 HW10 Q5, $\exists \epsilon > 0$ such that $f'(x) = \Phi(x, f(x))$ has a unique solution f 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 f 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.

If U and V are continuous but not differentiable, one can show that the solutions guaranteed to exist, but the uniqueness can fail. One example is given by Math 395 HW11 Q5, which considers the 1-form $\omega = y^{1/3} dx - dy$ with $f'(x) = \sqrt[3]{f(x)}$

Conversely given differential equation:

$$f'(x) = \Phi(x, f(x)) \tag{D1}$$

then the graphs of solutions of (D1) are integral curves for $\omega := -\Phi(x, y)dx + dy$, and also for $\tilde{\omega} := B(x, y)\omega = -B(x, y)\Phi(x, y)dx + B(x, y)dy$ with some function $B(x, y)$. If $\tilde{\omega}$ is exact, that is, if $\tilde{\omega} = dg$ for some function g , then B is called an integrating factor for ω . For constant c , since the level set $g^{-1}(c)$ is integral manifolds for $dg = \tilde{\omega}$, assuming B is not the zero function, then $g^{-1}(c)$ is also an integral manifold for $\omega = \Phi(x, y)dx + dy$, and hence $g^{-1}(c)$ is a graph of solutions of (D1).

Now suppose f solves the differential equation (D1), then we can write the following:

$$\begin{aligned} \frac{d}{dx} \left(g \begin{bmatrix} x \\ f(x) \end{bmatrix} \right) &= dg \begin{pmatrix} x \\ f(x) \end{pmatrix} \cdot \begin{bmatrix} 1 \\ f'(x) \end{bmatrix} \\ &= \left[-B \begin{pmatrix} x \\ f(x) \end{pmatrix} \right] \Phi \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) + B \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \cdot \begin{bmatrix} 1 \\ f'(x) \end{bmatrix} \\ &= -B \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \Phi \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) + B \left(\begin{bmatrix} x \\ f(x) \end{bmatrix} \right) f'(x) = 0 \end{aligned}$$

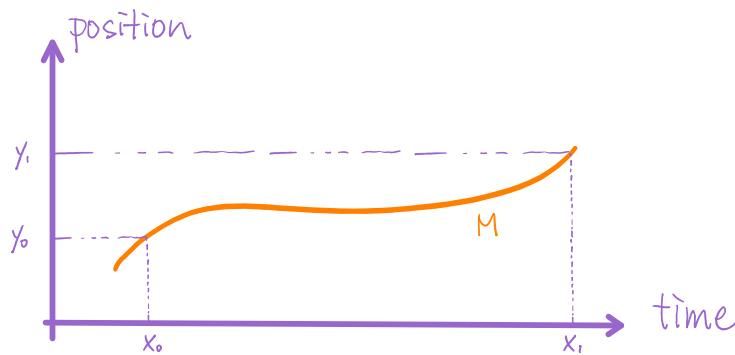
Here we see that $g(x, f(x))$ is locally constant on each component of domain of definition for $g(x, f(x))$. In conclusion, the good news are, such B always exists, looking for B is often useful when solving differential equation (D1), but the bad news is that finding such B is not always easier than trying to solve (D1).

Here are two classes of examples:

1. Consider $f'(x) = \beta(f(x))$, that is, here we suppose the velocity of a particle is given by the predetermined function of its position, and the position function $f(x)$ of the particle is a function of time x . Here we want $-B(x, y)\beta(y) dx + B(x, y) dy$ to be exact with some integrating factor $B(x, y)$, one can choose $B(x, y) = \frac{1}{\beta(y)}$. Then the solution to $f'(x) = \beta(f(x))$ are integration curves for the 1-form:

$$-dx + \frac{dy}{\beta(y)} = d\left(-x + \int \frac{dy}{\beta(y)}\right)$$

Hence the solutions $f(x)$ for the differential equation $f'(x) = \beta(f(x))$ satisfies $-x + \int \frac{dy}{\beta(y)} = C$. Solving $y = f(x)$ could be done implicitly. In the following, we will try to calculate the travel time from position y_0 to y_1 . the path is traced out by the parametrized manifold M :



$$x_1 - x_0 = \int_M dx = \int_M dx + 0 + \int_M dx + \int_M \left(-dx + \frac{dy}{\beta(y)}\right) = \int_M \frac{dy}{\beta(y)} = \int_{y_0}^{y_1} \frac{dy}{\beta(y)}$$

2. Consider $f''(t) = \beta(f(t))$ with $f(t_0) = x_0$ and $f'(t_0) = x_1$, that is, here we suppose we have a unit mass particle in 1-dimensional universe, subject to force field $\beta(x)$. Let $h(t) = f'(t)$, and let $h'(t) = \beta(f(t)) = f''(t)$. Then we can write the following:

$$\begin{bmatrix} f(t_0) \\ h(t_0) \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

with the following:

$$\begin{bmatrix} f'(t) \\ h'(t) \end{bmatrix} = \begin{bmatrix} h(t) \\ \beta(f(t)) \end{bmatrix} := \Psi \begin{bmatrix} f(t) \\ h(t) \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} f' \\ h' \end{bmatrix} = \Psi \circ \begin{bmatrix} f \\ h \end{bmatrix} \quad (\text{D2})$$

Math 395 HW10 Q6 suggests that there is a unique local solution (f, h) for such Ψ if $\beta \in C^1$. Suppose f and h solves the system (D2). Consider $\alpha : x \mapsto (x, f(x), h(x))$, and consider the manifold Y_α parametrized by α . Denote $(t, x, v) \in \mathbb{R}^3$, we have $x = f(t)$, $v = h(t) = f'(t)$ for $(t, x, v) \in Y_\alpha$. Here we see that Y_α is a integral manifold for $\omega_1 = dx - h dt$, and also for $\omega_2 = dv - \beta dt$, where ω_1 and ω_2 are not closed, hence not exact. One can show that ω_1 has no integral 2-manifold, and ω_1 has no non-zero integrating factors.

Exercise: If M is a integral manifold for the 1-form ω_1 and for the 1-form ω_2 , then M is integral for the 1-form $f_1\omega_1 + f_2\omega_2$, where f_1 and f_2 are scalar functions.

Applying to specific situation, $\omega_3 := h \cdot \omega_2 - \beta \cdot \omega_1 = h dv - \beta dx = d(\frac{h^2}{2} - \int \beta dx)$. Note that β is a function of x for $(t, x, v) \in Y_\alpha$. Hence $\frac{h^2}{2} - \int \beta dx$ is constant for any solution of the system (D2). In physics interpretation, at any point $(t, x, v) \in Y_\alpha$, $\frac{(h(t))^2}{2} = \frac{(f'(t))^2}{2} = \frac{v^2}{2}$ is the kinetic energy of an object of unit mass, and $-\int \beta dx$ is the potential energy.

$$E := \frac{h^2}{2} - \int \beta dx \quad (\text{E})$$

here E gives the total energy of the system, which is a constant. Also β can be interpreted as a function of force field divided by the constant mass of a particle, and the path of the particle is traced out by $f(t)$ as a function of time t . Rearranging equation (E) we can write:

$$f'(t) = \pm \sqrt{2 \left(E + \int \beta dx \right)}$$

which is a autonomous first order equation, and by the first class of example in this section, one can get the following implicit solution to the system (D2):

$$t + C = \pm \int \frac{dx}{\sqrt{2(E + \int \beta dx)}}$$

with $x = f(t)$, and some constant C determined by the initial conditions.

Example: (Simple Harmonic Motion) Take $\beta(x) = -x$.

We get $E = \frac{v^2+x^2}{2}$, and we have:

$$t + C = \pm \int \frac{dx}{\sqrt{2E - x^2}} = \pm \arcsin \left(\frac{x}{\sqrt{2E}} \right) \Rightarrow x = \sqrt{2E} \sin(t + C)$$

which is called a simple harmonic motion of a particle.

Example: (The Motion of a Frictionless Pendulum) Take $\beta(x) = -\sin(x)$.

We get $E = \frac{v^2}{2} - \cos(x)$, hence $v^2 = 2(E + \cos(x))$, so $v = \pm \sqrt{2(E + \cos(x))}$.

$$t + C = \pm \int \frac{dx}{\sqrt{2(E + \cos(x))}} = \pm \frac{\sqrt{2} \cdot \text{Elliptic F} \left[\frac{x}{2}, \frac{2}{1+E} \right]}{\sqrt{1+E}}$$

Example: (The Motion of a Pendulum under Friction)

Take $f''(t) = -\sin(f(t)) - \gamma \cdot f'(t)$ for some constant γ . One can show that $E'(x) < 0$.

5 | Complex Analysis I

Integration of Complex-valued Functions

Definition 21.0.0.0.1

Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{C}$ be a function, with $u : A \rightarrow \mathbb{R}$ $\vec{a} \mapsto \Re(f(\vec{a}))$ and $v : A \rightarrow \mathbb{R}$ $\vec{a} \mapsto \Im(f(\vec{a}))$, the integral of f over A is defined to be the following:

$$\int_A f := \int_A u + i \int_A v$$

here extended integral can be used when $\int_A u_+$, $\int_A u_-$, $\int_A v_+$, and $\int_A v_-$ are all finite.

Note: Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{C}$ be a function, we write $f = u + iv$ for functions u and v provided that we have $u : A \rightarrow \mathbb{R}$ $\vec{a} \mapsto \Re(f(\vec{a}))$ and $v : A \rightarrow \mathbb{R}$ $\vec{a} \mapsto \Im(f(\vec{a}))$.

Definition 21.0.0.0.2

Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{C}$ $\vec{a} \mapsto f(\vec{a})$ with $f = u + iv$.

For parametrized manifold $Y_\alpha \subseteq A$, we define the following:

$$\int_{Y_\alpha} f dV = \int_{Y_\alpha} u dV + i \int_{Y_\alpha} v dV$$

Exercise: Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{C}$. For $\lambda \in \mathbb{C}$, we have $\int_A \lambda f = \lambda \int_A f$.

Proposition 21.0.1

Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{C}$.

$$\left| \int_A f \right| \leq \int_A |f|$$

Proof. Note that $|\int_A f| \in \mathbb{C}$, and hence $|\int_A f| = e^{i\Phi} \cdot \int_A f$ for some $\Phi \in \mathbb{R}$. Then we obtain:

$$\left| \int_A f \right| = e^{i\Phi} \int_A f = \int_A e^{i\Phi} f$$

Here $\int_A e^{i\Phi} f \in \mathbb{R}$. That is, if we write $v : A \rightarrow \mathbb{R}$ $\vec{a} \mapsto \Im(f(\vec{a}))$, then $i \int_A v = 0$. Then we obtain the following by the triangle inequality:

$$\int_A e^{i\Phi} f = \int_A \Re(e^{i\Phi} f) \leq \int_A |e^{i\Phi} f| = \int_A |e^{i\Phi}| \cdot |f| = \int_A |f|$$

This completes the proof. \square

Holomorphic Functions

Definition 22.0.0.0.1

Let A be an open subset of \mathbb{R}^n , $\omega : A \rightarrow \mathbb{C}_{\text{row}}^n$ is called a \mathbb{C} -valued 1-form. Denote $\omega = \omega_1 + i\omega_2$, where $\omega_1 : A \rightarrow \mathbb{R}_{\text{row}}^n$ $\vec{a} \mapsto \Re(\omega(\vec{a}))$ and $\omega_2 : A \rightarrow \mathbb{R}_{\text{row}}^n$ $\vec{a} \mapsto \Im(\omega(\vec{a}))$, and let $Y_\alpha \subseteq A$ be a parametrized manifold, we define:

$$\int_{Y_\alpha} \omega := \left(\int_{Y_\alpha} \omega_1 \right) + \left(i \int_{Y_\alpha} \omega_2 \right)$$

Exercise: In the settings of Definition 21.0.1.0.1, we get $\int_{Y_\alpha} \lambda \omega = \lambda \int_{Y_\alpha} \omega$ for $\lambda \in \mathbb{C}$. If Y_α is a parametrized 1-manifold, we get the following:

$$\left| \int_{Y_\alpha} \omega \right| \leq (\text{The length of } Y_\alpha) \cdot \sup_{\vec{x} \in Y} \|\omega(\vec{x})\|$$

Definition 22.0.0.0.2

Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{C}$ with $f = u + iv$ for functions u and v . The function f is said to be of C^r type, denoted as $f \in C^r(A, \mathbb{C})$, provided that $u, v \in C^r(A, \mathbb{R})$. Here we define $D_j f := D_j u + i D_j v$ if $f \in C^1(A, \mathbb{C})$ is given.

Note: Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{C}$ with $f = u + iv$, we have $df = du + idv$.

Note: Consider $z : \mathbb{R}^2 \rightarrow \mathbb{C}$ $(x, y) \mapsto x + iy$,¹ we write $\bar{z} : \mathbb{R}^2 \rightarrow \mathbb{C}$ $(x, y) \mapsto x - iy$, and we get:

$$dz = dx + idy \quad d\bar{z} = dx - idy$$

Rearranging we have:

$$dx = \frac{dz + d\bar{z}}{2} \quad dy = \frac{dz - d\bar{z}}{2i}$$

Euler's Formula

With $dz = dx + idy$, we proceed to investigate the 1-form $\frac{dz}{z}$ defined on $\mathbb{C} \setminus \{0\}$.

$$\begin{aligned} \frac{dz}{z} &= \frac{dx + idy}{x + iy} \\ &= \frac{x - iy}{x - iy} \cdot \frac{dx + iy}{x + iy} \\ &= \left(\frac{x dx}{x^2 + y^2} + \frac{y dy}{x^2 + y^2} \right) + i \left(-\frac{y dx}{x^2 + y^2} + \frac{x dy}{x^2 + y^2} \right) \end{aligned} \tag{*}$$

Let log function defined by the followings:

$$\log : (0, \infty) \rightarrow \mathbb{R} \quad m \mapsto \int_1^m \frac{1}{t}$$

Let $X \subseteq \mathbb{C}$ be a ray starting at $\vec{0}$ making an angle $\theta_0 = 0$ from the x -axis.

One can define polar coordinates (r, Φ_X) on $\mathbb{C} \setminus X$ with $0 < \Phi_X < 2\pi$:

$$r : \mathbb{C} \setminus X \rightarrow \mathbb{R} \quad z \mapsto \sqrt{(\Re(z))^2 + (\Im(z))^2}$$

¹In this chapter, Complex Analysis I, and the chapter Complex Analysis II, if not specified, one might consider the notation z is referred to the function $z : \mathbb{R}^2 \rightarrow \mathbb{C}$ $(x, y) \mapsto x + iy$, where we denote $z = x + iy$

$$\Phi_X : \mathbb{C} \setminus X \rightarrow (0, 2\pi) \quad z \mapsto \cos^{-1} \left(\frac{\Re(z)}{\sqrt{(\Re(z))^2 + (\Im(z))^2}} \right)$$

In general, one can propose a definition for the polar coordinate (r, Φ_X) on $\mathbb{C} \setminus X$, with X being a ray starting at $\vec{0} \in \mathbb{C}$ making an angle $\theta_0 \in \mathbb{R}$ from the x -axis, in which case the codomain of Φ_X is defined to be on $(\theta_0, \theta_0 + 2\pi)$.

For nonzero $z \in \mathbb{C}$, the polar coordinate of z is denoted as $(r(z), \Phi_X(z))$. Here $r(z)$ is called the absolute value, or the modulus, of z , and $\Phi_X(z)$ is called the argument, or phase, of z . Here we denote $|z| := r(z)$, and $\arg(z) := \Phi_X(z)$.

One can show that the equation (*) is exact on $\mathbb{C} \setminus X$.

Equation (*) suggests that we can define the following:

$$\log_X : \mathbb{C} \setminus X \rightarrow \mathbb{C} \quad \omega \mapsto \int_1^\omega \frac{dz}{z} = \int_1^\omega d\log(r) + i d\Phi_X = \log(r(\omega)) + i \Phi_X(\omega)$$

Solve $z = \log_X \omega$ for $\omega \in \mathbb{C}$, we get $\Re(z) = \log r(\omega)$ and $\Im(z) = \Phi_X(\omega)$, so we can write:

$$e^z := \omega = r(\omega)(\cos(\Phi_X(\omega)) + i \sin(\Phi_X(\omega))) = e^{\Re(z)}(\cos(\Im(z)) + i \sin(\Im(z)))$$

Here $e^z = e^{\Re(z)}(\cos(\Im(z)) + i \sin(\Im(z)))$ is known as the Euler's Formula.

For $z = x + iy$, we also have the following holds:

$$\begin{aligned} de^z &= de^{x+iy} \\ &= d(e^x(\cos(y) + i \sin(y))) \\ &= e^x(\cos(y) + i \sin(y))dx + e^x(-\sin(y) + i \cos(y))dy \\ &= e^x(\cos(y) + i \sin(y))(dx + idy) \\ &= e^z dz \end{aligned}$$

Let A be an open subset of \mathbb{C} , let $f : A \rightarrow \mathbb{C}$ with $f = u + iv$ for functions u and v .² We can write:

$$f dz = (u + iv)(dx + idy) = (u + iv)dx + (iu - v)dy$$

If the 1-form $f dz$ is closed, we need $D_1(if) = D_2(f)$. Or equivalently, we need the following holds:

$$\begin{cases} D_1u = D_2v \\ D_2u = -D_1v \end{cases} \quad (\text{CR})$$

The system (CR) of equations is called the Cauchy-Riemann Equations.

Definition 22.0.0.3

Let A be an open subset of \mathbb{C} , a function $f : A \rightarrow \mathbb{C}$ of C^1 type, with $f = u + iv$, is holomorphic provided that $f dz$ is closed. Or equivalently, f is holomorphic provided that the Cauchy-Riemann Equations hold for the function f .

²Hereafter, we consider $\mathbb{C} \cong \mathbb{R}^2$ when defining a function $f : \mathbb{C} \rightarrow \mathbb{C}$, with $f = u + iv$, in which case it makes sense to write D_1u , D_2u , D_1v , and D_2v . That is, D_1u is the derivative of u with respect to the real part of the input, and D_2u is the derivative of u with respect to the complex part of the input.

Theorem 22.1

Let f be a holomorphic function defined on an open subset A of \mathbb{C} .

If A is diffeomorphic to a convex set, then $f dz$ is exact.

Proof. The proof of this theorem follows similarly from M.8 in Math 396 Supplement Materials. \square

Corollary 22.1.1 (Cauchy Integral Theorem)

Given a holomorphic function f defined on an open subset A of \mathbb{C} , where A is diffeomorphic to a convex set, and given $\alpha : [a, b] \rightarrow A$ being a piecewise C^1 function with $\alpha(a) = \alpha(b)$, we have:

$$\int_{Y_\alpha} f dz = 0$$

Example: Let A be an open subset of \mathbb{C} , the followings are holomorphic functions:

1. $f : A \rightarrow \mathbb{C} \quad z \mapsto z$
2. $f : A \rightarrow \mathbb{C} \quad z \mapsto \log_X z$
3. $f : A \rightarrow \mathbb{C} \quad z \mapsto c$ with $c \in \mathbb{C}$
4. $f : A \rightarrow \mathbb{C} \quad z \mapsto z^{-1}$, that is, we have $f(x + iy) = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}$
5. $f : A \rightarrow \mathbb{C} \quad z \mapsto e^z$, that is, we have $f(x + iy) = e^x \cos(y) + ie^x \sin(y)$

Example: Let A be an open subset of \mathbb{C} , the followings are non-holomorphic functions:

1. $f : A \rightarrow \mathbb{C} \quad z \mapsto \bar{z}$, that is, we have $f(x + iy) = x - iy$
2. $f : A \rightarrow \mathbb{C} \quad z \mapsto \Re(z)$, that is, we have $f(x + iy) = x$
3. $f : A \rightarrow \mathbb{C} \quad z \mapsto \Im(z)$, that is, we have $f(x + iy) = y$

Exercise: Let f and g be holomorphic functions defined on an open subset A of \mathbb{C} .

Then $f + g$, $f - g$ are holomorphic functions on A

Lemma 22.1.2

Let f and g be holomorphic functions defined on an open subset A of \mathbb{C} . Then $f \cdot g$ is a holomorphic function on A .

Proof.

$$D_2(f \cdot g) = f D_2 g + g D_2 f = i f D_1 g + i g D_1 f = i(D_1(f \cdot g))$$

The result follows. \square

Lemma 22.1.3

Let f and g be holomorphic functions such that $g \circ f$ is well defined. Then $g \circ f$ is a holomorphic function.

Proof. For $z \in \mathbb{C}$, Cauchy-Riemann Equations require the followings hold:

$$D \begin{bmatrix} \Re(f)(z) \\ \Im(f)(z) \end{bmatrix} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \quad D \begin{bmatrix} \Re(g)(z) \\ \Im(g)(z) \end{bmatrix} = \begin{bmatrix} \gamma & -\delta \\ \delta & \gamma \end{bmatrix}$$

for some scalars $\alpha, \beta, \gamma, \delta$. Here we get:

$$D \begin{bmatrix} \Re(g \circ f)(z) \\ \Im(g \circ f)(z) \end{bmatrix} = \begin{bmatrix} \gamma & -\delta \\ \delta & \gamma \end{bmatrix} \cdot \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} = \begin{bmatrix} \sigma & -\tau \\ \tau & \sigma \end{bmatrix}$$

for some scalars σ, τ . The result follows. \square

Lemma 22.1.4

Let f and g be holomorphic functions defined on an open subset A of \mathbb{C} , with $g \neq 0$. Then $\frac{1}{g}$ and $\frac{f}{g}$ are holomorphic functions on A .

Corollary 22.1.4.1

The rational functions $\frac{P(z)}{Q(z)}$ are holomorphic on where defined.

Here $P(z)$ and $Q(z)$ are polynomial functions.

Example:

The function $f : \mathbb{C} \rightarrow \mathbb{C}$ $z \mapsto e^{-z^2}$ is holomorphic on \mathbb{C} .

The function $f : \mathbb{C} \setminus \{\vec{0}\} \rightarrow \mathbb{C}$ $z \mapsto \frac{e^{iz}-1}{z^2}$ is holomorphic on $\mathbb{C} \setminus \{\vec{0}\}$.

Proposition 22.1.5

Let A, B be open subsets of \mathbb{C} . If the function $f : A \rightarrow B$ is a holomorphic diffeomorphism.

Then the function f^{-1} is holomorphic.

Proof. Let $g := f^{-1}$, by Cauchy-Riemann Equation, we can write:

$$D \begin{bmatrix} \Re(f(z)) \\ \Im(f(z)) \end{bmatrix} = \begin{bmatrix} \alpha & -\beta \\ \beta & -\alpha \end{bmatrix}$$

with the following holds:

$$D \begin{bmatrix} \Re(g(f(z))) \\ \Im(g(f(z))) \end{bmatrix} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{\alpha}{\alpha^2+\beta^2} & \frac{\beta}{\alpha^2+\beta^2} \\ \frac{-\beta}{\alpha^2+\beta^2} & \frac{\alpha}{\alpha^2+\beta^2} \end{bmatrix}$$

which agrees with the form required by Cauchy-Riemann Equations. The result follows. \square

Example: Note that we have the following:

$$\int_{(x,y) \in \mathbb{R}^2} e^{-x^2-y^2} = \int_{(x,y) \in \mathbb{R}^2} e^{-x^2} e^{-y^2} = \int_{y \in \mathbb{R}} e^{-y^2} \left(\int_{x \in \mathbb{R}} e^{-x^2} \right) = \left(\int_{x \in \mathbb{R}} e^{-x^2} \right)^2$$

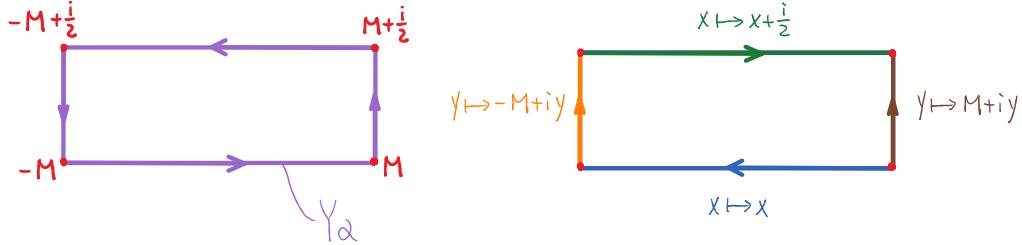
Or we can apply the Change of Variable Theorem:

$$\begin{aligned} \int_{(x,y) \in \mathbb{R}^2} e^{-x^2-y^2} &= \int_{\mathbb{R}^2 \setminus \{(x,0) | x \leq 0\}} e^{-x^2-y^2} \\ &= \int_0^\infty \int_{-\pi}^\pi e^{-r^2} r d\Phi dr \\ &= 2\pi \int_0^\infty e^{-r^2} r dr \\ &= \pi \int_0^\infty e^{-u} du \\ &= \pi \end{aligned}$$

So we see that we have:

$$\int_{x \in \mathbb{R}} e^{-x^2} = \sqrt{\pi}$$

Example: Consider the closed 1-form $e^{-z^2} dz$ on \mathbb{C} and the parametrized 1-manifold given by:



By Cauchy Integral Theorem, we have $\int_{Y_\alpha} e^{-z^2} dz = 0$, but here we can also write:

$$\begin{aligned} \int_{Y_\alpha} e^{-z^2} dz &= \int_{-M}^M e^{-x^2} dx - \int_{-M}^M e^{-(x+i/2)^2} dx + \int_0^{1/2} e^{-(M+iy)^2} dy - \int_0^{1/2} e^{(-M+iy)^2} dy \\ &= \int_{-M}^M e^{-x^2} dx - \int_{-M}^M e^{-x^2-i2My+y^2} dx + \int_0^{1/2} e^{-M^2-i2My+y^2} dy - \int_0^{1/2} e^{-M^2+i2My+y^2} dy \end{aligned}$$

Now suppose M approaches infinity, $\int_0^{1/2} e^{-M^2-i2My+y^2} dy - \int_0^{1/2} e^{-M^2+i2My+y^2} dy$ vanishes.

In which case we obtain the following:

$$\int_{-\infty}^{\infty} e^{-x^2} dx = e^{1/4} \int_{-\infty}^{\infty} e^{-x^2} (\cos(x) - i \sin(x)) dx$$

where we have:

$$\sqrt{\pi} = \int_{-\infty}^{\infty} e^{-x^2} dx$$

Hence we see that:

$$0 = -e^{1/4} \int_{-\infty}^{\infty} e^{-x^2} \sin(x) dx \quad \sqrt{\pi} = e^{-1/4} \int_{-\infty}^{\infty} e^{-x^2} \cos(x) dx$$

Rearranging we can write the following:

$$\int_{-\infty}^{\infty} e^{-x^2} \cos(x) dx = e^{-1/4} \sqrt{\pi}$$

Exercise: ³

$$\int_{-\infty}^{\infty} e^{-x^2} \cos(ax) dx = e^{-a^2/4} \sqrt{\pi}$$

Definition 22.1.5.0.1

The orientation of a 1-manifold in \mathbb{R}^n can be viewed as the direction of motion on the manifold, and one might argue the positive tangent vectors point in the direction of forward motion. While now we suppose a compact 1-manifold in \mathbb{R}^2 , and suppose $\mathbb{R}^2 \setminus M$ is disconnected and has two components, the default orientation of M is set by the condition that if we travel along M , then the bounded component of $\mathbb{R}^2 \setminus M$ lies to the left.

³In Fourier analysis, one can build generation functions from $\cos(ax)$ and $\sin(ax)$. That is, the decomposition of a function $f(x)$ relies on the followings:

$$\int_{-\infty}^{\infty} f(x) \cos(ax) dx \quad \int_{-\infty}^{\infty} f(x) \sin(ax) dx$$

Theorem 22.2 (Cauchy Integral Theorem)

Let C_1 and C_2 be disjoint circles in \mathbb{C} with C_2 lying inside C_1 , let A be an open set of points lying inside C_1 and outside of C_2 , let U be an open subset of \mathbb{C} containing $A \cup C_1 \cup C_2$, and let f be a holomorphic function on U , then we have the following holds:

$$\int_{C_1} f dz = \int_{C_2} f dz$$

Proof. This theorem is a consequence of Theorem 37.2 on Munkres. While alternatively, we can consider the following proof. First construct a regular P -gon inscribed in C_2 so that P and all points lying between P and C_2 are contained in U . Use the n rays from the center of C_2 through vertices of P to divide A into n pieces with each piece bounded by a piecewise C^1 curve Y_j consisting of an arc of C_1 , an arc of C_2 and two ray segments. Note that by pushing the arc of C_2 out to the corresponding side of P we obtain a compact convex set $K \subseteq U$ containing Y_j . Slightly pushing out the pieces of $Bd(K_j)$ we obtain an open convex set $V_j \subseteq U$ containing Y_j . Then by Corollary 22.1.1, we know that $\int_{Y_j} f dz = 0$. Summing over Y_j we obtain the desired result. \square

Corollary 22.2.1

Let U be an open subset of \mathbb{C} with some $z_0 \in U$, let f be a holomorphic function defined on $U \setminus \{z_0\}$, then for $K = \{z \in \mathbb{C} \mid \|z - z_0\| = r\} \subseteq U$, we have $\frac{1}{2\pi i} \int_K f dz$ being independent of r .

Definition 22.2.1.0.1

Let U be an open subset of \mathbb{C} with some $z_0 \in U$, let f be a holomorphic function defined on $U \setminus \{z_0\}$, then for $K = \{z \in \mathbb{C} \mid \|z - z_0\| = r\} \subseteq U$. The residue of $f dz$ at z_0 , denoted as $Res(f dz, z_0)$, is defined by the following:

$$Res(f dz, z_0) := \frac{1}{2\pi i} \int_K f dz$$

Theorem 22.3

Let U be an open subset of \mathbb{C} , let D be a closed disc in U with $z_0 \in Int(D)$, and let f be a holomorphic function defined on $U \setminus \{z_0\}$. We have the following holds:

$$\int_{Bd(D)} f dz = 2\pi i Res(f dz, z_0)$$

Proof. Theorem 22.3 follows immediately from the Fundamental Theorem of Laziness. \square

Let g be a holomorphic function defined on an open subset U of \mathbb{C} with $z_0 \in U$.

We have the following holds:

$$\begin{aligned}
 Res\left(\frac{g(z)}{z - z_0} dz, z_0\right) &= \frac{1}{2\pi i} \int_{||z-z_0||=r} \frac{g(z)}{z - z_0} dz \\
 &= \frac{1}{2\pi i} \int_0^{2\pi} \frac{g(z_0 + re^{it})}{re^{it}} ire^{it} dt \\
 &= \frac{1}{2\pi} \int_0^{2\pi} g(z_0 + re^{it}) dt \\
 &= \frac{1}{2\pi} \left(\lim_{r \rightarrow 0} \int_0^{2\pi} g(z_0 + re^{it}) dt \right) \\
 &= \frac{1}{2\pi} \int_0^{2\pi} g(z_0) dt \\
 &= g(z_0)
 \end{aligned}$$

Note that, since g is continuous at z_0 , $g(z_0 + re^{it})$ converges uniformly to $g(z_0)$ as r approaches 0.

ML-Estimate

Proposition 22.3.1 (ML-estimate in \mathbb{R}^n)

Let $Y_\alpha \subseteq \mathbb{R}^n$ be a parametrized 1-manifold parametrized by $\alpha : [a, b] \rightarrow Y_\alpha$. Let ω be a 1-form defined on an open subset of \mathbb{R}^n containing Y_α . Let $\text{length}(Y_\alpha)$ denote the 1-volume of the manifold Y_α . Then we have the following holds:

$$\left\| \int_{Y_\alpha} \omega \right\| \leq \sup_{\vec{v} \in Y_\alpha} \|\omega(\vec{v})\| \cdot \text{length}(Y_\alpha)$$

Proof. Here we can write the following:

$$\begin{aligned} \left\| \int_{Y_\alpha} \omega \right\| &= \left\| \int_a^b (\omega \circ \alpha) \cdot \alpha' \right\| \leq \int_a^b \|(\omega \circ \alpha) \cdot \alpha'\| \\ &\leq \int_a^b \|\omega \circ \alpha\| \cdot \|\alpha'\| \\ &\leq \sup \|\omega\| \cdot \int_a^b d\psi \\ &= \sup \|\omega\| \cdot \text{length}(Y_\alpha) \end{aligned}$$

The result follows. \square

Theorem 22.4 (ML-estimate in \mathbb{C})

Let $Y_\alpha \subseteq \mathbb{C}$ be a parametrized 1-manifold parametrized by $\alpha : [a, b] \rightarrow Y_\alpha$. Let $f : A \rightarrow \mathbb{C}$ be a continuous function with A being an open open subset of \mathbb{C} containing Y_α . Let $\text{length}(Y_\alpha)$ denote the 1-volume of Y_α . Then we have the following holds:

$$\left\| \int_{Y_\alpha} f dz \right\| \leq \sup_{z \in Y_\alpha} |f(z)| \cdot \text{length}(Y_\alpha)$$

Proof. Note that we have $dz = [1 \ i]$, we can write the following:

$$\int_{Y_\alpha} f dz = \int_a^b f(t) [1 \ i] \begin{bmatrix} \alpha'_1(t) \\ \alpha'_2(t) \end{bmatrix} dt = \int_a^b f(t) \cdot (\alpha'_1(t) + i\alpha'_2(t)) dt$$

Hence we have:

$$\left\| \int_{Y_\alpha} f dz \right\| \leq \int_a^b |f(t)| \cdot \|\alpha'(t)\| dt \leq \sup_{Y_\alpha} |f| \cdot \text{length}(Y_\alpha)$$

The result follows. \square

Complex Differentiability

Consider $dz = dx + idy$ and $d\bar{z} = dx - idy$, we have:

$$dx = \frac{dz + d\bar{z}}{2} \quad dy = \frac{dz - d\bar{z}}{2}$$

Let α, β be \mathbb{C} -valued functions. We first show that $\alpha dx + \beta dy$ has one-to-one correspondence with some function \mathbb{C} -valued function g, h such that $\alpha dx + \beta dy = g dz + h d\bar{z}$. Here we can write:

$$\alpha dx + \beta dy = \frac{\alpha - i\beta}{2} dz + \frac{\alpha + i\beta}{2} d\bar{z} \quad (\text{E})$$

Here we let g, h be defined such that equation (E) holds, then the result follows.

Consider $f \in C^1(U, \mathbb{C})$ with $f = u + iv$, where U is an open subset of \mathbb{C} , first note that we can write:

$$Df = [D_1 f \quad D_2 f] = [D_1 u + i D_1 v \quad D_2 u + i D_2 v]$$

Here we use $\frac{\partial f}{\partial x}$ to denote $D_1 f$, and $\frac{\partial f}{\partial y}$ to denote $D_2 f$. Then we can write the following:

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right) dz + \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right) d\bar{z}$$

Here we let:

$$\frac{\partial f}{\partial z} := \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right) \quad \frac{\partial f}{\partial \bar{z}} := \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right)$$

Note that we have:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial z} + \frac{\partial f}{\partial \bar{z}} \quad \frac{\partial f}{\partial y} = i \frac{\partial f}{\partial z} - i \frac{\partial f}{\partial \bar{z}}$$

Definition 23.0.0.0.1

Let U be an open subset of \mathbb{C} , let $f : U \rightarrow \mathbb{C}$ be a function. f is said to be differentiable in real sense at $t \in U$ provided that $u : U \rightarrow \mathbb{R}$ $z \mapsto \Re(f(z))$ and $v : U \rightarrow \mathbb{R}$ $z \mapsto \Im(f(z))$ are both differentiable at t . Here we consider $\mathbb{C} \cong \mathbb{R}^2$ when evaluating the differentiability of u and v .

Here we note that, for U being an open subset of \mathbb{C} , if a function $f : U \rightarrow \mathbb{C}$ is of C^1 type, then f is differentiable in the real sense, but in the following discussion we will investigate how we can define the differentiability of the function f in complex sense, given the additional algebraic structures of \mathbb{C} .

Now let f be a function of C^1 type defined on U , where U is an open subset of \mathbb{C} . Here we consider affine approximation of f at $z_0 \in U$, for $z = x + iy$ in a small neighborhood of z_0 , we can write:

$$f(z) \approx f(z_0) + \frac{\partial f}{\partial x}(z_0)(x - x_0) + \frac{\partial f}{\partial y}(z_0)(y - y_0) = f(z_0) + \frac{\partial f}{\partial z}(z_0)(z - z_0) + \frac{\partial f}{\partial \bar{z}}(z_0)(\bar{z} - \bar{z}_0)$$

By the differentiability of f in real sense, we have the following holds:

$$\lim_{(z-z_0) \rightarrow 0} \frac{f(z) - f(z_0) - \frac{\partial f}{\partial z}(z_0)(z - z_0) - \frac{\partial f}{\partial \bar{z}}(z_0)(\bar{z} - \bar{z}_0)}{z - z_0} = 0 \quad (\text{P})$$

We would like to evaluate the following limit and define it as the complex derivative of f :

$$\lim_{(z-z_0) \rightarrow 0} \frac{f(z) - f(z_0)}{z - z_0} \quad (\text{L})$$

Suppose that we have $\frac{\partial f}{\partial \bar{z}}(z_0) = 0$, then by (P) we get the following:

$$\lim_{(z-z_0) \rightarrow 0} \frac{f(z) - f(z_0)}{z - z_0} = \frac{\partial f}{\partial z}(z_0)$$

On the other hand, if $\frac{\partial f}{\partial \bar{z}}(z_0) \neq 0$, we can write:

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{\partial f}{\partial z}(z_0) + \frac{\partial f}{\partial \bar{z}}(z_0) \frac{\bar{z} - \bar{z}_0}{z - z_0} + \frac{f(z) - f(z_0) - \frac{\partial f}{\partial z}(z_0)(z - z_0) - \frac{\partial f}{\partial \bar{z}}(z_0)(\bar{z} - \bar{z}_0)}{z - z_0} \quad (\text{F})$$

Here $\frac{\partial f}{\partial z}(z_0)$ does not depend on $z - z_0$, while $\frac{\partial f}{\partial \bar{z}}(z_0) \frac{\bar{z} - \bar{z}_0}{z - z_0}$ depends on $z - z_0$.

Denote $z - z_0 = he^{i\Phi}$ with $h = ||z - z_0||$, we can rewrite (F) as the following:

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{\partial f}{\partial z}(z_0) + \frac{\partial f}{\partial \bar{z}}(z_0)he^{-i2\Phi} + \frac{f(z) - f(z_0) - \frac{\partial f}{\partial z}(z_0)(z - z_0) - \frac{\partial f}{\partial \bar{z}}(z_0)(\bar{z} - \bar{z}_0)}{z - z_0}$$

we see that if $\frac{\partial f}{\partial \bar{z}} \neq 0$, the limit (L) depends on Φ , hence the limit is not defined.

When the limit (L) exists and is finite, f is said to be complex differentiable, and we define the complex derivative of f at z_0 as $f'_{\mathbb{C}}(z_0) := \frac{\partial f}{\partial z}(z_0)$. Notice that, f is of C^1 type, so $\frac{\partial f}{\partial z}(z_0)$ guarantees to exist, but f might not be complex differentiable because $\frac{\partial f}{\partial \bar{z}}(z_0)$ might not necessarily be zero, and hence limit (L) might not necessarily exists.

Definition 23.0.0.0.2

Let f be a function of C^1 type defined on U , where U is an open subset of \mathbb{C} . f is said to be complex differentiable, denoted as \mathbb{C} -differentiable, at $z_0 \in U$ provided that $\frac{\partial f}{\partial \bar{z}}(z_0) = 0$. If f is \mathbb{C} -differentiable, $f'_{\mathbb{C}}(z_0) = \frac{\partial f}{\partial z}(z_0)$ is called the derivative of f at z_0 .⁴

Theorem 23.1

Let f be a function defined on U , where U is an open subset of \mathbb{C} . The followings are equivalent:

1. f is holomorphic on U
2. f is of C^1 type on U , and $\frac{\partial f}{\partial \bar{z}} = 0$
3. f is \mathbb{C} -differentiable at each point in U , and $f'_{\mathbb{C}}$ is continuous.

⁴ $f'_{\mathbb{C}}(z_0)$ might be denoted as $f'(z_0)$ thereafter.

Corollary 23.1.1

Let U be an open subset of \mathbb{C} , let $D \subseteq U$ be a closed disc with $z_0 \in \text{Int}(D)$, and let g be a holomorphic function defined on U . Then we have the followings:

$$\int_{z \in Bd(D)} \frac{g(z)}{z - z_0} dz = 2\pi i g(z_0) \quad g'_{\mathbb{C}}(z_0) = \frac{1}{2\pi i} \int_{z \in Bd(D)} \frac{g(z)}{(z - z_0)^2} dz$$

Proof. We will show that $\frac{g(z_0+h)-g(z_0)}{h}$ approaches $\frac{1}{2\pi i} \int_{z \in Bd(D)} \frac{g(z)}{(z - z_0)^2} dz$ as h approaches 0:

$$\frac{g(z_0 + h) - g(z_0)}{h} = \frac{1}{2\pi i h} \int_{z \in Bd(D)} g(z) \left(\frac{1}{z - z_0 - h} - \frac{1}{z - z_0} \right) dz$$

Hence with ML-estimate, we see that:

$$\begin{aligned} (*) &:= \left\| \frac{g(z_0 + h) - g(z_0)}{h} - \frac{1}{2\pi i} \int_{z \in Bd(D)} \frac{g(z)}{(z - z_0)^2} dz \right\| \\ &\leq \frac{\text{length}(\text{Bd}(D))}{2\pi} \cdot |h| \cdot \left(\sup_{z \in Bd(D)} |g(z)| \right) \cdot \frac{1}{(d(z_0, \text{Bd}(D)) - |h|) \cdot d(z_0, \text{Bd}(D))^2} \end{aligned}$$

as h approaches 0, the fraction $(*)$ approaches 0. The result follows. \square

Corollary 23.1.2

Let U be an open subset of \mathbb{C} , let $D \subseteq U$ be a closed disc with $z_0 \in \text{Int}(D)$, and let g be a holomorphic function defined on U . Then we have:

$$g''_{\mathbb{C}}(z_0) = \frac{1}{2\pi i} \int_{Bd(D)} \frac{2g(z) dz}{(z - z_0)^3}$$

That is, $g''_{\mathbb{C}}(z_0)$ exists.

Corollary 23.1.3 (Differentiated Cauchy Integral Formula)

Let U be an open subset of \mathbb{C} , let $D \subseteq U$ be a closed disc with $z_0 \in \text{Int}(D)$, and let g be a holomorphic function defined on U . Then we have:

$$g_{\mathbb{C}}^{(m)}(z_0) = \frac{m!}{2\pi i} \int_{z \in Bd(D)} \frac{g(z) dz}{(z - z_0)^{m+1}}$$

The function g is infinitely \mathbb{C} -differentiable.

Complex Series

Theorem 24.1 (Taylor's Theorem)

Let $z_0 \in \mathbb{C}$, let f be a holomorphic function defined on an open subset Ω of \mathbb{C} that contains z_0 .

For all $z \in \mathbb{C}$ that satisfies $|z - z_0| < \rho$ for some $d(z_0, \text{Bd}(\Omega)) > \rho > 0$, we have:

$$f(z) = \sum_{k=0}^{\infty} \frac{f_{\mathbb{C}}^{(k)}(z_0)}{k!} (z - z_0)^k$$

Here we denote $f^{(0)} := f$, $0! := 1$, and $(z - z_0)^0 := 1$ for $z = z_0$.

Note: ALTIR⁵

1. In \mathbb{R} , the Taylor series could converge to the real-valued function f , example includes:

$$f(x) = \begin{cases} e^{1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Note here $f^{(k)}(0) = 0$ for all k .

2. In \mathbb{R} , the Taylor series might not converge to f except at the point of interest z_0 . Examples of these are connected with Stirling's Formula.

Remark: All complex differentiable functions defined on an open subset of \mathbb{C} is analytic.

Proof of Theorem 24.1. Pick $0 < r < \tilde{r} < \rho$. We will first prove the following Lemma:

Lemma 24.1.1

For $\zeta, z \in \mathbb{C}$, the series

$$\sum_{k=0}^{\infty} \left(\frac{z - z_0}{\zeta - z_0} \right)^k$$

converges uniformly to the limit:

$$\frac{1}{1 - \frac{z - z_0}{\zeta - z_0}}$$

for ζ and z satisfies $|\zeta - z_0| = \tilde{r}$ and $|z - z_0| \leq r$.

Proof. Here we can write:

$$\left| \frac{1}{1 - \frac{z - z_0}{\zeta - z_0}} - \sum_{k=0}^m \left(\frac{z - z_0}{\zeta - z_0} \right)^k \right| = \left| \frac{\left(\frac{z - z_0}{\zeta - z_0} \right)^{m+1}}{1 - \frac{z - z_0}{\zeta - z_0}} \right| \leq \frac{\left(\frac{r}{\tilde{r}} \right)^{m+1}}{1 - \frac{r}{\tilde{r}}} := (*)$$

Note that $(*)$ tends to 0 when m tends to infinity. The result of the Lemma follows. \square

⁵ "Ain't like this in \mathbb{R} ."

-David Barrett 1/24/2022

By Cauchy Integral Formula, we have the following holds:

$$|z - z_0| \leq r \quad \Rightarrow \quad f(z) = \frac{1}{2\pi i} \int_{|\zeta - z_0|=\tilde{r}} \frac{f(\zeta)d\zeta}{\zeta - z}$$

Here we have:

$$\begin{aligned} \frac{1}{2\pi i} \int_{|\zeta - z_0|=\tilde{r}} \frac{f(\zeta)d\zeta}{\zeta - z} &= \frac{1}{2\pi i} \int_{|\zeta - z_0|=\tilde{r}} \frac{1}{\zeta - z_0} \frac{1}{1 - \frac{z-z_0}{\zeta-z_0}} f(\zeta) d\zeta \\ &= \frac{1}{2\pi i} \int_{|\zeta - z_0|=\tilde{r}} \frac{1}{\zeta - z_0} \sum_{k=0}^{\infty} \left(\frac{z-z_0}{\zeta-z_0} \right)^k f(\zeta) d\zeta \\ &= \sum_{k=0}^{\infty} \frac{(z-z_0)^k}{2\pi i} \int \frac{f(\zeta) d\zeta}{(\zeta-z_0)^{k+1}} \\ &= \sum_{k=0}^{\infty} (z-z_0)^k \frac{f_{\mathbb{C}}^{(k)}(z_0)}{k!} \end{aligned}$$

This completes the proof. \square

Theorem 24.2

Let f be a holomorphic function defined on a open subset Ω of \mathbb{C} . Denote $E := \bigcap_{k=0}^{\infty} (f_{\mathbb{C}}^{(k)})^{-1}(0)$. If we have Ω being connected, then we have either $E = \emptyset$ or $f(z) = 0$ for all $z \in \Omega$.

Proof. Here we can write:

$$E = \{z_0 \in \Omega \mid f_{\mathbb{C}}^{(k)}(z_0) = 0 \ \forall k \in \mathbb{N} \cup \{0\}\} = \bigcap_{k=0}^{\infty} (f_{\mathbb{C}}^{(k)})^{-1}(0)$$

Since f is continuous, then the preimage of a close set is relatively closed. Hence $(f_{\mathbb{C}}^{(k)})^{-1}(0)$ is closed for all k , and hence E is relatively closed. Note that E is open because if some $z_0 \in \Omega$ satisfies $f_{\mathbb{C}}^{(k)}(z_0) = 0$ for all $k \in \mathbb{N} \cup \{0\}$, then by Taylor's Theorem, there is a open neighborhood B of z_0 such that $f_{\mathbb{C}}^{(k)}(z) = 0$ for all $z \in U$. Here we see that E is both open and relatively closed, hence we must have either $E = \emptyset$ or $E = \Omega$. The result follows. \square

Corollary 24.2.1

Let Ω be a connected open subset of \mathbb{C} that contains z_0 , let f_1 and f_2 be holomorphic functions on Ω , with $f_1^{(k)}(z_0) = f_2^{(k)}(z_0)$ for all k . Then we have $f_1(z) = f_2(z)$ for all $z \in \Omega$.

Corollary 24.2.2

Let $Holo(A)$ denote the set of holomorphic functions defined on a set A . Let V be an open connected subset of \mathbb{C} , let U be a nonempty open proper subset of V . The restriction map from $Holo(V)$ to $Holo(U)$ is injective.

Exercise: ALTIR. Let V be an open connected subset of \mathbb{C} , let U be a nonempty open proper subset of V . The restriction map from $Holo(V)$ to $Holo(U)$ is not surjective.⁶

Theorem 24.3

For $z_0 \in \mathbb{C}$, consider the following power series:

$$\sum_{k=0}^{\infty} a_k(z - z_0)^k \tag{S}$$

If $\sum_{k=0}^{\infty} a_k(z - z_0)^k$ converges pointwise on $|z - z_0| < r$ for some $r > 0$. Then the function defined by the sum (S) is holomorphic on the set $\{z \in \mathbb{C} \mid |z - z_0| < r\}$.

Proof. The proof of this theorem is given on Section 27 Theorem 6 in Spivak. \square

Theorem 24.4

Let $z_0 \in \mathbb{C}$, let Ω be a connected open subset of \mathbb{C} that contains z_0 , let f be a holomorphic function defined on Ω and being not all zero on Ω . Then there exists $m \in \mathbb{N} \cup \{0\}$ such that, for $z \in \Omega$, $f(z) = (z - z_0)^m h(z)$ with some holomorphic function h defined on Ω and $h(z_0) \neq 0$.

⁶Sometimes in several complex variables analysis, this is surjective

Proof. Let $u = d(z_0, \text{Bd}(\Omega))$. On $B_u(z_0)$, by Theorem 24.1, we have the Taylor series expansion for f around z_0 , here we can write the following, for $z \in \Omega$ that satisfies $|z - z_0| < u$, with $m = \min\{k \in \mathbb{N} \cup \{0\} \mid f_{\mathbb{C}}^{(k)}(z_0) \neq 0\}$ such that $a_m \neq 0$:

$$f(z) = \sum_{k=m}^{\infty} a_k (z - z_0)^k = (z - z_0)^m \sum_{k=m}^{\infty} a_k (z - z_0)^{k-m} = (z - z_0)^m \sum_{j=0}^{\infty} a_{j+m} (z - z_0)^j$$

Here we can define h on Ω by the following:

$$h : \Omega \rightarrow \mathbb{C} \quad z \mapsto \begin{cases} \sum_{j=0}^{\infty} a_{j+m} (z - z_0)^j & z \in B_u(z_0) \\ \frac{f(z)}{(z - z_0)^m} & z \in \Omega \setminus \{z_0\} \end{cases}$$

The result follows. \square

Definition 24.4.0.0.1

In the settings of Theorem 24.4, m is called the order of f at z_0 , denoted as $\text{ord}_{z_0} f := m$.

Corollary 24.4.1

Let $z_0 \in \mathbb{C}$, let Ω be a connected open subset of \mathbb{C} that contains z_0 , let f be a holomorphic function defined on Ω with $f(z) \neq 0$ for some $z \in \Omega$. Then there exists $r > 0$ such that $f(z) \neq 0$ for all $z \in \Omega$ that satisfies $0 < |z - z_0| < r$.

Proof. One can make use of Theorem 24.4. Pick r such that the function h defined in Theorem 24.4 is nonzero on $|z - z_0| < r$. The result follows immediately. \square

Corollary 24.4.2

Let K be a compact set that is contained in some connected open subset of \mathbb{C} , let f be a holomorphic function defined on Ω with $f(z) \neq 0$ for some $z \in \Omega$. Then $\#(K \cap f^{-1}(0)) < \infty$.

Proof. For each $z_i \in K$, since $f(z) \neq 0$ for some $z \in \Omega$, then by Corollary 24.4.1, there is a ball $B_{\epsilon_{z_i}}(z_i)$ such that $\#(B_{\epsilon_{z_i}}(z_i) \cap f^{-1}(0)) \in \{0, 1\}$. Here $K \subseteq \bigcup_{i=1}^M B_{\epsilon_{z_i}}(z_i)$ for some $M \in \mathbb{N}$, then we have $\#(K \cap f^{-1}(0)) \leq M$. The result follows. \square

Corollary 24.4.3

Let f_1 and f_2 be holomorphic functions on an open connected subset Ω of \mathbb{C} with $f_1 = f_2$ on some infinite subset of a compact subset of Ω . Then $f_1 = f_2$ on Ω .

Corollary 24.4.3.1 (Persistence of Relations)

Let f_1 and f_2 be holomorphic functions defined on an open connected subset Ω of \mathbb{C} that satisfies $\Omega \cap \mathbb{R} \neq \emptyset$. If with $f_1(z) = f_2(z)$ for all $z \in \Omega \cap \mathbb{R}$, then we have $f_1 = f_2$ on Ω .

Definition 24.4.3.1.1

A sequence of functions (f_j) defined on $\Omega \subseteq \mathbb{C}$ is said to converge almost uniformly to a function f defined on Ω provided that the sequence (f_j) converges uniformly to f on each compact subset K of Ω .

Theorem 24.5 (Weierstrass Convergence Theorem)

The limit of a almost uniformly convergent sequence of holomorphic functions is holomorphic.

Proof. Let (f_j) be a sequence of functions defined on an open subset Ω of \mathbb{C} . Suppose (f_j) converges almost uniformly to a function f defined on Ω . It suffices to show that f is a holomorphic function on $\text{Int}(D)$ where D is a closed disc of Ω . Note that by definition $f_j dz$ is closed in Ω , hence by Theorem M.5 from Math 396 Supplement materials, $f_j dz$ is exact on $\text{Int}(D)$, which, by Math 395 HW14 Q1, implies $f dz$ is exact on $\text{Int}(D)$. Then we can write:

$$f dz = dg = \frac{\partial g}{\partial z} dz + \frac{\partial g}{\partial \bar{z}} d\bar{z}$$

Here we see that $\frac{\partial g}{\partial \bar{z}} = 0$, so g is holomorphic, and hence $g'_{\mathbb{C}} = \frac{\partial g}{\partial z} = f$. By Differentiated Cauchy Integral Formula, we know that f is holomorphic on $\text{Int}(D)$. This completes the proof. \square

Given power series:

$$\sum_{n=0}^{\infty} a_n z^n$$

with $a_n, z \in \text{Complex}$, let R be defined by:

$$R := \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}}$$

and if $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 0$, let $R := \infty$, if $\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \infty$, let $R = 0$. One can show that the series converges almost uniformly when $|z| < R$, and diverges when $|z| > R$.

Exercise: Let $z_0 \in \mathbb{C}$, let (f_j) be a sequence of functions defined on an open set that contains $D = \{z \in \mathbb{C} \mid |z - z_0| < \rho\}$. (f_j) converges almost uniformly to some function f on D if and only if (f_j) converges uniformly to f on each $D_r = \{z \in \mathbb{C} \mid |z - z_0| \leq r\}$ for all $r < \rho$.

Example:

$$1 - x^2 + x^4 - x^6 + \dots = \frac{1}{1 + x^2} \quad \text{for } -1 < x < 1$$

6 | Differentiable Manifolds II

Tensors

Definition 25.0.0.0.1

Let V be a vector space, denote $V_i = V$. For $k \in \mathbb{N}$, $V^k := V_1 \times V_2 \times \cdots \times V_k$. V^k is the set of k -tuples $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)$ for vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in V$.

Definition 25.0.0.0.2

Let V be a vector space, a function $f : V^k \rightarrow \mathbb{R}$ is said to be linear in the i -th variable provided that for fixed $\vec{v}_j \in V$ with $j \neq i$, the function $T : V \rightarrow \mathbb{R} \quad \vec{v} \mapsto f(\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}, \vec{v}_{i+1}, \dots, \vec{v}_k)$ is linear.

Definition 25.0.0.0.3

Let V be a vector space. A function $f : V^k \rightarrow \mathbb{R}$ is said to be multilinear provided that f is linear in the i -th variable for all $1 \leq i \leq k$. If f is multilinear, then f is called a k -tensor on V .

Example: The function f defined by the following is a 2-tensor:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad \left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right) \mapsto x_1 y_2 + x_2 y_1$$

Example: The function f defined by the following is not a 2-tensor:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad \left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right) \mapsto x_1 y_1 + x_2 y_2$$

Definition 25.0.0.0.4

A k -tensor f defined on a vector space V is symmetric provided that the following holds:

$$f(\vec{v}_1, \dots, \vec{v}_{i+1}, \vec{v}_i, \dots, \vec{v}_k) = f(\vec{v}_1, \dots, \vec{v}_i, \vec{v}_{i+1}, \dots, \vec{v}_k)$$

Definition 25.0.0.0.5

A k -tensor f defined on a vector space V is alternating provided that the following holds:

$$f(\vec{v}_1, \dots, \vec{v}_{i+1}, \vec{v}_i, \dots, \vec{v}_k) = -f(\vec{v}_1, \dots, \vec{v}_i, \vec{v}_{i+1}, \dots, \vec{v}_k)$$

Note: Let f be a tensor defined on a vector space V . Let $\vec{v}_i = \vec{v}_j$ for some $i \neq j$.

f is alternating if and only if we have $f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_j, \dots, \vec{v}_k) = 0$.

In other words, for $f \in \mathcal{A}^k(V)$, if we have $\#\{i_1, i_2, \dots, i_k\} < k$, then $f(\vec{v}_{i_1}, \vec{v}_{i_2}, \dots, \vec{v}_{i_k}) = 0$.

Definition 25.0.0.6

Let V be a vector space. We define the followings:

$\mathcal{L}^k(V)$ is the set of all k -tensor defined on V .

$\text{Symm}^k(V)$ is the set of all symmetric k -tensor defined on V .

$\text{Alt}^k(V)$, or denoted as $\mathcal{A}^k(V)$, is the set of all alternating k -tensor defined on V .

Note: For a vector space V , the spaces $\mathcal{L}^k(V)$, $\text{Symm}^k(V)$, and $\mathcal{A}^k(V)$ are all vector spaces.

Note: For a vector space V , $\mathcal{L}^1(V) = \text{Symm}^1(V) = A^1(V)$ gives the dual space V^* of V .

Let V be a finite dimensional vector space with basis $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$.

For $\vec{a} \in V$, one can write $\vec{a} = \sum_n c_j \vec{a}_j$, then for a k -tensor f , we can write the following:

$$\begin{aligned} f\left(\sum_{j_1=1}^n c_{1,j_1} \vec{a}_{j_1}, \sum_{j_2=1}^n c_{2,j_2} \vec{a}_{j_2}, \dots, \sum_{j_k=1}^n c_{k,j_k} \vec{a}_{j_k}\right) &= \sum_{j_1=1}^n \left(c_{1,j_1} \cdot f\left(\vec{a}_{j_1}, \sum_{j_2=1}^n c_{2,j_2} \vec{a}_{j_2}, \dots, \sum_{j_k=1}^n c_{k,j_k} \vec{a}_{j_k}\right) \right) \\ &= \sum_{1 \leq j_1, j_2, \dots, j_k \leq n} c_{1,j_1} \cdot c_{2,j_2} \cdots c_{k,j_k} \cdot f(\vec{a}_{j_1}, \vec{a}_{j_2}, \dots, \vec{a}_{j_k}) \end{aligned}$$

so the value of f is completely determined by $f(\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_k})$, where $(\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_k})$ is some combination of the basis vectors $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$.

Just as a linear transformation from vector spaces V to W can be defined by specifying its values arbitrarily on basis elements for V , a k -tensor defined on V can be defined by specifying its values arbitrarily on k -tuples of basis elements. This fact is a consequence of the next theorem:

Theorem 25.1

Let V be an n -dimensional vector space with a basis $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$. Let $I = (i_1, i_2, \dots, i_k)$ be a k -tuple of integers from the set $\{1, 2, \dots, n\}$. There exists a unique k -tensor Φ_I on V such that for every k -tuple $M = (m_1, m_2, \dots, m_k)$ of integers from the set $\{1, 2, \dots, n\}$, we have:

$$\Phi_I(\vec{a}_{m_1}, \vec{a}_{m_2}, \dots, \vec{a}_{m_k}) = \begin{cases} 0 & I \neq M \\ 1 & I = M \end{cases}$$

Proof. One can define Φ_I by the following:

$$\Phi_I : V^k \rightarrow \mathbb{R} \quad \left(\sum_{j_1=1}^n c_{1,j_1} \vec{a}_{j_1}, \sum_{j_2=1}^n c_{2,j_2} \vec{a}_{j_2}, \dots, \sum_{j_k=1}^n c_{k,j_k} \vec{a}_{j_k} \right) \mapsto c_{1,i_1} \cdot c_{2,i_2} \cdots \cdot c_{k,i_k} \quad (\text{I})$$

Here one can show that Φ_I satisfies the conditions and is a k -tensor. The rest of the proof of this theorem follows from Theorem 26.3 from on Munkres. \square

If one define Φ_I with equation (I) by using the basis $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ of V , then we have the following proposition, with the exact same settings as in Theorem 25.1:

Proposition 25.1.1

Let V be an n -dimensional vector space with a basis $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$. Let $I = (i_1, i_2, \dots, i_k)$ be a k -tuple of integers from the set $\{1, 2, \dots, n\}$. For $f \in \mathcal{L}^k(V)$, we have $f = \sum_I f(\vec{a}_I) \Phi_I$, where we write $\vec{a}_I := (\vec{a}_{m_1}, \vec{a}_{m_2}, \dots, \vec{a}_{m_k})$.

Proof. The proposition follows immediately if one expand the equation $f = \sum_I f(\vec{a}_I) \Phi_I$. \square

Given an n -dimensional vector space V and a basis $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$, let Φ_I be defined with equation (I), we conclude that the set defined by the following forms a basis for $\mathcal{L}^k(V)$:

$$\mathbb{P} := \{\Phi_I \mid I = (i_1, i_2, \dots, i_k) \text{ is a } k\text{-tuple of integers from the set } \{1, 2, \dots, n\}\} \quad (\text{P})$$

Remark: The definition of \mathbb{P} depends on the choice of basis of V .

Definition 25.1.1.0.1

Let V be a vector space. The tensor product operator \otimes is defined to be the following:

$$\otimes : \mathcal{L}^k(V) \times \mathcal{L}^l(V) \rightarrow \mathcal{L}^{k+l}(V) \quad (f, g) \mapsto f \otimes g$$

where, for $f \in \mathcal{L}^k(V)$ and $g \in \mathcal{L}^l(V)$, we write:

$$f \otimes g : V^{k+l} \rightarrow \mathbb{R} \quad (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k, \vec{v}_{k+1}, \dots, \vec{v}_{k+l}) \mapsto f(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) \cdot g(\vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_{k+l})$$

Let V be a vector space, for $f \in \mathcal{L}^k(V)$, $h \in \mathcal{L}^m(V)$, and $g \in \mathcal{L}^l(V)$, and $c \in \mathbb{R}$, we have:

1. $(f \otimes g) \otimes h = f \otimes (g \otimes h)$
2. $(c \cdot f) \otimes g = c \cdot (f \otimes g) = f \otimes (c \cdot g)$
3. $(f + g) \otimes h = f \otimes h + g \otimes h$
4. $f \otimes (g + h) = f \otimes g + f \otimes h$

Let V be an n -dimensional vector space over the field F , with basis $\mathbb{A} = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$, let $\mathbb{D} = (\Phi_1, \Phi_2, \dots, \Phi_n)$ be the dual basis of V associated to the basis \mathbb{A} , we can write:

$$\Phi_i : V \rightarrow F \quad \sum_{j=1}^n t_j \vec{a}_j \mapsto t_i$$

Note that \mathbb{D} forms a basis of $\mathcal{A}^1(V)$. Moreover, let \mathbb{P} be the basis of $\mathcal{L}^k(V)$ as defined by equation (P), for $\Phi_I \in \mathbb{P}$, we can write the following, with $I = (i_1, i_2, \dots, i_k)$:

$$\Phi_I := \Phi_{i_1} \otimes \Phi_{i_2} \otimes \cdots \otimes \Phi_{i_k}$$

Definition 25.1.1.0.2

Let V and W be vector spaces, let $T : V \rightarrow W$ be a linear transformation. We define the dual transformation T^* by the following:

$$T^* : \mathcal{L}^k(W) \rightarrow \mathcal{L}^k(V) \quad f \mapsto T^* f$$

where, for $f \in \mathcal{L}^k(W)$, and $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in V$, we have:

$$T^* f : V^k \rightarrow \mathbb{R} \quad (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) \mapsto f(T(\vec{v}_1), T(\vec{v}_2), \dots, T(\vec{v}_k))$$

Note: In the settings of Definition 25.1.1.0.2, it is immediate that $T^* f$ is multilinear because T is linear and f is multilinear. It is also true that T^* itself is linear, as map of tensors. Let X be a vector space, and let $S : W \rightarrow X$ be a linear transformation, then we have the followings hold:

1. $T^*(f \otimes g) = T^* f \otimes T^* g$ for all $f, g \in \mathcal{L}^k(W)$
2. $(S \circ T)^* f = T^*(S^* f)$ for all $f \in \mathcal{L}^k(W)$

The Permutation Group

Definition 26.0.0.0.1

Let $k \geq 2$. A permutation of the set of integers $K = \{1, 2, \dots, k\}$ is a bijection function σ from K to K . We denote the set of all such permutations by S_k .

Note: For $k \geq 2$, if $\sigma, \tau \in S_k$, then $\sigma \circ \tau, \sigma^{-1} \in S_k$. The set S_k equipped with function composition operation forms a group called the symmetric group, or the permutation group, on the set $\{1, 2, \dots, k\}$. There are $k!$ elements in such group.

Definition 26.0.0.0.2

For $k \geq 2$, let $K := \{1, 2, \dots, k\}$, and given $1 \leq i \leq k$, let e_i be the element of S_k defined by:

$$e_i : K \rightarrow K \quad j \mapsto \begin{cases} j & j \neq i \text{ and } j \neq i+1 \\ i+1 & j = i \\ i & j = i+1 \end{cases}$$

The element e_i is called an elementary permutation on K .

Note: In the settings of Definition 26.0.0.0.2, we have $e_i \circ e_i$ being the identity permutation, so e_i is its own inverse in the group S_k equipped with function composition operation.

Lemma 26.0.1 (Lemma 27.1 on Munkres)

Let $\sigma \in S_k$, σ is a composition of elementary permutations in S_k .

Definition 26.0.1.0.1

Let $\sigma \in S_k$. Consider the set of all pairs of integers (i, j) from the set $K := \{1, 2, \dots, k\}$, for which $i < j$ and $\sigma(i) > \sigma(j)$. Each such pair is called an inversion in σ . We define the sign of σ , denoted as $\text{sgn}(\sigma)$, or $\text{sgn}\sigma$ by the following:

$$\text{sgn} : S_k \rightarrow \{0, 1\} \quad \sigma \mapsto \begin{cases} 1 & \text{the number of inversion in } \sigma \text{ is even} \\ -1 & \text{the number of inversion in } \sigma \text{ is odd} \end{cases}$$

We call σ an odd or even permutation according to the sign of σ equals to -1 or $+1$, respectively.

Lemma 26.0.2 (Lemma 27.2 on Munkres)

Let $\sigma \in S_k$. If σ is a composition of m elementary permutations, then $\text{sgn}(\sigma) = (-1)^m$.

Lemma 26.0.3 (Lemma 27.2 on Munkres)

Let $\sigma, \tau \in S_k$, we have $\text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau)$, and $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$.

Lemma 26.0.4 (Lemma 27.2 on Munkres)

For $p, q \in \{1, 2, \dots, k\}$ with $p \neq q$, if τ is a permutation in S_k that exchanges p and q and leaves all other integers fixed, then $\text{sgn}(\tau) = -1$.

The Wedge Product

Definition 27.0.0.0.1

For a family $I \in \{1, 2, \dots, n\}^k$, $I = (i_1, i_2, \dots, i_k)$ is said to be ascending provided that we have:¹

$$1 \leq i_1 < i_2 < \dots < i_k \leq n$$

Note: For a family $J = (j_1, j_2, \dots, j_k) \in \{1, 2, \dots, n\}^k$. If $\#(\{j_1, j_2, \dots, j_k\}) = k$, then there exists a unique σ , and a unique ascending $I = (i_1, i_2, \dots, i_k) \in \{1, 2, \dots, n\}^k$ such that we have:

$$J = I_\sigma := (i_{\sigma(1)}, i_{\sigma(2)}, \dots, i_{\sigma(k)})$$

Let V be a vector space with basis $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$, and let \mathbb{P} denote the basis of $\mathcal{L}^k(V)$ as defined in section 25 equation (P) in this text. For $f \in \mathcal{A}^k(V)$, we can write the following:

$$f = \sum_{\substack{I \in \{1, 2, \dots, n\}^k \\ I \text{ has no duplicate}}} f(\vec{a}_I) \Phi_I = \sum_{\substack{I \in \{1, 2, \dots, n\}^k \\ I \text{ is ascending}}} \sum_{\sigma \in S_k} f(\vec{a}_{I_\sigma}) \Phi_{I_\sigma}$$

Note that $f(\vec{a}_{I_\sigma}) = \text{sgn}(\sigma)f(\vec{a}_I)$ because f is alternating, then we can write:

$$f = \sum_{\substack{I \in \{1, 2, \dots, n\}^k \\ I \text{ is ascending}}} \sum_{\sigma \in S_k} \text{sgn}(\sigma)f(\vec{a}_I) \Phi_{I_\sigma} = \sum_{\substack{I \in \{1, 2, \dots, n\}^k \\ I \text{ is ascending}}} f(\vec{a}_I) \sum_{\sigma \in S_k} \text{sgn}(\sigma) \Phi_{I_\sigma} \quad (\text{A})$$

For ascending I , we define the following:

$$\Psi_I := \sum_{\sigma \in S_k} \text{sgn}(\sigma) \Phi_{I_\sigma}$$

From equation (A), we observe that the set defined by the following forms a basis for $\mathcal{A}^k(V)$:

$$\mathbb{S} := \left\{ \Psi_I \mid I \text{ is an ascending } k\text{-tuple in } \{1, 2, \dots, n\}^k \right\}$$

Note here the definition of \mathbb{S} depends on the definition of \mathbb{P} , and the definition of \mathbb{P} depends on the choice of the basis of V , so we know that the definition of each elements in the basis \mathbb{S} of $\mathcal{A}^k(V)$ depends on the choice of basis of V .

In such setting, if we have $k > n$, then $A^k(V) = \{\vec{0}\}$. If we have $1 \leq k \leq n$, then the dimension of $\mathcal{A}^k(V)$ equals to the cardinality of ascending k -tuples in $\{1, 2, \dots, n\}^k$, which is equals to C_n^k . Moreover, we have $\mathcal{A}^n(V) = \text{span}(\{\Psi_{(1, 2, \dots, n)}\})$. We further notice that, for $f \in \mathcal{A}^k(V)$ and $g \in \mathcal{A}^l(V)$, $f \otimes g$ might not necessarily belongs to $\mathcal{A}^{k+l}(V)$.

Note: Let V be a finite dimensional vector space with basis $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$. Let $(\Phi_1, \Phi_2, \dots, \Phi_n)$ be the dual basis of $V^* = \mathcal{L}^1(V) = \mathcal{A}^1(V)$. For $I = (i_1, i_2, \dots, i_k) \in \{1, 2, \dots, n\}^k$, let $\Phi_I = \Phi_{i_1} \otimes \Phi_{i_2} \otimes \dots \otimes \Phi_{i_k}$, here $\{\Phi_I \mid I = (i_1, i_2, \dots, i_k) \in \{1, 2, \dots, n\}^k\}$ forms a basis for $\mathcal{L}^k(V)$. For ascending $I = (i_1, i_2, \dots, i_k)$, set $\Psi_I = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \Phi_{I_\sigma}$, here $I_\sigma = (i_{\sigma(1)}, \dots, i_{\sigma(k)})$, and $\Psi_I \in \mathcal{A}^k(V)$. $\{\Psi_I \mid I = (i_1, i_2, \dots, i_k) \text{ is ascending}\}$ forms a basis for $\mathcal{A}^k(V)$.

¹Here $\{1, 2, \dots, n\}^k$ is viewed as a set of families. Each family in this set has entries from $\{1, 2, \dots, n\}$.

Theorem 27.1

Let V be a vector space, there exists a function $W : \mathcal{A}^k(V) \times \mathcal{A}^l(V) \rightarrow \mathcal{A}^{k+l}(V)$ $(f, g) \mapsto f \wedge g$ such that $f \wedge g \in \mathcal{A}^{k+l}(V)$ for $f \in \mathcal{A}^k(V)$, $g \in \mathcal{A}^l(V)$, and satisfies all of the followings:

1. For $f \in \mathcal{A}^k(V)$, $g \in \mathcal{A}^l(V)$, and $h \in \mathcal{A}^m(V)$, we have $f \wedge (g \wedge h) = (f \wedge g) \wedge h$
2. For $f \in \mathcal{A}^k(V)$, $g \in \mathcal{A}^l(V)$, and scalar c , we have $(c \cdot f) \wedge g = c \cdot (f \wedge g) = f \wedge (c \cdot g)$
3. For $f, g \in \mathcal{A}^k(V)$ and $h \in \mathcal{A}^l(V)$, we have $h \wedge (f + g) = h \wedge f + h \wedge g$, and $(f + g) \wedge h = f \wedge h + g \wedge h$
4. For $f \in \mathcal{A}^k(V)$ and $g \in \mathcal{A}^l(V)$, we have $g \wedge f = (-1)^{kl} \cdot f \wedge g$
5. Given a finite basis of V , let $(\Phi_i \mid 1 \leq i \leq n)$ be the corresponding dual basis for V^* , and let $(\Psi_I \mid I \text{ is an ascending } k\text{-tuple of integers from the set } \{1, 2, \dots, n\})$ be the corresponding family of elementary alternating tensors. For any ascending k -tuple $I = (i_1, i_2, \dots, i_k)$ of integers from the set $\{1, 2, \dots, n\}$, we can write $\Psi_I = \Phi_{i_1} \wedge \Phi_{i_2} \wedge \dots \wedge \Phi_{i_k}$.
6. Let $T : V \rightarrow W$ be a linear transformation with W being a vector space, let f and g be alternating tensors on W , then we have $T^*(f \wedge g) = T^*f \wedge T^*g$.

Proof. Let $[I]$ denote the set of ascending k -tuple with entry from $Z = \{1, 2, \dots, n\}$, and let $[J]$ denote the set of ascending l -tuple with entry from Z , let $f = \sum_{I \in [I]} \alpha_I \Psi_I$ and $g = \sum_{J \in [J]} \beta_J \Psi_J$. Then we can define:

$$f \wedge g := \sum_{\substack{I \in [I], J \in [J] \\ \text{no duplicates}}} \alpha_I \beta_J \Psi_I \wedge \Psi_J := \sum_{\substack{I \in [I], J \in [J] \\ \text{no duplicates}}} \alpha_I \beta_J \operatorname{sgn}(I, J) \Psi_{\operatorname{sort}(I, J)}$$

here no duplicate means $I \cap J = \emptyset$, (I, J) denote the $(k+l)$ -tuple with the first k entries being I and then the last l entries being J , sort denote the function that sorts entries in the concatenated $(k+l)$ -tuple (I, J) , which is equivalent to apply a corresponding permutation σ_{IJ} to the tuple (I, J) that sorts (I, J) , and $\operatorname{sgn}(I, J)$ denotes $\operatorname{sgn}(\sigma_{IJ})$.

Here (2), (3), and (5) are cleared from such definition.

For (4), we consider the special case $f = \Psi_I$ and $g = \Psi_J$. It takes kl steps to go from (J, I) to (I, J) , that is, we have $\operatorname{sgn}(J, I) = (-1)^{kl} \operatorname{sgn}(I, J)$, so we have $\Psi_J \wedge \Psi_I = (-1)^{kl} \Psi_I \wedge \Psi_J$. For a more general case, one can decompose the tensor using Ψ_I and Ψ_J .

For (1), consider the special case where $f = \Psi_I$, $g = \Psi_J$, and $h = \Psi_K$, with I, J, K being disjoint. Here we can write the following:

$$\operatorname{sgn}(I, J) \cdot \operatorname{sgn}(\operatorname{sort}(I, J), K) = \operatorname{sgn}(I, \operatorname{sort}(J, K)) \cdot \operatorname{sgn}(J, K)$$

so $(\Psi_I \wedge \Psi_J) \wedge \Psi_K = \Psi_I \wedge (\Psi_J \wedge \Psi_K)$. The general case is left of exercise.

For (6), we first define the following function:

$$\hat{A} : \mathcal{L}^k(V) \rightarrow \mathcal{L}^k(V) \quad f \mapsto \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) f^\sigma$$

where $f^\sigma(\vec{v}_1, \dots, \vec{v}_k) := f(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(k)})$. It follows from Munkres Theorem 28.1 that we have $\hat{A}(f) \in \mathcal{A}^k(V)$ for all $f \in \mathcal{L}^k(V)$. Note that the definition for \hat{A} does not rely

on a basis of V . For ascending tuple I , we have $\hat{A}\Phi_I = \Psi_I$. Moreover, for $f \in \mathcal{A}^k(V)$, we have $f = \text{sgn}(\sigma)f^\sigma$, hence we can write the following:

$$\hat{A}(f) = \sum_{\sigma \in S_k} \text{sgn}(\sigma)\text{sgn}(\sigma)f = \sum_{\sigma \in S_k} f = k!f$$

To completes the proof for (6), here we will make use of the following proposition:

Proposition 27.1.1

Let $f \in \mathcal{A}^k(V)$ and $g \in \mathcal{A}^l(V)$, we have the following holds:

$$f \wedge g = \frac{1}{k!l!} \hat{A}(f \otimes g)$$

Proof. Consider the special case where $f = \Psi_I$ and $g = \Psi_J$ where I is an ascending k -tuple and J is an ascending l -tuple. It suffices to check that the following holds:

$$(\Psi_I \wedge \Psi_J)(\vec{a}_S) = \frac{1}{k!l!} \hat{A}(\Psi_I \otimes \Psi_J)(\vec{a}_S) \quad (*)$$

for S ascending $(k+l)$ -tuples, with $\vec{a}_S = (\vec{a}_{s_1}, \vec{a}_{s_2}, \dots, \vec{a}_{s_{k+l}})$. Here $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ is the basis of V associated to the basis $(\Psi_K \mid K \text{ is ascending})$ of $\mathcal{A}^{k+l}(V)$. Note here we have:

$$\text{LHS of equation } (*) = \begin{cases} 0 & I \cup J \neq S \\ \text{sgn}(I, J) & I \cup J = S \end{cases}$$

Here we write $S_\sigma = (S'_\sigma, S''_\sigma)$, where S'_σ denotes the first k entries of S_σ and S''_σ denotes the last l entries of S_σ , then we can write the following:

$$\text{RHS of equation } (*) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma)(\Psi_I \otimes \Psi_J)(\vec{a}_{S_\sigma}) \quad (**)$$

where we have, with some $\sigma' \in S_k$ and $\sigma'' \in S_l$:

$$\text{sgn}(\sigma)(\Psi_I \otimes \Psi_J)(\vec{a}_{S_\sigma}) = \begin{cases} 0 & I \neq S'_\sigma \text{ or } J \neq S''_\sigma \\ \text{sgn}(\sigma) \cdot \text{sgn}(\sigma') \cdot \text{sgn}(\sigma'') = \text{sgn}(I, J) & S_\sigma = (I_{\sigma'}, J_{\sigma''}) \end{cases}$$

Computing equation (**), the result follows. \square

From Proposition 27.1.1, we also get the following property of the wedge product:

Corollary 27.1.2

In the setting of Theorem 27.1, for $f \in \mathcal{A}^k(V)$ and $g \in \mathcal{A}^l(V)$, the value of $f \wedge g$ does not depend on the choice of basis of V .

With Proposition 27.1.1, we get the following result:

$$\begin{aligned} T^*(f \wedge g) &= \frac{1}{k!l!} T^* \left(\sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma)(f \otimes g)^\sigma \right) \\ &= \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma) T^*(f \otimes g)^\sigma \\ &= \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma) (T^*f \otimes T^*g)^\sigma \\ &= T^*f \wedge T^*g \end{aligned}$$

The results of Theorem 27.1 follows. \square

Example:

In the settings of Theorem 27.1 with $n = 3$, let $\alpha, \beta, \gamma, \tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}, A, B, C \in \mathbb{R}$, we have:

$$(\alpha\Psi_1 + \beta\Psi_2 + \gamma\Psi_3) \wedge (\tilde{\alpha}\Psi_1 + \tilde{\beta}\Psi_2 + \tilde{\gamma}\Psi_3) \wedge (\hat{\alpha}\Psi_1 + \hat{\beta}\Psi_2 + \hat{\gamma}\Psi_3) = \det \begin{pmatrix} \alpha & \beta & \gamma \\ \tilde{\alpha} & \tilde{\beta} & \tilde{\gamma} \\ \hat{\alpha} & \hat{\beta} & \hat{\gamma} \end{pmatrix} \Psi_{(1,2,3)}$$

The wedge product of n 1-tensor gives determinant of the coefficients.

$$(\alpha\Psi_1 + \beta\Psi_2 + \gamma\Psi_3) \wedge (A\Psi_{(2,3)} - B\Psi_{(1,3)} + C\Psi_{(1,2)}) = (\alpha A + \beta B + \gamma C)\Psi_{(1,2,3)}$$

The wedge product of an 1-tensor with an $(n-1)$ -tensor gives, sort of, dot product of the coefficients.

$$(\alpha\Psi_1 + \beta\Psi_2 + \gamma\Psi_3) \wedge (\tilde{\alpha}\Psi_1 + \tilde{\beta}\Psi_2 + \tilde{\gamma}\Psi_3) = (\alpha\tilde{\beta} - \beta\tilde{\alpha})\Psi_{(1,2)} + (\beta\tilde{\gamma} - \gamma\tilde{\beta})\Psi_{(2,3)} + (-\alpha\tilde{\gamma} + \gamma\tilde{\alpha})\Psi_{(3,1)}$$

wedge product of 1-tensor with 1-tensor gives cross product of the coefficients.

Differential Forms and Pullbacks

Let $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ be the standard basis for \mathbb{R}^n . By our discussion in previous sections, we have $\mathcal{L}^1(\mathbb{R}^n) = \mathcal{A}^1(\mathbb{R}^n) = (\mathbb{R}^n)_{row}$, and the basis for $\mathcal{L}^1(\mathbb{R}^n)$ is given by $\{\vec{e}_j^T \mid 1 \leq j \leq n\}$. Here we will, for most of the discussion, restrict our view to the standard basis of \mathbb{R}^n .

Definition 28.0.0.0.1

Let $[I]$ denote the set of ascending k -tuples of integers from $\{1, 2, \dots, n\}$. A k -form defined on an open subset U of \mathbb{R}^n is a continuous function defined by:

$$\omega : U \rightarrow \mathcal{A}^k(\mathbb{R}^n) \quad \vec{x} \mapsto \sum_{I \in [I]} b_I(\vec{x}) \Psi_I$$

where b_I are continuous functions from U to \mathbb{R} . The degree of a k -form is k , denoted as $\deg(\omega)$.

Let ω be defined as in Definition 28.0.0.0.1, with standard basis of \mathbb{R}^n , here we have:

$$\omega(\vec{x}) = \sum_{[I]} b_I(\vec{x}) \Psi_I = \sum_{[I]} b_I(\vec{x}) \Psi_{i_1} \wedge \Psi_{i_2} \wedge \dots \wedge \Psi_{i_k} = \sum_{[I]} b_I(\vec{x}) dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k} \quad (\text{W})$$

Definition 28.0.0.0.2

Let U be an open subset of \mathbb{R}^n , let k -form ω be defined by:

$$\omega : U \rightarrow \mathcal{A}^k(\mathbb{R}^n) \quad x \mapsto \sum_{I \in [I]} b_I(\vec{x}) \Psi_I$$

where b_I are continuous functions from U to \mathbb{R} .

ω is said to be of C^r type provided that each b_I is of C^r type.

Let U be an open subset of \mathbb{R}^n :

Remark: A scalar function defined on U can be viewed as a 0-form defined on U .

Note: The sum of two k -forms defined on U is a k -form defined on U

Note: A scalar function multiplied by a k -form defined on U is a k -form defined on U .

Note: A k -form defined on U wedge with a l -form defined on U is a $(k+l)$ -form defined on U ,

Definition 28.0.0.0.3

Let U be a subset of \mathbb{R}^n and let V be a subset of \mathbb{R}^l , let $\Phi : U \rightarrow V$ be a C^1 function, let ω be a k -form defined on V , then $\Phi^*\omega$ is a k -form defined on U given by:

$$\Phi^*\omega : U \rightarrow \mathcal{A}^k(U) \quad \vec{x} \mapsto (D\Phi(\vec{x}))^* \omega(\Phi(\vec{x}))$$

where we have:

$$(D\Phi(\vec{x}))^* \omega(\Phi(\vec{x})) : U^k \rightarrow \mathbb{R} \quad (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k) \mapsto \omega(\Phi(\vec{x}))(D\Phi(\vec{x})(\vec{u}_1), D\Phi(\vec{x})(\vec{u}_2), \dots, D\Phi(\vec{x})(\vec{u}_k))$$

Now we need an exterior derivative operator d that takes C^r k -forms on U to C^r $(k+1)$ -forms on U . Intuitively, we expect the operator d to satisfy the following conditions:

1. d is linear

2. $(d \circ d)$ should be a zero function.

This is because the boundary of the boundary of a manifold get us an empty set

3. d should satisfy some sort of product rule

These expectation suggests a definition that we can write:

Definition 28.0.0.0.4

Let ω be the k -form defined by equation (W) and of C^1 type, ω is defined on an open subset U of \mathbb{R}^n . $d(\omega)$ is defined to be the following:²

$$d(\omega) : U \rightarrow \mathcal{A}^{k+1}(\mathbb{R}^n) \quad \vec{x} \mapsto \sum_{[I]} db_I(\vec{x}) \wedge dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}$$

Example: Let $\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\beta : \mathbb{R}^2 \rightarrow \mathbb{R}$ be continuous functions.

Here $\omega = \alpha dx_1 + \beta dx_2$ is an 1-form, and we can write the following:

$$\begin{aligned} d(\alpha dx_1 + \beta dx_2) &= \left(\frac{\partial \alpha}{\partial x_1} dx_1 + \frac{\partial \alpha}{\partial x_2} dx_2 \right) \wedge dx_1 + \left(\frac{\partial \beta}{\partial x_1} dx_1 + \frac{\partial \beta}{\partial x_2} dx_2 \right) \wedge dx_2 \\ &= \left(\frac{\partial \beta}{\partial x_1} - \frac{\partial \alpha}{\partial x_2} \right) dx_1 \wedge dx_2 \end{aligned}$$

Example: Let U be an open subset of \mathbb{R}^n , and let $\omega : U \rightarrow \mathbb{R}_{row}^n$ $x \mapsto \sum_{j=1}^n b_j(\vec{x}) dx_j$ be a 1-form of C^1 type, then by definition we can write the following:

$$d \left(\sum_j b_j(\vec{x}) dx_j \right) = \sum_j \sum_k \frac{\partial b_j}{\partial x_k} dx_k \wedge dx_j = \sum_{j < k} \left(\frac{\partial b_j}{\partial x_j} - \frac{\partial b_j}{\partial x_k} \right) dx_j \wedge dx_k$$

Definition 28.0.0.0.5

A k -form ω is said to be closed provided that we have $d\omega = 0$.

Lemma 28.0.1

For k -forms ω_1 and ω_2 , we have $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$.

Proof. This lemma should follow from the definition. \square

Lemma 28.0.2

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^{\deg(\omega_1)} \omega_1 \wedge d\omega_2$$

Proof. First note that for 0-forms f and g , we have $d(fg) = f dg + df g$ for functions f and g .

Consider the key special case where $\omega_1 = \alpha dx_I$ and $\omega_2 = \beta dx_J$, where α and β are 0-forms:

$$\begin{aligned} d(\alpha dx_I \wedge \beta dx_J) &= d(\alpha \beta dx_I \wedge dx_J) \\ &= \beta d\alpha \wedge dx_I \wedge dx_J + \alpha d\beta \wedge dx_I \wedge dx_J \\ &= d(\alpha dx_I) \wedge (\beta dx_J) + (-1)^{\deg(\alpha dx_I)} (\alpha dx_I) \wedge d(\beta dx_J) \end{aligned}$$

\square

Lemma 28.0.3

For a k -form ω of C^2 type, we have $d(d\omega) = 0$.

² $d(\omega)$ will be denoted as $d\omega$ thereafter.

Proof. If the degree of ω is zero, $d\omega$ is exact, and exact 1-forms are closed, and hence the result follows. If the degree of ω is non-zero, we can write the following, with α_I being 0-forms:

$$d \left(d \left(\sum_{[I]} \alpha_I dx_I \right) \right) = d \left(\sum_{[I]} d\alpha \wedge dx_I \right) = \sum_{[I]} d(d\alpha) \wedge dx_I - \sum_{[I]} d\alpha_I \wedge d(dx_I)$$

Note here $d\alpha$ is an exact 1-form, and dx_j is exactly one form for all j , then we see that:

$$d \left(d \left(\sum_{[I]} \alpha_I dx_I \right) \right) = \sum_{[I]} d(d\alpha) \wedge dx_I - \sum_{[I]} d\alpha_I \wedge d(dx_I) = 0$$

The result follows. \square

Lemma 28.0.4

Let U be a subset of \mathbb{R}^l and let V be a subset of \mathbb{R}^n , let $\Phi : U \rightarrow V$ be a C^1 function, let ω be a k -form defined on V given by equation (W), we have $d(\Phi^*\omega) = \Phi^*d\omega$.

Proof. A detailed proof of this Lemma is given on Theorem 32.3 in Munkres. If the degree of ω is zero, then we can rewrite $d(\Phi^*\omega)$ through Chain Rule and the result will follow. If the degree of ω is non zero, then we can write the following, summing over $[I]$:

$$\begin{aligned} d(\Phi^*\omega) &= \sum d\Phi^*(b_I dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}) \\ &= \sum d(\Phi^*b_I \cdot \Phi^*(dx_{i_1}) \wedge \Phi^*(dx_{i_2}) \wedge \cdots \wedge \Phi^*(dx_{i_k})) \\ &= \sum d(\Phi^*b_I \cdot d(\Phi^*x_{i_1}) \wedge d(\Phi^*x_{i_2}) \wedge \cdots \wedge d(\Phi^*x_{i_k})) \\ &= \sum d(\Phi^*b_I) \wedge d(\Phi^*x_{i_1}) \wedge d(\Phi^*x_{i_2}) \wedge \cdots \wedge d(\Phi^*x_{i_k}) + 0 \\ &= \sum \Phi^*db_I \wedge \Phi^*dx_{i_1} \wedge \Phi^*dx_{i_2} \wedge \cdots \wedge \Phi^*dx_{i_k} \\ &= \Phi^*d\omega \end{aligned}$$

\square

Integration of Differential Forms

Definition 29.0.0.0.1

Let U be a subset of \mathbb{R}^k that is open in either \mathbb{R}^k or \mathbb{H}^k , and let ω be a k -form defined on an open subset U of \mathbb{R}^k given by the following, with f being a zero form defined on U :

$$\omega : U \rightarrow \mathcal{A}^k(\mathbb{R}^k) \quad \vec{x} \mapsto f(\vec{x}) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k \quad (\text{WF})$$

Then we define the following:

$$\int_U \omega := \int_U f$$

whenever $\int_U f$ exists.

Note: In the settings of Definition 29.0.0.0.1, $\int_U \omega$ guarantees to exist if the support of f is compact.

Let U, V be subsets of \mathbb{R}^k that are open in either \mathbb{R}^k or in \mathbb{H}^k , such that there is a function Φ which defines a diffeomorphism from U to V , with $\Phi_1, \Phi_2, \dots, \Phi_k$ being component functions of Φ , let ω be a k -form defined by equation (WF). Here we can write the following:

$$\begin{aligned} \int_U \Phi^* \omega &= \int_U \Phi^* f \cdot \Phi^*(dx_1) \wedge \Phi^*(dx_2) \wedge \cdots \wedge \Phi^*(dx_k) \\ &= \int_U (\Phi^* f) \cdot d\Phi_1 \wedge d\Phi_2 \wedge \cdots \wedge d\Phi_k \\ &= \int_U f \circ \Phi \cdot h(D\Phi) \cdot dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k \end{aligned}$$

Note here $h(D\Phi)$ is an alternating multilinear function of rows of $D\Phi$ with $h(I) = 1$ where I is the identity matrix. One can show that we have $h(D\Phi) = \det(D\Phi)$, details can be found on Theorem 32.2 on Munkres, and then we can write the following:

$$\begin{aligned} \int_U \Phi^* \omega &= \int_U f \circ \Phi \cdot h(D\Phi) \cdot dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k \\ &= \int_U f \circ \Phi \cdot \det(D\Phi) \cdot dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k \end{aligned}$$

If U is connected, or equivalently, if V is connected, by the Change of Variable Theorem, we have:

$$\begin{cases} \int_U \Phi^* \omega = \int_V f = \int_V \omega & \det(D\Phi(\vec{x})) > 0 \quad \forall \vec{x} \in U \\ \int_U \Phi^* \omega = - \int_V f = - \int_V \omega & \det(D\Phi(\vec{x})) < 0 \quad \forall \vec{x} \in U \end{cases}$$

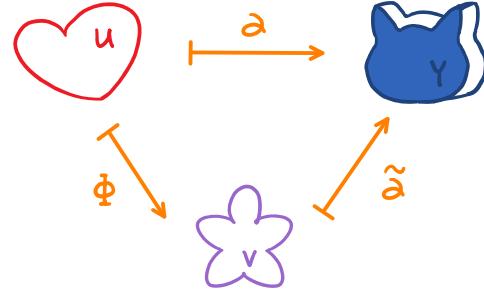
Note here $\det(D\Phi(\vec{x}))$ is nonzero for all $\vec{x} \in U$ because Φ is a diffeomorphism, and $\det(D\Phi(\vec{x}))$ does not change sign for all $\vec{x} \in U$ if U is connected.

Definition 29.0.0.0.2

Let U be an open subset of \mathbb{R}^k . Let Y be a parametrized k -manifold in \mathbb{R}^n parametrized by $\alpha : U \rightarrow Y$, let ω be a k -form defined on some open subset of \mathbb{R}^n containing Y , we define:

$$\int_{Y_\alpha} \omega := \int_U \alpha^* \omega$$

Consider a k -manifold $Y \in \mathbb{R}^n$ parametrized by $\alpha : U \rightarrow Y$, and can be reparametrized by $\tilde{\alpha} : V \rightarrow Y$, with U, V being open subsets of \mathbb{R}^k . Let Φ denote a diffeomorphism between U and V .

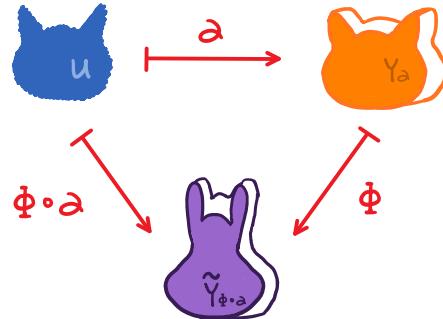


From argument in the discussion above, here we can write:

$$\int_{Y_\alpha} \omega = \int_V \tilde{\alpha}^* \omega = \pm \int_U \Phi^* \tilde{\alpha}^* \omega = \pm \int_U \alpha^* \omega = \pm \int_{Y_\alpha} \omega$$

The sign of $\pm \int_{Y_\alpha} \omega$ depends on the whether Φ is orientation preserving or orientation reversing, which we will discuss in the following section.

Now consider a k -manifold $Y \in \mathbb{R}^n$ parametrized by $\alpha : U \rightarrow Y$, with U being open subsets of \mathbb{R}^k , let Φ denote a diffeomorphism between Y and $\tilde{Y} \subseteq \mathbb{R}^n$, that is, \tilde{Y} can be parametrized by $\Phi \circ \alpha$.



Here we can write the following:

$$\int_{\tilde{Y}_{\Phi \circ \alpha}} = \int_U (\Phi \circ \alpha)^* \omega = \int_U \alpha^* \Phi^* \omega = \int_{Y_\alpha} \Phi^* \omega$$

Orientations of Manifolds

Here we see that the integral of a k -form over a parametrized k -manifold in \mathbb{R}^n , defined by Definition 29.0.0.0.2, is invariant under reparametrization up to a sign difference. That is, if α and β are two parametrizations for a k -manifold Y , with k -form defined on an open subset of \mathbb{R}^n containing Y , we have the following holds:

$$\int_{Y_\beta} \omega = \pm \int_{Y_\alpha} \omega$$

Hence it is natural to discuss the orientation of a manifold to resolve this issue. Intuitively, the orientation on a manifold M is the division of coordinate patches with connected domains into Group A and Group B, determined by the following conditions:

For coordinate patches α_1 and α_2 of M with $\alpha_2^{-1} \circ \alpha_1$ being well defined:

If we have $\det(D(\alpha_2^{-1} \circ \alpha_1)) > 0$, then α_1 and α_2 belong to the same group.

If we have $\det(D(\alpha_2^{-1} \circ \alpha_1)) < 0$, then α_1 and α_2 belong to opposite groups.

We call one of the two groups, say Group A, to be orientation preserving group, and the other to be orientation reversing group. Any coordinate patch in the orientation preserving group is said to be orientation preserving, and any coordinate path in the orientation reversing group is said to be orientation reversing.

Exercise: Given an oriented k -manifold M in \mathbb{R}^n , there exists a k -form ω defined on an open subset of \mathbb{R}^n containing M such that, for each point $\vec{p} \in M$ along with a coordinate patch $\alpha_{\vec{p}}$ for $\vec{p} \in M$, $(\alpha_{\vec{p}})^* \omega(\alpha_{\vec{p}}^{-1}(\vec{p}))$ is positive, or negative, multiple of $dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k$, for orientation preserving $\alpha_{\vec{p}}$, or orientation reversing $\alpha_{\vec{p}}$, respectively.

Definition 30.0.0.0.1

Let V be an n -dimensional vector space, let $A = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ be an n -tuple of independent vectors in V . A is called an n -frame in V . If $V = \mathbb{R}^n$, A is said to be right-handed provided that:

$$\det([\vec{a}_1 \ \vec{a}_2 \ \cdots \ \vec{a}_n]) > 0$$

and A is said to be left-handed provided that we have:

$$\det([\vec{a}_1 \ \vec{a}_2 \ \cdots \ \vec{a}_n]) < 0$$

Definition 30.0.0.0.2

The collection of all right-handed frames in \mathbb{R}^n is called an orientation of \mathbb{R}^n , and so is the collection of left-handed frames.

Definition 30.0.0.0.3

To define the orientation of an n -dimensional vector space V , one can choose a linear isomorphism $T : \mathbb{R}^n \rightarrow V$ and define one orientation of V to consist of all frames of the form:

$$(T(\vec{a}_1), T(\vec{a}_2), \dots, T(\vec{a}_n))$$

for which $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ is a right-handed frame in \mathbb{R}^n , and the other orientation of V to consist of all such frames for which $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ is left-handed.

Remark: For an n -dimensional vector space V , from Definition 30.0.0.0.3, we see that V has two orientations, each is called the reverse, or the opposite, of the other.

Note: It is easy to see that, in the settings of Definition 30.0.0.0.3, the orientation of V is independent of the choice of T . Note that in an arbitrary n -dimensional vector space, there is no well-defined notion of "right-handed," although there is a well-defined notion of orientation.

Theorem 30.1 (Theorem 20.3 on Munkres)

Let C be an invertible $n \times n$ matrix, let $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ $\vec{x} \mapsto C \cdot \vec{x}$ be a linear transformation, and let $A = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ be a frame in \mathbb{R}^n . If $\det(C) > 0$, then the frames A and $H = (h(\vec{a}_1), h(\vec{a}_2), \dots, h(\vec{a}_n))$ belong to the same orientation of \mathbb{R}^n . If $\det(C) < 0$, then the frames A and H belong to opposite orientations of \mathbb{R}^n .

Definition 30.1.0.0.1

Let C be an invertible $n \times n$ matrix, let $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ $\vec{x} \mapsto C \cdot \vec{x}$ be a linear transformation. h is said to be orientation preserving provided that $\det(C) > 0$. h is said to be orientation reversing provided that $\det(C) < 0$.

More generally, we can write the following:

Definition 30.1.0.0.2

Let A, B be open subsets of \mathbb{R}^k , let $g : A \rightarrow B$ be a diffeomorphism, g is said to be orientation preserving provided that $\det(Dg(\vec{x})) > 0$ for all $\vec{x} \in A$, and g is said to be orientation-reversing provided that $\det(Dg(\vec{x})) < 0$ for all $\vec{x} \in A$.

In order for Definition 30.1.0.0.2 to make sense, we here formulate a definition for the tangent spaces, following the notation from Munkres:

Definition 30.1.0.0.3

Let $\vec{x} \in \mathbb{R}^n$. For $\vec{v} \in \mathbb{R}^n$, the pair $(\vec{x}; \vec{v})$ is called a tangent vector to \mathbb{R}^n at \vec{x} .

Definition 30.1.0.0.4

Let $\vec{x} \in \mathbb{R}^n$. For $\vec{v}, \vec{w} \in \mathbb{R}^n$, and $c \in \mathbb{R}$, we define:

$$(\vec{x}; \vec{v}) + (\vec{x}; \vec{w}) = (\vec{x}; \vec{v} + \vec{w}) \quad c(\vec{x}; \vec{v}) = (\vec{x}; c\vec{v})$$

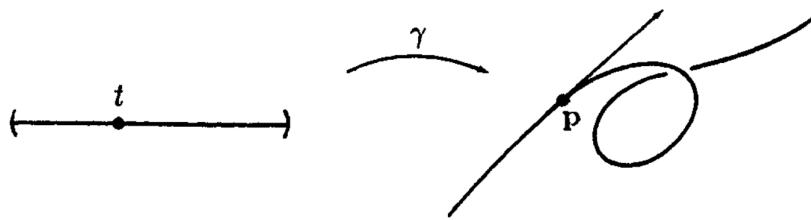
The set of all tangent vectors to \mathbb{R}^n at \vec{x} forms a vector space, called the tangent space to \mathbb{R}^n at \vec{x} , denoted as $\mathcal{T}_{\vec{x}}(\mathbb{R}^n)$

Note: Although both \vec{x} and \vec{v} given in Definition 30.1.0.0.4 are element of \mathbb{R}^n , they play different roles in such definition. One can think of \vec{x} as a point in \mathbb{R}^n and picture it as a dot, think of \vec{v} as an element in \mathbb{R}^n and picture it as an arrow. We picture $(\vec{x}; \vec{v})$ as an arrow with its initial point at \vec{x} . The set $\mathcal{T}_{\vec{x}}(\mathbb{R}^n)$ is pictured as the set of all arrows with their initial points at \vec{x} , which equivalently, can be written as $\{\vec{x}\} \times \mathbb{R}^n$.

Note: We do not attempt to form the sum $(\vec{x}; \vec{v}) + (\vec{y}; \vec{w})$ for $\vec{x}, \vec{y}, \vec{v}, \vec{w} \in \mathbb{R}^n$ if $\vec{x} \neq \vec{y}$.

Definition 30.1.0.0.5

Let (a, b) be an open interval in \mathbb{R} , let $\gamma : (a, b) \rightarrow \mathbb{R}^n$ be a function of C^r type. The velocity vector of γ at $t \in (a, b)$ is defined to be $(\gamma(t); D\gamma(t))$.



(Figure 29.1 from Munkres Section 29)

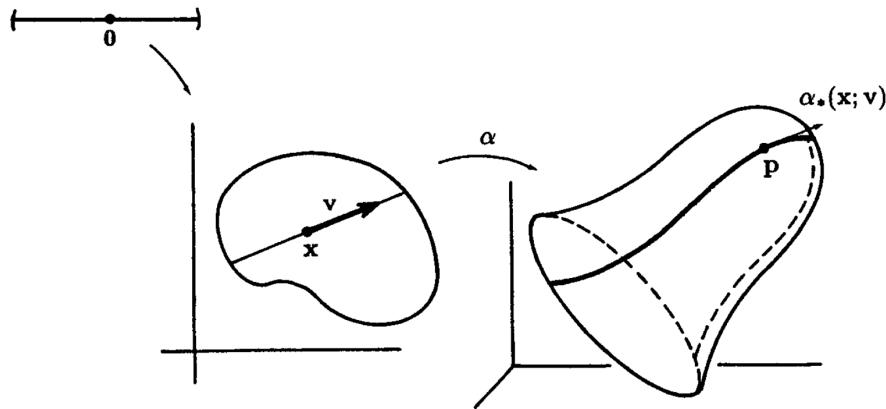
Note: In the settings of Definition 30.1.0.0.5, $(\gamma(t); D\gamma(t))$ is viewed as an arrow in \mathbb{R}^n with its initial point at $\vec{p} = \gamma(t)$, pointing towards the direction of motion.

Definition 30.1.0.0.6

Let A be a subset of \mathbb{R}^k , open in \mathbb{R}^k , or in \mathbb{H}^k , let $\alpha : A \rightarrow \mathbb{R}^n$ be a function of C^r type, let $\vec{x} \in A$, and let $\vec{p} = \alpha(\vec{x})$. We define a linear transformation:

$$\alpha_* : \mathcal{T}_{\vec{x}}(\mathbb{R}^k) \rightarrow \mathcal{T}_{\vec{p}}(\mathbb{R}^n) \quad (\vec{x}, \vec{v}) \mapsto (\vec{p}; D\alpha(\vec{x}) \cdot \vec{v})$$

α_* is said to be the transformation induced by the differentiable map α .



(Figure 29.2 from Munkres Section 29)

Note: Given $(\vec{x}; \vec{v})$ in the definition 30.1.0.0.6, the chain rule implies that the vector $\alpha_*(\vec{x}; \vec{v})$ is the velocity vector of the curve $\gamma(t) = \alpha(\vec{x} + t\vec{v})$, corresponding to the parameter value $t = 0$.

Note: From Lemma 29.1 on Munkres, Let A, B be an subsets of \mathbb{R}^k , open in either \mathbb{R}^k or in \mathbb{H}^k , let $\alpha : A \rightarrow \mathbb{R}^m$ be a function of C^r type, with B containing $\alpha(A)$, let $\beta : B \rightarrow \mathbb{R}^n$ be a function of C^r type. Then we have $(\beta \circ \alpha)_* = \beta_* \circ \alpha_*$.

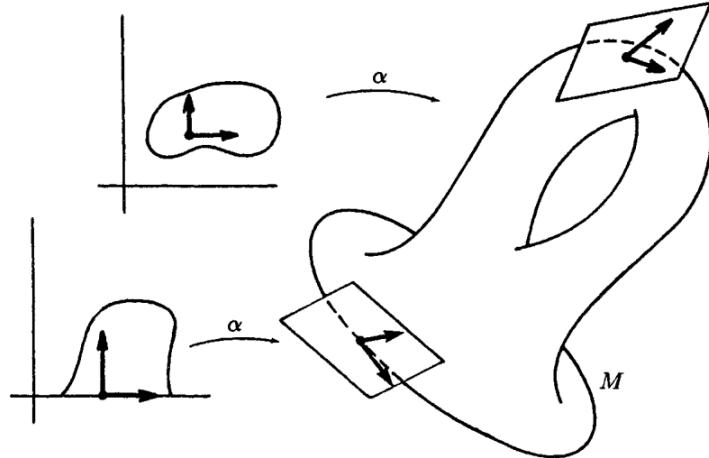
Definition 30.1.0.0.7

Let A be an open subset of \mathbb{R}^n , a tangent vector field in A is a continuous function $F : A \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ such that $F(\vec{x}) \in \mathcal{T}_{\vec{x}}(\mathbb{R}^n)$ for each $\vec{x} \in A$.

Note: In the settings of Definition 30.1.0.0.7, the function F has the form $F(\vec{x}) = (\vec{x}; f(\vec{x}))$ where $f : A \rightarrow \mathbb{R}^n$. If F is of C^r type, we say that it is a tangent vector field of C^r type.

Definition 30.1.0.0.8

Let M be a k -manifold of class C^r in \mathbb{R}^n , let $\vec{p} \in M$ with a coordinate patch $\alpha : U \rightarrow V$ for $\vec{p} \in M$, where U is a subset of \mathbb{R}^k , open in either \mathbb{R}^k or \mathbb{H}^k , let $\vec{x} \in U$ such that $\alpha(\vec{x}) = \vec{p}$. The set of all vectors of the form $\alpha_*(\vec{x}; \vec{v})$, where $\vec{v} \in \mathbb{R}^k$, is called the tangent space to M at \vec{p} , denoted as $\mathcal{T}_{\vec{p}}(M)$. That is, we have $\mathcal{T}_{\vec{p}}(M) = \alpha_*(\mathcal{T}_{\vec{x}}(\mathbb{R}^k))$.



(Figure 29.3 from Munkres Section 29)

From discussion in section 16 in this text, it is easy to show that $\mathcal{T}_{\vec{p}}(M)$ given in Definition 30.1.0.0.8 is a linear subspace of $\mathcal{T}_{\vec{p}}(\mathbb{R}^n)$ that is well-defined, independent of the choice of α . Since \mathbb{R}^k is spanned by the standard basis $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$, the space $\mathcal{T}_{\vec{p}}(M)$ is spanned by the vectors in the family given by:

$$T = \left((\vec{p}; D\alpha(\vec{x}) \cdot \vec{e}_j) = \left(\vec{p}; \frac{\partial \alpha}{\partial x_j} \right) \middle| j \in \mathbb{N}, 1 \leq j \leq k \right)$$

Since $D\alpha$ has rank k , vectors in T are linearly independent, hence they form a basis for $\mathcal{T}_{\vec{p}}(M)$.

Now for the function g defined in Definition 30.1.0.0.3, there is an associated linear transformation of tangent space:

$$g_* : \mathcal{T}_{\vec{x}}(\mathbb{R}^k) \rightarrow \mathcal{T}_{g(\vec{x})}(\mathbb{R}^k) \quad (\vec{x}; \vec{v}) \mapsto (g(\vec{x}); Dg(\vec{x}) \cdot \vec{v})$$

Then g is orientation preserving provided that for each \vec{x} in the domain of g , the linear transformation of \mathbb{R}^k , whose matrix is Dg , is orientation-preserving in the sense of Definition 30.1.0.0.1.

Definition 30.1.0.0.9

Let M be a k -manifold in \mathbb{R}^n . Given coordinate path $\alpha_i : U_i \rightarrow V_i$ on M for $i = 0, 1$, we say α_1, α_0 overlap if $V_0 \cap V_1 \neq \emptyset$. We say α_1, α_0 overlap positively provided that the transition function $\alpha_1^{-1} \circ \alpha_0$ is orientation preserving.

Definition 30.1.0.0.10

Let M be a k -manifold in \mathbb{R}^n . M is said to be orientable provided that M can be covered by a collection of coordinate patches such that each pair of coordinate patches overlap positively, if they overlap at all. M is said to be non-orientable if it cannot be covered by such collection of coordinate patches.

Definition 30.1.0.0.11

Let M be an orientable k -manifold in \mathbb{R}^n . Given a collection of coordinate patches covering M that overlap positively, adjoin to this collection all other coordinate patches on M that overlap these patches positively, denote such collection as O . O is called an orientation on M .

Note: In the settings of Definition 30.1.0.0.11, it is easy to see that coordinate patches in O overlap one another positively.

Definition 30.1.0.0.12

A manifold M together with an orientation of M is called an oriented manifold. Given an oriented manifold M along with its orientation O , a coordinate patch α on M is said to be orientation preserving provided that α overlaps any one of the coordinate patches in O positively. Otherwise α is said to be orientation reversing.

Note: If a k -manifold M in \mathbb{R}^n is orientable, let N denote the number of connected components of M , then M has 2^N orientations.

Orientable Manifolds

First we note that, by the definition of orientable manifold, the discussion makes no sense for a 0-manifold, which is just a discrete collection of points. A details about orientation of 0-manifold is discussed in Munkres.

In certain dimensions, the notion of orientation of a k -manifold has a geometric interpretation that is easily described. This situation occurs when $k = 1$, $k = n - 1$, or $k = n$.

In the case $k = 1$, we can picture an orientation in terms of a tangent vector field, as we proposed in the following definition:

Definition 31.0.0.0.1

Let M be an oriented 1-manifold in \mathbb{R}^n . We define a corresponding unit tangent vector field \vec{T} on M as follows: Given $\vec{p} \in M$, choose a coordinate patch $\alpha_{\vec{p}} : U \rightarrow V$ on M about \vec{p} belonging to the given orientation of M , and define the following:

$$\vec{T} : M \rightarrow \mathbb{R}^n \times \mathbb{R}^n \quad \vec{p} \mapsto \left(\vec{p}; \frac{D\alpha_{\vec{p}}(t_0)}{\|D\alpha_{\vec{p}}(t_0)\|} \right)$$

where t_0 is the parameter value such that $\alpha_{\vec{p}}(t_0) = \vec{p}$. Here \vec{T} is called the unit tangent field corresponding to the orientation of M .

Note that, in the settings of Definition 31.0.0.0.1, $(\vec{p}; D\alpha(t_0))$ is the velocity vector of the curve α corresponding to the parameter value $t = t_0$. Then we know that $\vec{T}(\vec{p})$ is such velocity vector divided by its length.

In the following we will show that \vec{T} given in the settings of Definition 31.0.0.0.1 is well defined.

Proof. Let β be another coordinate patch on M about \vec{p} belonging to the orientation of M , and we denote $\vec{p} = \beta(t_1)$, let $g = \beta^{-1} \circ \alpha$. Then g is a diffeomorphism between an open neighborhood of t_0 and an open neighborhood of t_1 . Here we can write:

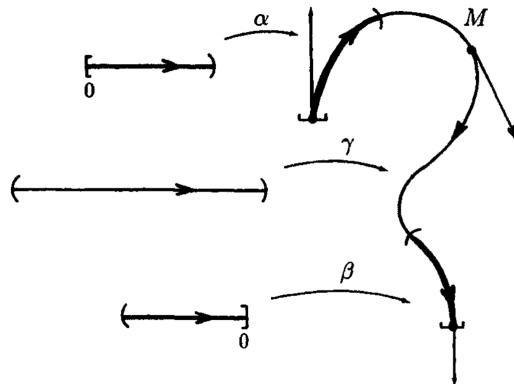
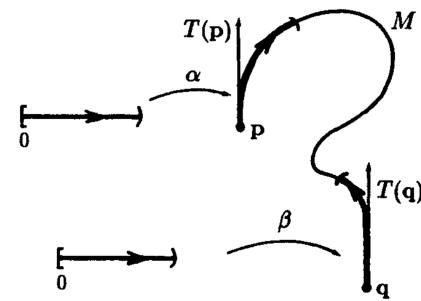
$$D\alpha(t_0) = D(\beta \circ g)(t_0) = D\beta(t_1) \cdot Dg(t_0)$$

Note that $Dg(t_0)$ is a 1×1 matrix, and since g is orientation preserving, we must have $Dg(t_0) > 0$, so we conclude that:

$$\frac{D\alpha(t_0)}{\|D\alpha(t_0)\|} = \frac{D\beta(t_1)}{\|D\beta(t_1)\|}$$

The result follows. □

Example: Note that a difficulty arises if a 1-manifold M has nonempty manifold boundary, in particular, if ∂M contains two distinct points \vec{p} and \vec{q} . The coordinate patches for \vec{p} and \vec{q} , defined on a subsets of \mathbb{H}^1 open in \mathbb{H}^1 , will induce tangent vectors pointing into M from the two points.



(Figure 34.2 from Munkres Section 34)

(Figure 34.3 from Munkres

Section 34)

In the 1-manifold indicated in the picture, there is no way to define a unit tangent field on M that points into M at both \vec{p} and \vec{q} . This problem disappears if we allow our coordinate patches whose domains are open in \mathbb{R}^1 , or \mathbb{H}^1 , or in the left half-line $\mathbb{L}^1 = \{x \mid x \leq 0\}$. With this extra degree of freedom, it is easy to cover the 1-manifold with two boundary points that overlap positively, and such that the unit tangent field of such manifold is well defined.

Note: In the case of 1-manifold M , we allow the domains of the coordinate patches on M to be open sets in \mathbb{R}^1 , \mathbb{H}^1 , or in \mathbb{L}^1 , such that every 1-manifold is orientable.

Now we consider the case where M is an $(n - 1)$ -manifold in \mathbb{R}^n . In this case, we can picture an orientation of M in terms of a unit normal vector field to M , as proposed in the following:

Definition 31.0.0.0.2

Let M be an oriented $(n - 1)$ -manifold in \mathbb{R}^n , let $\vec{p} \in M$, let $\alpha : U \rightarrow V$ be a coordinate patch on M about \vec{p} belonging to the given orientation of M , denote $\alpha(\vec{x}) = \vec{p}$, then the columns $D_i\alpha(\vec{x})$ of the matrix $D\alpha(\vec{x})$ gives basis for $\mathcal{T}_{\vec{p}}(M)$, the tangent space to M at \vec{p} . That is, the family of vectors defined by the following defines a basis of $\mathcal{T}_{\vec{p}}(M)$:

$$((\vec{p}; D_1\alpha(\vec{x})), (\vec{p}; D_2\alpha(\vec{x})), \dots, (\vec{p}; D_{n-1}\alpha(\vec{x})))$$

Let $(\vec{p}; \vec{n}(\vec{p}))$ be a unit vector in the n -dimensional vector space $\mathcal{T}_{\vec{p}}(\mathbb{R}^n)$ that is orthogonal to the $(n - 1)$ -dimensional linear subspace $\mathcal{T}_{\vec{p}}(M)$. Here $\vec{n}(\vec{p})$ is uniquely determined up to a sign, we specify the sign of $\vec{n}(\vec{p})$ by requiring the frame defined by the following to be right-handed:

$$(\vec{n}(\vec{p}), D_1\alpha(\vec{x}), D_2\alpha(\vec{x}), \dots, D_{n-1}\alpha(\vec{x}))$$

That is, the matrix $[\vec{n}(\vec{p}) \ D\alpha(\vec{x})]$ has positive determinant. The vector field defined by the following is called the unit normal field to M corresponding to the orientation of M :

$$\vec{N} : M \mapsto \mathbb{R}^n \times \mathbb{R}^n \quad \vec{p} \mapsto (\vec{p}; \vec{n}(\vec{p}))$$

Example: The 2-manifold Möbius Band in \mathbb{R}^3 is not orientable, because it has no continuous unit normal vector field.

Example: The 2-manifold Klein Bottle in \mathbb{R}^4 is also non-orientable.

Now we consider the case where M is an n -manifold in \mathbb{R}^n . In this case, M is always orientable, and it also has a natural orientation, as proposed in the following definition:

Definition 31.0.0.0.3

Let M be an n -manifold in \mathbb{R}^n . The natural orientation of M consists of all coordinate patches α on M for which $\det(D\alpha(\vec{x})) > 0$ for all \vec{x} in the definition of domain of α .

Proposition 31.0.1

All n -manifolds in \mathbb{R}^n are orientable.

Proof. Let $\vec{p} \in M$, let $\alpha : U \rightarrow V$ be a coordinate patch on M about \vec{p} . Here we note that U is either open in \mathbb{R}^n or open in \mathbb{H}^n , by shrinking U if necessary, we can assume that U is either an open ball, or the intersection of an open ball with \mathbb{H}^n . In either case, we know that U is connected, so $\det(D\alpha(\vec{x}))$ is either all positive, or all negative, for all $\vec{x} \in U$. If the former, then α is our desired coordinate patch about \vec{p} . If the latter, then $\alpha \circ r$ is our desired coordinate patch about \vec{p} , where r is defined by $r : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad (x_1, x_2, \dots, x_n) \mapsto (-x_1, x_2, \dots, x_n)$. Since \vec{p} is arbitrary, this completes the proof. \square

Definition 31.0.1.0.1

Let $r : \mathbb{R}^k \rightarrow \mathbb{R}^k \quad (x_1, x_2, \dots, x_n) \mapsto (-x_1, x_2, \dots, x_n)$ be the reflection map, let M be an oriented k -manifold in \mathbb{R}^n . For each coordinate patch on M , $\alpha_i : U_i \rightarrow V_i$, that belongs to the orientation of M , let β_i be the coordinate patch on M defined by:

$$\beta_i : r(U_i) \rightarrow V_i \quad \vec{x} \mapsto (\alpha_i \circ r)(\vec{x})$$

The set of all such β_i is called the reverse, or opposite, orientation to the orientation of M specified by the coordinate patches α_i

Note: In the settings of the Definition 31.0.1.0.1, β_i does not overlap α_i positively, so β_i does not belong to the orientation of M specified by the set of coordinate patches that contains α_i . By the definition, β_i overlap each other positively.

Note: The reflection map r defined in Definition 31.0.1.0.1 is its own inverse. r carries \mathbb{H}^k to \mathbb{H}^k if $k > 1$, and it carries \mathbb{H}^1 to \mathbb{L}^1 .

Note: Now it follows that every orientable k -manifold M has at least two orientations, a given one and its opposite. If M is connected, it has only two orientations, and otherwise it has more than two.

Note: If M is an oriented 1-manifold with corresponding tangent field \vec{T} , then reversing the orientation of M induces the corresponding tangent field $-\vec{T}$.

Note: If M is an $n - 1$ manifold in \mathbb{R}^n with corresponding normal field \vec{N} , reversing the orientation of M induces the corresponding normal field $-\vec{N}$.

Theorem 31.1 (Theorem 34.1 on Munkres)

Let $k > 1$, let M be an orientable k -manifold with non-empty boundary. ∂M is orientable.

Proof. Let $\vec{p} \in \partial M$, and let $\alpha : U \rightarrow V$ be a coordinate patch on M about \vec{p} belonging to a given orientation of M . Here we define:

$$b : \mathbb{R}^{k-1} \rightarrow \mathbb{R}^k \quad (x_1, x_2, \dots, x_{k-1}) \mapsto (x_1, x_2, \dots, x_{k-1}, 0)$$

The function $\alpha_0 := \alpha \circ b$ defines a coordinate patch on ∂M about \vec{p} . Here α_0 is said to be obtained by restricting α . Theorem 34.1 on Munkres shows in details that the set of such coordinate patches α_0 on ∂M , obtained by restricting coordinate patches α on M in the given orientation of M , constitutes an orientation of ∂M . \square

Theorem 31.1 shows that, given an orientation of a k -manifold M that has nonempty boundary, one can obtain an orientation of ∂M by simply taking restrictions of coordinate patches that belong to the orientation of M . However, this orientation of ∂M is not always the one we prefer. Here we propose the following definition:

Definition 31.1.0.0.1

Let M be an orientable k -manifold with nonempty manifold boundary ∂M . Given an orientation O of M , the corresponding induced orientation of ∂M is defined as follows:

1. If k is even, the corresponding induced orientation of ∂M is the orientation obtained by restricting coordinate patches belonging to O .
2. If k is odd, the corresponding induced orientation of ∂M is the opposite of the orientation of ∂M obtained by restricting coordinate patches belonging to O .

Example:

Let M be a naturally oriented 3-manifold in \mathbb{R}^3 with nonempty boundary, and let

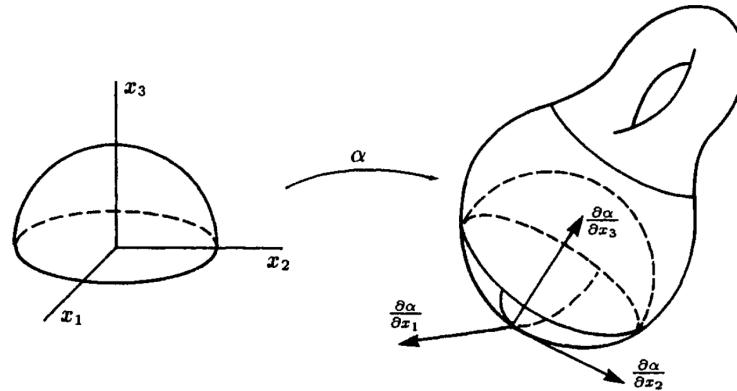
$\vec{p} \in \partial M$ with coordinate patch $\alpha : U \rightarrow V$ on M about \vec{p} belonging to the orientation of M . Since M is naturally oriented, then we can write, with $\alpha(\vec{q}) = \vec{p}$:

$$\det(D\alpha(\vec{q})) = \det([D_1\alpha(\vec{q}) \ D_2\alpha(\vec{q}) \ D_3\alpha(\vec{q})]) = \det([D_3\alpha(\vec{q}) \ D_1\alpha(\vec{q}) \ D_2\alpha(\vec{q})]) > 0$$

so we know that the frame $(D_3\alpha(\vec{q}), D_1\alpha(\vec{q}), D_2\alpha(\vec{q}))$ is right-handed. Restricting α in the orientation of M , we get an orientation of ∂M , but the induced orientation of ∂M induced by the orientation of M is, by definition, taken to be to opposite of the orientation of ∂M obtained by restricting α . Thus the normal field at point \vec{p} , denoted as $\vec{N}(\vec{p}) = (\vec{p}; \vec{n}(\vec{p}))$, corresponding to the induced orientation of ∂M satisfies the following condition:

The frame $(-\vec{n}(\vec{p}), D_1\alpha(\vec{q}), D_2\alpha(\vec{q}))$ is right-handed

Here we notice that $D_3\alpha(\vec{q})$, by definition, must point into M , and hence $-\vec{n}(\vec{p})$ points into M . Then we know that $\vec{n}(\vec{p})$, which is the unit normal vector at point \vec{p} specified by the induced orientation of ∂M by the orientation of M , points outward from M . One might denote $D_1\alpha(\vec{q}), D_2\alpha(\vec{q}), D_3\alpha(\vec{q})$ as $\frac{\partial \alpha}{\partial x_1}, \frac{\partial \alpha}{\partial x_2}, \frac{\partial \alpha}{\partial x_3}$ respectively.



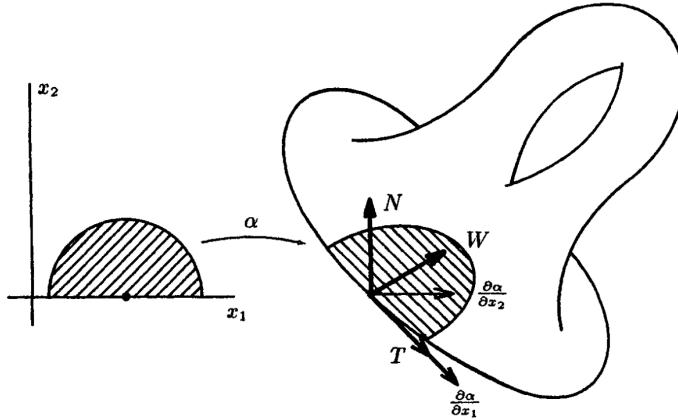
(Figure 34.9 from Munkres Section 34)

Example:

Now let M be an oriented 2-manifold with nonempty boundary in \mathbb{R}^3 , and give ∂M the induced orientation. Let \vec{N} be the unit normal field to M corresponding to the orientation of M , and let \vec{T} be the unit tangent field to ∂M corresponding to the induced orientation of ∂M . Note that given a coordinate patch on M about a point $\vec{p} \in \partial M$, with $\alpha(\vec{q}) = \vec{p}$, we note that $D_1\alpha(\vec{q})$ represents the velocity vector of the parametrized curve $\alpha \circ b$ where b is the function defined in Theorem 31.1, hence by definition it points in the same direction as the unit tangent vector \vec{T} . On the other hand, $D_2\alpha(\vec{q})$ points in the direction into M from ∂M by definition. Let $\vec{n}(\vec{p})$ be the unit normal vector at point $\vec{p} \in \partial M$, one can compute that:

$$\det([\vec{n}(\vec{p}) \ D_1\alpha(\vec{q}) \ D_2\alpha(\vec{q})]) > 0$$

Remark: Now let M be an oriented 2-manifold with nonempty boundary in \mathbb{R}^3 , and give ∂M the induced orientation. Let \vec{N} be the unit normal field to M corresponding to the orientation of M , and let \vec{T} be the unit tangent field to ∂M corresponding to the induced orientation of ∂M . Informally speaking, the relation between \vec{N} and \vec{T} is such that if one walks around ∂M in the direction specified by \vec{T} , with his head pointing in the direction specified by \vec{N} , then the manifold M is on his left, in the direction given by a vector \vec{W} orthogonal to both \vec{T} and \vec{N} .



(Figure 34.10 from Munkres Section 34)

Now let X be an oriented $(n-1)$ -manifold in \mathbb{R}^n , and suppose further that X is the boundary of a n -manifold M . Let α be a coordinate patch on X belonging to the given orientation of X , the tangent space of X at \vec{p} , denoted as $\mathcal{T}_{\vec{p}}(X)$ is given by the column space of $D\alpha(\vec{q})$, here $\dim(\mathcal{T}_{\vec{p}}(X)) = n-1$, and $\dim((\mathcal{T}_{\vec{p}}(X))^{\perp}) = 1$. Notice that, here we can define:

$$\mathcal{N}_{\vec{p}}(X) := (\mathcal{T}_{\vec{p}}(X))^{\perp}$$

Here we can pick $\vec{N}(\vec{p}) \in \mathcal{N}_{\vec{p}}(X)$ such that $\|\vec{N}(\vec{p})\| = 1$ and the following holds:

$$\det([\vec{N}(\vec{p}) \ D\alpha(\vec{q})]) > 0$$

Define $\vec{N} : X \rightarrow \mathbb{R}^n$ such that $\vec{N}(\vec{p}) \perp \vec{t}$ for all $\vec{t} \in \mathcal{T}_{\vec{p}}(X)$, and such that $\|\vec{N}(\vec{p})\| = 1$. Here we note that, such definition of the unit normal field \vec{N} on X corresponding to the

given orientation of X is consistent with the definition of normal field given in Definition 31.0.0.2.

Exercise: Let X be an orientable $(n-1)$ -manifold in \mathbb{R}^n , the definition of the unit normal field \vec{N} on X does not depend on the choice of coordinate patch up to the orientation of X which the coordinate patch belongs to.

Exercise: (Munkres Corollary 38.4)

If X is an $(n-1)$ -manifold in \mathbb{R}^n of class C^r , then $\vec{N} \in C^{r-1}(X, \mathbb{R}^n)$.

Exercise: For special case where n -manifold $M = \mathbb{H}^n$ in \mathbb{R}^n , we have $X = \partial M = \mathbb{R}^{n-1} \times \{0\}$, and we have the unit normal field on X has the property that $\vec{N}(\vec{p}) = -\vec{e}_n$, where \vec{e}_n is the last vector in the standard basis of \mathbb{R}^n .

Exercise: Let M be an oriented n -manifold in \mathbb{R}^n with nonempty boundary $X = \partial M$. Given the induced orientation of ∂M , the unit normal field on X at each point, $\vec{N}(\vec{p})$, points outwards from X , that is, we have $\vec{p} + t\vec{N}(\vec{p}) \notin M$ for $0 < t < \epsilon$ with some $\epsilon > 0$.

Conversely, given a unit normal vector field \vec{N} for an orientable $(n-1)$ -manifold X in \mathbb{R}^n , one can use the sign of following the determinant:

$$\det \begin{pmatrix} \vec{N}(\vec{p}) & D\alpha(\vec{q}) \end{pmatrix}$$

as a criteria to test the orientation of a coordinate patch α on X , or define the orientation of X .

Given an orientable $(n-1)$ -manifold X and a vector field $v : X \rightarrow \mathbb{R}^n$ such that $v(\vec{x}) \notin \mathcal{T}_{\vec{p}}(X)$, we obtain an orientation on X by declaring that each coordinate patch $\alpha : U \rightarrow V$ on X is orientation preserving provided that the following holds with $\vec{q} = \alpha^{-1}(\vec{p})$:

$$\det ([v(\vec{p}) \quad D\alpha(\vec{q})]) > 0$$

The orientation of a 1-manifold M is equivalent to the choice of continuous forward pointing motion \vec{T} along the tangent space of M , here \vec{T} is the unit tangent field as we defined above. For 2-manifold M in \mathbb{R}^2 , let $X = \partial M$, we get the following for $\vec{p} \in X$:

$$\vec{N}(\vec{p}) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \vec{T}(\vec{p})$$

In this sense of orientable manifolds, the orientation of 0-manifold can be defined with the direction of traveling towards the positive and away from the negative.

Definition 31.1.0.0.2

Let M be an oriented k -manifold in \mathbb{R}^n , let $\alpha : U \rightarrow V$ be a coordinate patch on M belonging to the given orientation, with $\alpha(\vec{q}) = \vec{p} \in M$, let ω be a k -form defined on an open subset of \mathbb{R}^n containing M . We can write $\alpha^*\omega = f(\vec{x}) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k$ for some 0-form f defined on the definition of domain of ω . ω is said to be positive for M at \vec{p} provided that $f(\vec{a}) > 0$, ω is said to be negative for M at \vec{p} provided that $f(\vec{a}) < 0$, and ω is said to be integral for M at \vec{p} provided that $f(\vec{a}) = 0$.

Exercise: In the settings of Definition 31.1.0.0.2, ω is integral at \vec{p} if and only if we have:

$$\omega(\vec{p})(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) = 0 \quad \forall \vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathcal{T}_{\vec{p}}(M)$$

Exercise: Let α and β be 0-form defined on \mathbb{R}^2 .

$\alpha dx + \beta dy$ is integral for the 1-manifold $\mathbb{R} \times \{0\}$ in \mathbb{R}^2 at $\vec{0}$ if and only if we have $\alpha(0, 0) = 0$.

Definition 31.1.0.0.3

M is integral manifold for ω provided that ω is integral for M at \vec{p} for all $\vec{p} \in M$.

Exercise: If M is an oriented k -manifold, then there exists an k -form ω being positive for M at all $\vec{p} \in M$. Conversely, given an k -form ω that is nowhere integral on M , there is a unique orientation of M such that ω is positive for M at all $\vec{p} \in M$.

Theorem 31.2 (Theorem 36.2 on Munkres)

Let M be a compact oriented k -manifold in \mathbb{R}^n , let ω be a k -form defined in a open subset of \mathbb{R}^n containing M , and let λ be the scalar function on M defined by the following:

$$\lambda : M \rightarrow \mathbb{R} \quad \vec{p} \mapsto \omega(\vec{p})((\vec{p}; \vec{a}_1), (\vec{p}; \vec{a}_2), \dots, (\vec{p}; \vec{a}_k))$$

where, for $\vec{p} \in M$, the family $((\vec{p}; \vec{a}_1), (\vec{p}; \vec{a}_2), \dots, (\vec{p}; \vec{a}_k))$ forms an orthonormal frame in the linear space $\mathcal{T}_{\vec{p}}(M)$ belonging to the given orientation of M . Then λ is continuous, and we have:

$$\int_M \omega = \int_M \lambda dV$$

Proof. A general proof is given in Theorem 36.2 on Munkres. For the case where $k = 1$, we can show the following holds:

$$\int_M \omega = \int_M \omega \cdot \vec{T} dV$$

here \vec{T} is the unit tangent field corresponding to the orientation of M . \square

Note: One can use Theorem 31.2 to define the extended integral $\text{ext } \int_M \omega$ for integrating k -form over oriented k -manifold M in \mathbb{R}^n .

Definition 31.2.0.0.1

Let M be an n -manifold in \mathbb{R}^n with nonempty manifold boundary. The inward normal to ∂M at $\vec{p} \in \partial M$ is the velocity vector of a curve that begins at \vec{p} and moves into M as parameter value increases, the outward normal is the negative velocity vector of a curve that begins at \vec{p} and moves into M as parameter value increases.

Lemma 31.2.1 (Lemma 38.7 on Munkres)

Let M be an orientable n -manifold in \mathbb{R}^n . If M is oriented naturally, then the induced orientation of ∂M corresponds to the unit normal field \vec{N} to ∂M that points outwards from M at all $\vec{p} \in \partial M$.

Proof. Let $\vec{p} \in \partial M$, let $\alpha : U \rightarrow V$ be a coordinate patch on M about \vec{p} belonging to the orientation of M , such that $\alpha(\vec{q}) = \vec{p}$. Here $\alpha_0 = \alpha \circ b$ is a coordinate patch on ∂M about \vec{p} , which belongs to the induced orientation of ∂M if n is even, and to the opposite if n is odd, with the function b as defined in Theorem 31.1. Let \vec{N} be the unit normal field to ∂M corresponding to the induced orientation of ∂M , denote $\vec{N}(\vec{p}) = (\vec{p}; \vec{n}(\vec{p}))$, here we can write:

$$\det([(-1)^n \vec{n}(\vec{p}) \quad D\alpha_0]) > 0 \quad \Rightarrow \quad 0 > \det([D\alpha_0 \quad \vec{n}(\vec{p})]) = -\det([D\alpha_0 \quad D_n\alpha(\vec{q})])$$

$D_n\alpha(\vec{q})$ is the velocity vector of a curve that begins at \vec{p} moving into M , hence it follows that $\vec{n}(\vec{p})$ is the outward normal to ∂M at \vec{p} . The result follows. \square

The Generalized Stokes' Theorem

Lemma 32.0.1 (Lemma 25.2 on Munkres)

Let M be a compact k -manifold in \mathbb{R}^n of class C^r . Given a covering \mathcal{C} of M by coordinate patches, there exists a finite collection of C^∞ functions from \mathbb{R}^n to \mathbb{R} , denoted as $P = \{\phi_1, \phi_2, \dots, \phi_l\}$, such that the followings hold:

1. For each $1 \leq i \leq l$, ϕ_i has compact support and there exists a coordinate patch $\alpha_i : U_i \rightarrow V_i$ in the collection \mathcal{C} such that we have $\text{supp}(\phi_i) \cap M \subseteq V_i$
2. $\phi_i(\vec{x}) \geq 0$ for all $\vec{x} \in \mathbb{R}^n$
3. $\sum_{i=1}^l \phi_i(\vec{x}) = 1$ for all $\vec{x} \in M$

Note: In the settings of Lemma 32.0.1, the collection $P = \{\phi_1, \phi_2, \dots, \phi_l\}$ is called a partition of unity on M dominated by the collection \mathcal{C} of the coordinate patches.

Let M be a compact oriented k -manifold in \mathbb{R}^n , along with the orientation O , let ω be a k -form defined in an open subset of \mathbb{R}^n containing M , let $C = M \cap \text{supp}(\omega)$. Here we see that C is compact because both M and $\text{supp}(\omega)$ are compact. Suppose there exists a coordinate patch $\alpha : U \rightarrow V$ on M belonging to the orientation O such that $C \subseteq V$. By replacing U by a smaller open set if necessary, we can assume that U is bounded, and here we define the integral of ω over M by:

$$\int_M \omega := \int_{\text{Int}(U)} \alpha^* \omega \quad (\text{KM})$$

Here we note that $U = \text{Int}(U)$ if U is open in \mathbb{R}^k , and $\text{Int}(U) = U \cap \mathbb{H}_+^k$ if U is open in \mathbb{H}^k .

We note that the integral defined by equation (KM) is linear. More precisely, if k -forms ω and η have supports whose intersections with M can be covered by the single coordinate patch $\alpha : U \rightarrow V$ belonging to O , then we can write:

$$\int_M a\omega + b\eta = \left(a \int_M \omega \right) + \left(b \int_M \eta \right)$$

where $a, b \in \mathbb{R}$. We also note that, if $-M$ denotes the manifold M with the opposite orientation, then, given that all appropriate conditions in the discussion above hold, we have:

$$\int_{-M} \omega = - \int_M \omega$$

To define $\int_M \omega$ in general, we need to make use of a partition of unity:

Definition 32.0.1.0.1

Let M be a compact oriented k -manifold in \mathbb{R}^n , along with orientation O on M . Cover M by a collection \mathcal{C} of coordinate patches in O , then choose a partition of unity $\phi_1, \phi_2, \dots, \phi_l$ on M that is dominated by the collection \mathcal{C} . Let ω be a k -form defined in an open neighborhood of M . We define the integral of ω over M by the following:

$$\int_M \omega = \sum_{i=1}^l \left(\int_M \phi_i \omega \right)$$

Equivalently, since M is compact, then we can cover M by finitely many coordinate patches, that is, we can take \mathcal{C} to be a finite collection of coordinate patches in O , denoted as $C = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$. One can use partition of unity to write $\omega = \sum_{j=1}^N \omega_j$ such that the support of each ω_j is a subset of V_j , where V_j is the codomain of a coordinate patch $\alpha_j : U_j \rightarrow V_j$ in C . Here we define:

$$\int_M \omega = \sum_{j=1}^N \left(\int_{(V_j)_{\alpha_j}} \omega_j \right)$$

Note: In the settings of Definition 32.0.1.0.1, let $c \in \mathbb{R}$, ω_m and ω_n be k -forms defined in an open set in \mathbb{R}^n containing M , we have immediately that the following holds:

$$\int_M c\omega = c \int_M \omega \quad \int_M (\omega_1 + \omega_2) = \int_M \omega_1 + \int_M \omega_2$$

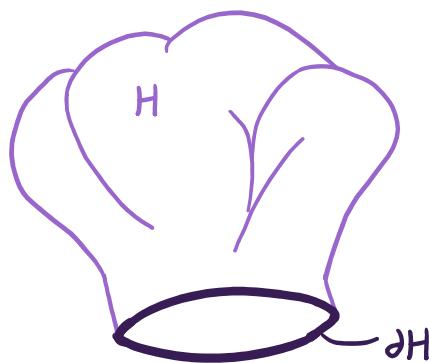
Remark: Let U be an open subset of \mathbb{R}^k , and let V be a k -manifold in \mathbb{R}^n , let $\alpha : U \rightarrow V$ be a orientation preserving coordinate patch for V , and let $\beta := \alpha^{-1}$. Here β is called a chart. For $g : V \rightarrow \mathbb{R}$, we can write the following:

$$\int_{V_\alpha} g d\beta_1 \wedge d\beta_2 \wedge \cdots \wedge d\beta_k = \int_U (g \circ \alpha) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k$$

Intuitively, the integral $\int_M dx \wedge dy$ of a 2-manifold M in \mathbb{R}^3 is the projection area of M on the xy -plane. Similarly, $\int_M dy \wedge dz$ is the projection area of M on the yz -plane, and $\int_M dx \wedge dz$ is the projection area of M on the xz -plane. Note that the orientation of the manifold M would affect the projection area, some of the regions would cancel each other due to the opposite sign specified the orientation.

Example:

Consider such chef hat as a 2-manifold H in \mathbb{R}^3 , and suppose the bottom of the chef hat, which is the manifold boundary of H , is a circle of radius 1 lies on the xy -plane.



Then we can write:

$$\int_H dx \wedge dy = \pi \quad \int_H dy \wedge dz = \int_H dx \wedge dz = 0$$

Example: Consider the sphere $S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$. Intuitively one can write:

$$\int_{S^2} dx \wedge dy = \int_{S^2} dy \wedge dz = \int_{S^2} dx \wedge dz = 0$$

$$\int_{S^2} x \, dx \wedge dy = \int_{S^2} y \, dx \wedge dy = 0$$

Let S_+^2 denote the upper hemisphere, we can write:

$$\int_{S^2} z \, dx \wedge dy = 2 \int_{S_+^2} \sqrt{1 - x^2 - y^2} \, dx \wedge dy = 4\pi \int_0^1 \sqrt{1 - r^2} r \, dr = \frac{4\pi}{3}$$

Now we will prove the generalized Stokes' Theorem, which is a general theorem about integrals of differential forms that includes the three basic theorems of vector integral calculus, the Green's theorem, the Stokes' theorem, and the divergence theorem, as special cases.

We will begin with a lemma that is in some sense a special case of the Stokes' Theorem:

Lemma 32.0.2

Let $I = [0, 1]$, and let I^k denote the unit k -cube in \mathbb{R}^k :

$$I^k = [0, 1]^k = [0, 1] \times [0, 1] \times \cdots \times [0, 1]$$

Let $k > 1$, let η be a $k-1$ form defined on an open subset U of \mathbb{R}^k containing I^k . If η vanishes at all points of $Bd(I^k)$ except possibly at points of the subset $(Int(I^{k-1})) \times \{0\}$, then we have:

$$\int_{Int(I^k)} d\eta = (-1)^k \int_{Int(I^{k-1})} b^* \eta \quad (*)$$

with the function $b : I^{k-1} \rightarrow I^k \quad (u_1, u_2, \dots, u_{k-1}) \mapsto (u_1, u_2, \dots, u_{k-1}, 0)$.

Proof. First we note that $Int(I^k)$ is the open cube $(0, 1)^k$, and we have $Bd(I^k) = I^k \setminus Int(I^k)$. Given $j \in \mathbb{N}$ with $1 \leq j \leq k$, let I_j denote the $(k-1)$ -tuple $I_j = (1, 2, \dots, j-1, j+1, \dots, k)$. We have:

$$dx_{I_j} := dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_k$$

Since the integrals in equation (*) are linear and the operator d and pullback are linear, it suffices to consider the case where $\eta = f \, dx_{I_j}$ with some 0-form f . First, by Fubini's

Theorem we can write the following:

$$\begin{aligned}
\int_{\text{Int}(I^k)} d\eta &= \int_{\text{Int}(I^k)} df \wedge dx_{I_j} \\
&= \int_{\text{Int}(I^k)} \left(\sum_{i=1}^k D_i f dx_i \right) \wedge dx_{I_j} \\
&= \int_{\text{Int}(I^k)} (-1)^{j-1} (D_j f) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k \\
&= (-1)^{j-1} \int_{\text{Int}(I^k)} D_j f \\
&= (-1)^{j-1} \int_{I^k} D_j f \\
&= (-1)^{j-1} \int_{\vec{v} \in I^{k-1}} \int_{x_j \in I} D_j f(x_1, x_2, \dots, x_k)
\end{aligned}$$

where $\vec{v} = (x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$. By the Fundamental Theorem of Calculus, we have:

$$\int_{x_j \in I} D_j f(x_1, x_2, \dots, x_k) = f(x_1, x_2, \dots, x_{j-1}, 1, x_{j+1}, \dots, x_k) - f(x_1, x_2, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_k)$$

By assumption, η , and hence f , vanish at all points of $\text{Bd}(I^k)$ except possibly at points of the subset $(\text{Int}(I^{k-1})) \times \{0\}$. Then we can write:

$$\int_{x_j \in I} D_j f(x_1, x_2, \dots, x_k) = \begin{cases} 0 & j \neq k \\ -f(x_1, x_2, \dots, x_{k-1}, 0) & j = k \end{cases}$$

Concluding the results we get:

$$\int_{\text{Int}(I^k)} d\eta = \begin{cases} 0 & j \neq k \\ (-1)^k \int_{I^{k-1}} (f \circ b) & j = k \end{cases}$$

On the other hand, let I_{k-1} denote the $(k-1) \times (k-1)$ identity matrix, we can write:

$$Db(\vec{u}) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} = \begin{bmatrix} I_{k-1} \\ \vec{0} \end{bmatrix} \quad \forall \vec{u} \in \mathbb{R}^{k-1}$$

Then by Theorem 32.2 on Munkres, we can write:

$$b^*(dx_{I_j}) = \det \left(\frac{\partial(b_1, b_2, \dots, b_{j-1}, b_{j+1}, \dots, b_k)}{\partial(x_1, x_2, \dots, x_{k-1})} \right) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{k-1}$$

with $\vec{u} \in \mathbb{R}^{k-1}$, here $\frac{\partial(b_1, b_2, \dots, b_{j-1}, b_{j+1}, \dots, b_k)}{\partial(x_1, x_2, \dots, x_{k-1})}(\vec{u})$ is the matrix formed by removing the j -th row of the matrix $Db(\vec{u})$. Combining with the value of $Db(\vec{u})$, we can write the

following:

$$b^*(dx_{I_j}) = \begin{cases} 0 & j < k \\ dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{k-1} & j = k \end{cases}$$

Here we can conclude that:

$$\int_{\text{Int}(I^{k-1})} b^* \eta = \begin{cases} 0 & j \neq k \\ \int_{\text{Int}(I^{k-1})} (f \circ b) & j = k \end{cases}$$

Here we see that the results follows. \square

Theorem 32.1 (The Generalized Stokes' Theorem)

Let $k > 1$, let M be a compact oriented k -manifold in \mathbb{R}^n , with ∂M having the induced orientation if ∂M is not empty, let ω be a $(k-1)$ -form defined in an open set of \mathbb{R}^n containing M , then we have the following holds if ∂M is not empty:

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

and we have the following holds if ∂M is empty:

$$\int_{\partial M} \omega = 0$$

Proof. Consider two cases: (1) consider a point $\vec{p} \in M \setminus \partial M$, and (2) consider a point $\vec{p} \in \partial M$.

For case (1), choose a coordinate patch $\alpha : U \rightarrow V$ belonging to the orientation of M , such that U is open in \mathbb{R}^k containing the unit cube I^k defined in Lemma 32.0.2, and such that α carries a point of I^k to the point \vec{p} . Here we note that, if we begin with an arbitrary coordinate patch $\alpha' : U' \rightarrow V$ about \vec{p} belonging to the orientation of M , one can obtain our desired coordinate patch α by a translation and possibly a stretching in \mathbb{R}^k . Let $W = \text{Int}(I^k)$, and let $Y = \alpha(W)$, then the map $\alpha : W \rightarrow Y$ is still a coordinate patch belonging to the orientation of M about the point \vec{p} , with $W = \text{Int}(I^k)$ open in \mathbb{R}^k . Since the operator d and the integrals involved are linear, it suffices to consider the special case ω being a $(k-1)$ -form such that the set $C := M \cap (\text{supp}(\omega))$ is a subset of Y . Note that the suppose of $d\omega$ is contained in the support of ω , hence the set $M \cap (\text{supp}(\omega))$ is contained in C , which is contained in Y . Let η denote the $(k-1)$ -form $\alpha^*\omega$, we see that the definition of domain of the $(k-1)$ -form η can be extended to an open subset of \mathbb{R}^k containing I^k , and we have η vanishes outside of $\text{Int}(I^k)$ as ω vanishes outside of W . Thus we now can apply Lemma 32.0.2, with $\alpha^*d\omega = d(\alpha^*\omega) = d\eta$:

$$\int_M d\omega = \int_{\text{Int}(I^k)} \alpha^* d\omega = \int_{\text{Int}(I^k)} d\eta = (-1)^k \int_{\text{Int}(I^{k-1})} b^* \eta$$

In this case, the form η vanishes outside of $\text{Int}(I^k)$, and hence η vanishes on $I^{k-1} \times \{0\}$, so we have:

$$\int_M d\omega = (-1)^k \int_{\text{Int}(I^{k-1})} b^* \eta = (-1)^k \int_{\text{Int}(I^{k-1})} 0 = 0$$

If ∂M is empty, the result follows, and if ∂M is nonempty, we note that the support of ω is disjoint from ∂M , hence $\int_{\partial M} \omega = \int_M \partial \omega$ as desired.

Now consider the second case where $\vec{p} \in \partial M$. Here we choose a coordinate patch $\alpha : U \rightarrow V$ belonging to the orientation of M , such that U is a subset of \mathbb{H}^k being open in \mathbb{H}^k , U contains I^k , and such that α carries a point of $\text{Int}(I^{k-1}) \times \{0\}$ to the point \vec{p} . Here we let $W = (\text{Int}(I^k)) \cup (\text{Int}(I^{k-1}) \times \{0\})$, and let $Y = \alpha(W)$, then the map $\alpha : W \rightarrow Y$ is still a coordinate patch belonging to the orientation of M about the point \vec{p} , with W being open in \mathbb{H}^k but not open in \mathbb{R}^k . Since the operator d and the integrals involved are linear, it suffices to consider the special case ω being a $(k-1)$ -form such that the set $C := M \cap (\text{supp}(\omega))$ is a subset of Y . Note that the suppose of $d\omega$ is contained in the support of ω , hence the set $M \cap (\text{supp}(\omega))$ is contained in C , which is contained in Y . Let η denote the $(k-1)$ -form $\alpha^*\omega$, we see that the definition of domain of the $(k-1)$ -form η can be extended to an open subset of \mathbb{R}^k containing I^k , and we have η vanishes at all points of $\text{Bd}(I^k)$ except possibly at points in $\text{Int}(I^{k-1}) \times \{0\}$, as ω vanishes outside of W . Thus we now can apply Lemma 32.0.2. First note that Y intersects ∂M non-trivially, and we have $\text{Int}(W) = \text{Int}(I^k)$. Here we can compute:

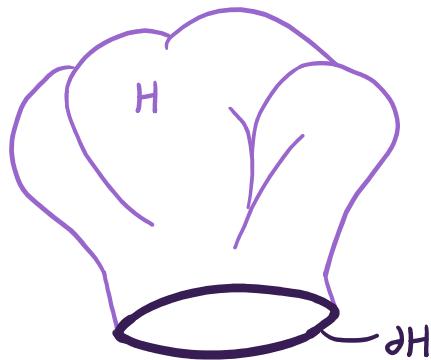
$$\int_M d\omega = \int_{\text{Int}(I^k)} d\eta = (-1)^k \int_{\text{Int}^{k-1}} b^* \eta \quad (**)$$

Now notice that $\partial M \cap (\text{supp}(\omega))$ is covered by the coordinate patch $\beta : \text{Int}^{k-1} \rightarrow Y \cap \partial M$ $\vec{x} \mapsto (\alpha \circ b)(\vec{x})$ on ∂M . Note that β belongs to the induced orientation of ∂M if k is even, and to the opposite orientation if k is odd. If we use β to compute the integral of ω over ∂M , we must reverse the sign of the integral when k is odd. Thus now we can write:

$$\int_{\partial M} \omega = (-1)^k \int_{\text{Int}(I^{k-1})} \beta^* \omega$$

Since we have $\beta^* \omega = (\alpha \circ b)^* \omega = b^* \alpha^* \omega = b^* \eta$, comparing with equation $(**)$ we see that the result of Theorem 31.1 follows. \square

Example: Let H be a chef hat in \mathbb{R}^3 , which is a manifold, the boundary of H is a circle, here we suppose further that the boundary of H is a circle of radius 1 that lies on the xy -plane.



Let $M = \{(x, y, 0) \mid x^2 + y^2 \leq 1\}$, here we have $\partial M = \{(x, y, 0) \mid x^2 + y^2 = 1\} = \partial H$. We can write:

$$\int_H dx \wedge dy = \int_{\partial H} x dy = \int_{\partial M} x dy = \int_M dx \wedge dy = \pi$$

Here we can also write:

$$\int_H dy \wedge dz = \int_{\partial H} (-z dy) = 0 \quad \int_H dx \wedge dz = \int_{\partial H} (-z \wedge dy) = 0$$

Example: For a sphere of radius 1, denoted as $S^2 \in \mathbb{R}^3$, we can write the following:

$$\int_{S^2} dx \wedge dy = \int_{\partial S^2} x dy = 0$$

Similarly, we can write:

$$0 = \int_{S^2} dy \wedge dz = \int_{S^2} dx \wedge dz$$

We can also write:

$$\int_{S^2} x dx \wedge dy = \int_{\partial S^2} \frac{x^2}{2} dy = 0$$

Similarly we have:

$$\int_{S^2} y dx \wedge dy = 0$$

Now let $B^3 := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1\}$, we know that $\partial B^3 = S^2$. Here we can write:

$$\int_{S^2} z dx \wedge dy = \int_{B^3} dz \wedge dx \wedge dy = \int_{B^3} dx \wedge dy \wedge dz = \frac{4\pi}{3}$$

Example: Let $M = \{(x, y, z) \mid x^2 + y^2 = 1, 0 \leq z \leq 1\}$. Here we define a 1-form ω :

$$\omega : \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3 \mid x = y = 0\} \rightarrow \mathbb{R}_{row}^3 \quad (x, y, z) \mapsto \frac{-y dx + x dy}{x^2 + y^2}$$

We have the following holds:

$$\int_M d\omega = \int_{\partial M} \omega = 2\pi - 2\pi = 0$$

Application in Vector Calculus

Definition 33.0.0.0.1

Let A be an open subset of \mathbb{R}^n , any function of the form $f : A \rightarrow \mathbb{R}$ is called a scalar field.

Definition 33.0.0.0.2

Let A be an open subset of \mathbb{R}^n , any function of the form $v : A \rightarrow \mathbb{R}^n$ is called a vector field.

Example: Velocity of fluid defines a vector field, while force field can be viewed as 1-form.

Definition 33.0.0.0.3

Let α be a local diffeomorphism, let v be a vector field defined on an open subset A of \mathbb{R}^n . The function $\alpha^*v := (D\alpha)^{-1} \cdot (v \circ \alpha)$ is called the pullback of vector field v by α .

For 1-form ω , we see that $\omega^T := (\omega(\vec{x}))^T$ is a vector field, here for α being a local diffeomorphism such that $\omega \circ \alpha$ is well defined, we can write the following, for all \vec{x} in the definition of domain of α :

$$((\alpha^*\omega)(\vec{x}))^T = (\omega(\alpha(\vec{x})) \cdot D\alpha(\vec{x}))^T = (D\alpha(\vec{x}))^T \cdot (\omega(\alpha(\vec{x})))^T$$

Here we see that:

$$(\alpha^*\omega)^T = \alpha^*(\omega^T) \iff (D\alpha(\vec{x}))^T = (D\alpha(\vec{x}))^{-1}$$

Note that $(D\alpha(\vec{x}))^T = (D\alpha(\vec{x}))^{-1}$ for all \vec{x} in the definition of domain of α if and only if α is orthogonal, and by Theorem 20.5 on Munkres, this is true if and only if α is an isometry.

Exterior Calculus in \mathbb{R}^3 works well with diffeomorphism. Differential forms in \mathbb{R}^3 has the following forms, with α, β, γ being scalar functions, or 0-forms defined on \mathbb{R}^3 :

1. 0-forms f
2. 1-forms $\alpha dx + \beta dy + \gamma dz$
3. 2-forms $\alpha dy \wedge dz + \beta dz \wedge dx + \gamma dx \wedge dy$
4. 3-forms $f dx \wedge dy \wedge dz$

Vector Calculus in \mathbb{R}^3 works well with isometry.

1. Scalar Fields f , corresponding to 0-forms or 3-forms defined on \mathbb{R}^3 .
2. Vector Fields v , corresponding to 1-forms, or 2-forms defined on \mathbb{R}^3 .

Note: Isometries defined on \mathbb{R}^2 , or \mathbb{R}^3 , consist of translations, rotation, and reflection.

In the following table, we compare objects in exterior calculus to objects in vector calculus. One will find that, on each row, the object on the right column can be interpreted in a way very similar to the object on the left column. First we restrict our view in \mathbb{R}^3 , and assume that $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a scalar field of C^1 type, α, β, γ are 0-forms, ω, ζ, ν are 1-forms, μ is 2-form, all defined on \mathbb{R}^3 , and $\vec{F}, \vec{F}_1, \vec{F}_2$ are vector fields defined on \mathbb{R}^3 , with \vec{F} defined by the following:

$$\vec{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \quad \vec{x} \mapsto \begin{bmatrix} \alpha(\vec{x}) \\ \beta(\vec{x}) \\ \gamma(\vec{x}) \end{bmatrix}$$

Here we get the table:

Exterior Calculus	Vector Calculus
Exterior derivative operator d	Del operator $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$
df	The gradient of f , $\text{grad}(f) := \nabla f$
$d(\alpha dx + \beta dy + \gamma dz)$	The curl of \vec{F} , $\text{curl}(\vec{F}) := \nabla \times \vec{F}$
$d(\alpha dy \wedge dz + \beta dz \wedge dx + \gamma dx \wedge dy)$	The divergence of \vec{F} , $\text{div}(\vec{F}) := \nabla \cdot \vec{F} = \langle \nabla, \vec{F} \rangle$
$\omega \wedge \mu$	Scalar field $\langle \vec{F}_1, \vec{F}_2 \rangle$
$\omega \wedge \zeta$	Vector field $\vec{F}_1 \times \vec{F}_2$
$\omega \wedge \zeta \wedge \nu$	Scalar field $\det [\vec{F}_1, \vec{F}_2, \vec{F}_3]$

Now in \mathbb{R}^2 , we note that exterior calculus in \mathbb{R}^2 works well with diffeomorphism, and vector calculus in \mathbb{R}^2 works well with isometry. Here we let f be a scalar field of C^1 type, α, β be 0-forms defined on \mathbb{R}^2 , $\omega = \alpha dx + \beta dy$ defined on \mathbb{R}^2 , vector field $\vec{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ $\vec{x} \mapsto (\alpha(\vec{x}), \beta(\vec{x}))$, and let M_1 be 1-manifold in \mathbb{R}^2 , M_2 be 2-manifold in \mathbb{R}^2 . Here we get the table:

Exterior Calculus	Vector Calculus
Exterior derivative operator d	Del operator $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$
0-form k define on \mathbb{R}^2	Scalar field k of C^1 type defined on \mathbb{R}^2
1-forms $\omega = \alpha dx + \beta dy$	Vector field \vec{F}
2-forms $f dx \wedge dy$ defined on \mathbb{R}^2	Scalar field f
1-form ω wedged with 1-form η	Scalar field $\det \begin{pmatrix} \vec{F}_1 \vec{F}_2 \end{pmatrix}$ with \vec{F}_1, \vec{F}_2 being vector fields
$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$	Gradient of f , $\text{grad}(f) := \nabla f = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$
$d(\alpha dx + \beta dy) = \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right) dx \wedge dy$	Curl of \vec{F} , $\text{curl}(\vec{F}) := \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right)$
$\int_{M_1} \omega$	$\int_{M_1} \langle \vec{F}, d\vec{l} \rangle = \int_{M_1} \langle \vec{F}, \vec{T} \rangle dV$ ¹
$\int_{M_1} df = \Delta_{M_1} f$	$\int_{M_1} \langle \nabla f, \vec{T} \rangle = \Delta_{M_1} f$ ²
$\int_{M_2} f dx \wedge dy = \int_{M_2} f$	$\int_{M_2} f$
$\int_{\partial M_2} \omega = \int_{M_2} d\omega$	Circulation of \vec{F} along ∂M_2 , $\int_{M_2} \text{curl}(\vec{F})$

¹ Here we define:

$$d\vec{l} := \begin{bmatrix} dx \\ dy \end{bmatrix}$$

Since we have $\vec{F}(\vec{x}) = (\alpha(\vec{x}), \beta(\vec{x}))$, so we have $d\vec{l} = \vec{T} dV$.

² The standard interpretation here is that f is called the potential energy.

The force associated to potential energy f is defined to be $-\nabla f$, here the negative sign for force

simplifies the conservation of energy, and work done by such force is then defined to be:

$$\int_{M_1} \langle -\nabla f, \vec{T} \rangle ds = -\Delta_{M_1} f$$

Definition 33.0.0.4

Let M be a compact 2-manifold in \mathbb{R}^2 with nonempty boundary, let \vec{F} be a vector field defined on an open subset of \mathbb{R}^2 containing M . Integrating the curl of \vec{F} over M is called the circulation of \vec{F} along ∂M . That is, we can write:

$$\text{Circulation of } \vec{F} \text{ along } \partial M := \int_M \text{curl}(\vec{F})$$

Consider a vector field $\vec{F} : A \rightarrow \mathbb{R}^2$ measuring the velocity of a fluid, with A being an open subset of \mathbb{R}^2 . Let D denote a disc in \mathbb{R}^2 , with a circular boundary X of radius r . Here D can be viewed as a 2-manifold in \mathbb{R}^2 with natural orientation. Here we give the 1-manifold $X = \partial D$ the induced orientation, and denote its unit tangent field

corresponding to such orientation as \vec{T} . The inner product defined by the following along the boundary of D is called the tangential velocity of the fluid at the point $\vec{x} \in X$:

$$\langle \vec{F}(\vec{x}), \vec{T}(\vec{x}) \rangle$$

With results from above, the following integral gives the circulation of \vec{F} on X :

$$\text{Circulation} := \int_X \langle \vec{F}, \vec{T} \rangle dV$$

The average tangential velocity on X is defined to be the following:

$$\text{Average Tangential Velocity} := \frac{\int_X \langle \vec{F}, \vec{T} \rangle dV}{\int_X dV}$$

note that the average tangential velocity has unit [Distance] / [Times].

Notice that in our settings, we have:

$$\int_X dV = 2\pi r$$

The average angular velocity is defined to be the following:

$$\text{Average Angular Velocity} := \frac{\int_X \langle \vec{F}, \vec{T} \rangle dV}{2\pi r^2}$$

We denote:

$$\vec{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \quad \vec{x} \mapsto \begin{bmatrix} \alpha(\vec{x}) \\ \beta(\vec{x}) \end{bmatrix} \quad \omega : \mathbb{R}^2 \rightarrow \mathbb{R}_{row}^2 \quad \vec{x} \mapsto \alpha(\vec{x}) dx + \beta(\vec{x}) dy$$

with α, β being scalar-valued differentiable functions defined on \mathbb{R}^2 .

Now by Stokes' Theorem, we can write the following:

$$\frac{\int_X \langle \vec{F}, \vec{T} \rangle}{2\pi r^2} = \frac{\int_X \alpha dx + \beta dy}{2\pi r^2} = \frac{\int_D \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right) dx \wedge dy}{2\pi r^2} = \frac{\int_D \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right)}{2\pi r^2} \quad (\text{AA})$$

Let \vec{o} denote the center of the disc D , the angular velocity of the fluid at \vec{o} is defined as the average angular velocity of the disc D_r , centered at \vec{o} , with radius r tends to zero. In such case, $(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y})$ is approximately constant, with equation (AA) we can write the following:

$$\text{Angular Velocity at } \vec{o} = \frac{1}{2} \left[\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right]_{(x,y)=\vec{o}} = \frac{1}{2} \operatorname{curl}(F)(\vec{o})$$

In general, for a compact 2-manifold M and a vector field defined by the following:

$$\vec{F} : A \rightarrow \mathbb{R}^2 \quad \vec{x} \rightarrow \begin{bmatrix} \alpha(\vec{x}) \\ \beta(\vec{x}) \end{bmatrix}$$

with A being open subset of \mathbb{R}^2 , α, β being scalar-valued differentiable functions defined on A . We can write the following by Stokes' Theorem:

$$\text{Circulation of } \vec{F} \text{ along } \partial M = \int_M \operatorname{curl}(\vec{F}) = \int_{\partial M} \langle \vec{F}, \vec{T} \rangle dV = \int_{\partial M} \alpha dx + \beta dy$$

This result is concluded as the Green's Theorem for vector calculus.

Let M be a compact oriented 2-manifold in \mathbb{R}^3 , by Theorem 38.9 on Munkres, the circulation of a vector field \vec{F} defined on \mathbb{R}^3 along ∂M is given by the following:

$$\int_{\partial M} \langle \vec{F}, \vec{T} \rangle dV = \int_M \langle \text{curl}(\vec{F}), \vec{N} \rangle dV$$

where \vec{N} is the unit normal field corresponding to the orientation of M , and \vec{T} is the unit tangent vector field corresponding to the induced orientation of ∂M . This result is concluded as the classical version of the Stokes' Theorem.

Now let $\vec{p} \in \mathbb{R}^3$ and $r \in \mathbb{R}$ with $r > 0$, let \vec{v}_1, \vec{v}_2 be orthonormal vectors, let $M_r \subseteq \mathbb{R}^3$ be the 1-manifold parametrized by the following function:

$$\alpha : [0, 2\pi) \quad \theta \mapsto \vec{p} + r \cos(\theta) \vec{v}_1 + r \sin(\theta) \vec{v}_2$$

Consider a vector field \vec{F} defined by the following, with α, β, γ being scalar valued differentiable functions defined on an open subset A containing M_r :

$$\vec{F} : A \rightarrow \mathbb{R}^3 \quad \vec{x} \mapsto \begin{bmatrix} \alpha(\vec{x}) \\ \beta(\vec{x}) \\ \gamma(\vec{x}) \end{bmatrix}$$

Here \vec{F} is equivalent to an 1-form defined by $\omega = \alpha dx + \beta dy + \gamma dz$ defined on A . The average angular velocity along M_r is then given by the following:

$$\text{Average angular velocity along } M_r = \left(\text{Circulation of } \vec{F} \text{ along } M_r \right) / (2\pi r^2) = \frac{\int_{M_r} \omega}{2\pi r^2}$$

And the angular velocity of \vec{F} at the point \vec{p} is given by:

$$\text{Angular velocity at } \vec{p} = \frac{1}{2} \left\langle \text{curl}(\vec{F})(\vec{p}), \vec{v}_1 \times \vec{v}_2 \right\rangle = \frac{1}{2} \det \left(\begin{bmatrix} \text{curl}(\vec{F})(\vec{p}) & \vec{v}_1 & \vec{v}_2 \end{bmatrix} \right)$$

The discussion above can be concluded more precisely following the notation from Munkres, given in the following:

Let $(\vec{e}_i \mid 1 \leq i \leq n)$ denote the standard basis in \mathbb{R}^n , let A be an open subset in \mathbb{R}^n . First we note that, each vector field $\vec{F} : A \rightarrow \mathbb{R}^n \quad \vec{x} \mapsto \sum_{i=1}^n F_i(\vec{x}) \vec{e}_i$, with F_i being component functions of \vec{F} , can be equivalently written as the following:

$$\vec{F} : A \rightarrow \mathbb{R}^n \times \mathbb{R}^n \quad \vec{x} \mapsto \left(\vec{x}; \sum_{i=1}^n F_i(\vec{x}) \vec{e}_i \right)$$

and such vector field corresponds to a 1-form defined on A defined by the following:

$$\omega : A \rightarrow \mathbb{R}_{row}^n \quad \vec{x} \mapsto \sum_{i=1}^n F_i(\vec{x}) dx_i$$

In such case, we say the 1-form ω corresponds to the vector field \vec{F} .

Lemma 33.0.1 (Lemma 38.1 on Munkres)

Let M be a compact oriented 1-manifold in \mathbb{R}^n , and let \vec{T} be the unit tangent vector to M corresponding to the given orientation of M . Let \vec{F} be a vector field defined in \mathbb{R}^n and let ω be the 1-form corresponds to \vec{F} . Then we have:

$$\int_M \omega = \int_M \langle \vec{F}, \vec{T} \rangle dV \quad (\text{sv})$$

Proof. Here it suffices to consider the case where $M \cap \text{supp}(\omega)$ is contained in the image of a single coordinate patch $\alpha : U \rightarrow V$ on M belonging to the orientation of M . Let F_i denote the component functions of \vec{F} . Since we have $\mathcal{V}(D\alpha) = \sqrt{\det((D\alpha)^T D\alpha)} = \|D\alpha\|$, we can write:

$$\begin{aligned} \int_M \omega &= \int_U \alpha^* \omega = \int_U \langle \vec{F} \circ \alpha, D\alpha \rangle \\ &= \int_U \left\langle \vec{F} \circ \alpha, \frac{D\alpha}{\|D\alpha\|} \right\rangle \mathcal{V}(D\alpha) \\ &= \int_U \langle \vec{F} \circ \alpha, \vec{T} \circ \alpha \rangle \mathcal{V}(D\alpha) \\ &= \int_M \langle \vec{F}, \vec{T} \rangle \end{aligned}$$

The result follows. \square

Remark: The notation dV in equation (sv) is sometimes denoted as ds for 1-volume.

Theorem 33.1 (The Gradient Theorem)

Let M be a compact oriented 1-manifold in \mathbb{R}^n , let \vec{T} be the unit tangent vector field to M corresponding to the orientation of M , and let f be a 0-form defined on an open subset of \mathbb{R}^n containing M . If ∂M is empty, then we have:

$$\int_M \langle \nabla f, \vec{T} \rangle dV = 0$$

If ∂M is not empty, and $\partial M = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m)$, we let $\epsilon_i = -1$ if $\vec{T}(\vec{x}_i)$ points into M , and let $\epsilon_i = 1$ if $\vec{T}(\vec{x}_i)$ points out of M . Then we have:

$$\int_M \langle \nabla f, \vec{T} \rangle dV = \sum_{i=1}^M \epsilon_i f(\vec{x}_i)$$

Proof. Note that the 1-form df correspond to the vector field ∇f , then by Theorem 31.1 on Munkres and Lemma 38.1 on Munkres, the result follows. \square

Let \vec{F} be a vector field defined on an open subset A of \mathbb{R}^n , by Munkres Theorem 31.1, \vec{F} corresponds to an $(n-1)$ -form defined by the following:

$$\omega = \sum_{i=1}^n (-1)^{i-1} F_i dx_1 \wedge dx_2 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n$$

where F_i are the component functions of \vec{F} .

Lemma 33.1.1 (Lemma 38.5 on Munkres)

Let M be a compact oriented $(n-1)$ -manifold in \mathbb{R}^n , and let \vec{N} be the corresponding unit normal vector field, let \vec{F} be a vector field defined on an open subset A of \mathbb{R}^n containing M , and let ω be the $(n-1)$ -form corresponds to \vec{F} , then we have:

$$\int_M \omega = \int_M \langle \vec{F}, \vec{N} \rangle dV$$

Proof. Here we note that, for the case where $n = 2$, it is equivalent to write the following:

$$\int_M \omega = \int_M \lambda dV$$

where we have:

$$\lambda : M \rightarrow \mathbb{R} \quad \vec{p} \mapsto \omega(\vec{p})(\vec{v}_1(\vec{p}), \vec{v}_2(\vec{p}))$$

where \vec{v}_1 and \vec{v}_2 are functions such that for $\vec{p} \in M$, $(\vec{v}_1(\vec{p}), \vec{v}_2(\vec{p}))$ constitutes an orthonormal frame in $\mathcal{T}_{\vec{p}}(M)$ belonging to the given orientation of M . Given the result of this lemma, we can write:

$$\int_M \lambda dV = \int_M \langle \vec{F}, \vec{v}_1 \times \vec{v}_2 \rangle dV = \int_M \langle \vec{F}, \vec{N} \rangle dV$$

The general proof of this Lemma is given in Lemma 38.5 on Munrkes. \square

In the settings of Lemma 33.1.1, given $n = 3$, one might consider \vec{F} as the velocity of fluid with unit density, here $\int_M \omega$ is called the flux of \vec{F} across M , which measures the rate of flow with cancellation, for which we get positive flow in a small region when $\langle \vec{F}, \vec{N} \rangle > 0$, and negative flow in such small region when $\langle \vec{F}, \vec{N} \rangle < 0$. If the fluid has density ρ , then fluid crosses M at a velocity given by $\rho \vec{F}$. Here ρ might not necessarily be a constant and can be a scalar-valued function defined on an open subset of \mathbb{R}^3 containing M . In general, \vec{F} and ρ could also be time dependent.

Let ω be an n -form defined on an open subset A of \mathbb{R}^n , we note that ω can be written as the following form:

$$\omega = h dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$$

with h being a scalar-valued function defined on A . Here h is called the scalar field associated to the n -form ω .

Lemma 33.1.2 (Lemma 38.6 on Munkres)

Let M be a compact n -manifold in \mathbb{R}^n , oriented naturally, and let $\omega = h dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$ be an n -form defined on an open set of \mathbb{R}^n containing M , with h being the scalar field corresponds to ω , then we have:

$$\int_M \omega = \int_M h dV$$

Proof. It suffices to consider the case where the set $M \cap \text{supp}(\omega)$ is contained in the image of a coordinate patch $\alpha : U \rightarrow V$ on M belonging to the orientation of M . Note

that $\det(D\alpha) > 0$ as α belongs to the natural orientation of M , hence we can write:

$$\begin{aligned}\int_M \omega &= \int_U \alpha^* \omega = \int_U (\omega \circ \alpha) \\ &= \int_U (h \circ \alpha) \det(D\alpha) \\ &= \int_U (h \circ \alpha) \sqrt{\det(D\alpha) \cdot \det((D\alpha)^T)} \\ &= \int_h h \mathcal{V}(D\alpha) \\ &= \int_U h dV\end{aligned}$$

□

Definition 33.1.2.0.1

Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{R}$ be a scalar field defined on A , the gradient of f , denoted as ∇f , or $\text{grad}(f)$, is a vector field defined on A defined by the following:

$$\nabla f : A \rightarrow \mathbb{R}^n \quad \vec{x} \mapsto (D_1 f(\vec{x}), D_2 f(\vec{x}), \dots, D_n f(\vec{x}))$$

One might also denote:

$$\nabla f(\vec{x}) = \text{grad}(f)(\vec{x}) = (\vec{x}; (D_1 f(\vec{x}), D_2 f(\vec{x}), \dots, D_n f(\vec{x})))$$

Definition 33.1.2.0.2

Let A be an open subset of \mathbb{R}^n , let $\vec{F} : A \rightarrow \mathbb{R}^n$ be a vector field defined on A , let F_i denote the component functions of \vec{F} , the divergence of \vec{F} , denoted as $\nabla \cdot \vec{F}$ or $\text{div}(\vec{F})$, is a scalar field defined on A defined by the following:

$$\nabla \cdot \vec{F} : A \rightarrow \mathbb{R} \quad \vec{x} \mapsto \sum_{i=1}^n D_i F_i(\vec{x})$$

Definition 33.1.2.0.3

Let A be an open subset of \mathbb{R}^3 , let $\vec{F} : A \rightarrow \mathbb{R}^3$ be a vector field defined on A , let F_i denote the component functions of \vec{F} , the curl of \vec{F} , denoted as $\nabla \times \vec{F}$ or $\text{curl}(\vec{F})$, is a vector field defined on A defined by the following:

$$\nabla \times \vec{F} : A \rightarrow \mathbb{R}^3 \quad \vec{x} \mapsto (D_2 F_3(\vec{x}) - D_3 F_2(\vec{x}), D_3 F_1(\vec{x}) - D_1 F_3(\vec{x}), D_1 F_2(\vec{x}) - D_2 F_1(\vec{x}))$$

One might also denote:

$$(\nabla \times \vec{F})(\vec{x}) = \text{curl}(\vec{F})(\vec{x}) = (\vec{x}; (D_2 F_3(\vec{x}) - D_3 F_2(\vec{x}), D_3 F_1(\vec{x}) - D_1 F_3(\vec{x}), D_1 F_2(\vec{x}) - D_2 F_1(\vec{x})))$$

Theorem 33.2 (The Divergence Theorem)

Let M be a compact n -manifold in \mathbb{R}^n , let \vec{N} be the unit normal vector field to ∂M that points outwards from M , and let \vec{F} be a vector field defined on an open subset of \mathbb{R}^n containing M , then we have:

$$\int_M \text{div}(\vec{F}) dV = \int_{\partial M} \langle \vec{F}, \vec{N} \rangle dV$$

Proof. Here we can orient M naturally such that ∂M has the induced orientation. Let ω be the $(n-1)$ -form corresponds to the vector field \vec{F} , then we can write the following by Lemma 33.1.1:

$$\int_{\partial M} d\omega = \int_{\partial M} \langle \vec{F}, \vec{N} \rangle dV$$

By Theorem 31.1 on Munkres, we have $d\omega = \operatorname{div}(\vec{F}) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$. Then by Lemma 33.1.2 and Stokes' Theorem, we get the following:

$$\int_{\partial M} \langle \vec{F}, \vec{N} \rangle dV = \int_{\partial M} d\omega = \int_M d\omega = \int_M \operatorname{div}(\vec{F}) dV$$

The result follows. \square

Note: In the settings of \mathbb{R}^3 , the Divergence Theorem is sometimes called the Gauss's Theorem.

Example: For compact oriented 3-manifold U in \mathbb{R}^3 , let \vec{F} be a vector field defined on an open subset of \mathbb{R}^3 containing U , representing the fluid velocity around U , let ρ be a scalar field defined on an open subset of \mathbb{R}^3 containing U , representing the density of the fluid around U . The flow out of U is the flux of $\rho\vec{F}$ across ∂U , by the Divergence Theorem, let ω denote the $(n-1)$ -form corresponds to the vector field $\rho\vec{F}$, here we can write:

$$\text{Flow out of } U = \int_{\partial U} \omega = \int_U d\omega = \int_U \operatorname{div}(\rho\vec{F})$$

If $\operatorname{div}(\rho\vec{F}) > 0$ on U , the fluid is said to be expanding. If $\operatorname{div}(\rho\vec{F}) < 0$ on U , the fluid is said to be contracting. For incompressible fluid, ρ is constant in space and time, and $\operatorname{div}(\rho\vec{F}) = 0$ on U .

Theorem 33.3 (Stokes' Theorem for 2-manifold in \mathbb{R}^3)

Let M be a compact oriented 2-manifold in \mathbb{R}^3 , let \vec{N} be a unit normal field to M corresponding to the orientation of M , and let \vec{F} be a vector field of C^∞ type defined on an open subset of \mathbb{R}^3 containing M . If ∂M is empty, then we have:

$$\int_M \langle \operatorname{curl}(\vec{F}), \vec{N} \rangle dV = 0$$

If ∂M is nonempty, let \vec{T} be the unit tangent vector field to ∂M chosen such that $\vec{N}(\vec{p}) \times \vec{T}(\vec{p})$ points into M from $\vec{p} \in \partial M$, then we have:

$$\int_M \langle \operatorname{curl}(\vec{F}), \vec{N} \rangle dV = \int_{\partial M} \langle \vec{F}, \vec{N} \rangle dV$$

Proof. Let ω be the 1-form corresponds to the vector field \vec{F} . Then by Theorem 31.2, the vector field $\operatorname{curl}(\vec{F})$ corresponds to the 2-form $d\omega$. By Lemma 33.1.1, one can write the following:

$$\int_M d\omega = \int_M \langle \operatorname{curl}(\vec{F}), \vec{N} \rangle dV$$

If ∂M is nonempty, then by Stokes' Theorem and Lemma 33.0.1, we have:

$$\int_M \langle \operatorname{curl}(\vec{F}), \vec{N} \rangle dV = \int_M d\omega = \int_{\partial M} \omega = \int_{\partial M} \langle \vec{F}, \vec{T} \rangle dV$$

The result follows. \square

For $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^3$, we get the followings:

1. $\vec{v}_1 \times \vec{v}_2 \in (\operatorname{span}\{\vec{v}_1, \vec{v}_2\})^\perp$
2. $\|\vec{v}_1 \times \vec{v}_2\| = \|\vec{v}_1\| \cdot \|\vec{v}_2\| \sin(\theta)$, where θ represents the angle between \vec{v}_1 and \vec{v}_2
3. Orthonormal \vec{v}_1, \vec{v}_2 induces the positively oriented orthonormal basis for \mathbb{R}^3 , $(\vec{v}_1, \vec{v}_2, \vec{v}_1 \times \vec{v}_2)$

Proposition 33.3.1 (Euler's Equations for Non-viscous Incompressible Fluid)

For non-viscous incompressible fluid, let \vec{F} denote the fluid velocity vector field, we have:

$$\operatorname{div}(\vec{F}) = 0 \quad \frac{\partial}{\partial t} \left(\operatorname{curl}(\vec{F}) \right) + \operatorname{curl} \left(\left(\operatorname{curl}(\vec{F}) \right) \times \vec{F} \right) = \vec{0}$$

Remark: Here non-viscous means that there is no fluid friction.

Note: One can include viscosity to get the Navier–Stokes equations.³

If we have $\operatorname{curl}(\vec{F}) = \vec{0}$, then the fluid is said to be irrotational, such incompressible irrotational fluid satisfies the Euler's Equations for Non-viscous Incompressible Fluid.

Consider the special case where we have:

$$\vec{F}(x, y, z) = \begin{bmatrix} \alpha(x, y) \\ \beta(x, y) \\ 0 \end{bmatrix}$$

for some differentiable scalar-valued functions α, β defined on \mathbb{R}^2 .

If \vec{F} satisfies $\operatorname{div}(\vec{F}) = \vec{0}$, then we get the following:

$$\frac{\partial \alpha}{\partial x} + \frac{\partial \beta}{\partial y} = 0 \quad \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} = 0$$

which is equivalent to say that the function $\alpha - i\beta$ is holomorphic.

Theorem 33.4

Let M be a compact 2-manifold in \mathbb{C} , let $f \in C^1(M, \mathbb{C})$ be holomorphic on $M \setminus \partial M$. We have:

$$\int_{\partial M} f dz = 0$$

Proof. Here we can write:

$$\int_{\partial M} f dz = \int_M d(f dz) = \int_M df \wedge dz = \int_M (f'(z) dz) \wedge dz = \int_M 0 = 0$$

The result follows. \square

³Get a million dollars by solving this or finding a counterexample!

7 | Differential Equations

The Cauchy-Lipschitz Theorem

Lemma 34.0.1

Let $\Phi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ be a continuous function that satisfies the Partial Lipchitz Estimate:

$$\|\Phi(\vec{x}_2, y) - \Phi(\vec{x}_1, y)\| \leq M\|\vec{x}_2 - \vec{x}_1\|$$

with some $M \in (0, \infty)$, for all $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$ and $y \in \mathbb{R}$. Let $\vec{a} \in \mathbb{R}^n$. Then there exists $\tau \in (0, \infty)$ such that the function T defined by the following is a contraction:

$$T : C([-\tau, \tau], \mathbb{R}^n) \rightarrow C([-\tau, \tau], \mathbb{R}^n) \quad (Tf)(t) = \vec{a} + \int_0^t \Phi(f(s), s) ds$$

here $C([-\tau, \tau], \mathbb{R}^n)$ is equipped with the norm $\|f\|_{sup} := \max_{t \in [-\tau, \tau]} \|f(t)\|$.

Proof. Take $\tau = \frac{1}{19M}$. Consider $f, g \in C([-\tau, \tau], \mathbb{R}^n)$, we can write the following:

$$\begin{aligned} \|T(f) - T(g)\|_{sup} &= \left\| \left(\int_0^t \Phi(f(s), s) ds \right) - \left(\int_0^t \Phi(g(s), s) ds \right) \right\|_{sup} \\ &= \left\| \int_0^t \Phi(f(s), s) - \Phi(g(s), s) ds \right\|_{sup} \\ &\leq |t - 0| \cdot \|\Phi(f(s), s) - \Phi(g(s), s)\|_{sup} \\ &\leq \tau \cdot \|\Phi(f(s), s) - \Phi(g(s), s)\|_{sup} \\ &\leq \tau M \|f - g\|_{sup} \\ &\leq \frac{1}{19} \|f - g\|_{sup} \end{aligned}$$

Hence we see that T is a Lipchitz map on $C([-\tau, \tau], \mathbb{R}^n)$ with Lipchitz constant $\frac{1}{19} < 1$. We conclude that T is a contraction when $\tau = \frac{1}{19M}$. This completes the proof. \square

Given the settings in Lemma 35.0.1, by the Contraction Mapping Theorem, we see that T has a unique fixed point $f \in C([-\tau, \tau], \mathbb{R}^n)$. That is, we have $T(f) = f$.

Corollary 34.0.1.1

Given the settings in Lemma 35.0.1, the function T has a unique fixed point $f \in C([-\tau, \tau], \mathbb{R}^n)$ such that f solves the following initial value problem:

$$f'(t) = \Phi(f(t), t) \text{ for } t \in (-\tau, \tau), \quad f(0) = \vec{a} \quad (\text{IVP})$$

Proof. By the Fundamental Theorem of Calculus, we can write the following:

$$f'(t) = (Tf)'(t) = \frac{d}{dt} \left(\vec{a} + \int_0^t \Phi(f(s), s) ds \right) = \frac{d}{dt} \vec{a} + \frac{d}{dt} \int_0^t \Phi(f(s), s) ds = \vec{0} + \Phi(f(t), t) = \Phi(f(t), t)$$

On the other hand, we have:

$$f(0) = \vec{a} + \int_0^0 \Phi(f(s), s) ds = \vec{a}$$

The result follows. \square

Corollary 34.0.1.2

Given the settings in Lemma 34.0.1, any function $f \in C([-\tau, \tau], \mathbb{R}^n)$ that solves the initial value problem given by (IVP) must be a fixed point of the function T .

Proof. Suppose $f \in C([-\tau, \tau], \mathbb{R}^n)$ solves the initial problem (IVP). Then we know that $f'(t) = \Phi(f(t), t)$. Note here Φ is a continuous function and f is also a continuous function, we know that composition of continuous functions is continuous, hence f' is continuous. Then we know that f' is integrable. Hence by the Fundamental Theorem of Calculus, we can write the following:

$$\int_0^t f' = f(t) - f(0) = f(t) - \vec{a} \quad \Rightarrow \quad f(t) = \vec{a} + \int_0^t f' = \vec{a} + \int_0^t \Phi(f(t), t) = (Tf)(t)$$

Hence we see that $(Tf)(t) = f(t)$ for all $t \in [-\tau, \tau]$, it follows that $Tf = f$, and hence f is a fixed point of T . This completes the proof. \square

Now combining Lemma 34.0.1, Corollary 34.0.1.1, and Corollary 34.0.1.2 with a translation on \mathbb{R} , it is easy to obtain the following result, called the Cauchy-Lipschitz Theorem, or the Fundamental Theorem of Existence and Uniqueness for Ordinary Differential Equations, or simply the Existence and Uniqueness Theorem:

Theorem 34.1 (Cauchy-Lipschitz Theorem)

Let $\vec{a} \in \mathbb{R}^n$, let V be an open neighborhood of \vec{a} , let I be a an open neighborhood of $t_0 \in \mathbb{R}$, and let $\Phi : V \times I \rightarrow \mathbb{R}^n$ be a continuous function that satisfies the Partial Lipchitz Estimate:

$$||\Phi(\vec{x}_2, y) - \Phi(\vec{x}_1, y)|| \leq M ||\vec{x}_2 - \vec{x}_1||$$

with some $M \in (0, \infty)$, for all $\vec{x}_2, \vec{x}_1 \in V$ and $y \in I$. Then there exists $\tau > 0$ and a function $F : B_\tau(\vec{a}) \times (t_0 - \tau, t_0 + \tau) \rightarrow V$ uniquely determined by the following conditions:

$$D_{n+1}F(\vec{z}, t) = \Phi(F(\vec{z}, t), t) \quad F(\vec{z}, t_0) = \vec{z}$$

Definition 34.1.0.0.1

For 0-forms f_1 and f_2 defined on an open subset of \mathbb{R} , we have:

$$f_1 df_2 - f_2 df_1 = (f_1 f'_2 - f'_1 f_2) dx$$

Here $(f_1 f'_2 - f'_1 f_2)$ is called the Wronskian of f_1 and f_2 , denoted as $\mathcal{W}(f_1, f_2)$.

Lemma 34.1.1

If the 0-forms f_1, f_2 are linearly dependent, then $\mathcal{W}(f_1, f_2) = 0$, while $\mathcal{W}(f_1, f_2) = 0$ does not imply f_1 and f_2 are linearly dependent in general. $\mathcal{W}(f_1, f_2) = 0$ implies f_1 and f_2 are linearly dependent in the cases where: (1) f_1, f_2 both solve the differential equation $f'' = a f' + b f$ with scalar-valued functions a and b defined on \mathbb{R} , and f_1, f_2 have no common zeros, or (2) f_1, f_2 are holomorphic if they were defined on some convex subset of \mathbb{C} .

Proof. Assume that A is the connected domain of definition for f_1 and f_2 , and we assume that f_1 is non zero on B , with B being nonempty connected subset of A . Then we can write:

$$d\left(\frac{f_2}{f_1}\right) = \frac{f_1 df_2 - f_2 df_1}{(f_1)^2}$$

Here we see that $f_1 df_2 - f_2 df_1 = 0$ on B if and only if f_2/f_1 is constant on B , if and only if $f_2 - cf_1 = 0$ on B with constant c , if and only if $f_2 - cf_1 = 0$ on A , if and only if f_2 and f_1 are linearly independent. \square

Retractions

Definition 35.0.0.0.1

Let A be a set, let B be a subset of A . A retraction from A to B is a function $R : A \rightarrow B$ that satisfies $R(x) = x$ for all $x \in B$.

Example: The function defined by the following is a retraction from \mathbb{H}^k to $\partial\mathbb{H}^k$:

$$R : \mathbb{H}^k \rightarrow \partial\mathbb{H}^k \quad (x_1, x_2, \dots, x_{k-1}, x_k) \mapsto (x_1, x_2, \dots, x_{k-1}, 0)$$

Lemma 35.0.1

Let M be a nonempty compact orientable k -manifold in \mathbb{R}^n with nonempty manifold boundary ∂M . There exists an $(k-1)$ -form of C^∞ type defined on \mathbb{R}^n such that $\int_{\partial M} \omega \neq 0$.

Proof. The lemma should follow from Math 396 HW6 Q1.

Lemma Y.1 on Math 396 Supplement martial also gives a proof. \square

Lemma 35.0.2

Let M be a k -manifold in \mathbb{R}^n with nonempty manifold boundary ∂M , and let ω be an $(k-1)$ -form defined on an open subset of \mathbb{R}^n containing M . If $R : M \rightarrow \partial M$ is a C^1 retraction, then ∂M is integral for the $(k-1)$ -form $\omega - R^*\omega$.

Proof. Let $\alpha : U \rightarrow V$ be a coordinate patch on ∂M , then we can write:

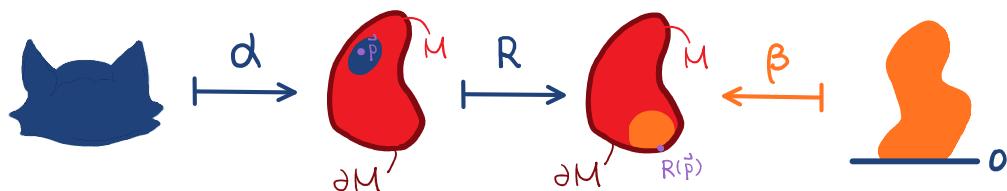
$$\alpha^*(\omega - R^*\omega) = \alpha^*\omega - \alpha^*R^*\omega = \alpha^*\omega - (R \circ \alpha)^*\omega = \alpha^*\omega - \alpha^*\omega = 0$$

The result follows. \square

Lemma 35.0.3

Let M be a k -manifold in \mathbb{R}^n with nonempty manifold boundary ∂M . Let $R : M \rightarrow \partial M$ be a function of C^1 type and let η be a k -form defined on an open subset of \mathbb{R}^n containing ∂M , then M is an integral for $R^*\eta$.

Proof. Here we want to show $\alpha^*R^*\eta = 0$ for coordinate patch α on M , and let β denote a regional coordinate patch on the codomain of $R \circ \alpha$, which is a subset of ∂M .



Here we can write:

$$\beta^*\eta = f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k$$

for some scalar-valued function f , hence we can write the following:

$$\begin{aligned}
 \alpha^* R^* \eta &= \alpha^* R^*(\beta^{-1})^* \beta^* \eta = \alpha^* R^*(\beta^{-1})^*(f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_k) \\
 &= \alpha^* R^*(\beta^{-1})^*(f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{k-1}) \wedge \alpha^* R^*(\beta^{-1})^* dx_k \\
 &= \alpha^* R^*(\beta^{-1})^*(f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{k-1}) \wedge d(\alpha^* R^*(\beta^{-1})^* x_k) \\
 &= \alpha^* R^*(\beta^{-1})^*(f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{k-1}) \wedge d(\alpha^* R^* 0) \\
 &= 0
 \end{aligned}$$

Then the result follows. \square

Theorem 35.1 (Non-retraction Theorem)

Let M be a nonempty compact orientable k -manifold in \mathbb{R}^n .

There is no retraction of C^1 type from M to ∂M .

Proof. For the case where $\partial M = \emptyset$, the result follows immediately. For the case where $\partial M \neq \emptyset$, we proceed by contradiction, suppose there exists a C^1 retraction $R : M \rightarrow \partial M$, then by the Stokes' Theorem, Lemma 35.0.2, Lemma 35.0.3, and Lemma 35.0.4, let ω be a $(k-1)$ -form defined on an open subset of \mathbb{R}^n containing ∂M such that $\int_{\partial M} \omega \neq 0$, we can write:

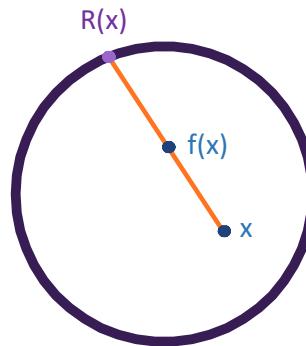
$$0 = \int_M R^* d\omega = \int_M d(R^* \omega) = \int_{\partial M} R^* \omega = \int_{\partial M} \omega \neq 0$$

Here we reach a contradiction, the result follows. \square

Theorem 35.2 (Brouwer Fixed Point Theorem)

Let $B^n := \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\| \leq 1\}$ be the closed unit ball in \mathbb{R}^n . If $f : B^n \rightarrow B^n$ is a function of C^1 type, then there exists $\vec{x} \in B^n$ such that $f(\vec{x}) = \vec{x}$, and such \vec{x} is called a fixed point of f .

Proof. Let $U = \{\vec{x} \in B^n \mid f(\vec{x}) \neq \vec{x}\}$. Connecting \vec{x} and $f(\vec{x})$ for $\vec{x} \in U$ through a line, which intersects ∂B^n at two points, pick one of the two, and get C^1 retraction R from U to ∂B^n .



Math 396 HW9 proposes a formula for such function R , and one can check that R is a C^1 retraction. Non-retraction theorem implies that $U \subsetneq B^n$, and $\vec{x} \in B^n \setminus U$ is given fixed point of f . \square

Remark: The result of Theorem 35.2 also holds for if the function f defined is continuous.

Remark: For the case where $n = 1$, Spivak Chapter 7 gives a proof.

Remark: Contraction mapping theorem is also a fixed point theorem.

Example: For discrete time dynamical system: given $f : X \rightarrow X$ defined similarly as in Theorem 35.2, one can construct a sequence $(x_1, f(x_1), f(f(x_1)), \dots)$ with $x_1 \in X$, the fixed point of f is the equilibrium. Denote $f_i = f$, define $f^{\circ k} := f_1 \circ f_2 \circ \dots \circ f_k$, we get a semi-group homomorphism:

$$H : \mathbb{N} \mapsto \{\text{self map of } f : X \rightarrow X\} \quad k \mapsto f^{\circ k}$$

Equilibrium and Flow of Vector Fields

Definition 36.0.0.0.1

Let M be a k -manifold in \mathbb{R}^n , and let v be a vector field in \mathbb{R}^n . For $\vec{p} \in M$, \vec{p} is called fixed point, or a equilibrium point, on M associated to v provided that $v(\vec{p}) = 0$.

Let v be a vector field in \mathbb{R}^n satisfying the Lipchitz estimate $\|v(\vec{x}_2) - v(\vec{x}_1)\| \leq M\|\vec{x}_2 - \vec{x}_1\|$ for some $M > 0$, and all $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$. By the Existence and Uniqueness Theorem, one can solve the following ODE for $f : A \rightarrow \mathbb{R}^n$ uniquely, where A is a subset of \mathbb{R} :

$$f'(t) = v(f(t)) \quad f(t_0) = \vec{z}$$

This result is related to the \mathbb{R} -action on \mathbb{R}^n . Equilibrium points for f are points where v vanishes, and this result suggests the theorem, while the proof of the theorem is left for the reader:

Theorem 36.1

Let B^n denote the closed unit ball in \mathbb{R}^n . If \vec{v} points inwards at all points \vec{p} on the boundary ∂B^n , then there is an equilibrium point in B^n .

Now consider a vector field $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfying the Lipchitz Estimate, that is, for $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$, we can write the following:

$$\|v(\vec{x}_2) - v(\vec{x}_1)\| \leq M\|\vec{x}_2 - \vec{x}_1\|$$

for $M \in (0, \infty)$. By Theorem 34.1, one can solve the following system with a unique function f :

$$f'(t) = v(f(t)) \quad f(t_0) = \vec{z} \tag{*}$$

with f defined on $(t_0 - \frac{1}{19M}, t_0 + \frac{1}{19M})$, one can check that, by manipulating the constant 19 in the proof of Lemma 34.0.1, that f is well defined on $(t_0 - \frac{1}{2M}, t_0 + \frac{1}{2M})$. Now we can also solve the following system with some function \tilde{f} :

$$\tilde{f}'(t) = v(\tilde{f}(t)) \quad \tilde{f}\left(t_0 + \frac{1}{4m}\right) = f\left(t_0 + \frac{1}{4m}\right)$$

where \tilde{f} is defined on $t \in (t_0 - \frac{1}{4M}, t_0 + \frac{3}{4M})$. By the uniqueness of f and \tilde{f} , we can write:

$$\tilde{f}(t) = f(t) \quad \forall t \in \left(t_0 - \frac{1}{4m}, t_0 + \frac{1}{2m}\right)$$

By redefining f on $(t_0 - \frac{1}{2m}, t_0 + \frac{3}{4m})$, here we get a solution for the initial value problem (*) on $(t_0 - \frac{1}{2m}, t_0 + \frac{3}{4m})$. Continue going forward and backward on the definition of domain of f , one can obtain a solution for (*) on all of \mathbb{R} , and get a function $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ defined for $\vec{z} \in \mathbb{R}^n$, $t \in \mathbb{R}$, such that the following holds:

$$D_{n+1}F(\vec{z}, t) = v(F(\vec{z}, t)) \quad F(\vec{z}, 0) = \vec{z}$$

Suppose further that $v \in C^2$, by Math 396 HW8 Q4, F is of C^1 type. More generally, if $v \in C^r$, one can show that $F \in C^r$, and if $v \in C^\infty$, then $F \in C^\infty$.

Now we denote $F^s(\vec{z}, t) := F(\vec{z}, t + s)$, then one can get the following:

$$D_{n+1}F^s(\vec{z}, t) = v(F^s(\vec{z}, t)) \quad F^s(\vec{z}, 0) = F(\vec{z}, s)$$

so we get $F(\vec{z}, t + s) = F^s(\vec{z}, t) = F(F(\vec{z}, s), t)$. Then we can define:

$$g^t : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \vec{z} \mapsto F(\vec{z}, t)$$

with the property $g^{t+s}(\vec{z}) = g^t(g^s(\vec{z}))$ for $\vec{z} \in \mathbb{R}^n$, $t, s \in \mathbb{R}$, and g^0 being the identity transformation.

Denote $\text{Diffeo}(\mathbb{R}^n) := \text{Diffeo}(\mathbb{R}^n, \mathbb{R}^n) := \{\text{Diffeomorphisms from } \mathbb{R}^n \text{ to } \mathbb{R}^n\}$. We define the function:

$$H : \mathbb{R} \rightarrow \text{Diffeo}(\mathbb{R}^n) \quad t \mapsto g^t$$

Here H is a group homomorphism from \mathbb{R} to $\text{Diffeo}(\mathbb{R}^n)$.

Conversely, one can show that, given a group homomorphism $H : \mathbb{R} \rightarrow \text{Diffeo}(\mathbb{R}^n)$ $t \mapsto g^t$, we get a vector field v defined on \mathbb{R}^n :

$$v : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \vec{x} \mapsto (D_{n+1}(g^t(\vec{x})))|_{t=0}$$

Concluding from above, let v be a C^∞ Partially Lipchitz vector field defined on \mathbb{R}^n , the definition of \vec{v} leads to a group homomorphism $H : \mathbb{R} \rightarrow \text{Diffeo}(\mathbb{R}^n)$ $t \mapsto g^t$ with $D_{n+1}(g^t(\vec{x})) = v(g^t(\vec{x}))$. Here, sometimes we denote $D_{n+1}F(\vec{z}, t)$ as $D_t F(\vec{z}, t)$.

A Lipschitz vector field $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ of C^2 type induces a unique group homomorphism from \mathbb{R} to $\text{Diff}_{C^1}(\mathbb{R}^n) := \{\text{Diffeomorphisms from } \mathbb{R}^n \text{ to } \mathbb{R}^n \text{ of } C^1 \text{ type}\}$, defined by the following:

$$H : \mathbb{R} \rightarrow \text{Diffeo}_{C^1}(\mathbb{R}^n) \quad t \mapsto g^t$$

such that $D_t(g^t(\vec{x}))|_{t=0} = v(\vec{x})$. Here H is called the flow of v .

Note: A function is Lipschitz and of C^2 type if and only if the function is of C^2 type with bounded first derivative.

Theorem 36.2 (Rectification Theorem)

Let $\vec{x}_0 \in A$ where A is an open subset of \mathbb{R}^n , let v be a vector field of C^∞ type defined on A , with $v(\vec{x}_0) \neq \vec{0}$. Then there exists an open neighborhood U of \vec{x}_0 contained in A , and a C^∞ -diffeomorphism $\alpha : U \rightarrow V$ such that $D\alpha(\vec{x})(v(\vec{x})) = \vec{e}_1$, where $\vec{e}_1 = (1, 0, 0, \dots, 0)$ is the first element in the standard basis of \mathbb{R}^n .

Note: Stated differently, Theorem 35.5 suggests that, given $\vec{x}_0 \in A$ where A is an open subset of \mathbb{R}^n , let v be a vector field of C^{r+1} type defined on A with $v(\vec{x}_0) \neq \vec{0}$, then there exists a C^r -diffeomorphism α such that $D\alpha(\vec{x}) \cdot \vec{e}_1 = v(\alpha(\vec{x}))$.

Theorem 36.3

Let M be a closed k -manifold without boundary in \mathbb{R}^n of C^∞ class, let v be a vector field that is tangent to M at all $\vec{p} \in M$, that is, we have $v(\vec{p}) \in \mathcal{T}_{\vec{p}}(M)$ for all $\vec{p} \in M$. Then each $g^t|_M$ induced by v belongs to $\text{Diffeo}(M)$, where $\text{Diffeo}(M) := \text{Diffeo}(M, M)$.¹

Proof. It suffices to show that $g^t(M) \subseteq M$, then $(g^t)^{-1} := g^{-t}$ also map M to M . Let $\vec{x} \in M$, let $E := \{t \in \mathbb{R} \mid g^t(\vec{x}) \subseteq M\}$, we want to show that $E = \mathbb{R}$. Note that we have $0 \in E$ and E is closed as M is closed. It suffices to show that E is open. For $\vec{x} \in M$, consider two cases, case (1) $v(\vec{x}) = 0$, and case (2) $v(\vec{x}) \neq 0$. For case (1), since $v(\vec{x}) = 0$, then $g^t(\vec{x}) = \vec{x}$, and so for all $t \in \mathbb{R}$, $g^t(\vec{x}) \in M$. For case (2), assume that $v(\vec{x}) \neq 0$, after the change of coordinate, locally, it suffices to consider the special case where we have $v(\vec{x}) = \vec{e}_1$, in which case we get $g^t(\vec{x}) = \vec{x} + t\vec{e}_1$ for small enough t . Note that M in such case, by Theorem 16.1, is equivalently given by $\text{Graph}(f)$, where $f \in C^1(U, \mathbb{R}^{n-k})$ with U being convex open subset of \mathbb{R}^k . By assumption, $v(\vec{x}) = \vec{e}_1$ is tangent to M at \vec{x} , so we must also have $D_1 f = 0$, so f is independent of q_1 at $\vec{q} = (q_1, q_2, \dots, q_k) \in U$ such that $\vec{x} = (\vec{q}, f(\vec{q}))$. Since U is convex, this implies that, if $\vec{x} + t_0 \vec{e}_1 \in M$, then $\vec{x} + t \vec{e}_1 \in M$ for $|t - t_0| < \epsilon$ for some $\epsilon > 0$, then we see that E is open, which completes the proof. \square

Corollary 36.3.1

Let M be a compact k -manifold without boundary in \mathbb{R}^n of C^∞ class, and let U be an open subset of \mathbb{R}^n containing M . For vector field $v \in C^2(U, \mathbb{R}^n)$ such that v is tangent to M , each $g^t|_M$ induced by v belongs to $\text{Diffeo}(M)$, and there exists Lipschitz $\hat{v} \in C^2(\mathbb{R}^n, \mathbb{R}^n)$ such that $\hat{v} = v$ on some neighborhood of M , here the flow for \hat{v} along M is also a flow for v along M .

Proof. Use a partition of unity on \mathbb{R}^n , $\{\phi_j\}$, such that $\text{supp}(\phi_j) \cap M \neq \emptyset$ if and only if $j \leq N$, let $\hat{v} = \sum_{j=1}^N (\phi_j v)$. One can check that \hat{v} has bounded first derivative. \square

¹Given a tangent vector field on a differentiable manifold X . The flow of the vector field is the group of diffeomorphisms of X that lets the points of the manifold flow along the vector field, hence which sends them along flow lines, or the integral curves, that are tangent to the vector field.

Example:

$v : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ $(x, y) \mapsto (y, -x)$ defines a non vanishing vector field tangent to S^1 .

Example:

$v : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ $(x, y, a, b) \mapsto (y, -x, b, -a)$ defines non vanishing vector field tangent to S^3 .

Example:

$v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ $(x_1, x_2, x_3) \mapsto (x_2, x_1, 0)$ defines non vanishing vector field tangent to torus T .

Theorem 36.4 (Hairy Billiard Ball Theorem)

Any tangential vector field of C^1 type defined on an even-dimensional sphere S^{2n} vanishes at some point \vec{p} in S^{2n} . That is, "you cannot comb a hairy billiard ball without creating a cowlick."

Proof. We proceed by contradiction, suppose we do have $v : S^{2n} \rightarrow \mathbb{R}^{2n+1} \setminus \{0\}$ with $\langle v(\vec{p}), \vec{p} \rangle = 0$ for all $\vec{p} \in S^{2n}$, then we can define the following function:

$$h : \mathbb{R}^{2n+1} \setminus \{0\} \rightarrow \mathbb{R}^{2n+1} \quad \vec{p} \mapsto \cos(\|\vec{p}\| - 1) \frac{\vec{p}}{\|\vec{p}\|} + \sin(\|\vec{p}\| - 1) \frac{v(\vec{p}/\|\vec{p}\|)}{\|v(\vec{p}/\|\vec{p}\|)\|}$$

one can show that the image of h belongs to S^{2n} . Moreover, for $\|\vec{p}\| = 1$, we can write $h(\vec{p}) = \vec{p}$, and for $\|\vec{p}\| = \pi + 1$, we can write $h(\vec{p}) = -\vec{p}/(\pi + 1)$. By Lemma 35.0.1, one can pick $(2n)$ -form ω with $\int_{S^{2n}} \omega \neq 0$. By Lemma Y.2, Y.3, and Y.4 on Math 396 Supplement Materials, we see that S^{2n} is an integral for $d\omega$, and h defines a smooth function from $\mathbb{R}^{2n+1} \setminus \{0\}$ to S^{2n} , then we have:

$$0 = \int_{1 \leq \|\vec{x}\| \leq \pi+1} h^* d\omega$$

and hence we can write the following:

$$0 = \int_{1 \leq \|\vec{x}\| \leq \pi+1} h^* d\omega = \int_{1 \leq \|\vec{x}\| \leq \pi+1} d(h^* \omega) = \int_{\|\vec{x}\|=\pi+1} h^* \omega - \int_{\|\vec{x}\|=1} h^* \omega$$

Here the diffeomorphism between $(2n)$ -sphere of radius 1 and $\pi + 1$ is induced from the diffeomorphism between $(2n + 1)$ -ball of radius 1 and of radius $\pi + 1$:

$$g : \{\|\vec{x}\| \leq \pi + 1\} \rightarrow \{\|\vec{x}\| \leq 1\} \quad \vec{x} \mapsto -\frac{\vec{x}}{\pi + 1}$$

Note that g is orientation reversing, hence we have:

$$0 = \int_{\|\vec{x}\|=\pi+1} h^* \omega - \int_{\|\vec{x}\|=1} h^* \omega = - \int_{S^{2n}} \omega - \int_{S^{2n}} \omega \neq 0$$

Here we reach a contradiction. \square

Corollary 36.4.1

S^2 is not diffeomorphic to the torus in \mathbb{R}^2 .

Definition 36.4.1.0.1

Let M be a compact naturally oriented n -manifold in \mathbb{R}^n , let v be a vector field defined on a neighborhood of M , and let \vec{N} denote the unit normal field corresponds to the induced orientation of ∂M . v is said to be inward-pointing to M provided that the following holds for all $\vec{p} \in \partial M$:

$$\langle v(\vec{p}), \vec{N}(\vec{p}) \rangle \leq 0$$

Definition 36.4.1.0.2

Let X be a subset of \mathbb{R}^n . A function $h : [0, \infty) \rightarrow C^1(X, X)$ is said to be a monoid homomorphism provided that we have $h(t) \circ h(s) = h(t+s)$ and $h(0) = Id$, where Id is the identity transformation.

Theorem 36.5

Let M be a compact n -manifold in \mathbb{R}^n , let v be a vector field defined on an open subset of \mathbb{R}^n containing M that is of C^2 type and inward-pointing to M . The function defined by the following is a monoid homomorphism:

$$h : [0, \infty) \rightarrow C^1(M, M) \quad t \mapsto g^t$$

where we have $(D_t(g^t(\vec{x})))|_{t=0} = v(\vec{x})$

Suppose we have an inward-pointing v of C^1 type on B^n . it is easy to show that each g^t induced by v has at least one fixed point, a natural question arises that whether or not all g^t in $\{g^t\}$ have a common fixed point, which is the point where v vanishes. The answer to this question is yes, as proposed in the following theorem:

Theorem 36.6

Let v be a C^1 inward-pointing vector field on B^n , then v must vanish some point on B^n .

Proof. By Math 396 HW Study Exercise, there exists some $\epsilon > 0$ such that the image of the function $h : B^n \rightarrow \mathbb{R}^n \vec{x} \mapsto \vec{x} + \epsilon v(\vec{x})$ is a subset of B^n . Now by Brouwer Fixed Point Theorem, there exists $\vec{x} \in B^n$ such that $h(\vec{x}) = \vec{x}$, which suggests that $v(\vec{x}) = \vec{0}$, which completes the proof. \square

Theorem 36.7 (Poincare-Bendixson Theorem)

Let v be a inward-pointing vector field on a compact 2-manifold M in \mathbb{R}^2 . Then we have either one of the two holds: (1) v has a equilibrium point in M , or (2) there exists $t_0 > 0$ such that g^{t_0} has a fixed point in M .

8 | Complex Analysis II

Residues and Contour Integrals

Theorem 37.1 (Cauchy Integral Theorem)

Let M be a naturally oriented compact 2-manifold in \mathbb{C} , let $f \in C^1(M, \mathbb{C})$ be holomorphic on $M \setminus \partial M$, and let ∂M obtain the induced orientation from M . then we have:

$$\int_{\partial M} f dz = 0$$

Proof. The result of this theorem follows immediately from Theorem 32.1, the Stokes' Theorem. \square

Corollary 37.1.1

Let M be a naturally oriented compact 2-manifold in \mathbb{C} , let $f \in C^1(M, \mathbb{C})$ be holomorphic on $M \setminus \partial M$. If ∂M is the disjoint union of two 1-manifolds C_1 and C_2 , each with the induced orientation from M , with C_2 lying in the bounded component of $\mathbb{C} \setminus C_1$, then we have the following holds:

$$\int_{C_1} f dz = \int_{C_2} f dz$$

Proof. This follows immediately from Theorem 37.1. Note that the orientation of ∂M is obtained from the standard orientation of M and the orientation of disjoint circles are opposite. \square

Theorem 37.2 (Residue Theorem)

Let M be a compact 2-manifold in \mathbb{C} , let $E = \{z_1, z_2, \dots, z_k\} \subseteq M \setminus \partial M$, let $f \in C^1(M \setminus E, \mathbb{C})$ be holomorphic on $M \setminus (\partial M \cup E)$, then we have the following holds:

$$\int_{\partial M} f dz = 2\pi i \sum_{j=1}^k \text{Res}(f dz, z_j)$$

Proof. For each z_j , one can choose a circle C_j centered at z_j with radius small enough such that the circle is contained in $M \setminus (\partial M \cup E)$, now apply Cauchy Integral Theorem, we get:

$$0 = \int_{\partial M} f dz - \sum_j \int_{C_j} f dz$$

The result follows. \square

Let Ω be a connected open subset of \mathbb{C} , let f, g be holomorphic on Ω that are not identically zero. By results from Chapter 5, $h = \frac{f}{g}$ is holomorphic on $\Omega \setminus g^{-1}(0)$, and $g^{-1}(0)$ is a discrete set, that is, it has no limit points in Ω . By easy computation we get the following:

$$\frac{h'_{\mathbb{C}}}{h} = \frac{f'_{\mathbb{C}}}{f} - \frac{g'_{\mathbb{C}}}{g}$$

Remark: For holomorphic function f , the derivative of f in complex sense, as previously denoted as $f'_{\mathbb{C}}$, might be denoted as f' thereafter.

By Math 396 HW4 Q2, we also have the following:

$$\operatorname{Res}\left(\frac{h'}{h} dz, z_0\right) = \operatorname{Res}\left(\frac{f'}{f} dz, z_0\right) - \operatorname{Res}\left(\frac{g'}{g} dz, z_0\right) = \operatorname{ord}_{z_0} f - \operatorname{ord}_{z_0} g$$

Now consider a compact 2-manifold $M \subseteq \Omega$ with $\partial M \cap (f^{-1}(0) \cup g^{-1}(0)) = \emptyset$, then we have:

$$\begin{aligned} \frac{1}{2\pi i} \int_{\partial M} \frac{h'}{h} dz &= \sum_{z_0 \in M \setminus \partial M} \operatorname{Res}\left(\frac{h'}{h} dz, z_0\right) \\ &= \sum_{z_0 \in M \setminus \partial M} \operatorname{Res}\left(\frac{f'}{f} dz, z_0\right) - \sum_{z_0 \in M \setminus \partial M} \operatorname{Res}\left(\frac{g'}{g} dz, z_0\right) \\ &= \sum_{z_0 \in M \setminus \partial M} \operatorname{ord}(f, z_0) - \sum_{z_0 \in M \setminus \partial M} \operatorname{ord}(g, z_0) \end{aligned} \quad (\text{SMS})$$

Note that since M is compact, then each of the sums in equation (SMS) is finite. We count the zeros of a holomorphic functions according to multiplicity, that is, a zero of order m counts as m zeros, then we can write:

$$\frac{1}{2\pi i} \int_{\partial M} \frac{h'}{h} dz = \#\{\text{zeros of } f \text{ in } M\} - \#\{\text{zeros of } g \text{ in } M\} \quad (\text{ARP})$$

Equation (ARP) is known as the argument principle.

Lemma 37.2.1

Let Ω be a connected open subset of \mathbb{C} , let M be a compact 2-manifold in \mathbb{C} , let f, g be holomorphic on Ω that are not identically zero, and let $h = \frac{f}{g}$ defined on $\Omega \setminus g^{-1}(0)$. We have the following holds, with some $N \in \mathbb{N}$:

$$\partial M = \bigcup_{j=1}^N C_j$$

where each C_j is diffeomorphic to $[0, 1]$, the C_j overlap only at the endpoints, and there exists a holomorphic function \log_{X_j} , with associated \arg_{X_j} on each neighborhood of C_j solving $\exp(\log_{X_j}(h(x))) = h(x)$ for x in the domain of definition of h .¹

¹ $\log_{X_j} : h(C_j) \rightarrow \mathbb{C}$ $z \mapsto \log|z| + i\arg_{X_j}(z)$, where X_j is a ray starting at $\vec{0} \in \mathbb{C}$, as defined in Chapter 8.

Proof. The proof of this lemma is left for exercise. \square

In the settings of Lemma 37.2.1, and with discussion above, we can write the following:

$$\begin{aligned}
\#\{\text{zeros of } f \text{ in } M\} - \#\{\text{zeros of } g \text{ in } M\} &= \frac{1}{2\pi i} \int_{\partial M} \frac{h'}{h} dz \\
&= \frac{1}{2\pi i} \sum_{j=1}^N \int_{C_j} \frac{h'}{h} dz \\
&= \frac{1}{2\pi i} \sum_{j=1}^N \int_{C_j} d(\log_{X_j} h) \\
&= \frac{1}{2\pi i} \sum_{j=1}^N \Delta_{C_j} \log_{X_j} h \\
&= \frac{1}{2\pi i} \sum_{j=1}^N \Delta_{C_j} \log |h| + \frac{1}{2\pi} \sum_{j=1}^N \Delta_{C_j} \arg_{X_j} h \\
&= \frac{1}{2\pi} \sum_{j=1}^N \Delta_{C_j} \arg_{X_j} h
\end{aligned}$$

Now we suppose further that $h(\partial M) \subseteq V$, where V is an open subset of $\mathbb{C} \setminus \{0\}$ being diffeomorphic to convex set, then there exists a holomorphic function $\log_X : V \rightarrow \mathbb{C}$ $z \mapsto \log|z| + i \arg_X(z)$ that is exact on V , so we get:

$$\#\{\text{zeros of } f \text{ in } M\} = \#\{\text{zeros of } g \text{ in } M\}$$

Theorem 37.3 (Rouche's Theorem)

Let M be a compact 2-manifold in Ω , where Ω is an open subset of \mathbb{C} , let f, h be holomorphic functions on Ω , with $|h(x)| < |f(x)|$ for $x \in \partial M$, then the number of zeros of $f + h$ in M is equal to the number of zeros of f in M .

Proof. Here we can write:

$$h + f = f \left(1 + \frac{h}{f} \right)$$

Note that $1 + \frac{h}{f}$ maps ∂M into $\{w \in \mathbb{C} \mid \|w - 1\| < 1\} \subseteq \{w \in \mathbb{C} \mid \Re(w) > 0\}$ because $|h(x)| < |f(x)|$ for $x \in \partial M$. Here $\{w \in \mathbb{C} \mid \|w - 1\| < 1\}$ is diffeomorphic to a convex set, then by Lemma 37.2.1 and discussion above, we can write the following with some $N \in \mathbb{N}$, C_j and \arg_{X_j} being defined as in Lemma 37.2.1 and discussion above:

$$\begin{aligned}
\#\{\text{zeros of } h + f \text{ in } M\} &= \frac{1}{2\pi} \sum_{j=1}^N \Delta_{C_j} \arg_{X_j} \left(f \left(1 + \frac{h}{f} \right) \right) \\
&= \frac{1}{2\pi} \sum_{j=1}^N \Delta_{C_j} \arg_{X_j}(f) = \#\{\text{zeros of } f \text{ in } M\}
\end{aligned}$$

The result of the theorem follows. \square

Consider the functions $f : \mathbb{C} \rightarrow \mathbb{C}$ $z \mapsto z^n$, and $h : \mathbb{C} \rightarrow \mathbb{C}$ $z \mapsto a_0 + a_1z + a_2z^2 + \cdots + a_{n-1}z^{n-1}$ with nonzero a_0, a_1, \dots, a_{n-1} , and consider 2-manifold $M_R\{z \in \mathbb{C} \mid |z| \leq R\}$ in \mathbb{C} , with some $R > 0$. It is trivial to show that there exists some $R > 0$ such that $|h(x)| < |f(x)|$ for $x \in \partial M_R$, and $f + h$ has no zero in $\mathbb{C} \setminus M_R$, and we have the following holds:

$$\#\{\text{zeros of } h + f \text{ in } M_R\} = \#\{\text{zeros of } f \text{ in } M_R\} = n$$

Note: For $z = re^{i\theta} \neq 0$ with $r, \theta \in \mathbb{R}$, we write $\arg(z) := \{\theta + 2\pi k \mid k \in \mathbb{Z}\}$.

Note: For $z_1, z_2 \in \mathbb{C}$ with $z_1 \neq \vec{0} \neq z_2$, we have $\arg(z_1 z_2) = \arg(z_1) + \arg(z_2)$ and $\arg(z_1^2) = 2\arg(z_1)$.

Example: The number of roots of $z^9 + z^5 - 8z^3 + 2z + 1$ in the set $\{z \in \mathbb{C} \mid 1 < |z| < 2\}$ is 6. Apply argument from the discussion above, one can find that there are 9 roots of the equation have absolute value less than 2, and 3 of the 9 roots of the equation have absolute value less than 1, hence there are 6 roots of the equation lie in the set $\{z \in \mathbb{C} \mid 1 < |z| < 2\}$.

Fourier Transformation

Definition 38.0.0.0.1

For nonzero $z = re^{i\theta} \in \mathbb{C}$ with $r \in \mathbb{R}_{>0}$ and $\theta \in \mathbb{R}$. We define the followings:

$$\arg(z) = \{\theta + 2k\pi \mid k \in \mathbb{Z}\} \quad \log(z) := \log|z| + i\arg(z) = \{\log|z| + \theta + 2k\pi i \mid k \in \mathbb{Z}\}$$

Here we will show that $\log(z)$ for nonzero $z \in \mathbb{C}$ is multiple-valued. We proceed by contradiction, suppose we had a single-valued holomorphic function \log defined on $\mathbb{C} \setminus \{0\}$, then we have:

$$\int_M d\log z = \Delta_M \log(z) = 0$$

on the closed path $M = \{|z| = 1\}$, but we also have:

$$\int_M \frac{dz}{z} = 2\pi i$$

Hence such single-valued holomorphic function \log is not well defined.

Let $z, w \in \mathbb{C}$ with $z \neq 0 \neq w$, we also get a set of values $z^w = (e^{\log(z)})^w = e^{w\log(z)}$, where we write:

$$z^w = \{e^{w\log|z| + iw\theta + iw2\pi k} \mid k \in \mathbb{Z}\} = \{e^{w\log|z| + iw\theta} e^{iw2\pi k} \mid k \in \mathbb{Z}\}$$

Here we have four cases:

1. If w is an integer, then $e^{iw2\pi k} = 1$ for all k , we get only one value of z^w .
2. If $w \in \mathbb{Q} \setminus \mathbb{Z}$, then we can write $w = \frac{m}{n}$ for $m, n \in \mathbb{Z}$ in the lowest term. In this case $\{e^{iw2\pi k} \mid k \in \mathbb{Z}\}$ consists of n evenly-spread points on the unit circle, so z^w has n values.
3. If $w \in \mathbb{R} \setminus \mathbb{Q}$, one can show that $\{e^{iw2\pi k} \mid k \in \mathbb{Z}\}$ is a countable dense subset of the unit circle.
4. If we have $w \in \mathbb{C} \setminus \mathbb{R}$, one can show that, if $\Im(w) > 0$, we get a sequence of values of z^w that tends to 0 as k tends ∞ , and $|z^w|$ tends to ∞ as k tends to $-\infty$. In the case where $\Im(w) < 0$, we get the reverse.

Definition 38.0.0.0.2

Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be a bounded function such that the set $\{x \in \mathbb{R} \mid f \text{ is discontinuous at } x\}$ has measure zero, the condition $\text{ext} \int_{\mathbb{R}} |f| < \infty$ holds if and only if all of the followings hold:

$$\text{ext} \int_{\mathbb{R}} (\Re(f))_+ < \infty \quad \text{ext} \int_{\mathbb{R}} (\Re(f))_- < \infty \quad \text{ext} \int_{\mathbb{R}} (\Im(f))_+ < \infty \quad \text{ext} \int_{\mathbb{R}} (\Im(f))_- < \infty$$

Definition 38.0.0.0.3

Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be a bounded function such that the set $\{x \in \mathbb{R} \mid f \text{ is discontinuous at } x\}$ has measure zero, and $\text{ext} \int_{\mathbb{R}} |f| < \infty$. The Fourier Transform of f , denoted as \hat{f} , is the function defined by:

$$\hat{f} : \mathbb{R} \rightarrow \mathbb{R} \quad t \mapsto \text{ext} \int_{\mathbb{R}} f(x) e^{-ixt} dx$$

Remark: In the settings of Definition 38.0.0.0.3, one can check that \hat{f} is well defined and finite, and \hat{f} can be used for comparison test for integrals.

Example:

Consider the function $f : \mathbb{R} \rightarrow \mathbb{C}$ $x \mapsto e^{-x^2}$, the Fourier transform of f is given by the following:

$$\hat{f} : \mathbb{R} \rightarrow \mathbb{R} \quad t \mapsto \int_{\mathbb{R}} e^{-x^2} e^{-ixt} dx$$

where we can evaluate $\hat{f}(t)$ for $t \in \mathbb{R}$:

$$\hat{f}(t) = \int e^{-x^2} e^{-ixt} dx = \int e^{-x^2} \cos(xt) dx - i \int e^{-x^2} \sin(xt) dx = \sqrt{\pi} e^{-t^2/4} + 0 = \sqrt{\pi} e^{-t^2/4}$$

Example:

Consider the function f defined by the following:

$$f : \mathbb{R} \rightarrow \mathbb{C} \quad x \mapsto \frac{1}{(x - \alpha)(x - \beta)}$$

with $\alpha, \beta \in \mathbb{C}$, $\Im(\alpha) < 0$, $\Im(\beta) < 0$, and $\alpha \neq \beta$.

We are interested in finding the Fourier transform of f , denoted as $\hat{f} : \mathbb{R} \rightarrow \mathbb{R}$. To evaluate $\hat{f}(t)$ for $t \in \mathbb{R}$, one can extend the definition of domain of f to \mathbb{C} so that we can make use of Contour integral in the complex field. Here we redefine f as the following:

$$f : \mathbb{C} \rightarrow \mathbb{C} \quad x \mapsto \frac{1}{(x - \alpha)(x - \beta)}$$

Note that with $t \in \mathbb{R}$ and $z = x + iy$, we have:

$$f(z) \cdot e^{-izt} = \frac{e^{yt-ixt}}{(z - \alpha)(z - \beta)}$$

We see that we get the following with some constants C and M :

$$\|f(z) \cdot e^{-izt}\| = \left\| \frac{e^{-ixt+yt}}{z^2 - z(\beta + \alpha) + \alpha\beta} \right\| \leq \frac{C \|e^{ixt}\| \cdot \|e^{yt}\|}{\|z^2\|} = \frac{Ce^{yt}}{\|z^2\|} \quad \text{for } |z| \geq M$$

For $t < 0$, let $Y_{up, R}$ denote the arc of the upper semi-circle in \mathbb{C} of radius $R > M$.

Note that for $z \in Y_{up, R}$, we can write the following:

$$\|f(z) \cdot e^{-izt}\| \leq \frac{Ce^{yt}}{R^2} \leq \frac{C}{R^2}$$

hence we see that, for all $z \in Y_{up, R}$, we have:

$$0 \leq \lim_{R \rightarrow \infty} \pi R \|f(z) \cdot e^{-izt}\| \leq \lim_{R \rightarrow \infty} \frac{C}{R} = 0$$

By ML-estimate and Cauchy Integral Theorem, for $t < 0$, and $L_R = \{z \in \mathbb{R} \mid r \leq R\}$, we have:

$$\begin{aligned} \hat{f}(t) &= \int_{\mathbb{R}} f(z) \cdot e^{-izt} dz \\ &= \lim_{R \rightarrow \infty} \int_{z \in L_R} f(z) e^{-izt} dz \\ &= \left(\lim_{R \rightarrow \infty} \int_{z \in (L_R \cup Y_{up, R})} f(z) e^{-izt} dz \right) - \left(\lim_{R \rightarrow \infty} \int_{z \in Y_{up, R}} f(z) e^{-izt} dz \right) \\ &= \left(\lim_{R \rightarrow \infty} \int_{z \in (L_R \cup Y_{up, R})} f(z) e^{-izt} dz \right) - \left(\lim_{R \rightarrow \infty} \pi R \cdot \sup \{ \|f(z) e^{-izt}\| \mid z \in Y_{up, R} \} \right) \\ &= 0 \end{aligned}$$

For $t \geq 0$, we proceed similarly but using the arc of the lower semi-circle in \mathbb{C} , the Residue Theorem, and ML-estimate. One can show that for $t \geq 0$ we get the following:

$$\hat{f}(t) = -2\pi i \left(\operatorname{Res} \left(\frac{e^{izt}}{(z - \alpha)(z - \beta)} dz, \alpha \right) + \operatorname{Res} \left(\frac{e^{izt}}{(z - \alpha)(z - \beta)} dz, \beta \right) \right) = 2\pi i \left(\frac{e^{-i\alpha t} - e^{-i\beta t}}{\beta - \alpha} \right)$$

Remark: For the function f defined in the two examples in section 38, we get the following:

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(t) \cdot e^{ixt} dt$$

Theorem 38.1 (Fourier Inversion Theorem)

Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be a bounded function such that the set $\{x \in \mathbb{R} \mid f \text{ is discontinuous at } x\}$ has measure zero, and $\text{ext} \int_{\mathbb{R}} |f| < \infty$. If $\int_{\mathbb{R}} |\hat{f}(t)| dt < \infty$, then we have:

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(t) \cdot e^{ixt} dt$$

The Cotangent Function

Definition 39.0.0.0.1

For $z \in \mathbb{C} \setminus \{w \mid \sin(w) = 0\}$, we define the following:²

$$\cot(z) = \frac{\cos(z)}{\sin(z)} = \frac{\sin'(z)}{\sin(z)}$$

Note: For $z = x + iy$, we can write the following:

$$\begin{aligned} \frac{e^{iz} - e^{-iz}}{2i} = \sin(z) = 0 &\iff e^{iz} = e^{-iz} \\ &\iff e^{2iz} = 1 \\ &\iff e^{2ix}e^{-2y} = 1 \\ &\iff y = 0, x \in \{k\pi \mid k \in \mathbb{Z}\} \\ &\iff z \in \{k\pi \mid k \in \mathbb{Z}\} \end{aligned}$$

Note: By Exercise V.7 on Math 396 Supplement Material, we can write:

$$\text{Res}(\cot(z) dz, n\pi) = \text{Res}\left(\frac{\sin'(z)}{\sin(z)} dz, n\pi\right) = (\text{order of } \sin \text{ at } n\pi) = 1$$

Exercise: For $n \neq 0$, we get the following:

$$\text{Res}\left(\frac{\cot(z)}{z^2} dz, n\pi\right) = \frac{1}{n^2\pi^2}$$

Example:

For $n = 0$, to calculate $\text{Res}(\cot(z)/z^2 dz, n\pi)$, we note that $\cos(z)$ is not well defined at $z = 0$, then by the Residue Theorem, we get the following:

$$\begin{aligned} \text{Res}\left(\frac{\cot(z)}{z^2} dz, 0\right) &= \frac{1}{2\pi i} \int_{|z|=1} \frac{\cot(z)}{z^2} dz \\ &= \lim_{\epsilon \rightarrow 0} \left(\frac{1}{2\pi i} \int_{|z|=1} \frac{\cot(z)}{z^2 - \epsilon^2} dz \right) \\ &= \lim_{\epsilon \rightarrow 0} \left(\text{Res}\left(\frac{\cot(z)}{z^2 - \epsilon^2}, \epsilon\right) + \text{Res}\left(\frac{\cot(z)}{z^2 - \epsilon^2}, -\epsilon\right) + \text{Res}\left(\frac{\cot(z)}{z^2 - \epsilon^2}, 0\right) \right) \end{aligned}$$

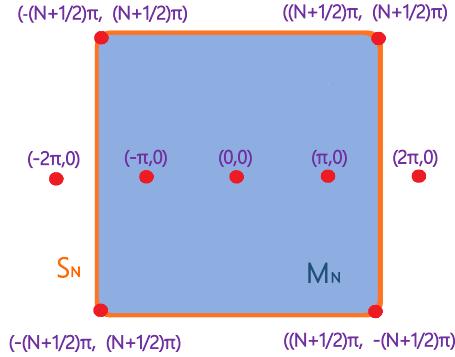
By Proposition U.11 and Exercise V.5 on Math 396 Supplement Materials, we get the

²Here we use \sin' to denote the derivative of \sin defined on \mathbb{C} .

following:

$$\begin{aligned}
 Res\left(\frac{\cot(z)}{z^2} dz, 0\right) &= \lim_{\epsilon \rightarrow 0} \left(Res\left(\frac{\cot(z)}{z^2 - \epsilon^2}, \epsilon\right) + Res\left(\frac{\cot(z)}{z^2 - \epsilon^2}, -\epsilon\right) + Res\left(\frac{\cot(z)}{z^2 - \epsilon^2}, 0\right) \right) \\
 &= \lim_{\epsilon \rightarrow 0} \left(\frac{\cos(\epsilon)}{2\epsilon \sin(\epsilon)} + \frac{\cos(-\epsilon)}{-2\epsilon \sin(-\epsilon)} - \frac{1}{\epsilon^2} \right) \\
 &= \lim_{\epsilon \rightarrow 0} \frac{\epsilon \cos(\epsilon) - \sin(\epsilon)}{\epsilon^2 \sin(\epsilon)} \\
 &= \lim_{\epsilon \rightarrow 0} \frac{\epsilon(1 - \frac{\epsilon^2}{2!} + \frac{\epsilon^4}{4!} + \dots) - (\epsilon - \frac{\epsilon^3}{3!} + \frac{\epsilon^5}{5!} + \dots)}{\epsilon^2(\epsilon - \frac{\epsilon^3}{3!} + \frac{\epsilon^5}{5!} + \dots)} \\
 &= \lim_{\epsilon \rightarrow 0} -\frac{\epsilon^3/3}{\epsilon^3} \frac{(1 - \frac{\epsilon^2}{10} + \frac{\epsilon^4}{280} + \dots)}{(1 - \frac{\epsilon^2}{3!} + \frac{\epsilon^4}{5!} + \dots)} \\
 &= -\frac{1}{3}
 \end{aligned}$$

Exercise:



Consider $M_N \in \mathbb{C}$ as described above. Note that for $z \in S_n := \text{Bd}(M_N)$, we have $|\cot(z)| \leq 2$. Here we round the corners of M_N to make it a 2-manifold. Note that we have $|\frac{\cot(z)}{z^2}| \leq \frac{2}{N^2}$ on S_N . Then applying ML-estimate, we get the following:

$$\left| \int_{S_N} \frac{\cot(z)}{z^2} dz \right| \leq \frac{2}{N^2} (\pi(8N + 4))$$

hence by Residue Theorem and ML-estimate, we have the following:

$$\begin{aligned} 0 &= \lim_{N \rightarrow \infty} \frac{1}{2\pi i} \int_{S_N} \frac{\cot(z)}{z^2} dz \\ &= \lim_{N \rightarrow \infty} \sum_{k=-N}^N \text{Res} \left(\frac{\cot(z)}{z^2} dz, k\pi \right) \\ &= \lim_{N \rightarrow \infty} \left(-\frac{1}{3} + 2 \sum_{k=1}^N \frac{1}{k^2 \pi^2} \right) \\ &= -\frac{1}{3} + 2 \sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} \end{aligned}$$

Rearranging we have:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$$

One can use the same technique as the example above to evaluate the following:

$$\sum_{k \in \mathbb{Z} \setminus F} \frac{P(k)}{Q(k)}$$

with P, Q being polynomials that satisfy $\deg(Q) \geq \deg(P) + 2$, and F is a finite subset of \mathbb{Z} .

One can show that we have the following holds:

$$\sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{1}{k^p} = \begin{cases} 2 \sum_{k=1}^{\infty} \frac{1}{k^p} & p \text{ is even} \\ 0 & p \text{ is odd} \end{cases}$$

Definition 39.0.0.2

For $p \in \mathbb{Z}$, $\zeta(p) = \sum_{k=1}^{\infty} \frac{1}{k^p}$ is the Riemann zeta function evaluated at p .

Theorem 39.1 (Apéry's Theorem)

$\zeta(3)$ is irrational.

- Roger Apéry (1978)

Theorem 39.2

At least one of $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational.

- W. V. Zudilin (2001)

Let $f : \mathbb{C} \rightarrow \mathbb{C}$ be a holomorphic function, let $z \in \mathbb{C}$, note here $(\zeta - z) \sin(\pi\zeta) = 0$ for $\zeta \in \mathbb{Z} \cup \{z\}$.

Applying the Residue Theorem and proceed similar to what we did above, we get the following:

$$\int_{\zeta \in M_N} \frac{f(\zeta) d\zeta}{(\zeta - z) \sin(\pi\zeta)} = 2\pi i \left(\frac{f(z)}{\sin(\pi z)} + \sum_{k=-N}^N \frac{(-1)^k f(k)}{\pi(k - z)} \right) \quad (\text{IN})$$

For some well-defined $M_N \subseteq \mathbb{C}$. If $|f(z)| \leq C e^{\sigma|\Im(z)|}$ for constant C and $0 < \sigma < \pi$, then *LHS* of equation (IN) tends to 0 when N tends to ∞ . Here we get:

$$\frac{f(z)}{\sin(\pi z)} = \sum_{k=-\infty}^{\infty} \frac{(-1)^{k+1} f(k)}{\pi(k - z)}$$

Rearranging we get:

$$f(z) = \sum_{k=-\infty}^{\infty} f(k) \frac{\sin(\pi(z - k))}{\pi(z - k)}$$

That is, we can explicitly recover f from $f|_{\mathbb{Z}}$, but we might not be able to do so if $|f(z)| \leq C e^{\sigma|\Im(z)|}$ fails. For details on information theory, check out Shannon's Theorem.

Winding Numbers

Definition 40.0.0.0.1

Let M be compact oriented 1-manifold without boundary in $\mathbb{R}^2 \simeq \mathbb{C}$. We define the following function:

$$\mathbb{W}_M : \mathbb{C} \setminus M \rightarrow \mathbb{C} \quad z \mapsto \frac{1}{2\pi i} \int_{\zeta \in M} \frac{d\zeta}{\zeta - z}$$

In the settings Definition 40.0.0.0.1, let M_1, M_2, \dots, M_n denote the components of M , one can show that we have:

$$\frac{1}{2\pi i} \int_{\zeta \in M_j} \frac{d\zeta}{\zeta - z} = \frac{1}{2\pi} (\Delta_{M_j} \widetilde{\arg}(\zeta - z)) \in \mathbb{Z}$$

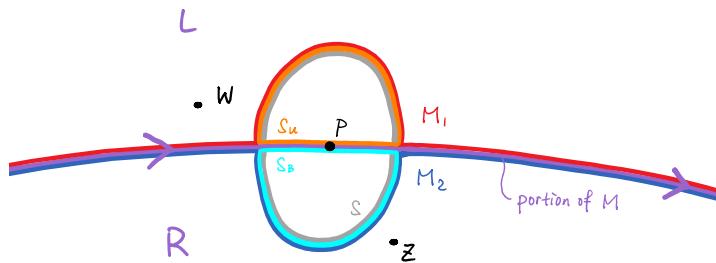
the term $\Delta_{M_j} \widetilde{\arg}(\zeta - z)$ measures the total angle of the path M_j winding around z , hence $\frac{1}{2\pi} \Delta_{M_j} \widetilde{\arg}(\zeta - z)$ is an integer and it measures the number of times M_j winding around z . Here $\mathbb{W}_M(z)$ is called the winding number of M with respect to z , which measures the total number of times the components of M winding around z .

Note: If M is a compact connected oriented 1-manifold without boundary in $\mathbb{R}^2 \simeq \mathbb{C}$, then M is diffeomorphic to S^1 .

In general, for compact oriented 1-manifold M without boundary in \mathbb{C} , one can show that $\mathbb{W}_M(z) \in \mathbb{Z}$ for all $z \in \mathbb{C}$, and \mathbb{W}_M is a continuous function on $\mathbb{C} \setminus M$, that is, \mathbb{W}_M is constant on each component of $\mathbb{C} \setminus M$.

On the other hand, for $z \in \mathbb{C}$, we observe that $\mathbb{W}(z)$ tends to 0 as $|z|$ tends to infinity. Then given $\epsilon > 0$, there exists $R > 0$ such that $\mathbb{W}_M(z) < \epsilon$ for $z \in \mathbb{C}$ that satisfies $|z| > R$, and hence $\mathbb{W}_M(z) = 0$ on the component containing $\{z \in \mathbb{C} \mid |z| > R\}$, such component of $\mathbb{C} \setminus M$ is called unbounded component.

Let M be an oriented connected compact 1-manifold without boundary in \mathbb{C} , consider now $p \in M$.



From the figure, by Cauchy Integral Theorem, we can write:

$$\mathbb{W}_M(z) - \mathbb{W}_{M_1}(z) = \frac{1}{2\pi i} \int_{\zeta \in S_u} \frac{d\zeta}{\zeta - z} = 0 \quad \Rightarrow \quad \mathbb{W}_{M_1}(z) = \mathbb{W}_M(z)$$

similarly we have:

$$\mathbb{W}_M(w) - \mathbb{W}_{M_2}(w) = \frac{1}{2\pi i} \int_{\zeta \in S_b} \frac{d\zeta}{\zeta - w} = 0 \quad \Rightarrow \quad \mathbb{W}_{M_2}(w) = \mathbb{W}_M(w)$$

Now we can write the following by following the orientation of M :

$$\mathbb{W}_M(z) - \mathbb{W}_M(w) = \mathbb{W}_{M_1}(z) - \mathbb{W}_{M_2}(w) = \mathbb{W}_{M_1}(p) - \mathbb{W}_{M_2}(p) = -\frac{1}{2\pi i} \int_{\zeta \in S} \frac{d\zeta}{\zeta - p} = -1$$

Exercise: There exists $U \subseteq \mathbb{C}$ containing M such that $U \setminus M$ has two components L, R . Denote the winding number of $w \in L$ as $\mathbb{W}(L)$ and denote the winding number of $z \in R$ as $\mathbb{W}(R)$, we have:

$$\mathbb{W}(L) = \mathbb{W}(R) + 1$$

We conclude that, for compact connected M without manifold boundary, $\mathbb{C} \setminus M$ has exactly 2 components, and hence $\mathbb{W}_M(\mathbb{C} \setminus M) = \{0, 1\}$ or $\{0, -1\}$. Reversing orientation of M if needed, we have $\mathbb{W}_M(\mathbb{C} \setminus M) = \{0, 1\}$. Here $\mathbb{W}_M^{-1}(1)$ is a connected open set forming the interior of a connected compact 2-manifold, and we have $\text{Bd}(\mathbb{W}_M^{-1}(1)) = M$. This result is known as Jordan Curve Theorem for differentiable case.

In general case, let M be a compact 1-manifold without boundary in \mathbb{C} , not necessarily connected.

Exercise: $\mathbb{W}_M(\mathbb{C} \setminus M)$ is a finite subset of \mathbb{Z} .

Exercise: For $j \in \mathbb{Z} \setminus \{0\}$, $X_j := W_M^{-1}(j)$ is either \emptyset , or is the interior of a compact 2-manifold in \mathbb{C} .

Theorem 40.1 (Residue theorem for Winding Numbers)

Let M be a compact 1-manifold without boundary in \mathbb{C} , let K denote the union of the bounded components on $\mathbb{C} \setminus M$, let U be an open subset of \mathbb{C} containing $M \cup K$, let $\{z_1, z_2, \dots, z_m\} \in U \setminus M$, and let f be a function being holomorphic on $U \setminus \{z_1, z_2, \dots, z_m\}$, then we have:

$$\int_M f dz = 2\pi i \sum_{j=1}^m \mathbb{W}_M(z_j) \cdot \text{Res}(f dz, z_j)$$

9 | Geometry

The Isoperimetric Inequality

Some problems related to infinite dimensional functional spaces make essential use of the fact that, the set of functions $C([a, b], \mathbb{R}^n)$ equipped with the metric defined by the following constitutes a complete normed vector space, or called a Banach space:

$$n_{\sup} : C([a, b], \mathbb{R}^n) \rightarrow \mathbb{R}_{\geq 0} \quad f \mapsto \max\{\|f(t)\| \mid a \leq t \leq b\}$$

Here $\|f\|_{\sup} := n_{\sup}(f)$ for $f \in C([a, b], \mathbb{R}^n)$.

Definition 41.0.0.0.1

$C_{per}^2(\mathbb{R}, \mathbb{R}^2) := \{f \in C^2(\mathbb{R}, \mathbb{R}^2) \mid f(t+1) = f(t)\}$ is a set of periodic functions of C^2 type on \mathbb{R} .

Exercise:

The set $C_{per}^2(\mathbb{R}, \mathbb{R}^2)$ equipped with the norm defined by the following constitutes a Banach space:

$$n : C_{per}^2(\mathbb{R}, \mathbb{R}^2) \rightarrow \mathbb{R}_{\geq 0} \quad f \mapsto \|f\|$$

where for $f \in C_{per}^2(\mathbb{R}, \mathbb{R}^2)$, we define:

$$\|f\| := \max\{\|f(t)\| \mid 0 \leq t \leq 1\} + \max\{\|f'(t)\| \mid 0 \leq t \leq 1\} + \max\{\|f''(t)\| \mid 0 \leq t \leq 1\}$$

Definition 41.0.0.0.2

For $f \in C_{per}^2(\mathbb{R}, \mathbb{R}^2)$, we denote:

$$f : \mathbb{R} \rightarrow \mathbb{R}^2 \quad t \mapsto \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$$

Here we define the following function:

$$l : C_{per}^2(\mathbb{R}, \mathbb{R}^2) \rightarrow \mathbb{R} \quad f \mapsto \int_0^1 \sqrt{(x'(t))^2 + (y'(t))^2} dt$$

Remark:

For $f \in C_{per}^2(\mathbb{R}, \mathbb{R}^2)$, $l(f)$ gives the 1-volume of a parametrized 1-manifold parametrized by f .

Note: For $f \in C_{per}^2(\mathbb{R}, \mathbb{R}^2)$, $l(cf) = |c| \cdot l(f)$ for $c \in \mathbb{R}$ and $f \in C_{per}^2(\mathbb{R}, \mathbb{R}^2)$.

Definition 41.0.0.3

For $h : C_{per}^2(\mathbb{R}, \mathbb{R}^2) \rightarrow \mathbb{R}$, we define the directional derivative of h at f in the direction of g as the following, if exists:

$$h'(f; g) := \lim_{u \rightarrow 0} \frac{h(f + ug) - h(f)}{u}$$

Now we get the following for $f, g \in C_{per}^2(\mathbb{R}, \mathbb{R}^2)$:

$$l'(f; g) := \lim_{u \rightarrow 0} \frac{l(f + ug) - l(f)}{u} = \left(\frac{d}{du}(l(f + ug)) \right) |_{u=0}$$

Note: For notation, $(\mu(u))|_{u=0}$ denote the function $\mu(u)$ evaluated at $u = 0$.

In this section, for notation, for $g, f \in C_{per}^2(\mathbb{R}, \mathbb{R}^2)$, we denote:

$$g : \mathbb{R} \rightarrow \mathbb{R}^2 \quad t \mapsto \begin{bmatrix} p(t) \\ q(t) \end{bmatrix} \quad f : \mathbb{R} \rightarrow \mathbb{R}^2 \quad t \mapsto \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$$

Then we can write:

$$l(f + ug) = \int_0^1 \sqrt{(x'(t) + up'(t))^2 + (y'(t) + uq'(t))^2} dt$$

By Leibniz Integral Rule, we get:

$$\begin{aligned} l'(f; g) &= \int_0^1 \left(\frac{\partial}{\partial u} \sqrt{(x'(t) + up'(t))^2 + (y'(t) + uq'(t))^2} \right) |_{u=0} dt \\ &= \int_0^1 \left(\frac{x'(t)p'(t) + y'(t)q'(t) + u(p'(t))^2 + u(q'(t))^2}{\sqrt{(x'(t) + up'(t))^2 + (y'(t) + uq'(t))^2}} \right) |_{u=0} dt \\ &= \int_0^1 \frac{x'(t)p'(t) + y'(t)q'(t)}{\sqrt{(x'(t))^2 + (y'(t))^2}} dt \end{aligned} \tag{LO}$$

Equation (LO) is valid if $(x'(t))^2 + (y'(t))^2 \neq 0$ for all t .

Definition 41.0.0.4

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

f is called an immersion provided that $Df(\vec{x})$ is injective for all $\vec{x} \in U$.

f is called a submersion provided that $Df(\vec{x})$ is surjective for all $\vec{x} \in U$.

In this section, for notation, we denote $V = C_{per}^2(\mathbb{R}, \mathbb{R}^2)$. The image of $f \in V$ might be a compact connected C^2 parametrized 1-manifold without boundary, but it depends on whether f is injection on $[0, 1]$. We denote $U_1 = \{f \in V \mid Df(\vec{x}) \neq 0\}$, for $f \in U_1$, we see that f is an immersion. Note that U_1 is an open subset of V , and $l'(f; g)$ exists for $f \in U_1$ and $g \in V$.

Exercise: $l'(f; f) = l(f)$ for $f \in U_1$.

Exercise: l is differentiable at each $f \in U_1$.

Exercise: $l'(f; cg) = c l'(f; g)$ for $c \in \mathbb{R}$, $f \in U_1$, and $g \in V$.

Definition 41.0.0.0.5

$$a : V \rightarrow \mathbb{R} \quad f \mapsto \frac{1}{2} \int_0^1 (x(t) y'(t) - x'(t) y(t)) dt$$

Note: Denote $Y = f([0, 1])$ for $f \in V$, we have:

$$a(f) = \frac{1}{2} \int_{Y_f} x dy - y dx$$

Note: We have $a(cf) = c^2 a(f)$ for $f \in V$ and $c \in \mathbb{R}$.

In this section, for notation, denote $U_2 := \{f \in U_1 \mid f \text{ is injective on } [0, 1]\}$. Note that U_2 is an open subset of U_1 . For $f \in U_2$, f parametrizes manifold without boundary.

Note: $f \in U_2$ implies $f(\mathbb{R}) = \partial\Omega$, where Ω is a compact 2-manifold in \mathbb{R}^2 . Here we can write:

$$a(f) = \pm \frac{1}{2} \int_{\partial\Omega} x dy - y dx = \pm \int_{\Omega} dx \wedge dy = \pm (\text{2-volume of } \Omega)$$

Define $U_3 = \{f \in U_2 \mid a(f) > 0\}$, in which case $a(f)$ is the 2-volume of Ω for $f \in U_3$ and $f(\mathbb{R}) = \partial\Omega$.

Exercise: For $f, g \in V$, by Integration by Part and noticing that f, g are periodic, one can get:

$$a'(f; g) = \frac{1}{2} \int_0^1 xq' - x'q - yp' + y'p = \int_0^1 xq' - yp' = \int_0^1 y'p - x'q \quad a'(f; f) = 2a(f)$$

Exercise: a is differentiable on V .

Note: Here we have:

$$\sup a(U_j) = \infty = \sup l(U_j) \quad \inf a(U_j) = 0 = \inf l(U_j)$$

Now consider the following with $\alpha, \beta \in \mathbb{R}$:

$$\frac{a^\alpha(cf)}{l^\beta(cf)} = |c|^{2\alpha-\beta} \frac{a^\alpha(f)}{l^\beta(f)} \quad (\text{RT})$$

the supremum of infimum of (RT) over U_j is still ∞ and 0 respectively unless $\beta = 2\alpha$.

Definition 41.0.0.0.6

$$h : U_1 \rightarrow \mathbb{R} \quad f \mapsto \frac{a(f)}{l^2(f)}$$

By results derived above, for $f \in U_1$, $a'(f; g) = 0$ for all $g \in V$ if and only if $x'(x) = y'(x) = 0$ for all $x \in U_1$, if and only if f is a constant function. Note that we also have, for $f \in U_1$ and $g \in V$:

$$l'(f; g) = \int_0^1 \frac{x'p' + y'q'}{\sqrt{(x')^2 + (y')^2}} = \int_0^1 (y'p - x'q) \frac{x'y'' - x''y'}{((x')^2 + (y')^2)^{3/2}}$$

Here $\kappa_f := \frac{x'y'' - x''y'}{((x')^2 + (y')^2)^{3/2}}$ is known as the curvature of f .

Here we see that, for $f \in U_1$, $l'(f; g) = 0$ for all $g \in U_1$ if and only if $y'(x'y'' - x''y') = 0 = x'(x'y'' - x''y')$, if and only if $x'y'' - x''y' = 0$. If x', y' are linearly dependent, then $x'y'' - x''y' = 0$ is guaranteed by Math 396 HW3 Q3, if $x'y'' - x''y' = 0$ holds on U_1 , then x', y' are linearly dependent.

Note: If x' and y' are linearly dependent, then $f(\mathbb{R})$ is a subset of a line in \mathbb{R}^2 .

For $f, g \in U_1$, we can write the following:

$$h'(f; g) = \frac{l(f)a'(f; g) - 2a(f)l'(f; g)}{l^3(f)}$$

Exercise: $h'(f, f) = 0$ for all $f \in U_1$

For $f \in U_1$, here we see that $h'(f; g) = 0$ for all $g \in V$ if and only if we have:

$$a'(f; g) = \frac{2a(f)}{l(f)}l'(f; g) \quad (*)$$

for all $g \in V$.

Suppose on the other hand, for $f \in U_1$, we get $a'(f; g) = \lambda l'(f; g)$ for all $g \in V$ for some constant λ . Set $g = f$, from result derived above we have $2a(f) = \lambda l(f)$, then $\lambda = \frac{2a(f)}{l(f)}$, in which case we see that $(*)$ holds. We conclude that $h'(f; g) = 0$ for all g if and only if there exists some λ such that $a'(f; g) = \lambda l'(f; g)$ for all g , if and only if f satisfies Lagrange condition for maximizing or minimizing a with l being a constant, if and only if f satisfies Lagrange condition for maximizing or minimizing l with a being a constant.

Also, for $f \in U_1$, and constant λ , $a'(f; g) = \lambda l'(f; g)$ for all $g \in V$ if and only if we have:

$$\int_0^1 (y'p - x'q) = \lambda \int_0^1 (y'p - x'q) \frac{x'y'' - x''y'}{((x')^2 + (y')^2)^{3/2}} \quad \forall p, q, \text{ s.t. } g \in V \quad (**)$$

Equation $(**)$ implies that we have:

$$\frac{1}{\lambda} = \frac{x'y'' - x''y'}{((x')^2 + (y')^2)^{3/2}} = \kappa_f = \text{curvature of } f$$

In general curvature of f is a function. For $f \in U_1$, by Math 396 HW11, if κ_f is a constant, then $f(\mathbb{R})$ is contained in a circle of radius $\frac{1}{\kappa_f} = \lambda$, or $f(\mathbb{R})$ is a line with $\kappa_f = 0$. ¹

Here if κ_f is a nonzero constant, f is a critical point of h as we have $\frac{1}{\lambda}$ being constant, and equation $(**)$ gives rise to a form of Lagrange-Euler Equation, solving for the critical points of h .

Here if f is a critical point of h , we get the following:

If $f \in U_3$, f winds once forward on a circle each period.

¹Here $f(\mathbb{R})$ are called circles

If $f \in U_2 \setminus U_3$, f winds once backward on a circle each period.

If $f \in U_1 \setminus U_2$, then f goes around a circle multiple times each period.

Remark: Many important PDE and ODE come from Lagrange-Euler Equation. For examples of applications of the Euler-Lagrange Equation of the form rising from equation (**), check out *Chapter 3 - Lagrangian and Hamiltonian Dynamics* in the text *Class Note: Physics 401 - Classical Mechanics*, transcribed by Jinyan Miao in Winter 2022. The course Physics 401 at the University of Michigan in Winter 2022 was taught by Professor Bing Zhou.

Note: The critical points for h in U_3 are circles. If $f \in U_3$ is a critical point for h , and f traces out a circle of radius r , then we have:

$$h(f) = \frac{\pi r^2}{(2\pi r)^2} = \frac{1}{4\pi}$$

Theorem 41.1 (Isoperimetric Inequality)

Let Ω be a compact connected 2-manifold in \mathbb{R}^2 with $\partial\Omega$ being connected, we have

$$\text{area}(\Omega) \leq \frac{1}{4\pi} (\text{length}(\partial\Omega))^2$$

with $\text{area}(\Omega) = \frac{1}{4\pi} (\text{length}(\partial\Omega))^2$ holds if and only if Ω is a disk. Here $\text{area}(\Omega)$ is the 2-volume of Ω , and $\text{length}(\partial\Omega)$ is the 1-volume of $\partial\Omega$.

Let Ω be a oriented compact connected 2-manifold in $\mathbb{R}^2 \simeq \mathbb{C}$, let \vec{N} be the unit normal vector field of $\partial\Omega$ corresponding to the induced orientation of $\partial\Omega$. One can show that there exists some $\epsilon_0 > 0$ such that the function $\Psi_\epsilon : \partial\Omega \rightarrow \mathbb{R}^2$ $\vec{z} \mapsto \vec{z} + \epsilon \vec{N}(\vec{z})$ satisfies the following conditions:

1. $\Psi_\epsilon(\partial\Omega) \cap \bar{\Omega} = \emptyset$ for $0 < \epsilon \leq \epsilon_0$
2. for $0 < \epsilon \leq \epsilon_0$ and $z \in \Omega$, we have:

$$\frac{1}{2\pi i} \int_{\zeta \in \Psi_\epsilon(\partial\Omega)} \frac{d\zeta}{\zeta - z} = 1$$

3. $l(\Psi_\epsilon(\partial\Omega))$ tends to $l(\partial\Omega)$ as ϵ tends to 0.

Theorem 41.2

Let Ω be a compact 2-manifold in \mathbb{R}^2 , we have the following holds:

$$\text{area}(\Omega) \leq \frac{1}{4\pi} (\text{length}(\partial\Omega))^2$$

Proof. Here we can write:

$$\begin{aligned}
\text{area}(\Omega) &= \int_{\Omega} dx \wedge dy = \int_{\Omega} \frac{i}{2} dz \wedge d\bar{z} \\
&= \int_{z \in \Omega} \left(\frac{1}{2\pi i} \int_{\zeta \in \Psi_{\epsilon}(\partial\Omega)} \frac{d\zeta}{\zeta - z} \frac{i}{2} \right) dz \wedge d\bar{z} \\
&= \frac{1}{4\pi} \int_{z \in \Omega} \left(\int_{\zeta \in \Psi_{\epsilon}(\partial\Omega)} \frac{d\zeta}{\zeta - z} \right) dz \wedge d\bar{z} \\
&= \frac{1}{4\pi} \int_{\zeta \in \Psi_{\epsilon}(\partial\Omega)} \left(\int_{z \in \Omega} \frac{dz \wedge d\bar{z}}{\zeta - z} \right) d\zeta \\
&= \frac{1}{4\pi} \int_{\zeta \in \Psi_{\epsilon}(\partial\Omega)} \left(\int_{z \in \Omega} d \left(\frac{\bar{\zeta} - \bar{z}}{\zeta - z} dz \right) \right) d\zeta \\
&= \frac{1}{4\pi} \int_{\zeta \in \Psi_{\epsilon}(\partial\Omega)} \left(\int_{z \in \partial\Omega} \frac{\bar{\zeta} - \bar{z}}{\zeta - z} dz \right) d\zeta
\end{aligned}$$

Note here we have $\left| \left| \frac{\bar{\zeta} - \bar{z}}{\zeta - z} \right| \right| = 1$, and area is non-negative, then we get:

$$\text{area}(\Omega) = \frac{1}{4\pi} \left| \int_{\zeta \in \Psi_{\epsilon}(\partial\Omega)} \left(\int_{z \in \partial\Omega} \frac{\bar{\zeta} - \bar{z}}{\zeta - z} dz \right) d\zeta \right| \quad (\text{AM})$$

Applying ML-estimate for both integrals in equation (AM), we get the following:

$$\text{area}(\Omega) \leq \frac{1}{4\pi} l(\partial\Omega) l(\Psi_{\epsilon}(\partial\Omega)) \quad (\text{AO})$$

Let ϵ in equation (AO) tends to 0, the result follows. \square

Riemannian Metrics

Consider $f \in C^1(U, \mathbb{R})$ where U is an open subset of \mathbb{R}^2 , let M be $Graph(f) \subseteq \mathbb{R}^3$, let Y be a parametrized 1-manifold in M parametrized by the following function:

$$\eta : V \rightarrow M \quad t \mapsto \begin{bmatrix} x(t) \\ y(t) \\ f\left(\begin{bmatrix} x(t) \\ y(t) \end{bmatrix}\right) \end{bmatrix}$$

where $V = [a, b] \subseteq \mathbb{R}$, and x, y are scalar-valued functions defined on V .

In such setting one can show that we get:

$$\text{length}(Y_\eta) = \int_a^b \sqrt{\alpha \left(\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \right) (x'(t))^2 + 2\beta \left(\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \right) x'(t) y'(t) + \gamma \left(\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \right) (y'(t))^2} dt$$

where $\alpha = 1 + (D_1 f)^2$, $\beta = D_1 f \cdot D_2 f$, $\gamma = 1 + (D_2 f)^2$. Note here we have:

$$\alpha \left(\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \right) (x'(t))^2 + 2\beta \left(\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \right) x'(t) y'(t) + \gamma \left(\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \right) (y'(t))^2 = [x'(t) \ y'(t)] \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}$$

which is positive as length must be positive, and the following matrix is always symmetric and positive with the set of eigenvalues $\{1, 1 + (D_1 f)^2 + (D_2 f)^2\}$:

$$\begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix}$$

Definition 42.0.0.0.1

$Symm(k)$ denotes the set of $k \times k$ symmetric matrix.

Note: $Symm(k)$ defines the space of symmetric 2-tensors on \mathbb{R}^k .

Definition 42.0.0.0.2

$Pos(k)$ is the set of positive symmetric $k \times k$ matrix.

Note: $Pos(k)$ is an ope subset of the set of symmetric $k \times k$ matrix.

Definition 42.0.0.0.3

Let A be an open subset of \mathbb{R}^k . A Riemannian metric G on A is a smooth function defined by:

$$G : A \rightarrow Pos(k)$$

Given $\psi \in C^2([a, b], A)$, we define the G -length of Y_ψ , or the length of Y_ψ in G metric, as:

$$l_G(Y_\psi) := \int_a^b \sqrt{\psi'(t)^T \cdot G(\psi(t)) \cdot \psi'(t)} dt$$

Theorem 42.1 (John Forbes Nash Jr. (1956), Mikhael Gromov (1986))

Let A be an open subset of \mathbb{R}^k . Let $G : A \rightarrow Pos(k)$ be a Riemannian metric on A . There exists diffeomorphism $\psi : A \rightarrow X$, where X is a k -manifold in $\mathbb{R}^{k^2+10k+3}$ such that, for all $\phi \in C^2([a, b], A)$, we have the following holds:

$$l_G(Y_\phi) = l(Y_{\psi \circ \phi})$$

Similarly, let A be an open subset of \mathbb{R}^k , there exists $\psi : A \rightarrow \mathbb{R}^{k+1}$ such that, for all $\phi \in C^2([a, b], A)$, we have $l_G(Y_\phi) = l(Y_{\psi \circ \phi})$. Conversely, for $A, U \subseteq \mathbb{R}^k$ with $A \subseteq U$, given a coordinate patch $\psi : U \rightarrow V$, then there exists Riemannian metric G on U such that $l_G(Y_\phi) = l(Y_{\psi \circ \phi})$ for all $\phi \in C^2([a, b], A)$.

Example: Consider $A \subseteq \mathbb{R}^2$, define a Riemannian metric G by the following:

$$G : A \rightarrow \text{Pos}(2) \quad (x, y) \mapsto \frac{4}{(1 + x^2 + y^2)^2} I$$

where I is the identity matrix, define a function $\psi : \mathbb{R}^2 \rightarrow S^2 \setminus \{\vec{e}_3\}$ by setting:

$$\psi(x, y) = \frac{1}{x^2 + y^2 + 1} (2x, 2y, x^2 + y^2 - 1)$$

ψ defines a diffeomorphism.

For $A = \mathbb{R}^2$, and graph curves $\phi_f : x \mapsto (x, f(x))$, suppose we have $\tilde{a}, \tilde{b} \in \mathbb{R}$, we set:

$$E = \{f \in C^2([a, b], \mathbb{R}) \mid f(a) = \tilde{a}, f(b) = \tilde{b}\}$$

Consider $L : E \rightarrow \mathbb{R}$ $f \mapsto l_G(Y_{\phi_f})$, we are interested in the critical points for L .

Here E is a space of functions, in particular, E is an affine subset of $C^2([a, b], \mathbb{R})$, then there exists an associate vector space $E' := \{f \in C^2([a, b], \mathbb{R}) \mid f(a) = 0 = f(b)\}$.

To find critical points of L , one can apply the two constraints $f(a) = 0 = f(b)$ and use the Lagrange Multiplier method, or one can look for $f \in E$ solving $L'(f; g) = 0$ for $g \in E'$. Here we write:

$$L'(f; g) = \left(\frac{d}{ds} (L(f + sg)) \right) |_{s=0} = \frac{d}{ds} \left(\int_a^b \sqrt{\begin{bmatrix} 1 & f' + sg' \\ f' + sg' & \end{bmatrix} \begin{bmatrix} \alpha \begin{bmatrix} t \\ f + sg \end{bmatrix} & \beta \begin{bmatrix} t \\ f + sg \end{bmatrix} \\ \beta \begin{bmatrix} t \\ f + sg \end{bmatrix} & \alpha \begin{bmatrix} t \\ f + sg \end{bmatrix} \end{bmatrix}} \begin{bmatrix} 1 \\ f' + sg' \end{bmatrix} \right) |_{s=0}$$

By Leibniz Integral Rule, boundary conditions, and Integration by Parts, one can show that we have the following holds:

$$L'(f; g) = \int_a^b (g \cdot \eta_1(f, f'; t) + g' \cdot \eta_2(f, f'; t)) dt = \int_a^b g \cdot (\eta_1(f, f'; t) - (\eta_2(f, f'; t))') dt$$

Here f is a critical point for $L|_E$ if and only if $\eta_1(f, f'; t) = (\eta_2(f, f'; t))'$, if and only if we have

$$f'' = \eta(f, f'; t) \quad (*)$$

if and only if we have:

$$\begin{bmatrix} f \\ f' \end{bmatrix}' = \begin{bmatrix} f' \\ \eta(f, f'; t) \end{bmatrix} \quad (**)$$

for some functions η, η_1, η_2 which can be computed.

Fundamental Theorem of ODEs implies that one can uniquely solve $(**)$ under condition $f(x_0) = y_0$ and $f'(x_0) = v_0$, which is equivalent to solve $(*)$ with condition $f(x_0) = y_0$ and $f'(x_0) = v_0$.

Example:

Let $A \subseteq \mathbb{R}^2$, consider a Riemannian metric G defined by the following:

$$G : A \rightarrow \text{Pos}(2) \quad (x, y) = I$$

where I is the 2×2 identity matrix, and here G defines the Euclidean metric.

For $f \in \{g \in C^2([a, b], \mathbb{R}) \mid g(a) = \tilde{a}, g(b) = \tilde{b}\}$, we can write the following with Leibniz Rule and Integration by Parts:

$$L'(f, g) = \left(\frac{d}{ds} \int_a^b \sqrt{1 + (f' + sg')^2} dt \right) |_{s=0} = \int_a^b \frac{f'g'}{\sqrt{1 + (f')^2}} dt = - \int_a^b \frac{f''g}{\sqrt{(1 + (f')^2)^{3/2}}} dt$$

$\frac{f''}{\sqrt{(1 + (f')^2)^{3/2}}} :=$ is called the curvature of graph curves $\text{Graph}(f)$

Here we see that f is a critical point for L if and only if $f'' = 0$ if and only if $\text{Graph}(f)$ is affine.

Exercise: Given Y_ϕ , rotating Y_ϕ does not change the value of $l_I(Y_\phi)$, where I is the identity matrix.

For Riemannian matrix $G = I$, we have shown that affine graph curves minimize L . Now consider joining point (a, \tilde{a}) to point (b, \tilde{b}) by a curve, here we can use rotation of the curves to reduce to the case where $\tilde{b} = \tilde{a}$. Let ϕ be a function in $C^1([a, b], \mathbb{R}^2)$. For $t \in [a, b]$, let $\phi(t) = (x(t), y(t))$, with $\phi(A) = (a, \tilde{a})$ and $\phi(B) = (b, \tilde{b}) = (b, \tilde{a})$. Here we write:

$$l_I(Y_\phi) = \int_A^B \sqrt{(x')^2 + (y')^2} \geq \int_A^B \sqrt{(x')^2} = \int_A^B |x'| \geq \int_A^B x' = x(B) - x(A) = b - a$$

Note here $b - a$ gives the length, under the Riemannian metric $G = I$, of a straight line segment joining the point (a, \tilde{a}) to the point (b, \tilde{b}) .

Example: (Poincare Metric)

Let $A = \mathbb{H}_+^2$, consider the Riemannian metric G defined by the following:

$$G : A \rightarrow \text{Pos}(k) \quad \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{bmatrix}$$

Here G here is called the Poincare metric, or hyperbolic metric.

For $\phi \in C^1([a, b], \mathbb{R}^2)$, denote $\phi(t) = (x(t), y(t))$, here we can write:

$$l_G(Y_\phi) = \int_a^b \frac{\sqrt{(x')^2 + (y')^2}}{y} = \int_a^b \frac{dV}{y}$$

For technical reasons, we will focus on sideways graphs, that is, we write:

$$\phi : C^1([a, b], \mathbb{R}) \quad y \mapsto \begin{bmatrix} f(y) \\ y \end{bmatrix}$$

then we have:

$$L(f) = \int_a^b \frac{\sqrt{1 + (f'(y))^2}}{y} dy$$

Using Leibniz Rule, and Integration by Parts, we get:

$$L'(f; g) = \int_a^b \frac{f'(y)g'(y)}{\sqrt{1 + (f'(y))^2}} dy = \int_a^b \frac{-yf''(y) + f'(y) + (f'(y))^3}{y(1 + (f'(y))^2)^{3/2}} g(y) dy$$

for f to be a critical point for constrained L , we must have the following holds:

$$L'(f; g) = 0 \quad \forall g \in E' \iff yf''(y) = f'(y) + (f'(y))^3 \quad (\text{C})$$

Equation (C) holds if and only if, by HW12, Y_ϕ is contained in, a circle centered on x -axis, or a vertical line. These are lines for Poincare's model of Bolyai-Lobachevsky non-Euclidean geometry. Any points \mathbb{H}_+^2 are joint by exactly one Poincare line.

Definition 42.1.0.0.1

We denote the Poincare length of a curve Y_ϕ as $l_{\text{Poin}}(Y_\phi)$.

Theorem 42.2

The followings minimize Poincare length among curves with the same end points:

1. Vertical segment in \mathbb{H}_+^2
2. Arc in \mathbb{H}_+^2 of circle centered on x -axis

Proof. Assume endpoints of the vertical segment as $(x_1, y_1), (x_1, y_2)$ with $y_1 < y_2$, and such vertical line segment is parametrized by ϕ defined by the following:

$$\phi : I \rightarrow \mathbb{H}_+^2 \quad t \mapsto \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$$

Note here we have the following holds:

$$l_{\text{Poin}}(Y_\phi) = \int_I \frac{\sqrt{(x')^2 + (y')^2}}{y} \geq \int_I \frac{\sqrt{(y')^2}}{y} = \int_I \frac{|y'|}{y} \geq \int_I \frac{y'}{y} = \Delta_I \ln(y) = \ln\left(\frac{y_2}{y_1}\right)$$

To complete the proof of Theorem 42.3, we will make use of the following exercise:

Exercise: The following diffeomorphisms of \mathbb{H}_+^2 preserve l_{Poin} :

1. $f : \mathbb{H}_+^2 \rightarrow \mathbb{H}_+^2 \quad (x, y) \mapsto (x + a, y)$ for $a \in \mathbb{R}$
which is equivalent to the holomorphic function $h : \mathbb{C} \rightarrow \mathbb{C} \quad z \mapsto z + a$
2. $f : \mathbb{H}_+^2 \rightarrow \mathbb{H}_+^2 \quad (x, y) \mapsto (\lambda x, \lambda y)$ with $\lambda > 0$
which is equivalent to the holomorphic function $h : \mathbb{C} \rightarrow \mathbb{C} \quad z \mapsto \lambda z$
3. $f : \mathbb{H}_+^2 \rightarrow \mathbb{H}_+^2 \quad (x, y) \mapsto (-\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2})$
which is equivalent to holomorphic $h : \mathbb{C} \rightarrow \mathbb{C} \quad z \mapsto 1/z$ for $z \in \mathbb{C}$

Note: the group of diffeomorphism generated by (1), (2), and (3) is given by the following:

$$\left\{ f : \mathbb{C} \rightarrow \mathbb{C} \quad z \mapsto \frac{az+b}{cz+d} \mid a, b, c, d, \in \mathbb{R}, ad - bc = 1 \right\} = SL(2, \mathbb{R})/\{\pm I\}$$

which is the group of all orientation preserving isometries of the metric space $(\mathbb{H}_+^2, \text{Poincare metric})$, and this is the group of holomorphic bijection from \mathbb{H}_+^2 to \mathbb{H}_+^2 .

Here we note that (3) maps the vertical line $x = 1$ to $(x + \frac{1}{2})^2 + y^2 = \frac{1}{4}$, and (1) and (2) maps $(x + \frac{1}{2})^2 + y^2 = \frac{1}{4}$ to any possible circles centered on the x -axis.

Finally, to complete the proof of Theorem 42.3, we will make use of the following exercise:

Exercise: Any $(x_1, y_1), (x_2, y_2) \in \mathbb{H}_+^2$ lie on such a circle, or a vertical line segment described above.

This completes the proof of Theorem 42.3. □

Remark: Here we defined the Riemannian metric to be $G : A \rightarrow \text{Pos}(k)$ where A is an open subset of \mathbb{R}^k . Given a Riemannian metric, the usual metric, or the point metric for

points in A , that we have defined previously is induced from such Riemannian metric, which is given by the following:

$$d : A \times A \quad (\vec{p}, \vec{q}) \mapsto \inf \{l_G(Y_\phi) \mid \phi \in C_{pw}^1([a, b], A), \phi(a) = \vec{p}, \phi(b) = \vec{q}\} \quad (***)$$

Note here it is not quite trivial to show that $d(\vec{p}, \vec{q}) = 0$ implies $\vec{p} = \vec{q}$ from such definition.

Note here in some cases, the infimum of the set $\{l_G(Y_\phi) \mid \phi \in C_{pw}^1([a, b], A), \phi(a) = \vec{p}, \phi(b) = \vec{q}\}$ in equation $(***)$ is in fact the minimum of the set, but this is not the case in general.

Now consider the three geometries defined on an open subset A of \mathbb{R}^2 :

1. The Euclidean metric defined on $A = \mathbb{R}^2$ by the following:

$$G_e : A \rightarrow \text{Pos}(2) \quad (x, y) \mapsto I$$

2. The Poincare metric defined on $A = \mathbb{H}_+^2$ by the following:

$$G_p : A \rightarrow \text{Pos}(2) \quad (x, y) \mapsto \frac{1}{y^2} I$$

3. The spherical metric defined on $A = \mathbb{R}^2$ by the following:

$$G_s : A \rightarrow \text{Pos}(2) \quad (x, y) \mapsto \frac{4}{(1 + x^2 + y^2)^2} I$$

Here I is the identity 2×2 matrix.

Note: Here we let $\bar{B}_r^*(\vec{p}) \subseteq \mathbb{R}^2$ denote a closed ball centered at $\vec{p} \in \mathbb{R}^2$ or radius r under the metric $*$. Spherical metric is abbreviated as *sph*, Euclidean metric is abbreviated as *Euc*, and Poincare metric is abbreviated as *Poin*.

For $\vec{p} \in \mathbb{R}^2$, and $r > 0$, we have:

Exercise: $l_{G_s}(\partial \bar{B}_r^{sph}(\vec{p})) = 2\pi \sin(r)$ for $0 < r < \pi$

Exercise: $l_{G_e}(\partial \bar{B}_r^{Euc}(\vec{p})) = 2\pi r$

Exercise: $l_{G_p}(\partial \bar{B}_r^{Poin}(\vec{p})) = 2\pi \sinh(r)$

Exercise: Note here we can write:

$$\lim_{r \rightarrow 0^+} 3 \left(\frac{2\pi r - l_*(\partial \bar{B}_r^*(\vec{p}))}{\pi r^3} \right) = \begin{cases} 1 & * = \text{spherical} \\ 0 & * = \text{Euclidean} \\ -1 & * = \text{Poincare} \end{cases}$$

10 | Differentiable Manifolds III

de Rham Cohomology

Let A be an open subset of \mathbb{R}^n , we define:

$$\begin{aligned}\Omega^k(A) &:= \{\omega \mid \omega \text{ is a } k\text{-form of } C^\infty \text{ type defined on } A\} \\ Cl^k(A) &:= \{\omega \in \Omega^k(A) \mid d\omega = 0\} \\ E^k(A) &:= \{d\eta \mid \eta \in \Omega^{k-1}(A)\}\end{aligned}$$

Note: For open subset A of \mathbb{R}^n , we have $E^k(A) \subseteq Cl^k(A)$.

Remark: Let A be an open subset of \mathbb{R}^n . On Munkres, $E^0(A)$ is defined to be the set of constant functions on A , but Professor Barrett defines $E^0(A)$ to be the set of zero function defined on A , denoted as $\{0\}$. In this text, if not specified, we will use Professor Barrett's definition for $E^0(A)$.

Definition 43.0.0.1

Let A be an open subset of \mathbb{R}^n , the k -th de Rham cohomology on A is defined by the following:

$$H_{dR}^k(A) := Cl^k(A)/E^k(A)$$

Here $H_{dR}^k(A)$ is the quotient of the vector space $Cl^k(A)$ by the subspace $E^k(A)$.

Exercise: Let A be an open subset of \mathbb{R}^n , we have the following holds:

$$\dim(H_{dR}^0(A)) = \begin{cases} (\text{number of connected components of } A) & \text{Barrett's definition of } E^0(A) \\ (\text{number of connected components of } A) - 1 & \text{Munkres's definition of } E^0(A) \end{cases}$$

Exercise: Given a diffeomorphism $g : A \rightarrow B$, we get a linear isomorphism from $H_{dR}^k(B)$ to $H_{dR}^k(A)$, defined by the following:

$$f : H_{dR}^k(B) \rightarrow H_{dR}^k(A) \quad [\omega] \mapsto [g^*\omega]$$

Here we see that the dimension of $H_{dR}^k(B)$ is the same as the dimension of $H_{dR}^k(A)$, when A is diffeomorphic to B , the dimension of $H_{dR}^k(A)$ called the k -th Betti number of A .

Let A be an open subset of \mathbb{R}^n , by Math 396 HW10 Q5, there exists a natural bilinear map from $H_{dR}^k(A) \times H_{dR}^k(A)$ to $H_{dR}^k(A)$ defined by:

$$L : H_{dR}^k(A) \times H_{dR}^k(A) \rightarrow H_{dR}^k(A) \quad ([\omega_1], [\omega_2]) \mapsto [\omega_1 \wedge \omega_2]$$

By Math 396 HW12 Q2, if $A \subseteq \mathbb{R}^n$ is an open box, then we have:

$$\dim(H_{dR}^k(A)) = \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

Theorem 39.3 on Munkres gives a similar result.

Exercise: Results from Math 396 HW7 Q2 imply that we have $\dim(H_{dR}^{n-1}(\mathbb{R}^n \setminus \{0\})) > 0$. In fact, the following function defines an isomorphism:

$$M : H_{dR}^{n-1}(\mathbb{R}^n \setminus \{0\}) \rightarrow \mathbb{R} \quad [\omega] \mapsto \int_{S^{n-1}} \omega$$

Theorem 43.1

Let E be a nonempty affine subset of \mathbb{R}^n , then we have the following holds:

$$\dim(H_{dR}^k(\mathbb{R} \setminus E)) = \begin{cases} 1 & k = n - \dim(E) \text{ or } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

with an exception where $\dim(E) = n - 1$ and $k = 0$, in which case $\dim(H_{dR}^0(\mathbb{R}^n \setminus E)) = 2$.

Corollary 43.1.1

Let E_1 and E_2 be nonempty affine subsets of \mathbb{R}^n , and if $\mathbb{R}^n \setminus E_1$ is diffeomorphic to $\mathbb{R}^n \setminus E_2$, then we have $\dim(E_1) = \dim(E_2)$.

Exercise:

For manifold M in \mathbb{R}^n , there exists open subset U of \mathbb{R}^n that contains M such that M is closed in U , and this exercise allows us to propose the following definitions:

Definition 43.1.1.0.1

Let M be a manifold in \mathbb{R}^n , and let U be an open subset of \mathbb{R}^n containing M such that M is closed in U . We define the following:

$$\Omega^k(M) := \Omega^k(U) / \{\omega \in \Omega^k(U) \mid M \text{ is integral for } \omega\}$$

Definition 43.1.1.0.2

Let M be a manifold in \mathbb{R}^n , and let U be an open subset of \mathbb{R}^n containing M such that M is closed in U . $\Omega^k(M)$ is a set consisting of smooth ω defined on U that maps $\vec{p} \in M$ to $\omega(\vec{p}) \in \mathcal{A}^k(\mathcal{T}_{\vec{p}}(M))$

Here we will show that Definition 43.1.1.0.1 and Definition 43.1.1.0.2 match. Consider the restriction map $\rho : \Omega^k(U) \rightarrow \Omega^k(M)$ where $\Omega^k(M)$ is defined by Definition 43.1.1.0.2, here one might want to show $\ker(\rho) = \{\omega \in \Omega^k(U) \mid M \text{ is integral for } \omega\}$. Note that, M is integral for ω , by Lemma Y.2 on Math 396 Supplement Material, if and only if $\int_N \omega = 0$ for all compact k -manifolds $N \subseteq M$, if and only if, by Theorem 31.2, $\int_{\vec{p} \in N} \omega(\vec{p})(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) dV = 0$ for all compact k -manifolds $N \subseteq M$, and here $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)$ is an orthonormal basis for $\mathcal{T}_{\vec{p}}(M)$.

For s -manifold M in \mathbb{R}^n , now we get the following fancy looking cochain complex:

$$0 \rightarrow \Omega^0(M) \xrightarrow{d_0} \Omega^1(M) \xrightarrow{d_1} \Omega^2(M) \xrightarrow{d_2} \dots \xrightarrow{d_{s-2}} \Omega^{s-1}(M) \xrightarrow{d_{s-1}} \Omega^s(M) \rightarrow 0$$

Here d_j denote the exterior derivative operator d , the subscript j denote the dimension of the manifold that the operator d_j operating on, and note here we have $d_{j+1} \circ d_j = 0$.

Definition 43.1.1.0.3

For s -manifold M in \mathbb{R}^n , we define the followings:

$$Cl^k(M) := \ker(d_k) \quad E^k(M) := Im(d_{k-1}) \subseteq Cl^k(M) \quad H_{dR}^k(M) := Cl^k(M)/E^k(M)$$

Exercise: Diffeomorphism $g : M_1 \rightarrow M_2$ induces linear bijection $b^* : H_{dR}^k(M_2) \rightarrow H_{dR}^k(M_1)$

Theorem 43.2

Let M be a compact oriented s -manifold without boundary, then $H_{dR}^s(M) \neq 0$.

Proof. By Lemma Y.1, there exists $\tilde{\omega}$ being a s -form defined on \mathbb{R}^n such that $\int_M \tilde{\omega} \neq 0$, here $\tilde{\omega}$ induces $\omega \in \Omega^s(M)$ by the restriction map $\rho : \Omega^k(U) \rightarrow \Omega^k(M)$, with $\int_M \omega \neq 0$. Suppose there exists $\eta \in \Omega^{s-1}(M)$ such that $d\eta = \omega$, then we have:

$$0 = \int_M (d\eta - \omega) = \int_{\partial M} \eta - \int_M \omega = 0 = \int_M \omega \neq 0$$

Here we reach a contradiction, which completes the proof. \square

The proof of the next few theorems in this section are left for exercises:

Theorem 43.3

Let M be a compact oriented s -manifold without boundary, then we have:

$$\dim(H_{dR}^s(M)) = \text{number of connected components of } M = \dim(H_{dR}^0(M))$$

Theorem 43.4

Let M be a compact connected non-orientable s -manifold, we have $H_{dR}^s(M) = 0$.

Theorem 43.5

Let M be a non-compact connected s -manifold, then $H_{dR}^s(M) = 0$

Theorem 43.6

Let M be a compact connected s -manifold with $\partial M \neq 0$, then $H_{dR}^s(M) = 0$.

Let M be a compact s -manifold without boundary in \mathbb{R}^n , let $\omega \in Cl^k(M)$, here we have:

$$[\omega] = \{\omega + d\eta \mid \eta \in \Omega^{k-1}(M)\}$$

$[\omega]$ is in fact an affine subset of $\Omega^k(M)$.

Definition 43.6.0.0.1

Let M be a compact s -manifold without boundary in \mathbb{R}^n . The following function defines a norm on $\Omega^k(M)$:

$$F : \Omega^k(M) \rightarrow \mathbb{R} \quad \omega \mapsto \|\omega\|^2 := \int_M \omega \wedge * \omega$$

Let M be a compact s -manifold without boundary in \mathbb{R}^n , for $\omega \in \Omega^k(M)$, and $d\eta \in E^k(M)$, let F be defined in Definition 43.6.0.0.1, one can show that we have:

$$F'(\omega, d\eta) = 2(-1)^{k+1} \int_M \eta \wedge d(*\omega)$$

Hence $F'(\omega, d\eta) = 0$ for all $d\eta \in E^k(M)$ if and only if $d(*\omega) = 0$.

Here ω minimizes the norm defined by F within $[\omega]$ if and only if we have $d\omega = 0$ and $d*\omega = 0$.

Theorem 43.7

Let M be a compact s -manifold without boundary in \mathbb{R}^n .

$\omega \in Cl^k(M)$ implies $[\omega]$ contains a unique norm-minimized element.

Proof. "Just believe it." □

Note that in the settings of Theorem 43.7, we have the followings hold:

1. If ω here is C^{k+1} , then the minimizer is C^k .
2. If ω here is C^∞ , then the minimizer is C^∞ .
3. If ω here is Holder α , then the minimizer is Holder α .

Equivalent to the statement of the theorem, each class $[\omega]$ in $H_{dR}^k(M)$ contains a unique solution of $d*\omega = 0$. And Equivalently, the following function is a linear bijection:

$$f : \{\omega \in \Omega^k(M) \mid d\omega = 0, d*\omega = 0\} \rightarrow H_{dR}^k(M) \quad \omega \mapsto [\omega]$$

Here $\{\omega \in \Omega^k(M) \mid d\omega = 0, d*\omega = 0\} := H_{Hodge}^k(M)$ is called the Hodge cohomology.

Proposition 43.7.1

Let M be a compact s -manifold without boundary in \mathbb{R}^n .

$* : H_{Hodge}^k(M) \rightarrow H_{Hodge}^{s-k}(M)$ defines an isomorphism.

Corollary 43.7.2

Let M be a compact s -manifold without boundary in \mathbb{R}^n . $\dim(H_{Hodge}^k(M)) = \dim(H_{Hodge}^{s-k}(M))$.

Corollary 43.7.3

Let M be a compact s -manifold without boundary in \mathbb{R}^n . $\dim(H_{dR}^k(M)) = \dim(H_{Hodge}^{s-k}(M))$

Let M be a subset of \mathbb{R}^n . One can use the following conditions to propose a definition for topological k -manifold M :

1. M is a T_3 space.
2. M has a countable many dense subsets.
3. M is a union of V_j with each V_j being relatively open in M , and V_j is homeomorphic to some U_j being open in \mathbb{R}^k or in \mathbb{H}^k .

Furthermore, if one can choose homeomorphism $\beta_j : V_j \rightarrow U_j$ such that $\beta_k \circ \beta_l^{-1}$ is of C^r type where defined. Then the above (3) conditions along with the existence of β_j characterizes the definition of an abstract k -manifold M of class C^r .

Theorem 43.8

*If M is a abstract C^r - k -manifold, then M is diffeomorphic to a differentiable k -manifold M' in \mathbb{R}^N for some $N \in \mathbb{N}$.*¹

Theorem 43.9 (Whitney Embedding Theorem)

If M is a abstract k -manifold of class C^r , then M is diffeomorphic to a differentiable k -manifold M' in \mathbb{R}^{2k}

Exercise: In the settings of Theorem 43.9, for $k = 1$, we have circle in \mathbb{R}^2

Exercise: In the settings of Theorem 43.9, for $k = 2$, we have Klein Bottle in \mathbb{R}^4 .

Remark: In the settings of Theorem 43.8, one can embed M in \mathbb{R}^{2k-1} unless k is a power of 2.

For $k = 1, 2, 3$, every topological k -manifold is homeomorphic to a differentiable k -manifold, but this statement is false for $k \geq 4$, by Freedman (1982).

For $k = 1, 2, 3$, two differentiable manifold are diffeomorphic if and only if they are homeomorphic, but this statement is false for $k \geq 4$. There exists open subset of \mathbb{R}^4 that is homeomorphic to \mathbb{R}^4 but not diffeomorphic to \mathbb{R}^4 . There are 28 non-equivalent smooth structures on S^7 . There is no exotic structure on S^1, S^2, S^3, S^5, S^6 .

¹The manifolds that we have been discussed in this text, and defined in chapter 4, are all differentiable manifolds.

The Hodge Star Operator

There is a very important operator in vector calculus, and in physics, that has not been discussed in this course, which is the Laplacian ∇^2 , we will define the Laplacian by first introducing the Hodge star operator $*$.

Definition 44.0.0.1

Consider an open subset A of \mathbb{R}^n , for ascending k -tuple I of integers in $\{1, 2, \dots, n\}$, let I' be the $(n - k)$ -tuple complementary to I , we define the Hodge star operator $*$ as the following:

$$*: \Omega^k(A) \rightarrow \Omega^{n-k}(A) \quad \left(\sum_{[I]} b_I(\vec{x}) dx_I \right) \mapsto \sum_{[I]} \text{sgn}(I, I') b_I(\vec{x}) dx_{I'}$$

Here the notation $\text{sgn}(I, I')$ denotes $\text{sgn}(\sigma_{II'})$, where $\sigma_{II'}$ is a permutation that sorts the concatenated n -tuple (I, I') .

Example: Consider 0-form f defined on \mathbb{R} , and 1-form $f dx$ defined on \mathbb{R} . $*(f) = f dx$, $*(f dx) = f$.

Example: Consider 0-form f , 1-form $\alpha dx + \beta dy$, and 2-form $f dx \wedge dy$ defined on \mathbb{R}^2 . Here we have $*(f) = f dx \wedge dy$, $*(\alpha dx + \beta dy) = -\beta dx + \alpha dy$, and $*(f dx \wedge dy) = f$.

Example: Consider 0-form f , 1-form $\alpha dx + \beta dy + \gamma dz$, 2-form $\alpha dy \wedge dz + \beta dz \wedge dx + \gamma dx \wedge dy$, and 3-form $f dx \wedge dy \wedge dz$ defined on \mathbb{R}^3 . Here we have the followings:

1. $*(f) = f dx \wedge dy \wedge dz$
2. $*(\alpha dx + \beta dy + \gamma dz) = \alpha dy \wedge dz + \beta dz \wedge dx + \gamma dx \wedge dy$
3. $*(\alpha dy \wedge dz + \beta dz \wedge dx + \gamma dx \wedge dy) = \alpha dx + \beta dy + \gamma dz$
4. $*(f dx \wedge dy \wedge dz) = f$

Lemma 44.0.1

For 0-form f and k -forms ω , l -form $\tilde{\omega}$ defined on an open subset of \mathbb{R}^n , we denote:

$$\omega = \sum_{[I]} b_I(\vec{x}) dx_I \quad \tilde{\omega} = \sum_{[J]} \tilde{b}_J(\vec{x}) dx_J$$

where $[I]$ is the set of ascending k -tuple, and $[J]$ is the set of ascending l -tuple. We have:

1. $*(f\omega) = f * \omega$
2. $*(\omega) = (-1)^{k(n-k)} \omega$
3. $*(\omega_1 + \omega_2) = *(\omega_1) + *(\omega_2)$
4. $\omega \wedge *(\omega) = \sum_{[I]} b_I^2 dx_1 \wedge \dots \wedge dx_n$
5. If $\deg(\omega) = \deg(\tilde{\omega})$, then $\omega \wedge *(\tilde{\omega}) = \tilde{\omega} \wedge *(\omega) = \sum_{[I]} b_I \tilde{b}_I dx_1 \wedge \dots \wedge dx_n$

Proof. Professor Barrett said that this lemma is an easy exercise. \square

Remark: There is no convenient rule for $*(\omega_1 \wedge \omega_2)$ in general.

Remark: For k -form ω , $*(\omega)$ is denoted as $*\omega$ hereafter.

Now we want to investigate in what condition we have $\Phi^*(\ast(\omega)) = \ast(\Phi^*(\omega))$ holds for k -form ω defined on \mathbb{R}^n , and a function $\Phi \in C^\infty(A, B)$ where A and B are connected open subsets of \mathbb{R}^n .

First we consider $k = 0$ and $n = 1$, ω can be written as a 0-form f , then we can write:

$$\ast(\Phi^* f) = f \circ \Phi \, dx \quad f \circ \Phi \cdot d\Phi = \Phi^*(\ast f)$$

here we require $d\Phi = dx$ such that $\Phi^*(\ast(\omega)) = \ast(\Phi^*(\omega))$, so we require $\Phi(x) = x + c$ for some constant c , which are orientation preserving isometries in \mathbb{R} .

Now consider $k = 0$ and n being general dimension, denote $\omega = f$, we can write the following:

$$\Phi^*(f \, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n) = (f \circ \Phi) \, d\Phi_1 \wedge d\Phi_2 \wedge \cdots \wedge d\Phi_n = f \circ \Phi \, \det(D\Phi) \, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$$

hence we need $\det(D\Phi) = 1$ in order to have $\Phi^*(\ast(\omega)) = \ast(\Phi^*(\omega))$, so we require Φ to preserve orientation and volume.

Now consider $k = 1$ and $n = 2$, denote $\omega = \alpha \, dx + \beta \, dy$ with α and β being scalar-valued functions, then we can write the following:

$$\ast(\Phi^*(\alpha \, dx + \beta \, dy)) = \ast((\alpha \circ \Phi) \, d\Phi_1 + (\beta \circ \Phi) \, d\Phi_2) = \alpha \circ \Phi \cdot \ast(d\Phi_1) + (\beta \circ \Phi) \cdot \ast(d\Phi_2)$$

On the other hand, we have:

$$\Phi^*(\ast(\alpha \, dx + \beta \, dy)) = \Phi^*(-\beta \, dx + \alpha \, dy) = -\beta \circ \Phi \, d\Phi_1 + (\alpha \circ \Phi) \, d\Phi_2$$

In order to have $\Phi^*(\ast(\omega)) = \ast(\Phi^*(\omega))$, here we need $d\Phi_2 = \ast(d\Phi_1)$ and $d\Phi_1 = -\ast(d\Phi_2)$, that is, we require the following holds:

$$\frac{\partial \Phi_1}{\partial x} = \frac{\partial \Phi_2}{\partial y} \quad \frac{\partial \Phi_1}{\partial y} = -\frac{\partial \Phi_2}{\partial x}$$

in other words, Φ is required to be holomorphic when we view $\mathbb{R}^2 \cong \mathbb{C}$.

Assuming working on a connected set, and consider $k = 1$ and $n > 2$, one can show that we have $\Phi^*(\ast(\omega)) = \ast(\Phi^*(\omega))$ if and only if we have $D\Phi \cdot (D\Phi)^T = \det(D\Phi) \cdot I$. Note here $D\Phi \cdot (D\Phi)^T$ is symmetric and semi-definite, then $\det(\det(D\Phi)) \geq 0$, and we also have $(\det(D\Phi))^2 = (\det(D\Phi))^n$, hence $\det(D\Phi) \in \{0, 1\}$. Then by Theorem 20.5 on Munkres, $\det(D\Phi) \in \{0, 1\}$ is guaranteed if Φ is either a constant map, or an isometry of the form $\vec{x} \mapsto A\vec{x} + \vec{b}$ with A being an orthogonal matrix, in which case Φ is an orientation preserving isometry.

Theorem 44.1

Let A, B be open subsets of \mathbb{R}^n , if $\Phi : A \rightarrow B$ defines an orientation preserving isometry, then we have $\Phi^(\ast(\omega)) = \ast(\Phi^*(\omega))$ holds for all k -forms ω defined on an open subset of \mathbb{R}^n containing B , with $k \leq n$.*

Proof. The proof of Theorem 44.1 is left for exercise. □

Euler's Equations for Non-viscous Incompressible Fluid Revisited

For non-viscous incompressible fluid, let \vec{F} denote the fluid velocity vector field, we wrote:

$$\operatorname{div}(\vec{F}) = 0 \quad \frac{\partial}{\partial t}(\operatorname{curl}(\vec{F})) + \operatorname{curl}((\operatorname{curl}\vec{F}) \times \vec{F}) = \vec{0} \quad (\text{EEF})$$

Here one can view \vec{F} as a possibly time dependent 1-form ω defined on \mathbb{R}^n .

Now we can rewrite equation (EEF) as the following:

$$d(*(\omega)) = 0 \quad d\left(\frac{\partial \omega}{\partial t}\right) + d * (*(\operatorname{d}\omega) \wedge \omega) = 0 \quad (\text{EEFF})$$

Here we consider a special case, in \mathbb{R}^n , a 1-form ω that satisfies the following conditions would satisfies equation (EEFF):

$$\frac{\partial \omega}{\partial t} = 0 \quad d(*\omega) = 0 \quad d\omega = 0$$

In such case ω is a closed 1-form, hence $\omega = df$ for some 0-form f , then we have:

$$\begin{aligned} 0 &= d(*df) = d\left(\sum_j D_j f * dx_j\right) \\ &= \sum_{j,k} D_k D_j f dx_k * dx_j \\ &= \sum_j D_j D_j f dx_j \wedge *dx_j \\ &= \left(\sum_j D_j D_j f\right) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n \end{aligned}$$

This gives us an intuition to define the Laplacian ∇^2 for 0-form f :

$$\sum_j D_j D_j f := \nabla^2 f := \Delta f$$

Then we can write the following:

$$0 = d(*df) \quad \Rightarrow \quad 0 = \nabla^2 f$$

The Laplace Operator

Definition 45.0.0.0.1

Let A be an open subset of \mathbb{R}^n , here we define the Laplace operator Δ as the following:

$$\Delta : \Omega^k(A) \rightarrow \Omega^k(A) \quad \omega \mapsto (-1)^{kn} * d * d\omega + (-1)^n d * d * \omega$$

Remark: The Laplace operator Δ , sometimes denoted as ∇^2 , is also called the Laplacian.

Example: For f being 0-form in \mathbb{R}^n , we can write:

$$\Delta f = (-1)^{kn} * d * df + (-1)^n d * d * f = (-1)^{kn} * d * df = \sum_j D_j D_j f$$

Exercise: For k -form $\omega = \sum_I b_I dx_I$, we have:

$$\Delta \left(\sum_I b_I dx_I \right) = \sum_I (\Delta b_I) dx_I$$

Definition 45.0.0.0.2

Let A be an open subset of \mathbb{C} . A harmonic function defined on A is a function $\Psi \in C^2(A, \mathbb{R})$ that satisfies the Laplace's Differential Equation on A :

$$D_1 D_1 \Psi + D_2 D_2 \Psi = 0$$

Example: Let f be a scalar-valued function defined on an open convex subset A of $\mathbb{C} \cong \mathbb{R}^2$, by Math 396 HW6 Q3, the function f is harmonic if and only if f is the real part of a holomorphic function defined on A .

Definition 45.0.0.0.3

Let A be an open subset of \mathbb{R}^n , a function $f \in C^2(A, \mathbb{R})$ is said to be harmonic provided that we have the following holds:

$$\sum_j D_j D_j f = 0$$

Exercise: If a function f is affine, then f is harmonic.

Exercise: Let A be an open subset of $\mathbb{R}^n \setminus \{\vec{0}\}$, consider $f : A \rightarrow \mathbb{R}$ $\vec{x} \mapsto \|\vec{x}\|^s$, then the function f is harmonic if and only if $*(d\|\vec{x}\|^s)$ is closed, if and only if $s = 0$ or $s = n - 2$ by Math 396 HW7 Q1.

Note: Let M be a compact n -manifold in \mathbb{R}^n , $\int_M f = \int_M *f$ for 0-form f .

Lemma 45.0.1 (Green's First Identities)

Let M be a compact n -manifold in \mathbb{R}^n , let $f, g \in C^2(M, \mathbb{R})$, then we have:

$$\int_M f \Delta g = \int_M f * \Delta g = \int_M f d * dg = \int_M d(f \wedge * dg) - df \wedge * dg = \int_{\partial M} f \wedge * dg - \int_M df \wedge * dg$$

Note: In the settings of Lemma 45.0.2, we note that we have:

$$\int_M df \wedge * dg = \int_M dg \wedge * df = \int_M \langle df, dg \rangle$$

Corollary 45.0.1.1 (Green's Second Identity)

Let M be a compact n -manifold in \mathbb{R}^n , let $f, g \in C^2(M, \mathbb{R})$, then we have:

$$\int_M (f \Delta g - g \Delta f) = \int_{\partial M} (f * dg - g * df)$$

Proof. This result is immediate from Lemma 45.0.1. □

Consider $n > 2$, let $\bar{B}_1(\vec{0})$ denote the closed ball in \mathbb{R}^n of radius 1 centered at $\vec{0} \in \mathbb{R}^n$, let $f \in C^2(\bar{B}_1(\vec{0}), \mathbb{R})$, and let g be defined by the following:

$$g : \mathbb{R}^n \setminus \{\vec{0}\} \rightarrow \mathbb{R} \quad \vec{x} \mapsto \|\vec{x}\|^{2-n} - 1$$

Note that g is harmonic on $\mathbb{R}^n \setminus \{\vec{0}\}$, $g(\vec{x}) \geq 0$ for $\vec{x} \in \bar{B}_1(\vec{0}) \setminus \{\vec{0}\}$, and $g(\vec{x}) = 0$ for $\vec{x} \in S^{n-1}$.

By Lemma 38.5 on Munkres, with \vec{N} being the unit normal field \vec{N} on rS^{n-1} , we get:

$$\begin{aligned} \int_{\|\vec{x}\|=r} f * dg &= (2-n)r^{1-n} \int_{\|\vec{x}\|=r} f * d(\|\vec{x}\|) \\ &= (2-n)r^{1-n} \int_{\|\vec{x}\|=r} \langle f\vec{N}, \vec{N} \rangle dV \\ &= (2-n)r^{1-n} \int_{\|\vec{x}\|=r} f dV \\ &= (2-n)r^{1-n} \cdot V_{n-1}(rS^{n-1}) \cdot \text{avg}_{\|\vec{x}\|=r} f \\ &= (2-n)r^{1-n} \cdot r^{1-n} \cdot V_{n-1}(S^{n-1}) \cdot \text{avg}_{\|\vec{x}\|=r} f \\ &= (2-n) \cdot V_{n-1}(S^{n-1}) \cdot \text{avg}_{\|\vec{x}\|=r} f \end{aligned}$$

Note that $*d\|\vec{x}\| = *(\vec{x}^T / \|\vec{x}\|)$ is the $(n-1)$ -form corresponds to the unit normal field \vec{N} , and $f * d\|\vec{x}\|$ corresponds to the vector field $f\vec{N}$. $V_{n-1}(S)$ denote the $(n-1)$ -volume of the set S .

Let $M_a = \{\vec{x} \mid a \leq \|\vec{x}\| \leq 1\}$ in \mathbb{R}^n , noting that $g(\vec{x}) = 0$ for $\vec{x} \in S^{n-1}$.

From Green's Second Identity we get the following:

$$\begin{aligned} \int_{M_a} g \Delta f &= - \int_{\|\vec{x}\|=1} f * dg + \int_{\|\vec{x}\|=a} f * dg - \int_{\|\vec{x}\|=a} g * df \\ &= (n-2) \cdot V_{n-1}(S^{n-1}) \cdot \text{avg}_{\|\vec{x}\|=1} f - (n-2) \cdot V_{n-1}(S^{n-1}) \cdot \text{avg}_{\|\vec{x}\|=a} f - (a^{2-n} - 1) \int_{\|\vec{x}\|=a} *df \end{aligned}$$

Note here we have the following by Stokes' Theorem:

$$(a^{2-n} - 1) \int_{\|\vec{x}\|=a} *df = (a^{2-n} - 1) \int_{\|\vec{x}\|\leq a} \Delta f = (a^2 - a^n) \cdot V(\bar{B}_1(\vec{0})) \cdot \text{avg}_{\|\vec{x}\|\leq a} \Delta f$$

Hence we see that, when a tends to 0, we get the following:

$$\lim_{a \rightarrow 0} \int_{M_a} g \Delta f = (n-2) \cdot V_{n-1}(S^{n-1}) \left(\text{avg}_{\|\vec{x}\|=1} f - f(\vec{0}) \right)$$

Concluding we get the following result for all $f \in C^2(\bar{B}_1(\vec{0}), \mathbb{R})$, with $n > 2$:

$$\text{avg}_{\|\vec{x}\|=1} f = f(\vec{0}) + \frac{\int_{\bar{B}_1(\vec{0}) \setminus \{0\}} ((\|\vec{x}\|^{2-n} - 1)\Delta f)}{(n-2)V_{n-1}(S^{n-1})}$$

we also observe that we have the followings hold:

1. If $\Delta f(\vec{x}) \geq 0$ for $\vec{x} \in \bar{B}_1(\vec{0})$, then $\text{avg}_{S^{n-1}} f > f(\vec{0})$, and f is said to be subaveraging at $\vec{0}$.

2. If $\Delta f(\vec{x}) \leq 0$ for $\vec{x} \in \bar{B}_1(\vec{0})$, then $\text{avg}_{S^{n-1}} f < f(0)$, and f is said to be superaveraging at $\vec{0}$.
3. If $\Delta f(\vec{x}) = 0$ for $\vec{x} \in \bar{B}_1(\vec{0})$, then $\text{avg}_{S^{n-1}} f = f(0)$.

When $n = 2$, one can get similar result using $g : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ $x \mapsto -\ln \|\vec{x}\|$.

One can also get corresponding results for other balls in \mathbb{R}^n by using translation and dilation.

The case for $n = 1$ and $n = 2$ is left for readers, and concluding all results above, in general dimension n , we get the following theorem:

Theorem 45.1 (Gauss' Mean Value Theorem)

Let $f \in C^2(A, \mathbb{R})$ where A is an open subset of \mathbb{R}^n , f is harmonic on A if and only if the mean value property holds for all closed balls in A :

$$\text{avg}_{\|\vec{x} - \vec{x}_0\|=r} f = f(\vec{x}_0)$$

Here $\{\vec{x} \mid \|\vec{x} - \vec{x}_0\| \leq r\} \subseteq A$ is a closed ball of radius r centered at $\vec{x}_0 \in A$.

Corollary 45.1.1

Let $\vec{a} \in \mathbb{R}^n$, let $f \in C^2(\bar{B}_r(\vec{a}), \mathbb{R})$, then we get the followings:

1. If $\Delta f(\vec{x}) \geq 0$ for $\vec{x} \in \bar{B}_r(\vec{a})$, and $\exists \vec{p} \in \bar{B}_r(\vec{a})$ such that $\Delta f(\vec{p}) > 0$, then $\text{avg}_{\partial \bar{B}_r(\vec{a})} f > f(\vec{a})$
2. If $\Delta f(\vec{x}) \leq 0$ for $\vec{x} \in \bar{B}_r(\vec{a})$, and $\exists \vec{p} \in \bar{B}_r(\vec{a})$ such that $\Delta f(\vec{p}) < 0$, then $\text{avg}_{\partial \bar{B}_r(\vec{a})} f < f(\vec{a})$
3. If $\Delta f(\vec{x}) = 0$ for $\vec{x} \in \bar{B}_r(\vec{a})$, then $\text{avg}_{\partial \bar{B}_r(\vec{a})} f = f(\vec{a})$

Definition 45.1.1.0.1

Let $f \in C^2(A, \mathbb{R})$ where A is an open subset of \mathbb{R}^n .

If $\Delta f \geq 0$ on A , then f is said to be subharmonic on A .

If $\Delta f \leq 0$ on A , then f is said to be superharmonic on A

Gravitational Force Fields

Gravitational force field is a conservative force field. A conservative force field can be defined as the negative gradient of a potential function. That is, let \vec{F} denote such conservative force field, and let P denote such potential function, we can write:

$$\vec{F} = \text{grad}(P) = -(dP)^T$$

For simplicity, here we denote the potential of a unit mass at $\vec{y} \in \mathbb{R}^3$ by the following:

$$P_{\vec{y}} : \mathbb{R}^3 \setminus \{\vec{y}\} \rightarrow \mathbb{R} \quad \vec{x} \mapsto -\frac{1}{4\pi \|\vec{x} - \vec{y}\|}$$

note here $P_{\vec{y}}$ is not defined at \vec{y} , but $P_{\vec{y}}$ is harmonic elsewhere.

In other words, we have $\nabla^2 P_{\vec{y}} = 0$ on $\mathbb{R}^3 \setminus \{\vec{y}\}$.

In this section, we will denote a closed ball in \mathbb{R}^3 of radius 1 centered at $\vec{0}$ as B^3 , and we denote a closed sphere in \mathbb{R}^3 of radius 1 centered at $\vec{0}$ as S^2 .

The potential of a unit mass spread over the unit sphere S^2 is given by:

$$P_{S^2} : \mathbb{R}^3 \rightarrow \mathbb{R} \quad \vec{x} \mapsto \text{avg}_{\vec{y} \in S^2} P_{\vec{y}}(\vec{x}) \quad \text{where } \text{avg}_{\vec{y} \in S^2} P_{\vec{y}}(\vec{x}) = -\frac{1}{16\pi^2} \int_{\vec{y} \in S^2} \frac{dV}{\|\vec{x} - \vec{y}\|}$$

The potential of a unit mass spread over a unit ball B^3 is given by:

$$P_{B^3} : \mathbb{R}^3 \rightarrow \mathbb{R} \quad \vec{x} \mapsto \text{avg}_{\vec{y} \in B^3} P_{\vec{y}}(\vec{x}) \quad \text{where } \text{avg}_{\vec{y} \in B^3} P_{\vec{y}}(\vec{x}) = -\frac{3}{16\pi^2} \int_{\vec{y} \in B^3} \frac{dV}{\|\vec{x} - \vec{y}\|}$$

For $\vec{x} \in \mathbb{R}^3$ with $\|\vec{x}\| > 1$, one can use Gauss' Mean Value Property and get the following:

$$P_{S^2}(\vec{x}) = P_{\vec{0}}(\vec{x})$$

That is, at a point outside S^2 , the shell S^2 of unit mass is indistinguishable from a unit point mass at the origin. Hence at a point outside the surface of the Earth, the gravitational potential of the Earth can be approximated as the gravitational potential of a point of the same mass as the Earth located at the center of the Earth.

Theorem 46.1

Given the potential of a unit mass spread over the unit sphere centered at $\vec{0}$, denoted as P_{S^2} , for $\vec{x} \in \mathbb{R}^3$ with $\|\vec{x}\| < 1$, we have $P_{S^2}(\vec{x}) = -\frac{1}{4\pi}$, in which case the gravitational force is zero at \vec{x} .

Proof. Consider $M_\epsilon = \{\vec{y} \in \mathbb{R}^3 \mid \epsilon \leq \|\vec{y}\| \leq 1, \|\vec{y} - \vec{x}\| \geq \epsilon\}$, one can choose small enough $\epsilon > 0$ such that we have the following holds:

$$\partial M = S^2 \cup \{\vec{y} \mid \|\vec{y}\| = \epsilon\} \cup \{\vec{y} \mid \|\vec{y} - \vec{x}\| = \epsilon\}$$

By Lemma AF.4 from Math 396 Supplement Material, and noting that $\frac{1}{\|\vec{y}\|} - 1 = 0$ for $\vec{y} \in S^2$, we get the following:

$$\begin{aligned} \int_{\vec{y} \in S^2} \frac{dV}{\|\vec{x} - \vec{y}\|} &= \int_{\vec{y} \in S^2} \left(\frac{-1}{\|\vec{x} - \vec{y}\|} * d\left(\frac{1}{\|\vec{y}\|} - 1\right) \right) \\ &= \int_{\vec{y} \in S^2} \left(\frac{-1}{\|\vec{x} - \vec{y}\|} * d\left(\frac{1}{\|\vec{y}\|} - 1\right) + \left(\frac{1}{\|\vec{y}\|} - 1\right) * d\left(\frac{1}{\|\vec{x} - \vec{y}\|}\right) \right) \quad (*) \end{aligned}$$

Note here $\beta : \mathbb{R}^3 \setminus \{\vec{0}\} \rightarrow \mathbb{R}$ $\vec{y} \mapsto \frac{1}{\|\vec{y}\|} - 1$ is harmonic on $\mathbb{R}^3 \setminus \{\vec{0}\}$, and $\alpha : \mathbb{R}^3 \setminus \{\vec{x}\} \rightarrow \mathbb{R}$ $\vec{y} \mapsto \frac{1}{\|\vec{x} - \vec{y}\|}$ is harmonic on $\mathbb{R}^3 \setminus \{\vec{x}\}$. Then by Green's Second Identity, and making use of M_ϵ defined previously, we get the following:

$$0 = \int_{M_\epsilon} \alpha \Delta \beta - \beta \Delta \alpha = \int_{\partial M_\epsilon} \alpha * d\beta - \beta * d\alpha$$

Combining with equation (*) we get the following:

$$\begin{aligned} \int_{\vec{y} \in S^2} \frac{dV}{\|\vec{x} - \vec{y}\|} &= + \int_{\|\vec{y} - \vec{x}\|=\epsilon} \frac{-1}{\|\vec{x} - \vec{y}\|} * d\left(\frac{1}{\|\vec{y}\|} - 1\right) \\ &\quad + \int_{\|\vec{y} - \vec{x}\|=\epsilon} \left(\frac{1}{\|\vec{y}\|} - 1\right) * d\left(\frac{1}{\|\vec{x} - \vec{y}\|}\right) \\ &\quad + \int_{\|\vec{y}\|=\epsilon} \frac{-1}{\|\vec{x} - \vec{y}\|} * d\left(\frac{1}{\|\vec{y}\|} - 1\right) \\ &\quad + \int_{\|\vec{y}\|=\epsilon} \left(\frac{1}{\|\vec{y}\|} - 1\right) * d\left(\frac{1}{\|\vec{x} - \vec{y}\|}\right) \end{aligned}$$

where we have the followings by Stokes' Theorem and Gauss' Mean Value Theorem:

$$\int_{\|\vec{y} - \vec{x}\|=\epsilon} \frac{-1}{\|\vec{x} - \vec{y}\|} * d\left(\frac{1}{\|\vec{y}\|} - 1\right) = -\frac{1}{\epsilon} \int_{\|\vec{y} - \vec{x}\| \leq \epsilon} d * d\left(\frac{1}{\|\vec{y}\|} - 1\right) = -\frac{1}{\epsilon} \int_{\|\vec{y} - \vec{x}\| \leq \epsilon} \Delta \left(\frac{1}{\|\vec{y}\|} - 1\right) = 0$$

$$\int_{\|\vec{y} - \vec{x}\|=\epsilon} \left(\frac{1}{\|\vec{y}\|} - 1\right) * d\left(\frac{1}{\|\vec{x} - \vec{y}\|}\right) = -\frac{1}{\epsilon^2} \int_{\|\vec{y} - \vec{x}\|=\epsilon} \left(\frac{1}{\|\vec{y}\|} - 1\right) dV = -4\pi \left(\frac{1}{\|\vec{x}\|} - 1\right)$$

$$\int_{\|\vec{y}\|=\epsilon} \frac{-1}{\|\vec{x} - \vec{y}\|} * d\left(\frac{1}{\|\vec{y}\|} - 1\right) = \frac{1}{\epsilon^2} \int_{\|\vec{y}\|=\epsilon} \frac{1}{\|\vec{x} - \vec{y}\|} dV = 4\pi \frac{1}{\|\vec{x}\|}$$

$$\int_{\|\vec{y}\|=\epsilon} \left(\frac{1}{\|\vec{y}\|} - 1\right) * d\left(\frac{1}{\|\vec{x} - \vec{y}\|}\right) = \left(\frac{1}{\epsilon} - 1\right) \int_{\|\vec{y}\| \leq \epsilon} \Delta \left(\frac{1}{\|\vec{x} - \vec{y}\|}\right) = 0$$

so we get:

$$\int_{\vec{y} \in S^2} \frac{dV}{\|\vec{x} - \vec{y}\|} = 4\pi \quad \Rightarrow \quad P_{S^2}(\vec{x}) = -\frac{1}{16\pi^2}(4\pi) = -\frac{1}{4\pi}$$

This completes the proof of Theorem 46.1 \square

Summarizing the results above, we get the following for $\vec{x} \in \mathbb{R}^3$:

$$\int_{\vec{y} \in S^2} \frac{dV}{\|\vec{x} - \vec{y}\|} = \begin{cases} \frac{4\pi}{\|\vec{x}\|} & \|\vec{x}\| > 1 \\ 4\pi & \|\vec{x}\| < 1 \end{cases}$$

Generalizing, for $\vec{x} \in \mathbb{R}^3$, we get the following:

$$\int_{\vec{y} \in \{\vec{p} \mid \|\vec{p}\|=r\}} \frac{dA}{\|\vec{x} - \vec{y}\|} = \begin{cases} \frac{4\pi r^2}{\|\vec{x}\|} & \|\vec{x}\| > r \\ 4\pi r & \|\vec{x}\| < r \end{cases}$$

Now for $\vec{x} \in \mathbb{R}^3$, to compute $P_{B^3}(\vec{x})$, one might use the Coarea Formula in section P on Math 396 Supplement material. Alternatively, we can compute the following integral:

$$P_{B^3}(\vec{x}) = -\frac{3}{16\pi^2} \int_0^1 \left(\int_{||\vec{y}||=r} \frac{dA}{||\vec{x} - \vec{y}||} \right) dr$$

For $\vec{x} \in \mathbb{R}^3$ with $||\vec{x}|| > 1$, we get the following:

$$P_{B^3}(\vec{x}) = -\frac{3}{16\pi^2} \int_0^1 \frac{4\pi r^2}{||\vec{x}||} dr = -\frac{1}{4\pi ||\vec{x}||} = P_0(\vec{x})$$

For $\vec{x} \in \mathbb{R}^3$ with $||\vec{x}|| < 1$, we get the following:

$$P_{B^3}(\vec{x}) = -\frac{3}{16\pi^2} \left(\int_0^{||\vec{x}||} \frac{4\pi r^2}{||\vec{x}||} dr + \int_{||\vec{x}||}^1 4\pi r dr \right) = -\frac{||\vec{x}||^2}{4\pi} - \frac{3}{4\pi} \left(\frac{1}{2} - \frac{1}{2} ||\vec{x}||^2 \right) = \frac{||\vec{x}||^2}{8\pi} - \frac{3}{8\pi}$$

So the gravitational force field inside the unit ball is given by the following:

$$\vec{F} : \mathbb{R}^3 \rightarrow \mathbb{R} \quad \vec{x} \mapsto -\frac{\vec{x}}{4\pi}$$

Consider a unit mass particle whose position is described by the vector-valued function $\vec{x} : [0, \infty) \rightarrow \mathbb{R}^3$, then the acceleration of the particle is given by the following:

$$\ddot{\vec{x}} = -\frac{1}{4\pi} \vec{x} \tag{GDE}$$

The solution to ODE (GDE) describes a simple harmonic motion of such particle.²

²For more details and discussion about simple harmonic motions, and motions under the gravitational acceleration, check out *Chapter 2 - Classical Physics Review* in the text *Class Note: Physics 401 - Classical Mechanics*, transcribed by Jinyan Miao in Winter 2022.

Electromagnetic Theory

Electric field in \mathbb{R}^3 can be defined by a vector-valued function $E : \mathbb{R}^4 \rightarrow \mathbb{R}^3$, here we denote $E = (E_1, E_2, E_3)$. Similarly, magnetic field in \mathbb{R}^3 can be defined by a vector-valued function $B : \mathbb{R}^4 \rightarrow \mathbb{R}^3$, denoted as (H_1, H_2, H_3) . Here the domains of E and B are \mathbb{R}^4 , which can be viewed as the Cartesian product $\{\text{position in } \mathbb{R}^3\} \times \{\text{time}\}$. Hence we consider the \mathbb{R}^4 space, with points in \mathbb{R}^4 denoted as (x, y, z, t) .

After normalizing the units, the differential form of the Maxwell's Equations is given by:³

$$\begin{aligned} \operatorname{div}(E) &= \rho & \operatorname{curl}(E) &= -\frac{\partial H}{\partial t} & \operatorname{div}(H) &= 0 & \operatorname{curl}(H) &= j + \frac{\partial E}{\partial t} \end{aligned} \quad (\text{MQ})$$

When computing the divergence and curl of E and B , we consider E and B as functions defined on \mathbb{R}^3 , with the position components (x, y, z) being the variables, and the time component is fixed at $t = t_0$. Specifically, $\operatorname{curl}(E) = -\frac{\partial H}{\partial t}$ is computed at a given time $t = t_0$, the RHS reads $\frac{\partial H}{\partial t}|_{t=t_0}$, and the LHS reads $\operatorname{curl}(E_{t_0})$ with $E_{t_0} : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \quad (x, y, z) \mapsto E(x, y, z, t_0)$. In the Maxwell's Equations, j is called the current density, ρ is called the charge density. Here $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a scalar-valued function defined on \mathbb{R}^4 , and $j : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ is a vector-valued function defined on \mathbb{R}^4 .

Now we define the following 2-forms:

$$\begin{aligned} \omega &= E_1 dx \wedge dt + E_2 dy \wedge dt + E_3 dz \wedge dt + H_1 dy \wedge dz + H_2 dz \wedge dx + H_3 dx \wedge dy \\ \eta &= H_1 dx \wedge dt + H_2 dy \wedge dt + H_3 dz \wedge dt - E_1 dy \wedge dz - E_2 dz \wedge dx - E_3 dx \wedge dy \end{aligned} \quad (\text{WN})$$

And a closed 3-form:

$$\lambda = j_1 dy \wedge dz \wedge dt + j_2 dz \wedge dx \wedge dt + j_3 dx \wedge dy \wedge dt - \rho dx \wedge dy \wedge dz \quad (\text{WN})$$

$\operatorname{curl}(E) = -\frac{\partial H}{\partial t}$, and $\operatorname{div}(H) = 0$ now requires $d\omega = 0$. $\operatorname{div}(E) = \rho$ and $\operatorname{curl}(H) = j + \frac{\partial E}{\partial t}$ now requires $d\eta = \lambda$. Hence Maxwell's Equations (MQ) now becomes the following:

$$d\omega = 0 \quad d\eta = \lambda \quad (\text{MQ})$$

Note that we must have necessarily $d\lambda = 0$ because $d\eta = \lambda$.

We observe that the relation between ω and η is reminiscent of the Hodge star operator, but in fact $\eta \neq \pm * \omega$. However, we can introduce the \circledast operator such that we have $\eta = \circledast \omega$.

In fact, we can define:

$$R : \mathbb{R}^4 \rightarrow \mathbb{R}^4 \quad \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \\ z \\ -t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}$$

³For more details and discussion about electric fields and potentials, magnetic fields and vector potentials, and the Maxwell's Equations, check out *Class Note: Physics 405 - Intermediate Electricity and Magnetism* transcribed by Jinyan Miao in Winter 2022.

In which case we get $\circledast\omega = -R^*(\omega)$.

Definition 47.0.0.0.1

For the following forms defined on \mathbb{R}^4 :

0-form f

1-form $n = \alpha dx + \beta dy + \gamma dz + \delta dt$

2-form $t = A dy \wedge dz + B dz \wedge dx + C dx \wedge dy + D dx \wedge dt + E dy \wedge dt + F dz \wedge dt$

3-form $h = \alpha dy \wedge dz \wedge dt + \beta dz \wedge dx \wedge dt + \gamma dx \wedge dy \wedge dt + \delta dx \wedge dy \wedge dz$

4-form $r = f dx \wedge dy \wedge dz \wedge dt$

we define the \circledast operator such that we have:

$$\circledast f = f dx \wedge dy \wedge dz \wedge dt$$

$$\circledast n = \alpha dy \wedge dz \wedge dt + \beta dz \wedge dx \wedge dt + \gamma dx \wedge dy \wedge dt + \delta dx \wedge dy \wedge dz$$

$$\circledast t = A dx \wedge dt + B dy \wedge dt + C dz \wedge dt - D dy \wedge dz - E dz \wedge dx - F dx \wedge dy$$

$$\circledast h = \alpha dx + \beta dy + \gamma dz + \delta dt$$

$$\circledast r = -f$$

Lemma 47.0.1

For k -forms μ, ζ defined on \mathbb{R}^4 , and 0-form f defined on \mathbb{R}^3 , we get the followings:

1. $\circledast(f\mu) = f \circledast \mu$
2. $\circledast(\mu + \zeta) = \circledast\mu + \circledast\zeta$
3. $\circledast \circledast \mu = (-1)^{1+\deg(\mu)} \mu$

Lemma 47.0.2

For $k \leq 4$, let μ be a k -form defined on \mathbb{R}^4 , for Φ defined on an open subset of \mathbb{R}^4 , we have $\circledast\Phi^*\mu = \Phi^* \circledast \mu$ holds if and only if Φ extends to an orientation preserving diffeomorphism of \mathbb{R}^4 satisfying the following condition:

$$\|\Phi(\vec{p}) - \Phi(\vec{q})\|_t^2 = \|\vec{p} - \vec{q}\|_t^2$$

for $\vec{p}, \vec{q} \in \mathbb{R}^4$ and we have:

$$d_t : \mathbb{R}^4 \rightarrow \mathbb{R} \quad (x, y, z, t) \mapsto \|(x, y, z, t)\|_t^2 := x^2 + y^2 + z^2 - t^2$$

It is clear that Φ satisfies the conditions above if and only if Φ is an orientation preserving affine map whose linear part preserves the distance defined by d_t . Here d_t defines the space-time interval.⁴

For η and ω defined by equation (WN), we now have $\eta = \circledast\omega$, and hence $\omega = -\circledast\eta$, then we get:

$$d\eta = \lambda \quad d \circledast \eta = 0$$

Hence Maxwell's Equations (MQ) now read s the following:

$$d\omega = 0 \quad d \circledast \omega = \lambda \tag{MQ}$$

⁴Space-time interval is closely related to the topic of Einstein's Special Relativity, for more discussion on Einstein's Special Relativity, check out *Class Note: Physics 360 - Honors Physics III*, transcribed by Jinyan in Fall 2021.

If we are working on a star-convex set, such as all of \mathbb{R}^4 , then by Poincare Lemma, or Theorem 39.3 on Munkres, we can write $\omega = d\phi$ for some 1-form ϕ defined on \mathbb{R}^4 , and hence Maxwell's Equations (MQ) reads the following in this special case:

$$\omega = d\phi \quad d \circledast d\phi = \lambda \quad (\text{MQS})$$

To simplify equation (MQS), one introduces the d'Alembertian \square by the following:

Definition 47.0.2.0.1

For k -form μ defined on \mathbb{R}^4 , we define $\square\mu$ as the following:⁵

$$\square\mu := - \circledast d \circledast d\mu - d \circledast d \circledast \mu$$

Exercise: For 1-form f defined on \mathbb{R}^4 , we have the following holds:

$$\square f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} - \frac{\partial^2 f}{\partial t^2}$$

Exercise: For k -form $\mu = \sum_{[I]} b_I dx_I$ defined on \mathbb{R}^4 , we have the following holds:

$$\square\mu = \sum_I (\square b_I) dx_I$$

Note here we also have $\eta = \circledast d\phi$. Now we get Maxwell's Equations, if working on a star-convex set in \mathbb{R}^4 , reads the following:

$$d \circledast \phi = 0 \quad \square\phi = - \circledast d \circledast d\phi = - \circledast \lambda \quad \eta = \circledast d\phi \quad (\text{MQS})$$

For the homogeneous situation where $\lambda = 0$, we would obtain vacuum solution from Maxwell's Equations (MQS). Here we write: $\phi = \alpha dx + \beta dy + \gamma dz + \delta dt$. To solve (MQS), here we require the followings hold:

$$\square\alpha = \square\beta = \square\gamma = \square\delta = 0, \quad \frac{\partial\alpha}{\partial x} + \frac{\partial\beta}{\partial y} + \frac{\partial\gamma}{\partial z} - \frac{\partial\delta}{\partial t} = 0$$

In particular, one could take $\phi = f(x+t) dy$ for some reasonable function f , and we get:

$$\omega = g'(x+t) dx \wedge dy - g'(x+t) dy \wedge dt$$

yielding the following:

$$E = (0, -g'(x+t), 0) \quad H = (0, 0, g'(x+t)) \quad (\text{WEQ})$$

Equation (WEQ) characterizes a electromagnetic wave traveling in \mathbb{R}^3 space of unit speed.

⁵"Box invented this notation"

- Xun Wang