Honors Analysis I

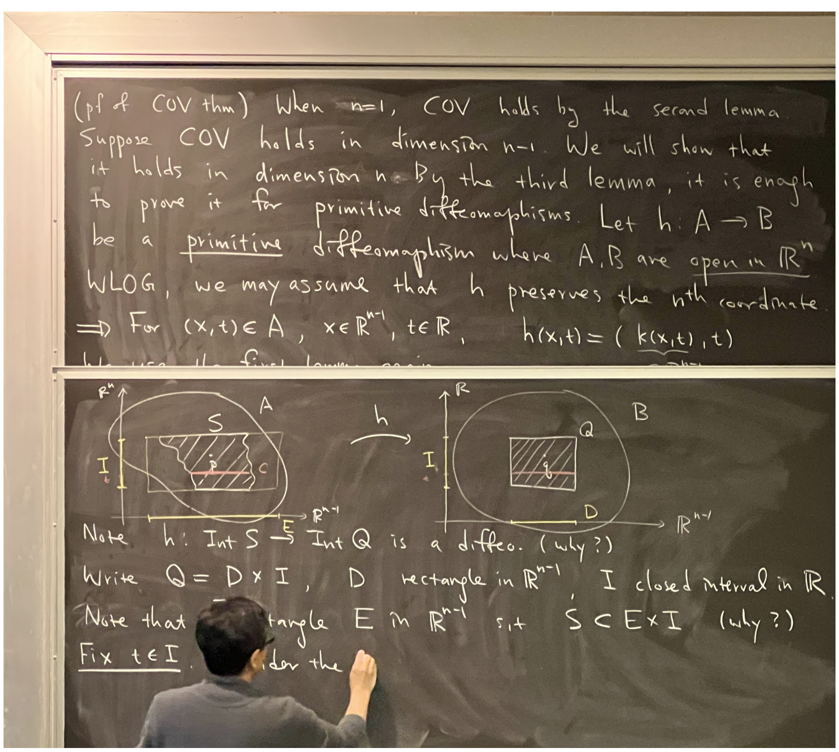
Math 395

University of Michigan

Harrison Centner

Prof. Jinho Baik

January 4, 2023



Contents

INTRODUCTION & MOTIVATION

Course Logistics and Syllabus

$\frac{2}{3}$ of the class will focus on real functions of several variables.

$\frac{1}{3}$ of thee class will be differential equations.

Motivation for Study of Several Variables

Functions of several variables describe many things.

$f : \mathbb{R}^3 \to \mathbb{R}$ assigns each triple of numbers to a number. Examples: air pressure, gravitational potential, contraction/movement of a robots joints, whether a jpeg image has a cat or not.

The robot arm is a function of θ_1 and θ_2 and the distance of the arms.

Vector-valued functions of a single variable $\vec{F} : \mathbb{R} \to \mathbb{R}^3$ can describe a curve in three dimensions (giving location of a satellite).

Vector-valued functions of several variables $g : \mathbb{R}^2 \to \mathbb{R}^2$ could describe a wind map over Ann Arbor.

$f : \mathbb{R}^3 \setminus \{0, 0, 0\} \to \mathbb{R}^3$ could describe gravitational force under a point mass. This gives a force field.

$$F(x, y, z) = -\frac{GM \cdot m}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

We want to study differentiation and integration theory for these functions.

TODO: Read chapter one of Munkres Text sections 1-4 by Labor Day.

Definition. (Product of two sets)
For two sets A and B, their product is a uniquely defined set $A \times B = \{(a, b) \mid a \in A, b \in B\}$.
The set product is associative.

LIMITS & CONTINUITY IN METRIC SPACES

Definition. (Metric Space)
A metric on X is a function $d : X \times X \to \mathbb{R}$ which satisfies

(i) $d(x, y) = d(y, x)$
(ii) $d(x, y) \geq 0$ and equality iff $x = y$
(iii) $d(x, y) + d(y, z) \geq d(x, z)$

The pair (X, d) is called the **metric space**.

Examples.

(1) $\mathbb{R}$; $d(x, y) = |x - y|$

(2) $\mathbb{R}$; $d(x, y) = |\int_x^y e^{-t}\, dt|$

(3) $\mathbb{R}^m$; $d(x, y) = \max\{|x_1 - y_1|, \ldots, |x_m - y_m|\}$[1] Sup metric

(4) $\mathbb{R}^m$; $d(x, y) = |x_1 - y_1| + \cdots + |x_m - y_m|$ L^1 metric

(5) $C([0, 1]) = \{f : [0, 1] \to [0, 1] \mid f \text{ is continuous}\}$; $d(f, g) = \max_{0 \le t \le 1}\{|f(t) - g(t)|\}$

(6) $C([0, 1]) = \{f : [0, 1] \to [0, 1] \mid f \text{ is continuous}\}$; $d(f, g) = \int_0^1 |f(t) - g(t)|\, dt$

(7) $X = S^2$; $d(x, y) =$ arc length of geodesic with minimal length.

Definition. (Topological Notions)
Let (X, d) be a metric space.

(a) For $\varepsilon > 0$, the **ε-neighborhood** of $x_0 \subseteq X$ is

$$\cup(x_0, \varepsilon) = \{u \in X \mid d(x_o, u) < \varepsilon\}$$

(b) A subset Ω of X is **open** in X if $\forall x \in \Omega$, $\exists \varepsilon = \varepsilon(x_0)$ such that $\cup(x_0, \varepsilon) \subset \Omega$

(c) A subset C of X is said to be **closed** in X if $X \setminus C$ is open.

(d) A metric space X is said to be **compact** if for every collection of open sets whose union is X, there is a finite subcollection of these sets whose union is again X.

(e) A metric space X is said to be **connected** if one cannot have $X = A \cup B$ for some $A, B \subseteq X$ where A, B are disjoint and nonempty. Equivalently, the only **clopen** (closed and open) sets in X are X and $\varnothing$.

Theorem. If $A \subseteq \mathbb{R}^n$ is open in the Euclidean metric, then A is open in any metric on $\mathbb{R}^n$. The converse is also true. Note that this is not true of $C([0, 1])$.

[1] This is the supremum of the Euclidean metric as $n \to \infty$

Limits and Continuity

Lemma. (Induced metric spaces)
Let (X, d) be a metric space and $Y \subseteq X$. Then $d \upharpoonright Y \times Y$ is a metric on Y.

Recall that open and closed are properties relative to a background set.

Lemma. (Induced metric spaces)
Let (X, d) be a metric space and $Y \subseteq X$. Suppose:

(a) Let $A \subseteq Y$. Then A be open in $Y \iff \exists \mathcal{O}$, open in X such that $A = \mathcal{O} \cap Y$.
(b) If "open" is replaced with "closed," then (a) also holds.
(c) Y is compact $\iff$ for every open (in X) cover of Y there is a finite subcover of X.

Theorem. (Heine-Borel & Bolzano Weierstrauss)
Let $K \subseteq \mathbb{R}^n$.
Then, K is compact $\iff$ K is closed and bounded in $\mathbb{R}^n$
$\iff$ K is sequentially compact.

Remark: In general metric space, sequential compactness $\iff$ compact $\implies$ closed & bounded.

Definition. (Convexity)
$S \subseteq \mathbb{R}^n$ is said to be **convex** if the chord between any two points in S is contained in S.

$$\forall p, q \in S, \quad (1-t)p + tq \subseteq S \qquad t \in [0,1].$$

Lemma. (Connectedness of Convex Sets)
A convex set in $\mathbb{R}^n$ is connected.

Definition. (Limits and Continuous Functions)
Let Z, Y be a metric spaces and $X \subseteq Z$.

(i) A point $p \in Z$ is said to be a **limit point** of X if every ε-neighborhood of p intersects X nontrivially.
(ii) Let $f : X \to Y$. Suppose p is a limit point of X and $q \in Y$. We say that $f(x) \to q$ as $x \to p$ (alternatively $\lim_{x\to p} f(x) = q$) if $\forall \varepsilon > 0, \exists \delta$ such that

$$0 < d_X(x, p) < \varepsilon \implies d_Y(f(x), q) < \varepsilon$$

$$x \in U_X(p, \delta) - \{p\} \implies f(x) \in U_Y(q, \varepsilon)$$

Remark: If f(x) converges as $x \to p$, the **limit** is unique.
Also

$$\lim_{x\to p} f(x) = q \iff \lim_{d_X(x,p)\to 0} d_Y(f(x), q) = 0.$$

Notation: We denote $|x - y|$ to be the distance between $x, y \in \mathbb{R}^n$ using the sup metric.

Examples: Let f be a function from the punctured Euclidean plane to $\mathbb{R}$.

(1) $f(x,y) = \frac{x^2y}{x^2+y^2}$ claim limit is 0.
Proof.
Fix $\varepsilon > 0$. Let $\delta = \varepsilon$. Suppose $0 < |(x,y)| < \delta$. Then, $|x|, |y| < \delta = \varepsilon$.

$$\left|\frac{x^2y}{x^2+y^2}\right| = \left|\frac{x^2|y|}{x^2+y^2}\right| \leq |y| < \varepsilon$$

(2) $f(x,y) = \frac{x^2}{x^2+y^2}$. In this case the limit does not exist because as we approach from the y-axis we find a different limit then an approach from the x-axis. Since limits are unique, the limit cannot exist at zero.

(3) $f(x,y) = \frac{x^2y}{x^4+y^2}$. If we follow the parabolic trajectory $f(t,t^2)$ we get a different limit than following $f(x,0)$. The limit does not exist at zero.

Lemma. (Properties of limits and continuous functions)

(a) For $f : X \subseteq Z \to \mathbb{R}^n$ with $f(x) = (f_1(x), \ldots, f_n(x))$,

$$\lim_{x\to p} f(x) = q \iff \forall i \in [n], \lim_{x\to p} f_i(x) = q_i$$

(b) Addition, multiplication, and division of limits behave as expected where they are defined.

Definition. (Continuous functions)
Let X, Y be metric spaces and $p \in X$ we say $f : X \to Y$ is **continuous at** p if $\forall \varepsilon, \exists \delta$ s.t.

$$x \in X \text{ and } d_Y(x,p) < \delta \implies d_Y(f(x), f(y)) < \varepsilon.$$

Lemma. (Properties of limits and continuous functions)
If p is an isolated point of X, then f is continuous at p.
Otherwise, f is continuous at $p \iff \lim_{x\to p} f(x) = f(p)$.

Continuous Functions

Definition. (Continuity)
We say that $f : X \to Y$ is **continuous** at $p \in X$ provided that $\forall \varepsilon > 0, \exists \delta > 0$ s.t.

$$x \in U_X(p,\delta) \implies f(x) \in U_Y(f(p), \varepsilon).$$

If p is not an isolated point then this is equivalent to

$$\lim_{x\to p} f(x) = f(p) \iff \lim_{x\to p} d_Y(f(x), f(p)) = 0.$$

Lemma. (Continuity Behaves as Expected)

(i) Let $f : X \to Y$ be continuous at $p \in X$ and $\phi : Y \to Z$ continuous at $f(p)$. Then $\phi \circ f : X \to Z$ is continuous.
(ii) Suppose $f : X \to \mathbb{R}^n$ is continuous at $p \in X$ iff $\forall i \in [n], f_i : X \to \mathbb{R}$ is continuous at $p \in X$.
(iii) If $f, g : X \to Y$ are continuous at $p \in X$ then $\alpha f(x) + \beta g(x), f(x)g(x), \frac{f(x)}{g(x)}$ are continuous where defined.
(iv) The projection map $\pi_i : \mathbb{R}^n \to \mathbb{R}$ given by $pi_i(x) = x_i$ is continuous on $\mathbb{R}^n$.
(v) Let $k < m$ and suppose $g : A \subseteq \mathbb{R}^k \to \mathbb{R}$ is continuous at $(p_1, \dots, p_k) \in A$. Let $f : A \times \mathbb{R}^{m-k} \subset \mathbb{R}^m$ be $f(x_1, \dots, x_n) = g(x_1, \dots, x_k)$. Then f is continuous at $(x_1, \dots, x_k, x_{k+1}, \dots, x_m)$ for every $x_{k+1}, \dots, x_n \in \mathbb{R}$. <u>We can extend a continuous function to a continuous function on more variables.</u>

Examples:

(a) $f : \mathbb{R}^3 \to \mathbb{R}^2$, $f(x, y, z) = (x^3yz, e^{x\sin(z)})$ is continuous.

(b) The function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given below is continuous on $\mathbb{R}^2$.

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

(c) The function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given below is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$.

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

Lemma. (Continuity via Topology)
Let X, Y be metric spaces. Then f is continuous on X
$\iff \forall \mathcal{O}$, open in Y, $f^{-1}(\mathcal{O})$ is open in X
$\iff \forall \mathcal{C}$, closed in Y, $f^{-1}(\mathcal{C})$ is closed in X .

<u>Note</u>: The last line follows since $f^{-1}(Y \setminus E) = X \setminus f^{-1}(E)$.
We also have the following facts:

(a)
$$f^{-1}(\cap_\alpha E_\alpha) = \cap_\alpha f^{-1}(E_\alpha)$$

(b)
$$f^{-1}(\cup_\alpha E_\alpha) = \cup_\alpha f^{-1}(E_\alpha)$$

Theorem. (Compactness, Connectedness, and Continuity)

(i) If $K \subseteq X$ is compact, then $f(K)$ is compact.
(ii) If $A \subseteq X$ is connected, then $f(A)$ is connected.
(iii) If X is compact, then f is uniformly continuous on X.

DIFFERENTIATION

Definition. (Directional Derivative)
Let $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ and $p \in \Omega$. For $u \in \mathbb{R}^m \setminus \{0\}$, the **directional derivative** of f at p **with respect to** u is

$$f'(p; u) = \lim_{t \to 0} \frac{f(p + tu) - f(p)}{t} = (g_1'(0), \dots g_m'(0))$$

if this limit exists.

Where $g_i(t) = f_i(p + tu)$.

Consider: $f : \mathbb{R}^2 \to \mathbb{R}$ and let $p = (a, b)$ and $u = (u_1, u_2) \neq (0, 0)$. The graph of f is a subset of $\mathbb{R}^3$. The intersection of the graph of f and the affine plane $p + \mathrm{Span}(u)$ yields a curve given by the graph $g(t) = f(p + tu)$. Then $g'(0)$ gives the directional derivative at p with respect to u.

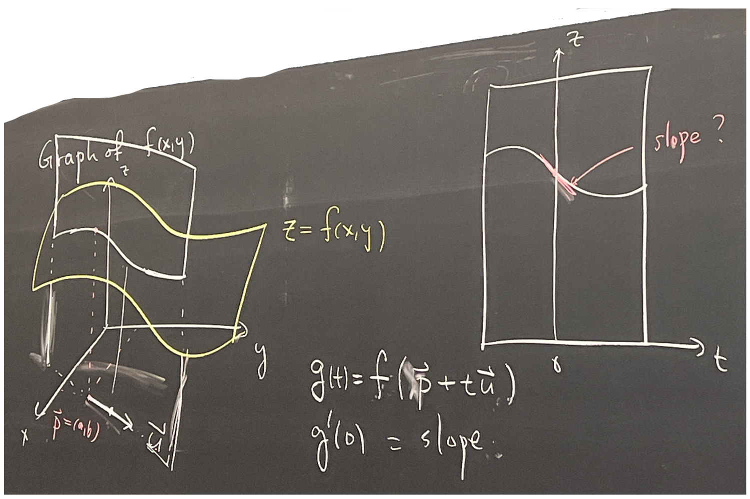


Figure 1: Intuition for a directional derivative.

Examples:

(a) $f : \mathbb{R}^2 \to \mathbb{R}^3$ given by $f(x, y) = (xy, e^{x+y}, x)$.
Then, $f'((x, y); (u_1, u_2)) = (xu_2, +yu_1, e^{x+y}(u_1 + u_2), u_1)$

(b) The function $f : \mathbb{R}^2 \to \mathbb{R}^2$ is not differentiable at $(0, 0)$.

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

$$g(t) = f(0 + tu_1, 0 + t_u 2) = \begin{cases} \frac{tu_1^2 u_2}{t^2 u_1^4 + u_2^2} & , t \neq 0 \\ 0 & , t = 0 \end{cases}$$

We compute the directional derivative at zero with respect to (u_1, u_2).

$$g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \begin{cases} \frac{u_1^2}{u_2} & \text{if } u_2 \neq 0 \\ 0 & \text{if } u_1 = 0 \end{cases}$$

The directional derivative at $(0,0)$ *exist* for all $u \neq 0$ even though f is not continuous at $(0,0)$.

Definition. (Differentiation) $(\star)$
Suppose $\phi : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ and Ω is open in $\mathbb{R}^m$.
We say that f is **differentiable** at $p \in \Omega$ if $\exists A \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ such that

$$\lim_{h\to 0} \frac{|f(p+h) - f(p) - Ah|}{|h|} = 0 \iff \lim_{h\to 0} \frac{f(p+h) - f(p) - Ah}{|h|} = \vec{0}$$

We say that A is the **total derivative** of f at p and we write $A = Df(p)$.

Notation: For $p \in \mathbb{R}^m$ we will denote

$$p = (p_1, \ldots, p_m) = \begin{bmatrix} p_1 \\ \vdots \\ p_m \end{bmatrix}$$

Lemma. (Properties of the Derivative)
Suppose $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at p.

(1) There is only one $A \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ such that $(\star)$ holds.
Proof.
It suffices to show that for $E \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$,

$$\lim_{h\to 0} \frac{|Eh|}{|h|}$$

(2) If f is continuous at p, then f is differentiable at p.
Proof.
Show

$$f(x) - f(p) = \frac{f(x) - f(p) - B(x-p)}{|x-p|}|x-p| + B(x-p)$$

Note that both terms on the RHS goes to 0 as $x \to p$. □

(3)

Examples:

(a) A constant function $f(x) = c$ is differentiable everywhere and $Df(x) = 0_{n\times n}$ for all $x \in \mathbb{R}^m$.

(b) A linear function $f(x) = \begin{bmatrix} M_{11}x_1 + \cdots + M_{1m}x_m + c_1 \\ \vdots \\ M_{n1}x_1 + \cdots + M_{nm} + c_N \end{bmatrix} = Mx + c$ is differentiable everywhere and $Df(x) = M$.

Theorem. (Directional Derivatives are Given by the Total Derivative)
If $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at p, then $f'(p;u)$ exist for all $u \neq 0$ and

$$\underbrace{f'(p;u)}_{n\times 1} = \underbrace{Df(p)}_{n\times m}\underbrace{u}_{m\times 1}$$

Proof.
Write $B = Df(p)$ since f is differentiable at p.

$$0 = \lim_{t\to 0} \frac{|f(p+tu) - f(p) - Btu|}{|tu|} = \lim_{t\to 0} \frac{|f(p+tu) - f(p) - Btu|}{|t||u|}$$

Multiply by $|u|$, to obtain

$$\lim_{t\to 0} \left| \frac{f(p+tu) - f(p) - Btu}{t} \right| = \lim_{t\to 0} \left| \frac{f(p+tu) - f(p)}{t} - Bu \right|$$

□

Notation: Let $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ be the j-th standard basis vector.

For a scalar function $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$, the j^{th} **partial derivative** of f at p is $f'(p;e_j)$.
We write $f;(p;e_j) = D_j f(p) = \frac{\partial f}{\partial x_j}(p)$.

Theorem. (Partial Derivatives are given by the Total Derivative)
If $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$, and we have $f(x) = (f_1(x), \ldots, f_n(x))$, then
(1) f is differentiable at $p \iff \forall i \in [n]$, f_i is differentiable at p.
(2) If f is differentaible at p, then

$$Df(p) = \begin{bmatrix} D_1 f_1(p) & \cdots & D_m f_1(p) \\ \vdots & & \vdots \\ D_1 f_n(p) & \cdots & D_m f_n(p) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_m}(p) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(p) & \cdots & \frac{\partial f_n}{\partial x_m}(p) \end{bmatrix} = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}$$

Proof.
Write $B = Df(p)$. By the previous theorem, $f'(p;e_j) = Be_j$.

$$\begin{bmatrix} f_1'(p;e_j) \\ \vdots \\ f_n'(p;e_j) \end{bmatrix} = \begin{bmatrix} B_{1j} \\ \vdots \\ B_{nj} \end{bmatrix}$$

So $B_{ij} = f_i(p;_j) = D_j f_i(p)$. □

Continuously Differentiable Functions

Recall:

Theorem. (Mean Value Theorem)
The **Mean Value Theorem** states that if $\phi : [a, b] \subseteq \mathbb{R} \to \mathbb{R}$, continuous on $[a, b]$ and differentiable on (a, b), then $\exists c \in (a, b)$ such that $\phi(c)(\phi(b) - \phi(a)) = \phi'(c)(b - a)$.

Proof.
Construct a function ψ such that $\psi(a) = \psi(b)$ and appeal to Rolle's Theorem. □

Definition. (Continuous Differentiability)
$f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is said to be **continuously differentiable** on Ω if $D_j f_i$ *exist* and *are continuous* on Ω, $\forall i, j$. We write that $f \in C^1(\Omega)$ or f is C^1 on Ω.

Theorem. ($C^1 \implies$ Differentiable)
If $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is differentiable on Ω then f is differentiable on Ω.

Proof.
It is enough to show that each component function is differentiable. WLOG we may assume $n = 1$. Fix $p \in \Omega$. We want to show $(\star)$ holds. Where the linear transformation is given by $B = \begin{bmatrix} D_1 f(p) & \cdots & D_m f(p) \end{bmatrix}$. Since Ω is open, there is a neighborhood U of p such that $U \subseteq \Omega$. If $|h|$ is small enough then $p + h \in \Omega$. We will use the mean value theorem (but we only can use the one variable version).

Let $p_0 = 0$, $p_1 = p + (h_1, 0, \ldots, 0)$ and $p_j = (h_1, h_2, \ldots, h_j, 0, \ldots, 0)$, so then $p_m = p + h$.

Then,

$$f(p + h) - f(p) = \sum_{i=1}^{m} (f(p_i) - f(p_{i+1})) = \sum_{i=1}^{m} D_i f(q_i) h_i$$

For some q_i that lies between p_{i-1} and p_i. Then, the last equality follows from MVT.

Now,

$$\begin{aligned} \frac{|f(p + h) - f(p) - Bh|}{|h|} &= \frac{|\sum_{i=1}^m D_i f(q_i) h_i - \sum_{i=1}^m D_i f(p) h_i|}{|h|} \\ &\leq \frac{\sum_{i=1}^m |D_i f(q_i) - D_i f(p)| \, |h_i|}{|h|} \\ &\leq \sum_{i=1}^{m} |D_i f(q_i) - D_i f(p)| \\ &= 0 \end{aligned}$$

Since $f \in C^1(\Omega)$, so $D_i f$ is continuous and $q_i \to p$ as $h \to 0$. □

<u>Examples:</u>

(a) $f(x, y) = x^2 y$. $D_1 f = 2xy$ and $D_2 f = x^2$. We can take further derivatives, $D_1 D_1 f = 2y$, $D_1 D_2 f = D_2 D_1 = 2x$, $D_2 D_2 f = 0$.

(b) $f(x,y) = (xy, e^{x+y}, x)$. We see that all the entries in Df are continuous. Explicitly,

$$Df = \begin{bmatrix} y & x \\ e^{x+y} & e^{x+y} \\ 1 & 0 \end{bmatrix}.$$

Since all six partial derivatives are continuous f is differentiable.

(c) The function $f : \mathbb{R}^2 \to \mathbb{R}^2$ is not continuous at $(0,0)$.

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

Note that $f(x,0) = 0$ for all $x \in \mathbb{R}$ which implies that $D_1 f(0,0) = 0$.
Similarly, $f(0,y) = 0$ for all $x \in \mathbb{R}$ which implies that $D_2 f(0,0) = 0$.
$D_1 f(x,y) = \frac{-2xy(x^4-y^2)}{(x^4-y^2)^2}$. Check that $\lim_{(x,y)\to(0,0)} D_1 f(x,y) \neq D_1 f(0,0)$.
But f is differentiable for all $v \in \mathbb{R}^2 \setminus \{(0,0)\}$. Thus, f is not C^1.

Notation: For $r \in \mathbb{N}$, we say that $f \in C^r(\Omega)$ if all partial derivatives of order r are continuous.

Theorem. (Clairout's Theorem)
If $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is C^2 on Ω, then $D_i D_j f = D_j D_i f$ for all $i, j \in [m]$ on Ω.
This implies that equality of mixed partials of order r holds for C^r functions on Ω.

Proof.
It is enough to consider $m = 2$. Fix $p = (a,b) \in \Omega$.

Let $h_1 h_2 > 0$. Consider

$$f(a+h_1, b+h_2) - f(a+h_1, b) - f(a, b+h_2) + f(a,b) \qquad (\heartsuit)$$

these four points form a rectangle.

Let $\phi(s) = f(s, b+h_2) - f(s,b)$ so that

$$(\heartsuit) = \phi(a+h_1) - \phi(a) = \phi'(s_1)h_1$$

by the MVT for some $s_1 \in (a, a+h_1)$.

Now

$$\phi'(s_1)h_1 = (D_1 f(s_1, b+h_2) - D_1 f(s_1, b))\, h_1 = D_2 D_1 f(s_1, t_1) h_1 h_2$$

for some $t_1 \in (b, b+h_2)$.

Let $\psi(t) = f(a+h_1, t) - f(a,t)$.

Following the same argument and applying MVT twice as before we see that $\psi(t) = D_1 D_2 f(s_2, t_2) h_1 h_2$ for some $s_2 \in (a, a+h_2)$ and $t_2 \in (b, b+h_2)$.

Since we obtained two ways of writing $(\heartsuit)$ we obtain

$$D_2 D_1 f(s_1, t_1) = D_1 D_2 f(s_2, t_2)$$

for some s_1, t_1, s_2, t_2 that depends on h and satisfies the box drawing.

Taking $h \to 0$ concludes the proof. $\square$

Definition. (Hessian Matrix)
For some $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$ (scalar valued), the **Hessian** of f is the $m \times m$ matrix

$$Hf = \begin{bmatrix} D_1D_1f & \cdots & D_1D_mf \\ \vdots & \ddots & \vdots \\ D_mD_1f & \cdots & D_mD_mf \end{bmatrix}$$

Hf is symmetric if f is C^2.

The Chain Rule

Theorem. (Chain Rule)
Suppose $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathcal{O} \subseteq \mathbb{R}^n \to \mathbb{R}^d$. Furthermore, $f(\Omega) \subseteq \mathcal{O}$ and f is differentiable at $p \in \Omega$ and g is differentiable at $q = f(p) \in \mathcal{O}$.
Then, $g \circ f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^d$ is differentiable at p and

$$\underbrace{D(g \circ f)(p)}_{d \times m} = \underbrace{Dg(f(p))}_{d \times n} \underbrace{Df(p)}_{n \times m}.$$

Corollary. If $f \in C^k(\Omega)$ and $g \in C^k(\Omega)$, then $g \circ f$ is C^k on Ω.
Proof.
Assume $f, g \in C^1(\Omega$. Let $h = g \circ f$. For each $i \in [d], j \in [n]$ we have

$$D_jh_i(p) = \sum_{k=1}^{n} D_kg_i(f(p))D_jf_k(p) = \sum_{k=1}^{n} (D_kg_i) \circ f(p)D_jf_k(p)$$

which is continuous (product, sum, composition).
Hence h is C^1.
Suppose the corollary holds for $k-1$. Let f, g be C^k. Let $h = g \circ f$. For all j, k D_kg_i is C^{k-1}.
Also f is C^{k-1}.
Now the induction hypothesis tells us that the composition is C^{k-1}. By the chain rule formula, D_jh_i is C^{k-1}. $\square$

<u>Notation</u>:
Let $(a, b)^m$ be the set product of $(a, b) \subseteq \mathbb{R}$.

Lemma. (Equivalent Differentiability Condition)
$f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at p iff there is a $B \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$, $\alpha > 0$ and $F : (-\alpha, \alpha)^m \to \mathbb{R}^n$ such that

(1) F is continuous at 0
(2) $F(0) = 0$
(3) $f(p+h) = f(p) + Bh + F(h)|h|$ for all $h \in (-\alpha, \alpha)^m$.

Proof.
($\Longleftarrow$) Do it yourself!

($\Longrightarrow$) Let $B = Df(p)$. Let $\alpha > 0$ such that $U(p;\alpha) \subseteq \Omega$. For $h \in (-\alpha, \alpha)^m$, let

$$F(h) = \begin{cases} \frac{f(p+h)-f(p)-Bh}{|h|} & h \neq 0 \\ 0 & h = 0 \end{cases}$$

For $h = 0$,

$$\lim_{h \to 0} F(h) = 0 = F(0).$$

since f is differentiable at p.
□

Lemma. (Approximation of Vector)
For $A \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n$ and vector $u \in \mathbb{R}^m$.

$$|Au| \leq m|A||u|$$

where $|A| = \max_{ij} |A_{ij}|$.

Proof.

$$(|Au|)_i = \sum_{j=1}^{m} A_{ij}u_j \leq \sum_{j=1}^{m} |A_{ij}||u_j| \leq m|A||u|$$

□

Proof. (**Chain Rule**)
Write $r = q = f(p)$, $A = Dg(r)$, and $B = Df(p)$.

$g(r+v) = g(r) + Av + G(v)|v|$ for all $v \in (-\alpha, \alpha)^n$ where G is given by the Lemma above. (1)

Similarly, $f(p+u) = f(p) + Bu + F(u)|u|$ for all $u \in (-\beta, \beta)^m$. (2)

Let $h = g \circ f$.

We want to compute $h(p+u) = g(f(p+u)) = g(r + f(p+u) - f(p))$. Since f is continuous at p, $|f(p+u) - f(p)| < \alpha$ if $|u| < \delta$. WLOG, we may assume $\delta \leq \beta$.

By (1), for $u \in (-\delta, \delta)^m$,

$$h(p+u) = g(r) + A(f(p+u) - f(p)) + G(f(p+u) - f(p))|f(p+u) - f(p)|$$

note that $f(p+u) - f(p)Bu + F(u)|u|$.

Using (2),

$$h(p+u) - h(p) - ABu = AF(u)|u| + G(Bu + F(u)|u|)\,|Bu + F(u)|u||$$

We want to show

$$\lim_{u\to 0} \frac{|h(p+u) - h(p) - ABu|}{|u|} = 0.$$

We use the triangle inequality to consider each part individually.

$$\frac{|AF(u)|u||}{|u|} = |AF(u)| \le n|A||F(u)| \to 0$$

as $u \to 0$ since $\lim_{u\to 0} F(u) = 0$.

The second part of the second term

$$\frac{|Bu + F(u)|u||}{|u|} \le \frac{|Bu| + |F(u)|u||}{|u|} \le m|B| + |F(u)| \to m|b|$$

as $n \to 0$.

Now we check the first part of the second term

$$\lim_{u\to 0} G(Bu + F(u)|u|) = 0$$

again by the lemmata.

Examples:

(a) $g(x,y,z) = (xyz, x^2+y^2+z^2)$ and $f(\theta,\phi) = (6\cos\theta\sin\phi, 6\sin\theta\sin\phi, 6\cos\theta)$.

Then,

$$g \circ f = \begin{bmatrix} yz & xz & xy \\ 2x & 2y & 2z \end{bmatrix} \begin{bmatrix} -6\sin\theta\sin\phi & 6\cos\theta\cos\phi \\ 6\cos\theta\sin\phi & 6\sin\theta\cos\phi \\ 0 & -6\sin\phi \end{bmatrix}$$

Then one should check,

$$\frac{\partial(g\circ g)_2}{\partial\phi} = 0.$$

since $(g \circ f)_2 = 6$.

(b) $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$ and $x : (a,b) \subseteq \mathbb{R} \to \mathbb{R}^n$. Then x defines a curve in $\mathbb{R}^n$.

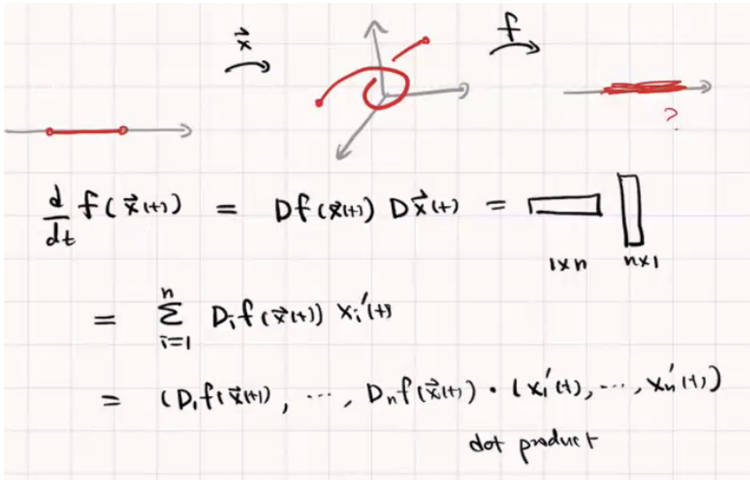


Figure 2: x is a parameterization of a curve.

<u>Notation</u>:
$\frac{d}{dx}(f(x(t)) = \nabla f(x(t)) \cdot x'(t)$.

LOCAL MAXIMA AND MINIMA

In this section we consider **SCALAR-VALUED** functions only.

Definition. (Extrema of Functions)
Let $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$. Let $p \in \Omega$. We say that
(i) p is a **local min** of f if $\exists U$-neighborhood of p such that $f(x) \geq f(p)$ $\forall x \in U$.
(ii) p is a **strict local min** if $\exists U$-neighborhood of p such that $f(x) > f(p)$ $\forall x \in U \setminus \{p\}$.

Definition. (Critical Point)
Let $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$ be differentiable on Ω.
We call p a **critical point** of f if $Df(p) = \vec{0} \in \mathbb{R}^{1\times m}$.

Proposition. (Critical Points Identify Local Maxima/Minima)
Let f be differentiable on $\Omega \subseteq \mathbb{R}^m$.
If p is a local max or a local min of f, then p is a critical point of f.

Proof.
Suppose p is a local min. For each $i \in [n]$, consider

$$D_i f(p) = \lim_{t\to 0} \frac{f(p+te_i) - f(p)}{t} \geq 0$$

Since $f(p+te_i) \geq f(p)$ when $t < 0$ the quantity is negative and positive for positive $t > 0$.

□

Proposition. (Second Derivative Test)
$f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$ if $Df(p) = 0$, then for small u, $f(p+u) \approx f(p)$.

$$f(p+u) \approx f(p) + Df(p)u + \frac{1}{2}\left(\sum_{i,j\in[m]} D_iD_jf(p)u_iu_j\right) = f(p) + 0 + \underbrace{u^T}_{1\times m}\underbrace{Hf(p)}_{m\times m}\underbrace{u}_{m\times 1}$$

We want to ensure that the quadratic function keeps the local structure in tact.

Definition. (Positive Definiteness)
Let H be an $m \times m$ symmetric matrix and define $Q : \mathbb{R}^m \to \mathbb{R}$ by

$$Q(u) = u^T H u = \sum_{i,j\in[m]} H_{ij}u_iu_j.$$

We say that H is (resp. **semi**) **positive definite** if $Q(u) > 0$ (resp. ≥ 0) for all $u \in \mathbb{R}^m \setminus \{0\}$.
And we write $H > 0$ (resp. $H \geq 0$).
We say that H is **negative definite** if $Q(u) < 0$ for all $u \in \mathbb{R}^m \setminus \{0\}$. And we write $H < 0$ (resp. $H \leq 0$).

Examples: I_m is positive definite, since $Q(u) = u_1^2 + u_2^2$ has global minimum at $u = 0$.

$A = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$ is positive definite.

Lemma. (Eigenvalues of Positive Definite Matrix)
Let H be an $m \times m$ symmetric matrix. Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of H, counted with multiplicities.
Then $H \ ? \ 0 \iff \lambda_i \ ? \ 0$ for all $i \in [m]$ holds for $? \in \{>, <, \geq, \leq\}$.

Proof.
By the spectral theorem for symmetric matrices, $H = ODO^{-1}$ for some orthonormal matrix O and diagonal D whose ii-th entry is λ_i. So then $u^T Hu = u^T ODO^T u = (Ou)^T D(O^T u) = \sum_{i \in [m]} \lambda_1 v_i^2$ where $v = O^T u$.

Finish the other direction of the proof on your own. □

Examples:

(a) $\det(A - I\lambda) = \begin{bmatrix} 1-\lambda & -2 \\ -1 & 2-\lambda \end{bmatrix} = \lambda^2 - (\mathrm{Tr}H)\lambda \det H = \lambda^2 - 3\lambda + 1 = 0$ which implies $\lambda_1, \lambda_2 > 0$.

Theorem. (Sylvester's Criterion)
Let $H = (H_{ij})_{i,j=1}^m$ be an $m \times m$ symmetric matrix. For each $k \in [m]$.

Let $H_k = (H_{ij})_{i,j=1}^k$ be the leading principal minor of H. Let $d_k = \det H_k$. So this is the top left $k \times k$ submatrix of H.

$d_k > 0, \forall k \in [m] \iff H > 0$.
$d_k < 0$, for all odd k and $d_k > 0$ for all even k iff $H < 0$.

Proof.
Do on homework. □

Now we try to convert what we have learned about quadratic functions to general functions.

Lemma. (Global Extrema on Convex Set)
The **second derivative test** states conditions for a critical point to be a global extremum.
Let $\Omega \subseteq \mathbb{R}^m$ be an open convex set. Suppose $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$ is C^2 and $p \in \Omega$ is a critical point of f.
If $Hf(x) > 0, \forall x \in \Omega$, $f(x) > f(p), \forall x \in \Omega \setminus \{p\}$.
Similarly, p is a global maxima if $Hf(x) < 0$ for all $x \in \Omega$.

Proof.
Fix arbitrary $y \in \Omega \setminus \{p\}$.

Then $\phi(t) = f(p + t(y - p))$ is a linear function of t, so ϕ is differentiable on $(-\delta, 1 + \delta)$ for some small δ by the **Chain Rule**.

We want to show $\phi(1) > \phi(0)$ since $\phi(1) = f(y)$ and $\phi(0) = f(p)$.

$$\phi'(t) = Df(p + ty - tp)(y - p) = \sum_{j\in[m]} D_j f(p + t(y - p))(y_j - p_j)$$

And

$$\begin{aligned}\phi''(t) &= \sum_{j\in[m]} D_i D_j f(p + t(y - p))(y_i - p_i)(y_j - p_j) \\ &= (y - p)^T Hf(p + t(y - p))(y - p) > 0\end{aligned}$$

Additionally $\phi'(0) = 0$ since p is a critical point. Since the derivative of $\phi'(t)$ is positive, then $\phi'(t)$ is an increasing function. This implies that $\phi(t)$ is also an increasing function and therefore $\phi(1) > \phi(0)$. □

Lemma. (Global Extrema on Compact Sets)
Let K be a compact set in $\mathbb{R}^m$. If $f : K \to \mathbb{R}$ is continuous, then f has global extrema on K.

Proof.
$f(K)$ is compact and a subset of $\mathbb{R}$. □

Lemma. (Perturbation of Positive Definite Matrix)
Let H be an $m \times m$ symmetric matrix such that $H > 0$. Then, $\exists \alpha > 0$ such that if M is an $m \times m$ symmetric matrix with $|M - H| < \alpha$, then $M > 0$.

Equivalently,
$\text{Sym}(m) = \{A \in R^{m\times m} \mid A \text{ symmetric}\} \cong \mathbb{R}^{\frac{m(m+1)}{2}}$ and $\text{Sym}_+(m) = \{H \in \text{Sym}(m) \mid H > 0\}$.
Then $\text{Sym}_+(m)$ is open in $\text{Sym}(m)$.

Proof.
We want to show $u^T M u > 0$ holds for all $u \in \mathbb{R}^m \setminus \{0\}$. Let $K = \{v \in \mathbb{R}^m : |v| = 1\}$. Since $u^T M v = |u|^2 \left(\frac{u}{|u|}\right)^T M \frac{u}{|u|}$, it is enough to check that $v^T M v > 0$ for all $v \in K$.

Now, since $Q(u) = u^T H u$ is continuous on $\mathbb{R}^m$, $Q(K)$ is compact and invoking the lemma above tells us there is a global minima of Q on K. Let this global minima be v_0 and let $Q(v_0) = v_0^T H v_0 = \delta > 0$ since H is positive definite.

Let $a = \frac{\delta}{2m^2} > 0$. Suppose M is symmetric and $|M - H| < \alpha$. Then $\forall v \in K$,

$$v^T M v = v^T H v + v^T (M - H)v$$

Note $v^T H v \geq \delta$. Also,

$$|v^T (M - H)v| \leq \sum_{i,j\in[m]} |(M_{ij} - H_{ij} v_i v_j| < \sum_{i,j\in[m]} am^2 = \frac{\delta}{2}$$

Substituting gives $v^T M v \geq \delta - \frac{\delta}{2} > 0$.

Theorem. (Second Derivative Test)
Let $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$ be a C^2 function. Suppose $p \in \Omega$ is a critical point of f.
(a) If $Hf(p) > 0$, then p is a strict local min of f.
(b) If $Hf(p) < 0$, then p is a strict local max of f.

Proof.
If f is a C^2 function then the Hessian of f has continuous components. Therefore, there is a neighborhood of p such that $|Hf(x) - Hf(p)|$ (sup norm of matrix) is smaller than ε for all x in the neighborhood.

Then we can make ε (possibly) smaller such that $Hf(x) > 0$ for all x in the ε-neighborhood. Then from the lemma above, p is a global minima of f on the ε-neighborhood.

So p is a local min of f on Ω. $\square$

THE INVERSE FUNCTION THEOREM

Statement

Consider $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^m$ (matching dimension!) then $y = f(x) \approx f(p) = Df(p)(x-p)$. Let $B = Df(p)$. We attempt to invert this approximation for $y = f(x)$ which gives us

$$B^{-1}(y - f(p)) + p \approx x$$

should be somewhat true if B is invertible.

Question: If $B = Df(p)$ is **non-singular**, can we find a neighborhood of p and function $g(y) = x$ such that $g \circ f(x) = x$ on the neighborhood?

Theorem. (Inverse Function Theorem)
Let $\Omega \subseteq \mathbb{R}^m$ be open. Let $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^m$ be C^r. Let $p \in \Omega$.
Suppose $Df(p)$ is **non-singular** (invertible). Then,

(1) There exist neighborhoods U of p and $\mathcal{O}$ of $f(p)$ such that $f : U \to \mathcal{O}$ is a bijection.
(2) The inverse function $g : \mathcal{O} \to U$ is C^r.
(3) $Dg(y) = \left[Df(x)\right]^{-1}$. Where $x = g(y)$.

Remarks:

(a) The property (1) is called **local invertibility**.

(b) Suppose we want to solve a system of m equations $f_1, \ldots, f_m$ with m unknowns $y_1, \ldots, y_m$. If $f_1, \ldots, f_m$ are C^r and

$$\det \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_m} \end{bmatrix} \neq 0$$

at $(x_1, \ldots, x_m) = (p_1, \ldots, p_m)$

Then we can solve the system of equations for **all** $(y_1, \ldots, y_m)$ that are close enough to $f(p)$. The solutions are unique and continuously differentiable.

(c) The IFT does not guarantee the global invertibility of f even if $\det Df(x) \neq 0$ for all $x \in \Omega$.

As an example consider $f : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ and $\Omega = (1,3) \times (-\pi, 3\pi)$.
Let $f(r, \theta) = (r\cos\theta, r\sin\theta)$ Then

$$\det Df(r, \theta) = r$$

is nonzero $\forall r \in \Omega$.

However, $f(2,0) = (2,0) = f(2, 2\pi)$.
So, f is not injective on Ω, so there is no global inverse of f on Ω.

Definition. (Open Map)
Let X, Y be metric spaces. We say $f : X \to Y$ is an **open map** (or just **open**) if for every $\mathcal{O}$, open in X, $f(\mathcal{O})$ is open in Y.

The Proof

Lemma. (Lemma 1)
Suppose $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^m$ is C^1 and $p \in \Omega$. If $\det Df(p) \neq 0$, then $\exists \alpha, \varepsilon > 0$ such that

$$|f(u) - f(v)| \geq \alpha|u - v|$$

$$\forall u, v \in U(p; \varepsilon).$$

Which implies that f is injective on $U(p; \varepsilon)$.
This lemma says if there is an inverse function, then it will be Lipschitz.

Proof.
Case 1: f is a linear function.
Suppose $f(x) = r + Ax$. Then f is C^∞, $Df(x) = A$. By assumption $\det A \neq 0$.

Then $\forall u, v \in \Omega$,

$$|f(u) - f(v)| = |A(u - v)|$$

$$|u - v| = |A^{-1}A(u - v)| \leq m|A^{-1}||A(u - v)|$$

Combining gives and setting $\alpha = \frac{1}{m|A^{-1}|}$ gives,

$$|f(u) - f(v)| \geq \alpha|u - v|$$

Case 2: f is a general function.
Let $B = Df(p)$ and $h(x) = f(x) - \ell(x)$ where $\ell(x) = f(p) + Df(p)(x - p)$. Then ℓ is a linear approximation of f near p, and h gives the error of the approximation.

Then $f(x) = \ell(x) + h(x)$. We want to make $h(x)$ small.

f is C^1 and ℓ is C^∞ so h is C^1. Also $Dh(x) = Df(x) - D\ell(x) = Df(x) - B$. Giving, $Dh(p) = 0$ and $Dh(x)$ is continuous. Continuous functions are locally bounded. Therefore we can find a neighborhood of p such that h is bounded.

Explicitly, $\exists \varepsilon > 0$ such that $|Dh(x)| \leq \frac{1}{2m^2|B^{-1}|}$ for all $x \in U(p; \varepsilon)$.

Now consider $\forall u, v \in U(p; \varepsilon)$. We have

$$\begin{aligned}
|f(u) - f(v)| &= |\ell(u) - \ell(v) + h(u) - h(v)| \\
&\geq |\ell(u) - \ell(v)| - |h(u) - h(v)| \\
&\geq \frac{1}{m|B^{-1}|}|u - v| - |h(u) - h(v)|
\end{aligned}$$

Where the last inequality follows from case 1.

Now we want to estimate $|h(u) - h(v)|$. We compute h componentwise. By the MVT for several variables $\forall i \in [m]$ and $\forall u, v \in U(p; \varepsilon)$, $\exists c_i$ between u and v such that

$$h_i(u) - h_i(v) = Dh_i(c_i)(u - v) \leq m|Dh_i(c_i)||u - v| \leq \frac{1}{2m|B^{-1}|}|u - v|$$

Now substituting into the final inequality for $|f(u) - f(v)|$ we find that

$$|f(u) - f(v)| \geq \frac{1}{m|B^{-1}|}|u-v| - |h(u) - h(v)| \geq \frac{1}{m|B^{-1}|}|u-v| - \frac{1}{2m|B^{-1}|}|u-v| = \frac{1}{2m|B^{-1}|}|u-v|$$

$\square$

Lemma. (Lemma 2)
Let M_m be the set of all $m \times m$ matrices. M_m can be identified with $\mathbb{R}^{m^2}$. We write $|A|$ for the sup norm of A under this identification.

(1) The map $A \mapsto \det A$ from M_m to $\mathbb{R}$ is C^∞.
(2) Let GL_m be the set of invertible matrices in M_m. Then GL_m is open and $A \mapsto A^{-1}$ from GL_m to GL_m is C^∞.

Proof.
The determinant is a polynomial in the entries of A.
All entries of A^{-1} are rational functions of the entries of A. $\square$

Proposition. (Proposition 3)
Let $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^m$ be C^1.
Suppose,

- f is injective on Ω.
- $\det Df(x) \neq 0$ for all $x \in \Omega$.

Then f is an open map.

Proof.
Let $\|x\|$ be the Euclidean norm of x. Note that $x \mapsto \|x\|^2$ is C^∞.

We write $B_a(p) = \{x \in \mathbb{R}^m \mid \|x - p\| < a\}$ be the Euclidean open call centered at p with radius a.

Let A be an open set in Ω. Then A is open in $\mathbb{R}^m$ as well. To show $f(A)$ is open in $\mathbb{R}^m$, fix arbitrary $q \in f(A)$. There is a unique $p \in A$ such that $f(p) = q$ be assumption.

Since A is open, $\exists \delta > 0$ such that $\overline{U(p;\delta)} \subset A$. Let $\mathcal{Q} = \overline{U(p;\delta)}$. Consider $\operatorname{Bd} \mathcal{Q} = \overline{\mathcal{Q}} \setminus \mathcal{Q}^\circ$.

So $\Gamma = f(\operatorname{Bd} \mathcal{Q})$ is compact. $q \notin \Gamma$ since f is injective. Moreover, $\operatorname{dist}(\Gamma, q) > 0$. Meaning $\exists \varepsilon > 0$ such that $B_{2\varepsilon}(q)$ is disjoint from Γ.

<u>Claim</u>: $B_\varepsilon(q) \subset f(A)$.

Suppose $\tilde{q} \in B_\varepsilon(q)$. We will show that $\exists \tilde{p} \in A$ such that $f(\tilde{p}) = \tilde{q}$.

Let $\phi : \mathcal{Q} \to \mathbb{R}$ by $\phi(x) = \|f(x) - \tilde{q}\|^2$. Certainly, ϕ has a global minimum by EVT since $\mathcal{Q}$ is compact and ϕ is continuous.

<u>Case 1</u>: Suppose that $\tilde{p} \in \operatorname{Bd} \mathcal{Q}$.
$\phi(\tilde{p})^{\frac{1}{2}} = \|f(\tilde{p}) - \tilde{q}\| \leq \|f(p) - \tilde{q}\| = \|q - \tilde{q}\|$.
$\phi(\tilde{p})^{\frac{1}{2}} = \|f(\tilde{p}) - \tilde{q}\| \geq \|f(p) - q\| - \|q - \tilde{q}\|$.
Together, we get $\|f(\tilde{p}) - \tilde{q}\| \leq 2\|q - \tilde{q}\| < 2\varepsilon$.

But $\tilde{p} \in \operatorname{Bd} \mathcal{Q} \implies f(\tilde{p}) \in \Gamma \implies \|f(\tilde{p}) - q\| \geq 2\varepsilon$. Contradiction!

<u>Case 2</u>: $\tilde{p} \in \mathcal{Q}^\circ$.

By chain rule ϕ is C^1 (Check!).

$\tilde{p}$ is a global min so it is a critical point of ϕ. Meaning $0 = D\phi(\tilde{p}) = 2(f(\tilde{p}) - \tilde{q})^T Df(p)$. But since $Df(p)$ is non-singular $f(\tilde{p}) - \tilde{q}$ is the zero vector.

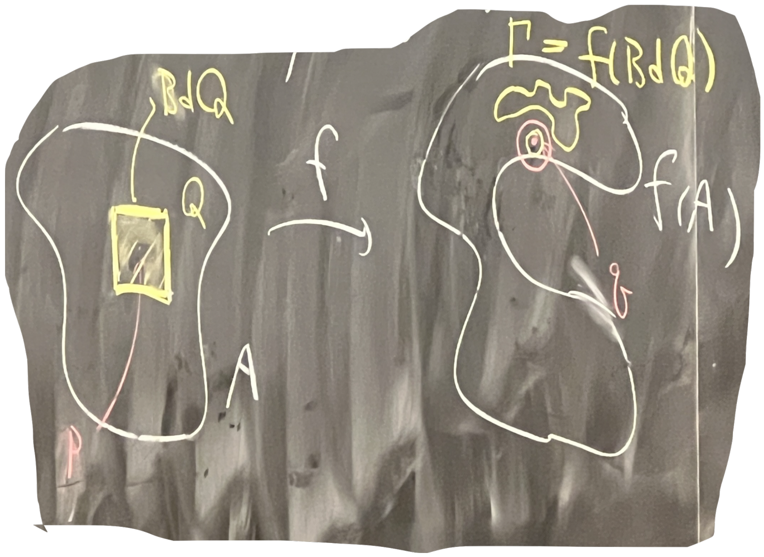


Figure 3: Γ is the image of $\operatorname{Bd} Q$ under f.

$\square$

Proposition. (Proposition 4)
Suppose

- $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^m$ is C^k and
- f is injective on Ω and
- $\det Df(x) \neq 0, \forall x \in \Omega$.

Then $g = f^{-1}$ is C^k.

Proof.
From Proposition 3, g is continuous.

Step 1: Show that g is differentiable.

Fix an arbitrary $r \in f(\Omega)$. Let $p = g(r)$. We write $B = Df(p)$. We will show

$$\lim_{y \to r} \frac{|g(y) - g(r) - B^{-1}(y - r)|}{|y - r|} = 0$$

Since f is injective, we can uniquely write $x = g(y)$ so that $f(x) = y$.

$$\begin{aligned}
\lim_{y\to r}\frac{|g(y)-g(r)-B^{-1}(y-r)|}{|y-r|} &= \frac{|x-p-B^{-1}(f(x)-f(p))}{|f(x)-f(p)|} \\
&= m|B^{-1}|\frac{|B(x-p)-(f(x)-f(p))|}{|f(x)-f(p)|} \\
&= m|B^{-1}|\underbrace{\frac{|B(x-p)-Bx-Bp)|}{|f(x)-f(p)|}}_{\text{(a)}}\underbrace{\frac{|x-p|}{|f(x)-f(p)|}}_{\text{(b)}}
\end{aligned}$$

Where the second inequality follows from $|B^{-1}Bv| \le m|B^{-1}||Bv|$.

By Lemma 1, $\exists \alpha > 0$ such that (b) $\le \frac{1}{\alpha}$ for all x close enough to p.

Since g is continuous as $y \to r$, we have $x \to p$. Thus (a)$\to 0$ by the differentiability of f at p.

Step 2: Now we show that g is C^k.

$Dg(y) = [Df(g(y))]^{-1}$ is a composition of three functions: (i) $y \mapsto g(y)$, (ii) $x \mapsto Df(x)$, and (iii) $A \mapsto A^{-1}$.

We know that (ii) is C^{k-1} and (iii) is C^∞. Since (i) is continuous we get that g is C^1 (since the composition will have the least smoothness of (i), (ii), (iii)).

We apply this many times to get that Dg is C^{k-1}. So g is C^k. This is a bootstrapping argument. $\square$

Proof. (**Inverse Function Theorem**)

$(1) \wedge (2) \implies (3)$ is in the book [Munkres, Theorem 7.4].

By combining Lemma 1, Lemma 2, and Proposition 3 $\implies$ (1) and g is continuous. Why?

The proof of (1) and (2) is given by Proposition 4. $\square$

THE IMPLICIT FUNCTION THEOREM

Consider $x^2+y^2=4$. This represents a circle in $\mathbb{R}^2$, but is not a graph of a single variable function.

The **Implicit function Theorem** says if there is a function $y=y(x)$ such that

$$x^2+(y(x))^2=4 \implies 2x+2yy'=0 \implies y'=-\frac{x}{y}$$

fails for $y=0$.

Consider $(xy^2+e^{xz}y,\ \sin(xy)+z^2x)=(0,0)$. Can we write x,y as a function of z? If the total derivative is invertible, then we can locally represent x,y as a function of z.

Notation:

- For $f:\Omega\subseteq\mathbb{R}^{k+n}\to\mathbb{R}^n$, we write $p\in\Omega$ as (x,y) for $x\in\mathbb{R}^k$ and $y\in\mathbb{R}^n$.
- $\underbrace{Df}_{n\times(n+k)}=\begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{bmatrix}$. Where $\frac{\partial f}{\partial x}$ is an $n\times k$ matrix and $\frac{\partial f}{\partial y}$ is an $n\times n$ matrix.
- Suppose $g:\mathcal{O}\subseteq\mathbb{R}^k\to\mathbb{R}^n$ is differentiable and $f(x,g(x))=0$, $\forall x\in\mathcal{O}$, then by differentiating we get $0=\frac{\partial f}{\partial x}+\left(\frac{\partial f}{\partial y}\right)\left(\frac{\partial g}{\partial x}\right)=\frac{\partial f}{\partial x}(x,g(x))+\frac{\partial f}{\partial y}(x,g(x))\frac{\partial g}{\partial x}(x)$.

Theorem. (Implicit Function Theorem)
Let $\Omega\subseteq\mathbb{R}^{k+n}$ be open and let $f:\Omega\subseteq\mathbb{R}^{k+n}\to\mathbb{R}^n$ be C^ℓ. Suppose $(a,b)\in\Omega$ such that

- $f(a,b)=0$ (not very important)
- $\det\frac{\partial f}{\partial y}(a,b)\neq 0$ (very important).

Then we can locally express y as a function of x near (a,b) as a C^ℓ function.
Meaning,
There is an open, connected set $\mathcal{O}\subseteq\mathbb{R}^k$ such that $a\in\mathcal{O}$, and $\exists! g:\mathcal{O}\to\mathbb{R}^n$ such that $g(a)=b$ and $f(x,g(x))=0$ for all $x\in\mathcal{O}$. Furthermore, g is C^ℓ.

Proof.
Define $F:\Omega\subseteq\mathbb{R}^{k+n}\to\mathbb{R}^{k+n}$ by $F(x,y)=(x,f(x,y))$.

- $F(a,b)=(a,0)$.
- F is C^ℓ since $f(x,y)$ is C^ℓ.
- $DF=\begin{bmatrix} I & 0\\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{bmatrix}\implies \det DF=\det\frac{\partial f}{\partial y}$. In particular $\det DF(a,b)\neq 0$.

By applying the **Inverse Function Theorem**, we obtain a neighborhood $U\times V$ around (a,b) that maps to some W under F. We get a map $G=F^{-1}$. We note F maps vertical lines in $U\times V$ to vertical lines in W, and F maps the curve given by $f(x,y)=0$ to the $\mathbb{R}^k$ axis.

We name $(t,s)=F(x,y)$ and $(x,y)=G(t,s)$. In general we have, $G(t,s)=(\alpha(t,s),\beta(t,s))\implies (t,s)+F(G(t,s))=F(\alpha(t,s),\beta(t,s)=(\alpha(t,s),f(\cdots)$.

This gives us that $\alpha(t,s)=t$. This shows us algebraically that f maps curves that are zero on $\mathbb{R}^k$ to curves in the image that are zero on $\mathbb{R}^k$.

Let $g(x) = \beta(x, 0)$. Check that this is a function that works!

We now need to show that g is unique. Check the textbook for the proof!

CONSTRAINED OPTIMIZATION

Question: find max/min of a function $f(x_1, \dots, x_m)$ given that $(x_1, \dots, x_m) \in \mathbb{R}^m$ satisfies n constraints given by $g_1(x_1, \dots, x_m) = 0; \cdots ; g_n(x_1, \dots, x_m) = 0$ where $m > n$.

Example:

(a) Let Σ be the intersection of the cone $x^2 + y^2 = z^2$ and the place $z = x + y + 2$. Find the distance between the origin and Σ. We want to find the minimum of $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraint $(x, y, z) \in \Sigma$.

(b) Find max/min of $f(x, y) = x^2 + y^2$ subject to $g(x, y) = 0$ where $g(x, y) = xy - 4$. We can draw the **level curves** where $f(x, y) = c$ for $c \in \mathbb{R}$. At the (local) max/min point, it looks like the level curve of f and the constraint curve $g = 0$ are tangential to each other.

In general, if $h(x, y) = 0$ and $y = y(x)$ then $h(x, y(x)) = 0$ implies $\frac{\partial h}{\partial x} + \left(\frac{\partial h}{\partial y}\right) y' = 0$ implies

$$y' = -\frac{\frac{\partial h}{\partial x}}{\frac{\partial h}{\partial y}}$$

For the slopes to agree,

$$\frac{\frac{\partial f}{\partial x}(p_c)}{\frac{\partial f}{\partial y}(p_c)} = \frac{\frac{\partial g}{\partial x}(p_c)}{\frac{\partial g}{\partial y}(p_c)}$$

This gives that $Df(p_c) = \lambda Dg(p_c)$ with $\lambda \in \mathbb{R}$ as a necessary condition for p_c to be a local max/min.

Theorem. (Lagrange Multipliers)
Let $m > n$ be positive integers. Let $g : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be C^1.
Define $\Sigma = \{x \in \mathbb{R}^m : g(x) = 0\} = g^{-1}(0)$. Then Σ is closed.
Let $f : \Omega \subseteq \mathbb{R}^m \to \mathbb{R}$ be C^1.
Suppose

- $f \upharpoonright \Sigma$ has a local max/min at $p \in \Sigma$.
- $Dg(p)$ has <u>rank n</u>.

Then, $\exists \lambda_1, \dots, \lambda_m \in \mathbb{R}$ such that

$$\underbrace{Df(p)}_{1\times m} = \sum_{i=1}^{n} \underbrace{\lambda_i Dg_i(p)}_{1\times m}$$

$(\lambda_1, \dots, \lambda_n) \neq \vec{0}$.

Proof.
We have the $m \times n$ matrix $Dg(p)$ has rank n for $m > n$, meaning there are n linearly independent column vectors in $Dg(p)$. By relabelling the variables if needed, we can assume the last n columns are independent.

Write $p = (a, b) \in \mathbb{R}^{m-n} \times \mathbb{R}^n$ and write $g(x, y)$ for $x \in \mathbb{R}^{m-n}$ and $y \in \mathbb{R}^n$. By assumption $\det \frac{\partial g}{\partial y}(a, b) \neq 0$.

By the **Implicit Function Theorem**, $\exists! h$ and $\exists \mathcal{O}$ such that $h : \mathcal{O} \subseteq \mathbb{R}^{m-n} \to \mathbb{R}^n$ is C^1, $a \in \mathcal{O}$, and $h(a) = b$. Letting $\Sigma = g^{-1}(0)$ tells us that $g(x, h(x)) = 0\ \forall x \in \mathcal{O}$.

Now we will be solving the constraint optimization problem theoretically and locally. Let $\psi(x) = f(x, h(x))$. We know that $f(x, y)$ has local max/min at p, implying that ψ has a local max/min at a. So a is a critical point of ψ. From the **Chain Rule** we obtain,

$$0 = Df(a, b) \begin{bmatrix} I_{m-n} \\ Dh(a) \end{bmatrix} \qquad (1)$$

On the other hand, $g(x, h(x)) = 0$ which yields that

$$Dg(x, h(x)) \begin{bmatrix} I_{m-n} \\ Dh(x) \end{bmatrix} = 0 \qquad (2)$$

$$\forall x \in \mathcal{O}.$$

Let $u = Df(p)^T \in \mathbb{R}^{m \times 1}$. Let $Dg(p)^T = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ where $v_i \in \mathbb{R}^{m \times 1}$. Let $M = \begin{bmatrix} I & Dh(a)^T \end{bmatrix}$. Now $(1) \wedge (2) \implies u \in \operatorname{Ker} M$. This means that $v_i \in \operatorname{Ker} M$ for $i \in [n]$.

We now search for a basis of $\operatorname{Ker} M$ to determine its dimension.

$$\dim \operatorname{Ker} M = m - \dim \operatorname{rank} M = m - (m - n) = n$$

since M is full rank.

$\operatorname{rank} Dg(p) = n \implies \operatorname{rank} Dg(p)^T = n \implies \{v_1, \ldots, v_n\}$ is linearly independent.
Thus $\{v_1, \ldots, v_n\}$ forms a basis of $\ker M$. So $u \in \operatorname{Span}(v_1, \ldots, v_n)$. Meaning,

$$u = \sum_{i=1}^{n} \lambda_i v_i$$

for some scalars λ_i, $i \in [n]$.

Take the transpose of u and note $u^T = Df(p),\ v_i^T = Dg_i(p)$.

$\square$

Example: Let $(A_{ij})_{i,j=1}^{m} = A \in \text{Sym}$. Let $\Sigma = S^{m-1}$ (the unit sphere in $\mathbb{R}^m$). Let $f(x_1, \ldots, x_m) = x^T A x = \sum_{i,j \in [m]} A_{ij} x_i x_j$. Then f is a homogeneous, symmetric, and quadratic function.
Find max of $f \upharpoonright \Sigma$ (this exists by EVT).

Constraining function $g(x_1, \ldots, x_m) = \sum_{i \in [m]} x_i^2 - 1$ and $\Sigma = g^{-1}(0)$.

$$Dg = \begin{bmatrix} 2x_1 & \cdots & 2x_m \end{bmatrix}$$

On Σ, at least one of $x_1, \ldots, x_m$ is nonzero. Hence Dg has rank 1.

Check that,

$$Df = 2x^T A$$

Lagrange's Theorem tells us that $\exists \lambda \in \mathbb{R}$ such that $2x^T A = \lambda 2x \implies Ax = \lambda x$. So the maximizing point is an eigenvector of A. This is a geometric interpretation of an eigenvector of a symmetric matrix. We have $\lambda_{\max} = \max_{\|x\|=1} x^T A x$ and we have a similar fact for $\lambda_{\min}$.

Explore on your own: What is a geometric interpretation of the other $m - 2$ eigenvalues? What about for linear functions, sums of linear and quadratic functions, cubics, tensors?

Example: Find the min of $f(x, y, z) = x^2 + y^2 + z^2$ subject to $g_1(x, y, z) = 0 = g_2(x, y, z)$ where $g_1(x, y, z) = x^2 + y^2 - z^2$, $g_2(x, y, z) = x + y + 2 - z$.

This will give us the point with minimum distance from the origin to $\Sigma = \{(x, y, z) : g_1(x, y, z) = 0 = g_2(x, y, z)\}$. Σ is compact which gives that f has a min on Σ.

$$Dg = \begin{bmatrix} 2x & 2y & -2z \\ 1 & 1 & -1 \end{bmatrix}$$

is not rank 2 $\iff x = y = z$.

But this case never happens on Σ. So Dg has full rank on all of Σ.

Lagrange Multipliers $\implies$ Find $(x, y, z, \lambda_1, \lambda_2)$.
Solving gives,

$$\begin{aligned} 2x &= \lambda_1 2x + \lambda_2 \\ 2y &= \lambda_1 2y + \lambda_2 \\ 2z &= \lambda_1(-2z) - \lambda_2 \\ 0 &= x^2 + y^2 - z^2 \\ 0 &= x + y + 2 - z \end{aligned}$$

Subtracting row 1 from row 2, shows that $(x - y)(\lambda_1 - 1) = 0$.
Case 1: $\lambda_1 = 1 \implies \lambda_2 = 0 \implies z = 0 \implies x = 0 = y$ but this violates row 5.
Case 2: $x = y$.
Row 4 becomes $2x^2 = z^2$ and row 5 becomes $2x + 2 = z$. This gives us a quadratic $x^2 + 4x + 2 = 0$, giving $x = \pm\sqrt{2}$.

This gives

$$(x, y, z) = (-2 + \sqrt{2}, -2 + \sqrt{2}, -2 + 2\sqrt{2}) \quad \vee \quad (x, y, z) = (-2 - \sqrt{2}, -2 - \sqrt{2}, -2 - 2\sqrt{2})$$

INTEGRATION OVER A RECTANGLE

Partitions

There is a huge gap in the generalization of the Riemann integral on the real line to higher dimensions. For example when you partition some subset K of $\mathbb{R}^2$, it might not be that all partitions are cubes.

Notation: We call $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ a rectangle in $\mathbb{R}^n$. Define $v(Q) = \prod_{i=1}^n (b_i - a_i)$ to be the volume of Q.

Definition. (Partition)
For a closed interval $[a, b]$ of $\mathbb{R}$, a **partition** of $[a, b]$ is a finite set

$$P = \{a = t_0 < t_1 < \cdots < t_n = b\}.$$

Then P **determines subintervals** $[t_0, t_1], [t_1, t_2], \ldots, [t_{n-1}, t_n]$.

For a rectangle $Q = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n$, a partition is

$$P = (P_1, \ldots, P_n)$$

where P_i is a partition of $[a_i, b_i]$.
Then P **determines subrectangles** R of Q where $R = I_1 \times \cdots \times I_n$ where I_i is a subinterval of $[a_i, b_i]$ for $i \in [n]$.

Definition. (Upper & Lower Darboux Sums)
Let Q be a rectangle in $\mathbb{R}^n$, $f : Q \to \mathbb{R}$ be a bounded function, and P a partition of Q.
Define the **lower Darboux sum** of f determined by P as

$$L(f, P) = \sum_R m_R(f) \cdot v(R)$$

where $m_R(f) = \inf_{x \in R} f(x)$.
Similarly, define the **upper Darboux sum** of f determined by P as

$$U(f, P) = \sum_R M_R(f) \cdot v(R)$$

where $M_R(f) = \sup_{x \in R} f(x)$.

Definition. (Refinement)
If P and P' are partitions of Q, we say P' is a **refinement** of P if $P' \supseteq P$.
For two partition of P and P' of Q, we say that $P \cup P'$ is a partition that is a refinement of both P and P'. This is called the **common refinement** of P and P'.

If P' is a refinement of P, then

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

This shows that for arbitrary partitions P and P'' of Q, we have

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

by setting the common refinement to P'.

Definition. (Upper & Lower Integrals)
Let Q be a rectangle in $\mathbb{R}^n$, $f : Q \to \mathbb{R}$ be a bounded function. Define the **lower integral** of f over Q as

$$\underline{\int_Q} f = \sup_P L(f, P)$$

and the **upper integral** as

$$\overline{\int_Q} f = \inf_P U(f, P)$$

Both of these quantities necessarily exist since f is bounded.

Lemma: We have

$$\underline{\int_Q} f \leq \overline{\int_Q} f$$

Proof.
$L(f, P) \leq U(f, P''),\ \forall P, P''$.
For fixed P'',

$$\underline{\int_Q} f = \sup_P L(f, P) \leq U(f, P'')$$

Since this holds for all P'',

$$\underline{\int_Q} f \leq \inf_{P''} U(f, P'') = \overline{\int_Q} f$$

□

Definition. (Darboux Integral)
Let Q be a rectangle in $\mathbb{R}^n$, $f : Q \to \mathbb{R}$ be a bounded function. We say that f is **Darboux (Riemann) integrable** over Q if

$$\underline{\int_Q} f = \overline{\int_Q} f$$

and we write this value as $\int_Q f$.

Remark: There is a theory of integration called **Lebesgue Integration**, which is more broadly applicable than the Riemann integral. Take Math 597.

Recall:
Let S be a bounded set in $\mathbb{R}$.

$$\alpha = \sup S \iff \forall x \in S,\ \alpha \geq x \ \wedge\ \forall \varepsilon > 0,\ \exists y \in S,\ \alpha - \varepsilon \leq y$$

Similarly,

$$\beta = \inf S \iff \forall x \in S,\ \beta \leq x \ \wedge\ \forall \varepsilon > 0,\ \exists y \in S,\ \beta + \varepsilon \geq y$$

The Riemann Condition

Definition. (The Riemann Condition)
Let Q be a rectangle in $\mathbb{R}^n$, $f : Q \to \mathbb{R}$ be a bounded function. Then, f is integrable over Q if and only if $\forall \varepsilon > 0$, $\exists$ partition P of Q such that

$$U(f,P) - L(f,P) \leq \varepsilon.$$

Proof.
($\Longrightarrow$) Suppose f is integrable over Q. Fix $\varepsilon > 0$. By the definition of lower integral, $\exists P'$ such that

$$\underline{\int_Q} f - \frac{\varepsilon}{2} \leq L(f,P')$$

Similarly $\exists P''$,

$$\overline{\int_Q} f + \frac{\varepsilon}{2} \geq U(f,P'')$$

Now take the common refinement $P \coloneqq P' \cup P''$, then

$$U(f,P) - L(f,P) \leq U(f,P''), -L(f,P') \leq \left(\overline{\int_Q} f + \frac{\varepsilon}{2}\right) - \left(\underline{\int_Q} f - \frac{\varepsilon}{2}\right) \leq \varepsilon.$$

($\Longleftarrow$) Suppose f is not integrable over Q. Set $2\varepsilon \coloneqq \overline{\int_Q} f - \underline{\int_Q} f$

For any partition P,

$$U(f,P) - L(f,P) \geq \overline{\int_Q} f - \underline{\int_Q} f = 2\varepsilon > \varepsilon$$

$\square$

Proposition. (Continuous Functions are Integrable)
Let Q be a rectangle in $\mathbb{R}^n$, if $f : Q \to \mathbb{R}$ is continuous on Q, then f is integrable over Q.

Proof.
Since Q is compact, f is uniformly continuous on Q. Fix arbitrary $\varepsilon > 0$. Put $\varepsilon' \coloneqq \frac{\varepsilon}{v(Q)}$. The uniform continuity of f furnishes a δ such that $|f(x) - f(y)| < \varepsilon'$ for all $x, y \in Q$ satisfying $|x - y| < \delta$. Choose P so that the "mesh size" is less than δ.

$$U(f,P) - L(f,P) = \sum_R (M_R(f) - m_R(f))\, v(R) \leq \sum_R (\varepsilon')\, v(R) = \varepsilon' v(Q) = \varepsilon$$

□

Note: For $c \in \mathbb{R}$, we have $\int_Q c = c \cdot v(Q)$.

But we want to find the integral of shapes that are not rectangular. To find the volume of a "loaf of bread" S we would want to compute

$$\int_Q 1_S = \int_S 1 = v(S).$$

where Q is a rectangle that bounds S.

The indicator function 1_S is not continuous on $\operatorname{Bd} S$.

Boundary sets can be quite complicated. Consider $S = \{(x, y) \in [0,1]^2 \mid x, y \text{ are rational.}\}$, then $\operatorname{Bd} S = [0,1]^2$. 1_S is not Riemann Integrable.

Example:
Let

$$f(x,y) = \begin{cases} 1 & x = y \\ 0 & x \neq y. \end{cases}$$

Then f is integrable over $Q = [0,1]^2$.

Fix $\varepsilon > 0$, let $n \in \mathbb{N}$. Let $P = \{\{\frac{k}{n} : k \in [n]\}, \{\frac{k}{n} : k \in [n]\}\}$. This gives that

$$U(f,P) - L(f,P) = (3n-2)\frac{1}{n^2}.$$

For large n, this quantity vanishes.

So $\int_Q f = \underline{\int_Q} f = 0$.

CHARACTERIZATION OF RIEMANN INTEGRABILITY

Measure Zero Sets

We will find a characterization of Riemann integrable functions, completely.

Definition. (Measure) A set $A \subseteq \mathbb{R}^n$ is said to be a set of **measure zero** in $\mathbb{R}^n$ if $\forall \varepsilon > 0$, $\exists$ countably many rectangles $Q_1, Q_2, \ldots$ in $\mathbb{R}^n$ such that

$$\bigcup_{i=1}^{\infty} \supseteq A \quad \text{and} \quad \sum_{i=1}^{\infty} v(Q_i) \leq \varepsilon$$

Remark: We allow $Q_i = \varnothing$, so a finite covering is allowed.

Examples:

(a) $A = (0,2] \times \{0\}$ has measure zero in $\mathbb{R}^2$.
Fix $\varepsilon > 0$. Let $Q_1 = [0,2] \times [-\frac{\varepsilon}{4}, \frac{\varepsilon}{4}]$ and $Q_k = \varnothing$ for $k \geq 2$.

(b) $B = (0,2]$ in $\mathbb{R}$. Then $\mathbb{R}$ does not have nonzero measure.

Proposition: If $\Omega \neq \varnothing$ is an open set in $\mathbb{R}^n$, then Ω is not a set of measure zero in $\mathbb{R}^n$.

Examples:
All of the following sets A have measure zero,

(a) $A = \mathbb{Q}$.
Then $A = \{q_i\}_{i=1}^{\infty}$. Fix $\varepsilon > 0$, Let

$$Q_i = [q_i - \frac{\varepsilon}{2^{i+1}}, q_i + \frac{\varepsilon}{2^{i+1}}]$$

Then $Q_1, Q_2, \ldots$ covers A and

$$\sum_{i=1}^{\infty} v(Q_i) = \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon$$

Theorem. (Measure Consequences)

(1) A subset of a set of measure zero, is a set of measure zero.
(2) The countable union of sets of measure zero is measure zero.
(3) A is a set of measure zero in $\mathbb{R}^n$ if and only if $\forall \varepsilon > 0$, $\exists$ rectangles $Q_1, Q_2, \ldots$ in $\mathbb{R}^n$ such that

$$\bigcup_{i=1}^{\infty} \operatorname{Int} Q_i \supset A \quad \text{and} \quad \sum_{i=1}^{\infty} v(Q_i) \leq \varepsilon$$

(4) Q rectangle in $\mathbb{R}^n \implies \operatorname{Bd} Q = \overline{Q} - \operatorname{Int} Q$ is a set of measure zero.

Proof.

(1) is easy (do it!)

(2) We will use a standard trick in analysis called the $\frac{\varepsilon}{2^h}$ trick. Suppose $A_1, A_2, \dots$ be sets of measure zero in $\mathbb{R}^n$. Let $A = \bigcup_{k=1}^{\infty} A_k$.
Fix arbitrary $\varepsilon > 0$. For each $k \in \omega$, since A_k has measure zero, $\exists$ rectangles $Q_{k,1}, Q_{k,2}, \dots$ such that

$$\bigcup_{i=1}^{\infty} Q_{k,i} \supset A_k \quad \text{and} \quad \sum_{i=1}^{\infty} v(Q_{k,i}) \le \frac{\varepsilon}{2^k}$$

This gives us that

$$\bigcup_{k=1}^{\infty} \bigcup_{i=1}^{\infty} Q_{k,i} \supset A, \quad \text{and} \quad \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} v(Q_{k,i}) \le \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon$$

Since the countable union of countable sets is countable and the summation is absolutely convergent (so any arrangement gives the same value).

(3) $(\Longleftarrow)$ Clear.
$(\Longrightarrow)$ Suppose A has measure zero in $\mathbb{R}^n$.
Fix $\varepsilon > 0$. $\exists Q_1', Q_2', \dots$ such that

$$\bigcup_{i=1}^{\infty} \operatorname{Int} Q_i' \supset A \quad \text{and} \quad \sum_{i=1}^{\infty} v(Q_i') \le \frac{\varepsilon}{2}$$

For each i, let Q_i be a rectangle such that $\operatorname{Int} Q_i \supset Q_i'$ and $v(Q_i) \le 2v(Q_i')$ (check!).

So then

$$\bigcup_{i=1}^{\infty} \operatorname{Int} Q_i' \supset A \quad \text{and} \quad \sum_{i=1}^{\infty} v(Q_i') \le 2 \sum_{i=1}^{\infty} v(Q_i') \le \varepsilon$$

(4) Do it yourself!

Example: Let $A = [0,1] - \mathbb{Q}$, the irrational points in $[0,1]$.
Suppose A has measure zero in $\mathbb{R}$. Then $[0,1] = A \cup B$, where $B = [0,1] \cap \mathbb{Q}$ (B has measure zero). The union of two measure zero sets has measure zero, so $[0,1]$ has measure zero. Contradiction!

Remark: If a set A is not a set of measure zero in $\mathbb{R}^n$, it does not mean that A is a set of positive measure. The reason being some sets are not measurable.

The Proof

Theorem. (Characterization of Riemann Integrability)
Let Q be a rectangle in $\mathbb{R}^n$ and $f : Q \to \mathbb{R}$ a bounded function.
Let $D = \{x_0 \in Q : f \text{ is not continuous at } x_0\}$.

Then f is Riemann integrable over Q iff D has measure zero.

We say that f is continuous **almost everywhere**.

Proof.
$(\Longleftarrow)$ Suppose D has measure zero in $\mathbb{R}^n$. Since f is a bounded function, $\exists \alpha > 0$ such that

$|f(x)| \le \alpha$ for all $x \in Q$.
Fix $\varepsilon > 0$. Let $\varepsilon' = \frac{\varepsilon}{2\alpha + 2v(Q)}$.
D has measure zero $\implies \exists_{i=1}^{\infty} Q_i$ in $\mathbb{R}^n$ such that $\bigcup_{i=1}^{\infty} \operatorname{Int} Q_i \supset D$ and $\sum_{i=1}^{\infty} v(Q_i) \le \varepsilon'$.

Since f is continuous on $Q - D$. We have that $\forall a \in Q - D$, there exists a rectangle Q_a with $a \in \operatorname{Int} Q_a$ such that $|f(x) - f(a)| \le \varepsilon'$ for all $x \in Q_a \cap (Q - D)$.

Then,

$$Q \subseteq \left(\bigcup_{i=1}^{\infty} \operatorname{Int} Q_i \right) \cup \left(\bigcup_{a \in Q - D} \operatorname{Int} Q_a \right)$$

This collection is an open cover of Q. Since Q is compact, $\exists k, l \in \omega$ such that

$$Q \subseteq \left(\bigcup_{i=1}^{k} \operatorname{Int} Q_i \right) \cup \left(\bigcup_{j=1}^{l} \operatorname{Int} Q'_j \right) \subseteq \left(\bigcup_{i=1}^{k} Q_i \right) \cup \left(\bigcup_{j=1}^{l} Q'_j \right)$$

where each Q_a in the countable subcover is labeled Q'_j.

Then we have

$$\sum_{i=1}^{k} v(Q_i) \le \sum_{i=1}^{\infty} v(Q_i) \le \varepsilon'$$

If $x, y \in Q'_i \cap (Q - D)$, then $|f(x) - f(y)| \le 2\varepsilon'$.

We will call $Q_i \cap Q$ by Q_i and call $Q'_j \cap Q$ by Q'_j.

Then

$$Q = \left(\bigcup_{i=1}^{k} \operatorname{Int} Q_i \right) \cup \left(\bigcup_{j=1}^{l} \operatorname{Int} Q'_j \right)$$

We define a partition P of Q using the vertices of Q_is and Q'_js. So the partition looks like extending all lines in each Q_i.

Then each Q_i or Q'_j is a finite union of subrectangles of P.
Let $\mathscr{R}$ be the collection of subrectangles R of P that intersects (equivalently is a subset of) some Q_i.
Let $\mathscr{R}'$ be all other subrectangles defined by P.

Now we compute

$$U(f, P) - L(f, P) = \sum_{R} (M_R(f) - m_R(f))\, v(R)$$

$$= \sum_{R \in \mathscr{R}} (M_R(f) - m_R(f))\, v(R) + \sum_{R \in \mathscr{R}'} (M_R(f) - m_R(f))\, v(R)$$

The LHS has small volume and the RHS has small difference. Which gives that

$$\le 2\alpha \sum_{R \in \mathscr{R}} v(R) + 2\varepsilon' \sum_{R \in \mathscr{R}'} v(R)$$

Noting that all Q_is cover all Rs gives that,

$$\le 2\alpha \sum_{i=1}^{k} v(Q_i) + 2\varepsilon' v(Q) = (2\alpha + 2v(Q))\varepsilon' = \varepsilon.$$

($\Longrightarrow$) Suppose f is integrable over Q. By the Oscillation Lemma, $D = \{p \in Q : \operatorname{osc}(f;p) > 0\}$. We will show that $D_\varepsilon = \{p \in Q : \operatorname{osc}(f;p) > \varepsilon\}$ has measure zero.
Since

$$D = \bigcup_{\varepsilon \in \mathbb{Q}^+} D_\varepsilon$$

D_ε has measure zero for arbitrary $\varepsilon > 0$, therefore implies that D has measure zero.

Fix m and consider $D_{\frac{1}{m}}$. Let $\varepsilon > 0$ be arbitrary. f is integrable $\implies$ there exists a partition P of Q such that $U(f,P) - L(f,p) \leq \frac{\varepsilon}{2m}$.

We write $D_{\frac{1}{m}} = D'_{\frac{1}{m}} \cup D''_{\frac{1}{m}}$ where $D'_{\frac{1}{m}} = \{p \in D_{\frac{1}{m}} : p \in \operatorname{Int} R \text{ for some } R \text{ determined by } P\}$ and $D''_{\frac{1}{m}} = \{p \in D_{\frac{1}{m}} : p \in \operatorname{Bd} R \text{ for some } R \text{ determined by } P\}$.

Let $R_1, R_2, \ldots, R_k$ be the subrectangles such that $D'_{\frac{1}{m}} \cap R_i \neq \varepsilon$ (i.e. $\operatorname{Int} R_i \cap D_{\frac{1}{m}} \neq \varnothing$). So then $\cup_{i=1}^k R_i \supset D'_{\frac{1}{m}}$ and

$$\frac{\varepsilon}{2m} \geq \sum_R \left(M_R(f)) - m_R(f)\right) v(R) \geq \sum_{R_i} \underbrace{\left(M_{R_i}(f)) - m_{R_i}(f)\right)}_{\geq \frac{1}{m}} v(R_i)$$

This gives that $\frac{\varepsilon}{2} \geq \sum_{i=1}^k v(R_i)$.

Now since $D''_{\frac{1}{m}} \subset \bigcup_R \operatorname{Bd} R$, we know that $D''_{\frac{1}{m}}$ has measure zero. The union of countable many boundaries of rectangles has measure zero.

Thus $D_{\frac{1}{m}} = D'_{\frac{1}{m}} \cup D''_{\frac{1}{m}}$ has measure zero. $\square$

Lemma. (Oscillation Lemma)
Let $f : Q \to \mathbb{R}$ be a bounded function. For $p \in Q$ and $r > 0$, let $\operatorname{osc}_r(f;p) = \sup f(x) - \inf f(y)$ for $x, y \in Q \cap B_r(p)$. Then we define $\operatorname{osc}(f;p) = \inf_{r>0} \operatorname{osc}_r(f;p) = \lim_{r \to 0} \operatorname{osc}_r(f;p)$.
Then,
(1) $\operatorname{osc}(f;p) \geq 0$.
(2) f is continuous at p if and only if $\operatorname{osc}(f;p) = 0$.

Proof.
(1) is clear.
(2) ($\Longrightarrow$) Suppose f is continuous at p. We give ourselves a room of epsilon. To show that $t = 0$, we show that $|t| \leq \varepsilon$ for arbitrary $\varepsilon > 0$.

Let $\varepsilon > 0$ be given. Since f is continuous at p we know that $\exists \delta > 0$ such that $|f(x) - f(p)| < \frac{\varepsilon}{2}$ whenever $x \in B_\delta(p) \cap Q$. Taking inf and sup over all x gives us that $|f(x) - f(y)| < y$ whenever $x, y \in B_\delta(p) \cap Q$. We thus have that $\operatorname{osc}_\delta(f;p) \leq \varepsilon \implies \operatorname{osc}(f;p) \leq \varepsilon$. Since $\operatorname{osc}(f;p) \leq \operatorname{osc}_\delta(f;p)$.

(2) ($\Longleftarrow$) Suppose $\operatorname{osc}(f;p) = 0$.
Let $\varepsilon > 0$ be given. Since

$$\operatorname{osc}(f;p) = \inf_{r>p} \operatorname{osc}_r(f;p) = 0, \quad \implies \quad \exists \delta \text{ s.t. } \operatorname{osc}_\delta(f;p) \leq \varepsilon$$

This gives that $\forall x, y \in Q \cap B_\delta(p),\ f(x) - f(y) \leq \varepsilon$. Meaning that $|f(x) - f(p)| \leq \varepsilon$. $\square$

Example: $1_{\mathbb{Q}}$ is not Riemann integrable, however it is Lebesgue integrable.

Theorem. (Zero Functions Almost Everywhere)
Let $f : Q \to \mathbb{R}$ be a bounded, integrable function over Q. Then,
(1) $f = 0$ almost everywhere, then $\int_Q f = 0$.
(2) Suppose $f \geq 0$ on Q and $\int_Q f = 0$, then f is zero almost everywhere.

Proof.
(1) Read the book!
(2) By the Characterization of Riemann Integrability, f is continuous almost everywhere. Let D be the measure zero set of discontinuities of f. Let $p \in Q - D$.

Suppose for the sake of contradiction that $f(p) > 0$. Since f is continuous at p, $\exists \delta > 0$ such that $f(x) \geq \frac{f(p)}{2}$ (why?) for all $x \in Q \cap B_\delta(p)$.

Now we build a partition P such that one of the subrectangles R_0 is $Q \cap \overline{B_\delta(p)}$. So then $L(f, P) \geq m_(R_0)v(R_0) \geq \frac{f(p)}{2}v(R_0) > 0$ which gives that

$$\int_Q f = \int_{\underline{Q}} f \geq L(f, P) \geq \frac{f(p)}{2}v(R_0) > 0$$

a contradiction!

We conclude that $f(p) = 0$. $\square$

Notation: Sometimes we will write

$$\int_Q f = \int_{x \in Q} f(x)$$

because f may have auxiliary variables.

FUBINI'S THEOREM

For $n = 1$, we can compute the integral by using the Fundamental Theorem of Calculus (evaluate an antiderivative at the bounds).

Fubini's Theorem tells us how to compute n-dimensional integrals. We want to know under what conditions can we rearrange the order of integration?

We do need to apply some conditions. For example the function

$$f(x,y) = \begin{cases} 1 & x = \frac{1}{2} \ \wedge \ y \in \mathbb{Q} \\ 0 & \text{otherwise} \end{cases}$$

Then $f : [0,1]^2 \to \mathbb{R}$ is Riemann integrable over $\mathbb{Q}$. But then $g(y) = f(\frac{1}{2}, y)$ is not Riemann integrable over $[0,1]$. So the integral with respect to y gives us trouble at a single point. But a single point is measure zero, so we should still be able to integrate f.

Theorem. (Fubini's Theorem)
Let $Q = A \times B$, A rectangle in $\mathbb{R}^k$. Let $f : Q \to \mathbb{R}$ be a bounded, integrable function over Q.

(1) The functions

$$\underline{I}(x) = \underline{\int_{y \in B}} f(x,y) \text{ and } \overline{I}(x) = \overline{\int_{y \in B}} f(x,y)$$

are integrable over A.

(2)

$$\int_Q f = \int_{x \in A} \underline{\int_{y \in B}} f(x,y) = \int_{x \in A} \overline{\int_{y \in B}} f(x,y).$$

Corollary $Q = I_1 \times \cdots \times I_n \subset \mathbb{R}^n$. Let $f : Q \to \mathbb{R}$ is continuous, then for all $\sigma \in S_n$,

$$\int_Q f = \int_{x_{\sigma(1)} \in I_{\sigma(1)}} \cdots \int_{x_{\sigma(n)} \in I_{\sigma(n)}} f(x_1, \ldots, x_n)$$

Proof.
(1) We will use the Riemann Condition. Let $P = (P_A, P_B)$ be a partition of Q.

Claim:

$$L(f,P) \leq L(\underline{I}, P_A)$$

Consider a subrectangle R_A of P_A and a subrectangle R_B of P_B. Fix $x_0 \in \mathbb{R}_A$.

$$m_{R_A \times R_B} = \inf_{(x,y) \in R_A \times R_B} f(x,y) \leq \inf_{y \in R_B} f(x_0, y) = m_{R_B}(f(x_0,y))$$

This implies

$$\sum_{R_B} m_{R_A \times R_B}(f) v(R_B) \leq \sum_{R_B} m_{R_B}(f(x_0,y)) v(R_B) = L(f(x_0,y), P_B) \leq \underline{\int_{y \in B}} f(x_0,y) = \underline{I}(x_0).$$

Which holds for all $x_0 \in \mathbb{R}_A$ and the LHS does not depend on x_0
(this is a uniform lower bound for all $x_0 \in R_A$).

$$\sum_{R_B} m_{R_A \times R_B}(f) v(R_B) \leq \inf_{x_0 \in R_A} \underline{I}(x_0) = m_{R_A}(\underline{I})$$

Thus,

$$\sum_{R_A}\left(\sum_{R_B}\sum_{R_B} m_{R_A\times R_B}(f)v(R_B)\right)v(R_A)=\sum_{R_A} m_{R_A}(\underline{I})v(R_A)$$

So then $L(f,P)\le L(\underline{I},P_A)$

A completely analogous argument gives

$U(f,P)\ge U(\overline{I},P_A)$.

The claim is proved.

$$L(f,P)\le L(\underline{I},P_A)\le \underline{\int_{x\in A}}\underline{I}$$

This holds for all partitions P. Which gives us that

$$\underline{\int_Q} f=\int_Q f\le \underline{\int_{x\in A}}\underline{I}.$$

Similarly,

$$\overline{\int_{x\in A}}\underline{I}\le U(\underline{I},P_A)\le U(\overline{I},P_A)\le U(f,P)$$

This holds for all partitions P. Which gives us that

$$\overline{\int_{x\in A}}\underline{I}\le \overline{\int_Q} f=\int_Q f$$

Now we have a sandwich of inequalities that shows

$$\int_Q f=\underline{\int_{x\in A}}\underline{I}=\overline{\int_{x\in A}}\underline{I}=\int_Q f$$

To show that the order of integration can be rearranged define $g(x,y)=f(y,x)$ and apply the previous proof to g. $\square$

INTEGRATION OVER BOUNDED SETS

Definition. (Integral on Bounded Sets)
Let S be a bounded set in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be a bounded function. We say f is **integrable over** S if there is a rectangle Q in $\mathbb{R}^n$ that contains S such that

$$\int_Q 1_S \tilde{f} = \int_Q \tilde{f}$$

exists.
Where 1_S is the indicator function of S and $\tilde{f}$ is the extension of f to $\mathbb{R}^n$ such that $\tilde{f}(x) = 0$ for all $x \notin S$.

Set

$$f_S(x) = \begin{cases} f(x) & x \in S \\ 0 & \text{otherwise.} \end{cases}$$

Lemma. (Integral is Well Defined)
Let Q, Q' be two rectangles containing S. Then, f_S is integrable over $Q \iff f_s$ is integrable over Q'. Furthermore, if f_s is integrable over Q then,

$$\int_Q f_S = \int_{Q'} f_S.$$

Therefore $\int_S f$ is well defined if f is integrable over S.

Lemma. (Max & Min Lemma)
For $f, g : S \to \mathbb{R}$, define $h(x) = \max\{f(x), g(x)\}$ and $k(x) = \min\{f(x), g(x)\}$.
If f and g are integrable, then so are h and k.

Proof. HW.

Theorem. (Properties of Integrals)
Integration properties like monotonicity and linearity on the real line extend naturally to $\mathbb{R}^n$.
Suppose $S \subset \mathbb{R}^n$ is a bounded set and $f, g : S \to \mathbb{R}$ are bounded, integrable functions. Then,
(a) $\alpha f + \beta g$ is integrable over S, $\forall \alpha, \beta \in \mathbb{R}$.
(b) $\int_S (\alpha f + \beta g) = \alpha \int_S f + \beta \int_S g$.
(c) If $f(x) \leq g(x)$, $\forall x \in S$, then $\int_S f \leq \int_S g$.
(d) $|f|$ is integrable over S.
(e) $|\int f| \leq \int |f|$ (triangle inequality for function space).

Proof.
HW7 asked us to prove when S is a bounded rectangle, and the proof for arbitrary bounded set S follows.

Theorem. (Properties of Integrals Over Varying Sets)

(a) Suppose $T \subseteq S \subset \mathbb{R}^n$, T and S are bounded sets. Suppose f is integrable over T and also over S. If $f(x) \geq 0$, $\forall x \in S$ then

$$\int_T f \leq \int_S f.$$

(b) Let S_1, S_2 be bounded sets in $\mathbb{R}^n$. Suppose f is integrable over S_1 and also S_2. Then f is integrable over $S_1 \cup S_2$ and $S_1 \cap S_2$, and

$$\int_{S_1 \cup S_2} f = \left(\int_{S_1} f + \int_{S_2} f\right) - \int_{S_1 \cap S_2} f.$$

Proof.
(1) It suffices to show $\int_Q f_T \leq \int_Q f_S$ for a bounding rectangle Q of S. Then since $f_T(x) \leq f_S(x)$, for all $x \in \mathbb{R}^n$ (there are three cases to consider here). Now by the monotonicity property, we obtain the result.

(2) Follows by a simple Exclusion-Inclusion argument.

First consider the case where $f(x) \geq 0$ for all $x \in S_1 \cup S_2$. Note that $f_{S_1 \cup S_2} = \max_x\{f_{S_1}(s), f_{S_2}(x)\}$, which is integrable. Similarly, $f_{S_1 \cap S_2} = \min\{f_{S_1}(s), f_{S_2}(x)\}$, which is integrable.

Now consider the case of a general f. Let

$$f_+(x) = \begin{cases} f(x) & f(x) > 0 \\ 0 & f(x) < 0 \end{cases}$$

define f_- similarly.

Then $f = f_+ + f_-$. Consider

$$\int_{S_1 \cup S_2} f(x) = \left(\int_{S_1} f(x) + \int_{S_2} f(x)\right) - \int_{S_1 \cap S_2} f(x)$$

for all $x \in S_1 \cup S_2$.

We check all four cases for the signs, then use linearity of the integral, finally extending to a bounding rectangle Q completes the result.

Alternatively, since f is bounded we can add the bound to f and use linearity of the integral to apply the first case. □

From this point on. We will restrict our study to functions $f : S \to \mathbb{R}$ that are continuous on S (in practice we can often use inclusion-exclusion to build back up to the more general case).

Note that S may have many isolated points, so this is not overly restrictive to only consider continuous functions on S.

Recall: For rectangle Q, if f is continuous on Q, then it is integrable over Q.

This does not necessarily hold for a general set S. For example, let $S = \mathbb{Q}^2 \cap [0,1]^2$ and $f : S \to \mathbb{R}$ defined by $f(x) = 1$. Then f is continuous on S, but f is not integrable. This follows since f extended to a bounding rectangle Q has $[0,1]^2$ as a set of discontinuities, which is not measure zero.

Recall: $\operatorname{Int} S \subseteq S$ is the topological interior of S. Similarly, $\operatorname{Ext} S$ is the topological interior of S the complement of S. Finally, $\operatorname{Bd} S = \overline{S} \setminus \operatorname{Int} S$ (the closure of S minus the interior).

Theorem. (Characterization of Integrability on Bounded Sets)
Suppose $S \subset \mathbb{R}^n$ is bounded. Let $f : S \to \mathbb{R}$. Let

$$E = \{x_0 \in \operatorname{Bd} S : x \in S, \lim_{x \to x_0 x \in S} f(x) \neq 0\}.$$

If E is a set of measure zero in $\mathbb{R}^n$, then f is integrable over S.

Proof.
We will show that f_S is integrable over Q. It is enough to show that the discontinuity set of f_S is a subset of E (which has measure zero). We will show that f_S is continuous on $\mathbb{R}^n \setminus E$.

There are three cases (1) $x_0 \in \operatorname{Int} S$, (2) $x_0 \in \operatorname{Ext} S$ and (3) $x_0 \in \operatorname{Bd} S$.

(a) Since f is continuous, then f_S is continuous at x_0.

(b) Since f_S is zero on an open set of x_0, f_S is continuous at x_0.

(c) Let $x_0 \in (\operatorname{Bd} S) \setminus E$. f_S is sequentially continuous at x_0 since $\lim_{x \to x_0} f_S(x) = 0$ for all $x \in S$ (by construction of E) and $x \notin S$ (since $f_S(x) = 0$ for $x \notin S$). We conclude that if f_S is discontinuous at $x_0 \in \operatorname{Bd} S$, then $x_0 \in E$.

$\square$

Example: Let $S = \{(x, y) \in \mathbb{R}^s : 0 \leq x \leq y \leq 1\}$ then $\int_S xy$ is integrable. We use Fubini's Theorem by extending the integrand to $[0, 1]^2$.

$$\begin{aligned}
\int_S xy &= \int_Q f_S \\
&= \int_{x \in [0,1]} \underline{\int_{y \in [0,1]}} f_S \\
&= \int_{x \in [0,1]} \int_{y \in [0,1]} f_S \\
&= \int_{x \in [0,1]} \left(\int_{y \in [0,x)} f_S + \int_{y \in [x,1]} f_S \right) \\
&= \int_{x \in [0,1]} \left(0 + \frac{1}{2}(x - x3) \right) \\
&= \int_{x \in [0,1]} \frac{1}{2}(x - x^3) \\
&= \frac{1}{8}
\end{aligned}$$

Where the third inequality follows from:
Fix $x \in [0, 1]$ and consider $g(y) = f_S(x, y) = xy$ if $y \in [0, 1]$ and zero otherwise. So then $g(y)$ is discontinuous at only (possibly) one point x.

The fourth inequality follows by splitting the domain. The fifth follows from the Fundamental Theorem of Calculus.

Theorem. (Boundary Does not Affect Integral)
Let $S \subseteq \mathbb{R}^n$ be bounded and $f : S \to \mathbb{R}$ a bounded, continuous function. Let $A = \operatorname{Int} S$. If f is integrable over S, then if is integrable over A and

$$\int_A f = \int_S f.$$

Corollary. $\Omega \subseteq \mathbb{R}^n$ is open, and $f : \Omega \to \mathbb{R}$ is a bounded function. If f can be extended to a bounded continuous function on $\overline{\Omega}$ that is integrable over $\overline{\Omega}$, then

$$\int_\Omega f = \int_{\overline{\Omega}} f.$$

Proof.
Read the book. $\square$

RECTIFIABLE SETS

These are sets in which the notion of volume makes sense.

Definition. (Rectifiable Sets)
A bounded set $S \subseteq \mathbb{R}^n$ is **rectifiable** provided that the function $f(x) = 1$ is integrable over S. For a rectifiable set S, define its **volume** as $v(S) = \int_S 1$.

Example: The unit disk is rectifiable.

Theorem. (Characterization of Rectifiable Sets)
$S \subseteq \mathbb{R}^n$ is rectifiable $\iff$ S is a bounded set such that $\operatorname{Bd} S$ has measure zero.

Proof.
Easy (follows from the theorem above). □

Example: There are non-rectifiable sets. The set of rational points is not rectifiable in $\mathbb{R}$.

Remark: This volume notion is usually called the **Jordan Content** of a set. Since the late 19-th century mathematicians have been trying to extend the notion of volume to all sets in $\mathbb{R}^n$. It turns out that such a function to construct is not possible which respects volume properties: (1) unit cube has volume ones, (2) finitely many disjoint sums volumes add, and (3) volume should be preserved under translation and rotations. A counterexample is the Tarski Paradox. Consider the interval $[0, 1]$ and partition it into uncountably many points (singletons), move each point from x to $2x$. People have been trying to fill in the gap (of assigning volume) between rectifiable sets and general sets. This is called Measure Theory. Fubini's Theorem can be generalized under milder conditions using Measure Theory.

Theorem. (Properties of Rectifiable Sets)
Let $S, S_1, S_2 \subset \mathbb{R}^n$ be rectifiable sets.

(1) $v(S) \geq 0$.
(2) $S_1 \subseteq S_2 \implies v(S_1) \leq v(S_2)$.
(3) $S_1 \cup S_2$ and $S_1 \cap S_2$ are also rectifiable and $v(S_1 \cup S_2) = v(S_1) + v(S_2) - v(S_1 \cap S_2)$.
(4) $v(S) = \iff S$ has measure zero.
(5) $\operatorname{Int}(S)$ is rectifiable and $v(\operatorname{Int} S) = v(S)$.
(6) If $f : S \to \mathbb{R}$ is a bounded, continuous functions, then f is integrable over S.

Proof.
Do it yourself! □

Definition. (Simple Regions)
A **simple region** in $\mathbb{R}^n$ is a set of the following form:

$$S = \{(x,t) \in \mathbb{R}^{n-1} \times \mathbb{R} : x \in C,\ \phi(x) \leq t \leq \psi(x)\}$$

where C is a compact, rectifiable in $\mathbb{R}^{n-1}$ and $\phi, \psi : C \to \mathbb{R}$ are continuous functions satisfying $\phi(x) \leq \psi(x), \forall x \in C$.

Examples:

(a) The closed unit disk in $\mathbb{R}^2$. $C = [-1,1]$ and $-\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}$ shows that the unit disk is a simple region.

(b) B is the closed unit ball. Then C is the closed unit disk and $-\sqrt{1-x^2-y^2} \leq z \leq \sqrt{1-x^2-y^2}$ shows that B is rectifiable.

(c) Now we can show inductively that the closed unit ball in $\mathbb{R}^n$ is a simple region.

Remark: t on the last coordinate is arbitrary.

Lemma. (Every Simple Region is Rectifiable)
Every simple region $S \subset \mathbb{R}^n$ is compact and rectifiable.

Proof.
First we check that S is compact. Continuous functions on a compact set shows that S is bounded and it is compact since it is also closed. Now we see that

$$\operatorname{Bd} S = G_\phi \cup G_\psi \cup D$$

where $D = \{(x,t) : x \in \operatorname{Bd} C \ \wedge \phi(x) \leq t \leq \psi(x)\}$. Now we check that all of these are measure zero.

(a) G_ϕ, G_ψ are measure zero. Let Q be a bounding rectangle in $\mathbb{R}^{n-1}$ of C. Fix $\varepsilon > 0$. Let $\varepsilon' = \frac{\varepsilon}{2v(Q)}$. Since C is compact, ϕ is uniformly continuous on C. Therefore $\exists \delta > 0$ such that $\forall x, y \in C$, $|x - y| < \delta \implies |\phi(x) - \phi(y)| < \varepsilon'$. Let P be a partition of Q such that every side length of a subrectangle is less than δ. If $R \cap C \neq \varnothing$ pick any point $x_R \in R \cap C$ and let $I_R = [\phi(x_R) - \varepsilon', \phi(x_R) + \varepsilon']$. Then,

$$G_\phi \subseteq \bigcup_{R \cap C \neq \varnothing} R \times I_R.$$

Then,

$$\sum_R v(R \times I_R) \leq \sum_R v(R) 2\varepsilon' = 2\varepsilon' v(Q) = \varepsilon.$$

□

(b) Now we use that C is rectifiable to show D is measure zero. Note that $\exists \alpha > 0$ such that $|\phi(x)|, |\psi(x)| \leq \alpha, \forall x \in C$. We know that C is rectifiable in $\mathbb{R}^{n-1} \implies \operatorname{Bd} C$ has measure zero in $\mathbb{R}^{n-1}$. Let $\varepsilon > 0$ be given. $\exists^\infty Q_i$ whose union covers $\operatorname{Bd} C$ and whose volume sum is $\leq \frac{\varepsilon}{2\alpha}$.

Then $\{Q_i \times [-\alpha, \alpha] : i \in \omega\}$ covers D and has volume sum less than ε.

all of which are measure zero, so S is measure zero. $\square$

The following theorem is a work horse!

Theorem. (Fubini's Theorem for Simple Regions)
Suppose $S \subset \mathbb{R}^n$ is a simple region and $f : S \to \mathbb{R}$ is continuous. Then f is integrable over S and

$$\int_S f = \int_C \int_{t\in[\phi(x),\psi(x)]} f(x,t).$$

Proof.
Read the book. $\square$

Example: Let S be the solid tetrahedron in $\mathbb{R}^3$ with vertices $(0,0,0),(1,0,1),(0,1,1),(0,0,1)$. Let $f(x,y,z) = z$ be the mass density function. Find the total mass $M = \int_S f$. Let C be the projection of the tetrahedron onto $\mathbb{R}^2$.

$$\begin{aligned}\int_S f &= \int_C \int_{z\in[x+y,1]} z \\ &= \int_{x=0}^{x=1} \int_{y=0}^{y=1-x} \int_{z=x+y}^{1} z \\ &= \frac{1}{2}\end{aligned}$$

There is no clear topological criterion to determine the rectifiability of a set.

Examples:

Example of an open bounded set that is not rectifiable. Write $\mathbb{Q} \cap [0,1] = \{q_1, q_2, \ldots\}$. Fix $0 < r < 1$, $\forall i \in \mathbb{N}$, $\exists$ open interval such that $q_i \in (a_i, b_i) \subseteq (0,1)$ and $b_i - a_i < \frac{r}{2^i}$. Let

$$A = \bigcup_{i\in\mathbb{N}} (a_i, b_i)$$

We will show that $A \neq [0,1]$. Note that A is open, bounded, and $A \subsetneq [0,1]$. Suppose for the sake of contradiction that A is rectifiable, then $\operatorname{Bd} A$ has measure zero. Let $\varepsilon = 1 - r$. Then $\exists^\infty Q_i$ which are open rectangles that cover $\operatorname{Bd} A$ and whose volume sum is less than ε. Note that $\overline{A} = [0,1]$ since $\mathbb{Q} \subset \mathbb{Q}$. So

$$\overline{A} = A \cup \operatorname{Bd} A \subseteq \left(\bigcup_{i\in\omega} Q_i\right) \cup \left(\bigcup_{i\in\omega} (a_i, b_i)\right)$$

Since $[0,1]$ is compact we have

$$\overline{A} \subseteq \left(\bigcup_{i\in[k]} Q_i\right) \cup \left(\bigcup_{i\in[\ell]} (a_i, b_i)\right)$$

$$1 \le \sum_{i\in[k]} (b_i - a_i) + \sum_{i\in[\ell]} v(Q_i) < r + \varepsilon = 1$$

a contradiction.

IBL 7(c) Diversion: Suppose $\{f_n\}$ is a Cauchy sequence of functions in $X = C([a,b],\mathbb{R}^n)$. and $f_n(x) \to f(x)$ for all $x \in [a,b]$ pointwise. Together these show $f_n \to f$ uniformly. *Proof.*

Let $\varepsilon > 0$. Being Cauchy, $\exists N$ such that $d(f,f_n) \leq \varepsilon$, for all $k, n \geq N$. Let $x \in [a,b]$ and $n \geq N$ be fixed. Consider $k \geq N$ and observe that

$$|f_k(x) - f_n(x)| \leq d(f_k, f_n) \leq \varepsilon.$$

Now $|f(x) - f_n(x)| = \lim_{k\to\infty} |f_k(x) - f_n(x)| \leq \varepsilon$. This holds for all $x \in [a,b]$ so N witnesses the uniform convergence of f_n to f. $\square$

IMPROPER INTEGRATION

We now turn our attention to Improper Integrals.

Definition. (Unbounded Integration)
Let Ω be an open set in $\mathbb{R}^n$ (not necessarily bounded). Let $f : \Omega \to \mathbb{R}$ is continuous (not necessarily bounded).

(a) If $f(x) \geq 0 \, \forall x \in \Omega$, define

$$\int_\Omega f = \sum_D \int_D f$$

where D are compact, rectifiable subsets of Ω. So this integral can be $+\infty$. We say that f is **extended integrable** if $\int_\Omega f$ is finite.

(b) For general f, we say that f is **extended integrable** if $\int_\Omega |f| < \infty$. When this occurs we define

$$\int_\Omega f = \int_\Omega f_+ - \int_\Omega |f_-|$$

Recall to show $\sup_D \alpha_D = M$ we need to show (1) $\forall D$, $\alpha_D \leq M$ and (2) $\exists^\infty D_i$, $\lim_{i\to\infty} \alpha_{D_i} = M$.

Example:
Let's do (2) first.

$$\begin{aligned}
\int_{(0,1)} \frac{1}{\sqrt{x}} &\geq \int_{[\frac{1}{k}, 1-\frac{1}{k}]} \frac{1}{\sqrt{x}} \\
&= 2\sqrt{1 - \frac{1}{k}} - 2\sqrt{\frac{1}{k}} \\
&\geq \lim_{k\to\infty} 2\sqrt{1 - \frac{1}{k}} - 2\sqrt{\frac{1}{k}} \\
&= 2
\end{aligned}$$

Now let D be an arbitrary compact, rectifiable subset of $(0,1)$. Being a compact set D has a maximum and minimum which we call a and b respectively. We have $D \subseteq [a,b] \subseteq (0,1)$. Then for every D, $\int_{[a,b]} \frac{1}{\sqrt{x}} = 2\sqrt{b} - 2\sqrt{a} \leq 2 \implies \int_D \frac{1}{\sqrt{x}} \leq 2$.

Combining these inequalities gives us that $\int_{(0,1)} \frac{1}{\sqrt{x}} = 2$.

In conclusion, this is terrible and cumbersome! We need a better way to compute! $\square$

Definition. (Exhaustion)
For an open set $\Omega \subseteq \mathbb{R}^n$, an exhaustion of Ω by compact, rectifiable subsets is a sequence $C_1, C_2, \ldots$ of subsets of $\mathbb{R}^n$ satisfying:

(a) $C_k \subseteq \Omega$ and C_k is compact, rectifiable for all $k \in \omega$.
(b) $\bigcup_{k\in\omega} C_k = \Omega$.
(c) $C_k \subseteq \operatorname{Int} C_{k+1}$ for all $k \in \omega$.

Example:

$$(0,1) = \bigcup_{k=100}^{\infty} [\frac{1}{k}, 1 - \frac{1}{k}]$$

$$\mathbb{R}^2 = \bigcup_{k=100}^{\infty} [-k, k]^2$$

Lemma. (Open Sets are Exhaustible)
Every open set in $\mathbb{R}^n$ has an exhaustion by compact, rectifiable subsets.

This is a topological statement. We will use this fact is several places down the road.
Proof.
Recall: If C is a closed set in $\mathbb{R}^n$ the function $h(x) = d(x, C)$ is continuous.

Let Ω be an open set in $\mathbb{R}^n$. For each $k \in \mathbb{N}$, define $D_k = \left\{x \in \mathbb{R}^k : |x| \leq k \ \wedge \ d(x, \mathbb{R}^n \setminus \Omega) \geq \frac{1}{k}\right\}$. Note that D_k is compact since it is the finite intersection of the preimages of closed sets under a continuous function (distance). Now check that $\bigcup_{k\in\mathbb{N}} D_k = \Omega$.

We will modify the D_ks to produce compact, rectifiable sets. Fix k and consider $x \in D_k$. Then $x \in \operatorname{Int} D_{k+1}$, so there is a closed cube Q_x such that $x \in \operatorname{Int} Q_x \subset Q_x \subset \operatorname{Int} D_{k+1}$. Then we have $D_k = \bigcup_{x\in D_k} Q_x$. Since D_k is compact we have $x_1, \ldots, x_m \in D_k$ such that $\bigcup_{i=1}^{\ell} Q_{x_i} = C_k \supseteq D_k$. Note that C_k is rectifiable (by inclusion exclusion), compact, and also $D_k \subseteq C_k \subseteq \operatorname{Int} D_{k+1} \implies C_k \subset \operatorname{Int} C_{k+1}$. $\square$

Theorem. (Criterion)
Let $\Omega \subseteq \mathbb{R}^n$ be open. Let $f : \Omega \to \mathbb{R}$ be continuous. Lt $\{C_k\}_{k\in\omega}$ be an exhaustion of Ω by compact rectifiable subsets.

(a) If $f \geq 0$ on Ω,
$$\int_{\Omega} f = \lim_{k\to\infty} \int_{C_k} f$$
no matter if this is finite or infinite.

(b) For general f, f is **extended integrable** on Ω iff the sequence $S_k = \int_{C_k} |f|$ is a bounded sequence (of reals). Furthermore if it is **extended integrable**, then
$$\int_{\Omega} f = \lim_{k\to\infty} \int_{C_k} f$$

Proof.[2]
(a) Note that since $f \geq 0$ on Ω, the sequence $t_k = \int_{C_k} f$, $k \in \mathbb{N}$ is an increasing sequence of real numbers. Therefore, $\lim_{k\to\infty} t_k$ must converge in $\mathbb{R}^+ \cup \{+\infty\}$.

We will show inequality in both directions. Since C_k is compact and rectifiable for all $k \in \mathbb{N}$, $\int_{C_K} f \leq \sup_D \int_D f = \int_\Omega f$. The inequality persists when taking the limit in k. So $\lim_{k\to\infty} \int_{C_k} f \leq \int_\Omega f$.

Let D be an arbitrary compact rectifiable subset of Ω. Note that $\bigcup_{k\in\mathbb{N}} \operatorname{Int} C_k = \Omega$ since $C_k \subset \operatorname{Int} C_{k+1}$. Therefore $\{\operatorname{Int} C_k : k \in \mathbb{N}\}$ is an open cover of D. Being compact we have $D \subseteq \bigcup_{k=1}^{\ell} \operatorname{Int} C_k = \operatorname{Int} C_\ell \subset C_\ell$. Therefore,

$$\int_D f \leq \int_{C_\ell} f \leq \lim_{k\to\infty} \int_{C_k} f$$

[2]You should be able to DO THIS proof!

Since the RHS does not include D, we have $\sup_D \int_D f \leq \lim_{k\to\infty} \int_{C_k} f$. ☐

Example:

(a)

$$\int_{(0,1)} \frac{1}{\sqrt{x}} = \lim_{k\to\infty} \int_{[\frac{1}{k},1-\frac{1}{k}]} \frac{1}{\sqrt{x}}\,dx = 2$$

(b) $f(x,y) = \frac{1}{x^2y^2}$. Let $\Omega = \{(x,y) \in \mathbb{R}^2 : (x,y) > (1,1)\}$.

$$\int_\Omega f = \lim_{k\to\infty} \int_{[1+\frac{1}{k},k]^2} \frac{1}{x^2y^2} = \lim_{k\to\infty} \int_{[1+\frac{1}{k},k]} \int_{[1+\frac{1}{k},k]} \frac{1}{x^2y^2}\,dy\,dx = \lim_{k\to\infty} \left(-\frac{1}{k} + \frac{1}{1+\frac{1}{k}}\right)^2 = 1$$

(c) $f(x,y) = \frac{1}{x^2y^2}$. Let $\Omega_0 = \{(x,y) \in \mathbb{R}^2 : 0 < x,\ y < 1\}$.

$$\int_{\Omega_0} f = \lim_{k\to\infty} \int_{[\frac{1}{k},1-\frac{1}{k}]^2} \frac{1}{x^2y^2} = \lim_{k\to\infty} \left(-k + \frac{1}{1-\frac{1}{k}}\right)^2 = \infty$$

Theorem. (Properties of Improper Integrals)
Let $\Omega \subseteq \mathbb{R}^n$ be open and $f, g : \Omega \to \mathbb{R}$ be continuous.

(a) f, g are extended integrable $\implies \alpha f + \beta g$ is extended integrable.

(b) If g dominates f, then the extended integral satisfies the monotonicity criterion.

$$\int_\Omega f \leq \int_\Omega g.$$

Also

$$\left|\int_\Omega f\right| \leq \int_\Omega |f|.$$

(c) Monotonicity of the domain for $\mathcal{U} \subseteq \Omega$.

(d) Inclusion-exclusion holds.

Proof.
See Theorem 15.3 in the book. ☐

Theorem. (Equivalence of Integrals)
Let Ω be a bounded, open set in $\mathbb{R}^n$ and $f : \Omega \to \mathbb{R}^n$ a bounded continuous function.

(1) Then f is extended integrable over Ω.

(2) If f is (ordinary) integrable over Ω, then the (ordinary) integral equals the extended integral.

Corollary. Suppose $S \subset \mathbb{R}^n$ is a bounded set and $f : S \to \mathbb{R}$ is a bounded, continuous function. If f is (ordinary) integrable over S, then

$$\int_{\operatorname{Int} S} f = \int_S f$$

where the LHS is the extended integral and the RHS is the ordinary integral.

Proof.
Take a look at the textbook. □

Theorem. (Open Exhaustion)
Suppose Ω is open in $\mathbb{R}^m$ and $\Omega_1 \subset \Omega_2 \subset \cdots$ are all open with $\bigcup_{k\in\mathbb{N}} \Omega_k = \Omega$. Further suppose that $f : \Omega \to \mathbb{R}$ is continuous.

(1) If $f \geq 0$ on Ω, $\mathrm{Int}_\Omega f = \lim_{k\to\infty} \int_{\Omega_k} f$.
(2) For general f, f is extended integrable over Ω iff $\alpha_k = \int_{\Omega_k} |f|$ is a bounded sequence. When this occurs $\mathrm{Int}_\Omega f = \lim_{k\to\infty} \int_{\Omega_k} f$.

Proof.
Read the textbook (the proof is quite short). □

We now have three ways to compute the integral:

(1) Exhaustion of compact rectifiable sets.

(2) Open Exhaustion.

(3) Simple regions?

<u>Notation</u>: We simply write $\int_\Omega f$ for the extended integral of f over Ω. □

CHANGE OF VARIABLES

We have not discussed yet how to compute the surface area of an object in $\mathbb{R}^3$ this is chapter 4. We also have not discussed integration over paths in vector fields this is chapter 5. We now discuss change of variables.

Example:

(a)

$$\int_1^2 2x(x^2+1)\,dx = \int_2^5 t\,dt$$

where $t = x^2+1$.

(b)

$$\int_0^1 \sqrt{1-x^2}\,dx = \int_0^{\pi/2} (\cos\theta)\underbrace{(\cos\theta)\,d\theta}_{dx}$$

where $x = \sin\theta$.

We will now study the n-dimensional analogue. We also want to do these computations for the extended integrals like $\int_{\mathbb{R}^2} e^{-x^2-y^2}$.

The Change of Variables Theorem

Recall:

Proposition. (Change of Variable in $\mathbb{R}$)
Let $I = [a,b]$. Suppose
(a) $g : I \to \mathbb{R}$ is C^1 and $g'(x) \neq 0,\ \forall x \in (a,b)$.
(b) $f : g(I) \to \mathbb{R}$ is continuous.
Then,

$$\int_{g(a)}^{g(b)} g(t)\,dt = \int_a^b (f\circ g)(x)g'(x)\,dx$$

when that $g(a) < g(b)$ $(g'(x) > 0)$.

Equivalently,

$$\int_{g(I)} g(t) = \int_a^b (f\circ g)\cdot |g'|$$

This latter statement extends naturally to higher dimensions.

Proof.
Read the proof in the textbook, again. □

Intuition for $n=2$:
Consider the area A of a parallelogram, the convex hull of $\{\vec{0},\vec{u},\vec{v},\vec{v}+\vec{u}\}$.
Let $\vec{v} = (c,d)$ and $\vec{u} = (a,b)$.

$$A = \|u\|h = \|u\|\|v\|\sin\theta$$

So

$$A^2 = \|u\|^2\|v\|^2(1-\cos^2\theta) = \|u\|^2\|v\|^2 - (u\cdot v)^2 = a^2b^2 + b^2c^2 - 2acbd = (ad-bc)^2$$

So

$$A = \left|\det \begin{bmatrix} a & c \\ b & d \end{bmatrix}\right|$$

$$\int_{g(A)} f = \sum_{i\in[k]} \sum_{j\in[m]} \int_{g(A_{ij})} f \approx \sum_{i\in[k]} \sum_{j\in[m]} f(g(x_i, y_i))\Delta g(A_{ij})$$

Note that if g is C^1 then $\Delta g(A_{ij})$ is approximately the image of a linear transformation. The columns of this linear transformation are given by $\begin{bmatrix} g(x,y) & g(x,y+\Delta y) \end{bmatrix}^T$, $\begin{bmatrix} g(x,y) & g(x+\Delta x, y) \end{bmatrix}^T$. For small $\Delta(x,y)$ then this is linear transformation is approximately $\begin{bmatrix} D_1 g(x,y)\Delta x & D_2 g(x,y)\Delta y \end{bmatrix}$. So the area of $g(A_{ij})$ is approximately $\left|\det \begin{bmatrix} D_1 g(x,y)\Delta x & D_2 g(x,y)\Delta y \end{bmatrix}\right| = |(\Delta A_{ij}) \det Dg|$ since det is multilinear. This is good intuition, now let's prove it!

Now we return to Rigor

Definition. (Diffeomorphisms in $\mathbb{R}^n$)
Let $A, B \subseteq \mathbb{R}^n$ be open sets. Let $r \geq 1$. We say $g : A \to B$ is a **diffeomorphism** (of class C^r) between A and B provided that

(i) g is a bijection.

(ii) g and g^{-1} are C^r.

We simply say diffeomorphism when $r = 1$. We call such a function a **homeomorphism** when $r \geq 0$.

Lemma. (Condition for Diffeomorphism)
Let $A \subseteq \mathbb{R}^n$ be open. Suppose $g : A \to \mathbb{R}^n$ is 1-1 (injective) and C^r on A. Then, $\det Dg(x) \neq 0$ for all $x \in A$ if and only if $g(A)$ is open in $\mathbb{R}^n$ and $g^{-1} : g(A) \to A$ is C^r.

Proof.
Do it! (Use the Inverse Function Theorem). $\square$

Theorem. (Change of Variables in $\mathbb{R}^n$)
Let $A, B \subseteq \mathbb{R}^n$ be open. Let $g : A \to B$ be a diffeomorphism and $f : B \to \mathbb{R}$ a continuous function.

(1) If $f \geq 0$ on B, then
$$\int_{g(A)} f = \int_A (f \circ g) \cdot |\det Dg|$$

(2) For general f, f is (extended) integrable over $g(A)$ if and only if $(f \circ g) \cdot |\det Dg|$ is (extended) integrable over A. If so, then (1) holds.

Note:
Letting $u = g(x)$, gives

$$Dg = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \cdots & \frac{\partial u_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_n}{\partial x_1} & \cdots & \frac{\partial u_n}{\partial x_n} \end{bmatrix} = \frac{\partial(u_1, \ldots, u_n)}{\partial(x_1, \ldots, x_n)}$$

You can imagine that

$$du_1\, du_2 \cdots du_n = \left|\det \frac{\partial(u_1, \ldots, u_n)}{\partial(x_1, \ldots, x_n)}\right| dx_1\, dx_2 \cdots dx_n$$

Even though this has no formal meaning.

Examples:

(a) Consider A bounded by $y = 2x, y = x, y = \frac{2}{x}, y = \frac{1}{x}$. Compute $\int_A \frac{2y}{x} e^{xy}$. Integral exists since A is a simple region (it is rectifiable) and f is continuous on A. Recall that the integral is not affected if we compute over $\operatorname{Int} A$ (the domain is now open). We will use the change of variables $u = xy$ and $v = \frac{y}{x}$. Then $g(x, y) = (xy, \frac{y}{x})$. Then $g(A) = [1, 2]^2$. Now we check the conditions of the Change of Variables function. $g \in C^r \checkmark$ since it is componentwise differentiable. We need to check g is 1-1. Also $\det Dg \neq 0$.

$$Dg = \begin{bmatrix} y & x \\ -\frac{y}{x^2} & -\frac{1}{x} \end{bmatrix}$$

So $\det Dg = \frac{2y}{x} \neq 0$ for $(x, y) \in A$.

$$\int_A \frac{2y}{x} e^{xy} = \int_{[1,2]^2} f = \int_{[1,2]^2} e^u\, du\, dv = \int_1^2 \int_1^2 e^u\, du\, dv = e^2 - e$$

where we guess $f(u, v) = e^u$.

Polar, Cylindrical, Spherical Coordinates.

In Polar coordinates we represent (x, y) by (r, θ) where r is the distance from the origin to (x, y) and θ is the angle from the x axis to (x, y). So $(x, y) = (r \cos\theta, r \sin\theta)$.

Take $g : (r, \theta) \mapsto (x, y)$ given by the Polar change of coordinates. This is C^∞. First we compute det Dg.

$$\left|\det \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}\right| = r$$

So "$dx\,dy = r\,dr\,d\theta$." We sometimes call this the **Jacobian**.

Example:

Let B be the upper half of the annulus bounded by the unit circle and the circle of radius 2 about the origin. Compute $\int_B x+y$. Then

$$\int_B x+y = \int_0^\pi \int_1^2 (r\cos\theta + r\sin\theta) r\,dr\,d\theta = \frac{2}{3}(8-1) = \frac{14}{3}$$

Example:

Let $B = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 < 9\}$. Compute $\int_B \sqrt{x^2+y^2}$. Converting to Polar gives $A = \{(r,\theta) : 0 \le r < 3 \wedge 0 \le \theta < 2\pi\}$. But now this is a problem since A is not open and the map $g : A \to B$ is not 1-1. We can remedy this by replacing A with $A' = \{(r,\theta) : 0 < r < 3 \wedge 0 \le \theta < 2\pi\} \cup \{(0,0)\}$. Then $g : A' \to B$ is 1-1, but A' is still not open. The Change of variables theorem is still not applicable.

Let $U = \operatorname{Int} A$. Let $V = B \setminus \{(x,0) : 0 \le x < 3\}$. Then U and V are both open and g is 1-1: g is a diffeomorphism.

$\int_B \sqrt{x^2+y^2} = \int_V \sqrt{x^2+y^2} = \int_U r \cdot r = \int_0^{2\pi} \int_0^3 r^2\,dr\,d\theta = 2\pi \cdot 9 = 18\pi$.

In Cylindrical coordinates, we represent (x,y,z) by (r,θ,z). Project (x,y,z) onto $\mathbb{R}^2$ then use polar coordinates and keep z the same.

Computing $|\det Dg| = r$. So "$dx\,dy\,dz = r\,dr\,d\theta\,dz$"

Example:

Let $a,h > 0$. Put $S = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 \le a^2,\ \frac{h}{a}\sqrt{x^2+y^2} \le z \le h\}$. Then S is a cone capped above by h with its apex pointing downward and located on the origin. Looks like an ice cream cone but someone ate the top.

Compute $\int_S x^2+y^2$. Let $B = \operatorname{Int} S \setminus \{(x,0,z) : x \ge 0, z \in \mathbb{R}\}$. Then $A = \{(r,\theta,x) : 0 < r < a, 0 < \theta < 2\pi,\}$. Then $A = \{(r,\theta,z) : 0 < r < a, 0 < \theta < 2\pi, \frac{h}{a} < z < h\}$.

So

$$\begin{aligned} \int_S x^2+y^2 = \int_B x^2+y^2 = \int_A r^2 \cdot r = \int_0^{2\pi}\int_0^a \int_{\frac{h}{a}r}^h r^3\,dz\,dr\,d\theta \\ = 2\pi\left(\int_0^a \left(hr^3 - \frac{h}{a}r^4\right) dr\right) = 2\pi h\left(\frac{1}{4} - \frac{1}{5}\right)a^4 = \frac{a^4\pi h}{10} \end{aligned}$$

In Spherical coordinates, we represent (x,y,z) by (ρ,θ,ϕ). ρ is the distance from the origin to (x,y,z). ϕ is the angle from the positive z-axis to (x,y,z), $0 \le \phi \le \pi$. θ is the angle from x-axis to (x,y,z) projected onto $\mathbb{R}^2$.

$$\begin{aligned} x &= \rho\sin\phi\cos\theta \\ y &= \rho\sin\phi\sin\theta \\ z &= \rho\cos\phi \end{aligned}$$

Computing $\det Dg = \rho^2 \sin\phi$. So "$dx\,dy\,dz = \rho^2 \sin\phi\, d\rho\, d\phi\, d\theta$". There is a well defined way to generalize spherical coordinate systems to n-dimensions. The spherical coordinates decompose $\mathbb{R}^3$ into a positive real number and a spherical coordinate: $\mathbb{R}^+ \times S^n$.

Example:

Let $B = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 < 1, x > 0, y > 0, z > 0\}$. So B is the positive part of the unit ball.

Then $A = \{(\rho,\phi,\theta) : 0 < \rho < 1,\ 0 < \phi < \frac{\pi}{2},\ 0 < \theta < \frac{\pi}{2}\}$.

Change of variables gives us

$$\int_B z = \frac{\pi}{16}$$

Example:

Let $a > 0$. Let $S = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 \leq a^2\}$. Compute $v(S)$.

Let $A = \{(\rho,\phi,\theta) : 0 < \rho < a,\ 0 < \phi < \pi,\ 0 < \theta < 2\pi\}$. Then A is open and $S \supset g(A) = B = (\operatorname{Int} S) \setminus \{((x,0,z) : x \geq 0, z \in \mathbb{R}\}$ is open.

Check that $S \setminus B$ is measure zero in $\mathbb{R}^3$.

$$v(S) = \int_S 1 = \int_B 1 = \int_A 1 \cdot \rho^2 \sin\phi = 2\pi\left(\frac{a^3}{3} \cdot 2\right) = \frac{4\pi a^3}{3}$$

We will do more applications (section 20) later.

Partitions of Unity

This is a technical preparation for the subsequent chapters.

Definition. (Support of a Function)
The **support** of a function $\phi : \mathbb{R}^n \to \mathbb{R}$ is the **closure** of the set $\{x \in \mathbb{R}^n : \phi(x) \neq 0\}$.

Definition. (Partition of Unity)
Let $A \subseteq \mathbb{R}^n$ be open. A **partition of unity** on A is a sequence of continuous functions $\{\phi_k\}_{k\in\mathbb{N}}$ with $\phi_k : \mathbb{R}^n \to \mathbb{R}$ satisfying

(1) $\phi_k(x) \geq 0\ \forall k \geq 0,\ \forall x \in \mathbb{R}^n$.
(2) $S_k := \operatorname{supp} \phi_k \subseteq A$.
(3) $\forall x \in A$, there is a neighborhood of x that intersects only finitely many sets in $\{S_k : k \in \mathbb{N}\}$.
(4)
$$\sum_{i\in\mathbb{N}} \phi_i(x) = 1$$
$\forall x \in A$.

Example: $A = \mathbb{R}$.

(a) Let $\phi_1(x) = 1$ and $\phi_k(x) = 0$ for all $k \geq 2$. So $\{\phi_k\}$ is a partition of unity.

(b) Sometimes we want functions with compact support. So we could define

$$\phi(x) = \begin{cases} x+1 & x \in [-1,0] \\ 1-x & x \in [0,1] \\ 0 & \text{otherwise.} \end{cases}$$

Then define $\phi_k(x) = \phi(x-k)$ for $k \in \mathbb{Z}$. Then $\{\phi_k\}$ is a partition of unity on A with compact support (but not differentiable).

(c) Let's do this again but require C^1 functions with compact support. A Put,

$$f(x) = \begin{cases} \frac{1}{2}(1+\cos x) & x \in [-\pi,\pi]. \\ 0 & \text{otherwise.} \end{cases}$$

Then define $\phi_k(x) = f(x-\pi k)$ for $k \in \mathbb{Z}$. Then $\{\phi_k\}$ is a partition of unity on A with compact support and each function is C^1.

Can we do the same thing for $A = (0,1)$ or $A =$ closed unit disk.

Lemma. (The Chad Bump Function)
Let A be a rectangle in $\mathbb{R}^n$, then there exist C^∞ functions $\phi : \mathbb{R}^n \to \mathbb{R}$ such that
(i) $\phi(x) > 0$, $\forall x \in \operatorname{Int} Q$.
(ii) $\phi(x) = 0$, $\forall x \notin \operatorname{Int} Q$.

Proof.
Let $f : \mathbb{R} \to \mathbb{R}$ be

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

We only need to check that this is C^∞ at 0. Note that the decay at zero is faster than any polynomial, this is enough to show that it is C^∞ (check!).

Let $g(x) = f(x) \cdot f(1-x)$. So then g is C^∞ and $g(x) = 0$ for $x \notin [0,1]$. Otherwise $g(x)$ is a positive smooth function.

Now it is clear that for any interval of finite measure in $\mathbb{R}$ we can stretch / collapse / translate this g to make it work.

For higher dimensions let

$$\phi(x) = \prod_{k \in [n]} g\left(\frac{x_k - a_k}{b_k - a_k}\right)$$

where $\mathbb{R}^n \supset Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$.

Lemma. (Cube Partition Lemma)
Let $\mathscr{A}$ be a collection of open sets in $\mathbb{R}^n$. Let

$$A = \bigcup_{S \in \mathscr{A}} S.$$

Then, $\exists^\infty Q_k$ (countable) closed cubes in $\mathbb{R}^n$ such that $Q_k \subseteq A$ and

(i)
$$\bigcup_{k \in \mathbb{N}} \operatorname{Int} Q_k = A$$

(ii) Each Q_k is contained in one element of $\mathscr{A}$.

(iii) Each point of A has a neighborhood that intersects only finitely many of the sets in $\{Q_k : k \in \mathbb{N}\}$.

Proof.
Let $D_1, D_2, \ldots$ be an exhaustion of A by compact subsets. For convenience, let $D_i = \varnothing$ for $i \leq 0$. Let $B_k = D_k \setminus \operatorname{Int} D_{k-1}$ (which is closed and bounded), so B_k is compact. Moreover, $B_k \cap D_{k-2} = \varnothing$. Since $B_k \subset A$, $\forall x \in B_k$ there exists a closed cube C_x centered at x such that

(1) $C_x \subset A$,

(2) $C_x \cap D_{k-2} = \varnothing$, and

(3) C_x is contained in an element of $\mathscr{A}$.

Since $\left(\bigcup_{x \in B_k} \operatorname{Int} C_k\right) \supset B_k$ and B_k is compact, there is a finite covering. Let $\mathscr{C}_k$ be the collection of these finitely many closed cubes and write $\mathscr{C}_k = \{C_{k,1}, \ldots, C_{k,N_k}\}$.

Let

$$\mathscr{C} = \bigcup_{k \in \mathbb{N}} \mathscr{C}_k.$$

Condition (i) and (ii) are satisfied by the construction. (iii) is satisfied because $x \in B_k$ intersects at most the elements in $\mathscr{C}_{k+1}$ and $\mathscr{C}_{k-1}$ but no other $\mathscr{C}_j$ for $j \in \mathbb{N} \setminus \{k-1, k, k+1\}$. $\square$

Theorem. (Partition of Unity Theorem)
Let $\mathscr{A}$ be a collection of open sets in $\mathbb{R}^n$. Let A be the union of the elements of $\mathscr{A}$. Then there is a partition of unity $\{\phi_k\}_{k \in \mathbb{N}}$ on A additionally satisfying

(5) ϕ_k is C^∞ for all $k \in \mathbb{N}$.

(6) The sets $S_k \coloneqq \operatorname{supp} \phi_k$ are compact and rectifiable (actually closed cubes).

(7) S_k is contained in an element of $\mathscr{A}$, $\forall k \in \mathbb{N}$.

We say that the partition of unity of A is **dominated** by $\mathscr{A}$.

*Proof.*Let $Q_1, Q_2, \ldots$ be the closed cubes from the **Cube Partition Lemma**. Then $\forall k$, $\exists \psi_k : \mathbb{R}^n \to \mathbb{R}$ that is C^∞, has compact support, is positive on $\operatorname{Int} Q_k$, zero outside Q_k. So $\operatorname{supp} \psi_k = Q_k$ and $\psi_1, \psi_2, \ldots$ satisfy (1)-(3) and (5)-(7). Now we have to normalize $\{\psi_k\}$ so they sum to 1.

By (3), $\forall x \in A$, $\sum i \in \mathbb{N} \psi_i(x)$ has only finitely many nonzero terms (so this is a finite sum). Thus, the series converges absolutely (trivially). Put

$$\lambda(x) = \sum_{i \in \mathbb{N}} \psi_i(x).$$

By (3) again, $\forall x \in A$, there is a neighborhood of x on which all but finitely many ψ_is vanishes identically. This gives us that λ is locally C^∞, so λ is C^∞ on A. Also λ vanishes outside A.

We have $\lambda(x) > 0$ for all $x \in A$. Define

$$\phi_k(x) = \begin{cases} \frac{\psi_k(x)}{\lambda(x)} & x \in A, \\ 0 & x \notin A. \end{cases}$$

Since ψ_k was chosen to be supported only on Q_k whose interior is a proper subset of A. Thus ψ_k vanishes outside of Q_k, so ϕ_k is C^∞ on $\mathbb{R}^n$ (important to check!).

Now $\{\phi_k\}_{k \in \mathbb{N}}$ satisfies (1)-(7). □

Remark: If $D \subset A$ is compact, then there is an $N \in \mathbb{N}$ such that ϕ_i vanishes identically on D for all $i > N$.

We can use partitions of unity to show an extended integral exists. The following result is not useful computationally, but very useful theoretically.

Theorem. (Integral via Partitions of Unity)
Let $A \subseteq \mathbb{R}^n$ be open and $f : A \to \mathbb{R}$ continuous. Let $\{\phi_i\}_{i \in \mathbb{N}}$ be a partition of unity on A with compact rectifiable support.

(1) If $f \geq 0$ on A,

$$\int_A f = \sum \int_A \phi_i \cdot f = \sum_{i=1}^{\infty} \int_{\operatorname{supp} \phi_i} \phi_i \cdot f$$

(2) For general f, f is extended integrable if and only if

$$\sum \int_A \phi_i \cdot |f| < \infty$$

if so, (1) holds.

Proof.
(a) Let D be an arbitrary compact, rectifiable subset of A.
Then $\exists N \in \mathbb{N}$ such that $\sum_{i=1}^{\infty} \phi_i(x) = \sum_{i=1}^{N} \phi_i(x)$ for all $x \in D$.

$$\begin{aligned} \int_D &= \int_D \sum_{i=1}^{N} \phi_i(x) f(x) \\ &= \sum_{i=1}^{N} \int_D \phi_i \cdot f \\ &\leq \sum_{i=1}^{\infty} \int_D \phi_i \cdot f \end{aligned}$$

The last inequality follows from monotonicity of the domain since $D \subseteq A$.

This implies that

$$\int_A f = \sup_D \int_D f \leq \sum_{i=1}^{\infty} \int_A \phi_i \cdot f.$$

(b) Let $N \in \mathbb{N}$ be arbitrary. Consider

$$K = \bigcup_{i=1}^{N} \operatorname{supp} \phi_i.$$

Then we have,

$$\begin{aligned}
\sum_{i=1}^{N} \int_A \phi_i \cdot f &= \sum_{i=1}^{N} \int_{\operatorname{supp} \phi_i} \phi_i \cdot f \\
&= \sum_{i=1}^{N} \int_K \phi_i \cdot f \\
&= \int_K \sum_{i=1}^{N} \phi_i \cdot f \\
&\leq \int_A \sum_{i=1}^{N} \phi_i \cdot f \\
&\leq \int_A \sum_{i=1}^{\infty} \phi_i \cdot f \\
&= \int_A f
\end{aligned}$$

Taking $N \to \infty$ gives that

$$\sum_{i=1}^{\infty} \int_A \phi_i \cdot f \leq \int_A f.$$

Taking (a) together with (b) gives the equality.

(2) follows from (1) (check this!) □

Diffeomorphisms in $\mathbb{R}^n$

Recall that we can decompose linear transformations into elementary transformations. We will show that we can do a similar thing with diffeomorphisms in general.

In this section we will show the following results:

(a) Diffeomorphisms take compact rectifiable sets to compact rectifiable sets.

(b) Every diffeomorphism is locally composable of "elementary / primitive" diffeomorphisms.

Lemma. (Covering by Closed Cubes)
If $S \subseteq \mathbb{R}^n$ has measure zero and $\varepsilon, \delta > 0$, then S can be covered by countably many closed cubes of width less than δ and volume-sum less than ε.

Proof.
Since S is a measure zero set there is a covering $\{Q_i\}$ of rectangles whose volume sum is less than

or equal to $\frac{\varepsilon}{2}$. Fix some $i \in \mathbb{N}$. Enlarge Q_i to a rectangle which has side lengths which are integer multiples of δ such that the volume is less than $2v(Q_i)$. Then uniformly partition Q_i into finitely many cubes. Do this process for each Q_i, then the resulting set is a covering that satisfies the requirements. □

Lemma. (Diffeomorphisms and Volume)
Let $C \subseteq \mathbb{R}^n$ be a closed cube, $\mathcal{U} \subseteq C$ open, and let $g : \mathcal{U} \to \mathbb{R}^n$ be C^1. Then $\exists M > 0$ such that $|Dg(x)| \leq M$ for all $x \in C$.
In turn, this implies that there is a closed cube $D \subseteq \mathbb{R}^n$ such that $g(C) \subseteq D$ and $v(D) \leq (nM)^n \cdot v(C)$.

Proof.
Let a be the center of C. Fix $i \in [n]$ and consider $|g_i(x) = g_i(a)|$. The MVT tells us that $\exists p$ between x and a such that

$$g_i(x) - g_i(a) = Dg_i(p)(x - a)$$

So,

$$|g_i(x) - g_i(a)| = |Dg_i(p)(x - a)| \leq n|Dg_i(p)| \cdot |x - a| \leq nM|x - a|$$

Such an M exists since g is C^1 and C is bounded.

This tells us that $2M|x - a|$ can be the width of the i-th side of D. □

Theorem. (C^1 Maps Preserve Measure Zero)
Let $A \subseteq \mathbb{R}^n$ be open and $g : A \to \mathbb{R}^n$ be C^1. Then g maps sets of measure zero to sets of measure zero.

Corollary.
If $n < m$, then there are no surjective C^1 maps from $\mathbb{R}^n$ to $\mathbb{R}^m$.

To prove, we will tessellate E into countably many cubes and use the previous lemma to create a covering of $g(E)$ whose volume sum is less than ε.
Proof.
DO PROOF

<u>Remark</u>: There are surjective continuous maps $\mathbb{R}^n \to \mathbb{R}^m$. These maps are called space filling curves. One of the most famous is the Peano space filling curve.

Theorem. (Set Theoretic Properties of Diffeomorphisms)
Suppose $A, B \subseteq \mathbb{R}^n$ are open and $g : A \to B$ is a diffeomorphism. Let $D \subset A$ be compact.

(1) $g\,(\operatorname{Int} D) = \operatorname{Int} g(D)$ and $g\,(\operatorname{Bd} D) = \operatorname{Bd} g(D)$.
(2) If D is rectifiable, then $g(D)$ is rectifiable. The result also holds if D is not a necessary condition. The necessary condition is $\operatorname{Bd} D \subset A \ \wedge \ \operatorname{Bd} g(D) \subseteq B$.

Proof.
For any bijection ϕ if $S_2 \subseteq S_1$, then $\phi\,(S_1 \setminus S_2) = \phi(S_2) \setminus \phi(S_1)$. Similar statements hold for set union and intersection.

Write $E = g(D)$. Note that g and g^{-1} are open maps. So $g(\operatorname{Int} D) \subseteq E$ is open. Similarly, $g^{-1}(\operatorname{Int} E) \subseteq \operatorname{Int} D$. Thus, $\operatorname{Int} E \subseteq g(\operatorname{Int} D)$ (by applying g to both sides).

Hence, $g(\operatorname{Int} D) = \operatorname{Int} E$.

$g\left(A \setminus \overline{D}\right)$ is an open subset of $g(A \setminus D) \implies g\left(A \setminus \overline{D}\right) \subseteq \operatorname{Int} g(A \setminus D) \implies B \setminus g(\overline{D}) \subseteq \operatorname{Int}(B \setminus E) = B \setminus \overline{E} \implies \overline{E} \subseteq g(\overline{D})$.

Now use g^{-1} in the logic above to obtain $\overline{D} \subseteq g^{-1}(\overline{E})$.

Hence, $g(\overline{D}) = \overline{E}$.

Together, $\operatorname{Bd} E = \overline{E} \setminus \operatorname{Int} E = g(\overline{D}) \setminus \operatorname{Int} g(D) = g(\overline{D}) \setminus g(\operatorname{Int} D) = g\left(\overline{D} \setminus \operatorname{Int} D\right) = g(\operatorname{Bd} D)$.

If D is rectifiable, then $\operatorname{Bd} D$ has measure zero. Since $\operatorname{Bd} g(D) = g(\operatorname{Bd} D)$ and g preserves measure zero sets, we see that $g(D)$ is rectifiable.

Definition. (Primitive/Elementary Maps)
We call $h : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ **primitive** or **elementary** if $\exists i \in [n]$ such that h preserves the i-th coordinate.

$$h(x) = \begin{bmatrix} h_1(x_1, \ldots, x_n) \\ \vdots \\ x_i \\ \vdots \\ h_n(x_1, \ldots, x_n) \end{bmatrix}$$

Example:

(a) The map h given by $h(x_1, x_2) = (x_1 + x_2^2, x_2)$ is elementary.

(b) $h(x, y) = (y, x)$ is not elementary.

(c) $h(x, y) = (x + y, x)$ is not elementary.

(d) $h(x) = A(x)$ is elementary if

$$A = \begin{bmatrix} * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \end{bmatrix}$$

Lemma. (Decomposition of Linear Diffeomorphism)
Let $n \geq 2$. Let $h : \mathbb{R}^n \to \mathbb{R}^n$ with $h(x) = Mx$ where M is a matrix of rank n. Then there exists elementary Diffeomorphisms such that $h = h_k \circ \cdots \circ h_1$. Meaning, $M = M_k \cdots M_1$ for elementary M_j with $j \in [k]$.

Proof.
Recall that Gaussian Elimination is done by multiplying elementary matrices which consist of three forms: (1) swap two rows, (2) multiply a row by a scalar, and (3) add to a row a scalar multiple of another row. (2) and (3) are elementary Diffeomorphisms. We need to check (1) when $n = 2$ separately. This case is satisfied since

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

So in fact (1) is not a needed type, (2) and (3) would suffice. From the row reduction process, we obtain a decomposition of every matrix into a product of elementary matrices and a row-echelon form matrix. Since M is rank n, the row-echelon form obtained from M is I_n.

Lemma. (Decomposition of Affine Linear Diffeomorphism)
Let $n \geq 2$ and $h : \mathbb{R}^n \to \mathbb{R}^n$ be a diffeomorphism of the form $h(x) = Mx + c$. Then h can be decomposed into elementary diffeomorphisms.

Proof.
In light of the previous lemma, we only need to show that a translating map can be decomposed. This is easy $h_1(x) = x + (0, c_2, \ldots, c_n)$ and $h_1(x) = x + (c_1, 0, \ldots, 0)$ works. □

Theorem. (Diffeomorphisms are Locally Decomposable)
Let $n \geq 2$ and $A, B \subseteq \mathbb{R}^n$ be open. Let $g : A \to B$ be a diffeomorphism and $a \in A$. Then there is an neighborhood of a $\mathcal{U}_0$ contained in A and a sequence of primitive diffeomorphisms $\{h_j\}_{j \in [k]}$ such that

$$\mathcal{U}_0 \xrightarrow{h_1} \mathcal{U}_1 \xrightarrow{h_1} \cdots \xrightarrow{h_k} \mathcal{U}_n$$

where $g = h_k \circ \cdots \circ h_1$ on $\mathcal{U}_0$.

Proof.
<u>Case 1</u>: (Linear Case)
Suppose $a = 0, g(0) = 0, Dg(0) = I$. Write $g(x)$ with its coordinate functions. Let $h : A \to \mathbb{R}^n$ be $h(x) = (g_1(x), \ldots, g_{n-1}(x), x_n)$. Then h is primitive. Furthermore h is a diffeomorphism since it is (1) C^1 in its coordinates, (2) invertible on an open $\mathcal{U}_0$ by the Inverse Function Theorem, and (3) differentiable since we assumed that $Dg(0) = I_n$ so this gives us that $Dh(0) = I_n$.

We elaborate on (2): We have $h : V_0 \to V_1$ is a diffeomorphism for some neighborhood V_0 of zero contained in A and $0 = g(0) \in V_1 \subseteq B$.

Let $k : V_1 \to \mathbb{R}^n$ given by $k(y) = (y_1, \ldots, y_{n-1}, g_n(h^{-1}(y))$. This map is primitive and a diffeomorphism. This follows since Dk is the identity matrix up to the last row. The last row is $D(g_n \circ h^{-1})(0) = (Dg_n(0))\,(Dh^{-1}(0)) = e_n^T I_n^{-1} = e_n^T I_n =$ by the Chain Rule.

So then k is locally a diffeomorphism by the Inverse Function Theorem. Let the neighborhood obtained from the IFT be W_1 and W_2. Then set $\mathcal{U}_0 \coloneqq h^{-1}(W_1)$.

<u>Case 2</u>: (General Case) Let $a \in A$ be arbitrary. Denote $b = g(a)$ and $Dg(a) = M$. Since g is a diffeomorphism, M is non-singular.

Define,

$$
\begin{aligned}
h_1(x) &= x + a \\
h_2(x) &= x - b \\
h_2(x) &= M^{-1}x
\end{aligned}
\qquad\qquad
\begin{aligned}
\tilde{g} &= h_3 \circ h_2 \circ g \circ h_1 \\
&\Downarrow \\
\tilde{g}(0) &= 0 \\
D\tilde{g}(0) &= M^{-1}IMII = I.
\end{aligned}
$$

By Case 1 $\tilde{g}$ is the composition of primitive diffeomorphisms near the origin.

Thus $g = h_2^{-1} \circ h_3^{-1} \circ \tilde{g} \circ h_1^{-1}$. But everything on the RHS are all locally decomposable, since g is their composition then g is also locally decomposable. □

The Proof

Now we prove the Change of Variables Theorem. Even though we developed many tools, the proof will still take considerable time.

Notation: For an open set $B \subseteq \mathbb{R}^n$, write $C(B, \mathbb{R}_+)$ to mean the collection of $f : B \to \mathbb{R}$ that are continuous and nonnegative valued. Write $C_c(B, \mathbb{R}_+)$ to mean the subcollection of $f \in C(B, \mathbb{R}_+)$ such that $\operatorname{supp} f$ is a compact subset of B.

We now show,

Theorem. (Change of Variables for Nonnegative Functions)
Let $g : A \to B$ be a diffeomorphism where $A, B \subseteq \mathbb{R}^n$ are open. For every $f \in C(B, \mathbb{R}_+)$,

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

Lemma. (COV Lemma 1)
Local change of variables for functions with compact support
$\implies$ global change of variables for continuous functions.

Let $g : A \to B$. Suppose $x \in A$, there is an open set $\mathcal{U}$ such that $x \in \mathcal{U} \subseteq A$ with $\mathcal{V} = g(\mathcal{U})$, and

$$\int_{\mathcal{V}} f = \int_{\mathcal{U}} (f \circ g) \cdot |\det Dg|$$

for all $f \in C_c(\mathcal{V}, \mathbb{R}_+)$.
Then the change of variables theorem for nonnegative functions holds for g.

Proof.
Let $\mathscr{B}$ be the collection of open sets $\mathcal{V} = g(\mathcal{U})$. So

$$B = \bigcup_{\mathcal{V} \in \mathscr{B}} \mathcal{V}.$$

Let $\{\phi_k\}_{k \in \mathbb{N}}$ be a partition of unity on B, with compact rectifiable support, and which is dominated by $\mathscr{B}$, meaning $\operatorname{supp} \phi_k$ is contained in some $\mathcal{V} \in \mathscr{B}$.

Let $f \in C(B, \mathbb{R}_+)$. We compute

$$
\begin{aligned}
\int_B f &= \sum_{k=1}^{\infty} \int_B \phi_k \cdot f && \text{Theorem on Integrals via POU.} \\
&= \sum_{k=1}^{\infty} \left(\int_{\operatorname{supp} \phi_k} \phi_k \cdot f \right) && \phi_k = 0 \text{ on } B \setminus \operatorname{supp} \phi_k. \\
&= \sum_{k=1}^{\infty} \left(\int_{\mathcal{V}} \phi_k \cdot f \right) && \exists \mathcal{V} \in \mathscr{B} \text{ with } \operatorname{supp} \phi_k \subseteq \mathcal{V}. \\
&= \sum_{k=1}^{\infty} \left(\int_{\mathcal{U}} (\phi_k \circ g)(f \circ g) \cdot |\det Dg| \right) && \text{assumption of lemma.} \\
&= \sum_{k=1}^{\infty} \left(\int_{A} (\phi_k \circ g)(f \circ g) \cdot |\det Dg| \right) && \phi_k \circ g = 0 \text{ on } A \setminus \mathcal{U}. \\
&= \sum_{k=1}^{\infty} \left(\int_{A} (\psi_k)(f \circ g) \cdot |\det Dg| \right) && \text{Put } \psi_k = \phi_k \circ g \text{, then } \{\psi_k\}_{k \in \mathbb{N}} \text{ is a POU on } A. \\
&= \int_A \sum_{k=1}^{\infty} (\psi_k)(f \circ g) \cdot |\det Dg| && (f \circ g) \cdot |\det Dg| \text{ is continuous, nonnegative.} \\
&= \int_A (f \circ g) \cdot |\det Dg| && \text{Definition of POU.}
\end{aligned}
$$

□

Recall,

Proposition. (Change of Variables on $\mathbb{R}$)
If $f : [c, d] \to \mathbb{R}$ is continuous and $g : [a, b] \to [c, d]$ is C^1 with non vanishing derivative, then

$$\int_c^d f(x)\, dx = \int_a^b f(g(t)) \cdot |g'(t)|\, dt$$

Proof.
We already know this! □

Lemma. (COV Lemma 2)
The change of variables theorem holds for $n = 1$.

Proof.
Let $g : A \to B$ be a diffeomorphism where A, B are open in $\mathbb{R}$.[3] We use COV Lemma 1. Let $I = [a, b]$ such that $x \in (a, b) \subseteq [a, b] \subset A$. Let $J = g(I)$. Then, $J = [c, d]$ (why?).

Let $F \in C_c(\operatorname{Int} J, \mathbb{R}_+)$ and consider

$$\int_{\operatorname{Int} J} F = \int_{\operatorname{supp} F} F = \int_J F$$

since $\operatorname{supp} F \subset J$.

[3]To prove, one could use that every open set in $\mathbb{R}$ can be written as a countable union of open intervals. We will not proceed with the proof in this direction. $\mathbb{R}$ is second countable.

Now the Change of Variable Theorem on $\mathbb{R}$ tells us that this is

$$\int_I (F \circ g) \cdot |g'|.$$

$\square$

Lemma. (COV Lemma 3)
Let $n \geq 2$, if the Change of Variable Theorem holds for all **primitive** diffeomorphisms, then the theorem holds for all diffeomorphisms.

Proof.
Let $g : A \to B$ be a diffeomorphism for open $A, B \subseteq \mathbb{R}^n$. We use COV Lemma 1. Fix $x \in A$. Then since diffeomorphisms are locally decomposable, there is an open $\mathcal{U}_0 \subseteq A$ and primitive diffeomorphisms

$$\mathcal{U}_0 \xrightarrow{h_1} \mathcal{U}_1 \xrightarrow{h_1} \cdots \xrightarrow{h_{k-1}} \mathcal{U}_{k-1} \xrightarrow{h_k} \mathcal{U}_k$$

such that $g = h_k \circ \cdots h_1$ on $\mathcal{U}_0$.
Let $\mathcal{U} = \mathcal{U}_0$. Then $\mathcal{V} = g(\mathcal{U}) = \mathcal{U}_k$. Suppose $F \in C_c(V, \mathbb{R}_+)$. By the assumption of COV Lemma 1,

$$\int_{\mathcal{V}} F = \int_{\mathcal{U}_k} F = \int_{\mathcal{U}_{k-1}} \underbrace{(F \circ h_k) \cdot |\det Dh_k|}_{\in\ C(\mathcal{U}_{k=1}, \mathbb{R}_+)} = \int_{\mathcal{U}_{k-2}} (F \circ h_k \circ h_{k-1}) \cdot |\det Dh_k \circ h_{k-1}| \cdot |\det Dh_{k=1}|.$$

The last equality follows from applying the change of variables again. Checking the assumptions, the integrand on the RHS is continuous.

By the Chain Rule,

$$D(h_k \circ h_{k-1})(v) = (Dh_k) \circ h_{k-1}(v) \cdot Dh_{k-1}(v)$$

This gives us that

$$\int_{\mathcal{U}_{k-2}} (F \circ h_k \circ h_{k-1}) \cdot |\det D(h_k \circ h_{k-1})|$$

Repeating this process gives,

$$\int_{\mathcal{U}_0} (F \circ h_k \circ h_{k-1} \circ \cdots \circ h_1) \cdot |\det(Dh_k \circ h_{k-1} \circ \cdots \circ h_1)| = \int_{\mathcal{U}} (F \circ g) \cdot |\det Dg|$$

$\square$

Proof. (Change of Variables Theorem)
When $n = 1$, the theorem holds by COV Lemma 2. Suppose COV holds in dimension $n-1$. We will show that it holds in dimension n. By COV Lemma 3, it is enough to check for primitive diffeomorphisms.

Let $h : A \to B$ be a **primitive** diffeomorphism on open $A, B \subseteq \mathbb{R}^n$. Without loss of generality we may assume that h preserves the n-th coordinate. So for $(x,t) \in A \subseteq \mathbb{R}^{n-1} \times \mathbb{R}$, $h(x,t) = (k(x,t), t)$ for $k(x,t) \in \mathbb{R}^{n-1}$. We apply COV Lemma 1 again. Let $p \in A$ and denote $q = h(p) \in B$. Let Q be a rectangle in B such that $q \in \operatorname{Int} Q$. Put $S = h^{-1}(Q)$. S is rectifiable since it is the image of a rectifiable set under a diffeomorphism. Also, $h : \operatorname{Int} S \to \operatorname{Int} Q$ is a diffeomorphism. h maps horizontal lines in the $\mathbb{R}^{n-1} \times \mathbb{R}$ plane to horizontal lines in B. Write $Q = D \times I$, where D is a rectangle in $\mathbb{R}^{n-1}$ and $I \subseteq \mathbb{R}$ a closed interval. $\pi_n(S) = I$. There is a rectangle E in $\mathbb{R}^{n-1}$ that contains the projection of S into $\mathbb{R}^{n-1}$.

Fix $t \in I$. Consider the t-slice of S, $C = \{x \in \mathbb{R}^{n-1} \mid (x,t) \in S\} \subseteq \mathbb{R}^{n-1}$. Let $\ell : C \to D$ given by $\ell(x) = k(x,t)$. Then,

(a) ℓ is a bijection (check surjectivity).

(b) ℓ is C^1.

(c) $\det D\ell(x) = \det Dh(x,t) \neq 0$. Hence, ℓ is a diffeomorphism.

$$Dh(x,t) = \begin{bmatrix} * & * & * & \heartsuit \\ * & * & * & \heartsuit \\ * & * & * & \heartsuit \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

Since the $*$ block[4] is $\frac{\partial k}{\partial x}$, cofactor expansion gives us

$$\det D(x,t) = \det \frac{\partial k}{\partial x}(x,t) = \det \frac{\partial \ell}{\partial x}(x) = \det \ell(x).$$

This will allow us to apply Fubini's Theorem and use the induction Hypothesis. Suppose $F \in C_c(\operatorname{Int} Q, \mathbb{R}_+)$.

$$\int Int_Q F = \int_Q F = \int_{t \in I} \int_{y \in D} F(y,t)$$

Let $G(y) = F(y,t) \implies G \in C_c(\operatorname{Int} D)$.

$$\begin{aligned} \int_{y \in D} F(y,t) = \int_D G &= \int_{\operatorname{Int} D} G \\ &= \int_{\operatorname{Int} C} (G \circ \ell) \cdot |\det D\ell| \\ &= \int_{\operatorname{Int} C} F(\ell(y),t) \cdot |\det D\ell| \\ &= \int_{z \in C} (F \circ h)(x,t) \cdot |\det Dh(x,t)| \end{aligned}$$

This gives us

$$\begin{aligned} \int Int_Q F &= \int_{x \in I} \int_{x \in E} (F \circ h)(x,t) \cdot |\det Dh(x,t)| \\ &= \int_{I \times E} (F \circ h) \cdot |\det Dh| \\ &= \int_{E \times I} (F \circ h) \cdot |\det Dh| \\ &= \int_{\operatorname{Int} S} (F \circ h) \cdot |\det Dh| \end{aligned}$$

$\square$

[4] The $\heartsuit$ block is $\frac{\partial k}{\partial t}$.

More Applications of Change of Variables

In this section, we will work to understand (1) the geometric meaning of the determinant and (2) isometries of $\mathbb{R}^n$.

Theorem. (Determinant Gives Volume Distortion)
Let $A \in \mathbb{R}^{n\times n}$ and let $h : \mathbb{R}^n \to \mathbb{R}^n$ be given by $h(x) = Ax$. Then, for arbitrary rectifiable $S \subseteq \mathbb{R}^n$,

$$v(h(S)) = |\det(S)| \cdot v(S)$$

Proof.
Write $T \coloneqq h(S)$. There are two cases.

(1) Assume $\det A \neq 0$. Then h is a diffeomorphism. So T is rectifiable because h maps $\operatorname{Int} T$ to $\operatorname{Int} T$.

$$v(T) = v(\operatorname{Int} T) = \int_{\operatorname{Int} T} 1 = \int_{\operatorname{Int} S} 1 \cdot |\det Dh| = |\det(S)| \cdot v(S)$$

(2) Assume $\det A = 0$. The image of $\mathbb{R}^n$ under h is a proper subspace W of $\mathbb{R}^n$. $T \subset W$. Then the closure of T is contained in W. So $\overline{T}$ has measure zero in $\mathbb{R}^n$.

$$v(T) = \int_T 1 = \int_{\mathbb{R}^n} 1_T$$

1_T is continuous outside $\overline{T}$ which has measure zero.

So, the theorem holds.

$\square$

Definition. (Parallelepiped)
Let $a_1, \ldots, a_k$ be linearly independent vectors in $\mathbb{R}^n$. The k-dimensional **parallelepiped** with edges $a_1, \ldots, a_k$ is the set

$$\mathcal{P} = \mathcal{P}(a_1, \ldots, a_k) = \left\{\sum_{i=1}^{n} c_i a_i : c_1, \ldots, c_k \in [0,1]\right\}$$

<u>Example</u>:
In $\mathbb{R}^3$ there are three cases: $k = 1, 2, 3$

Theorem. (Volume of Parallelepiped)
The volume of an n-dimensional parallelepiped $\mathcal{P} \subset \mathbb{R}^n$ generated by $a_1, \ldots, a_n$ is

$$\left|\det \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}\right|$$

Proof.
Let $h(x) = A(x)$. Then $h([0,1]^n) = \mathcal{P}$. So, $v(\mathcal{P}) = 1 \cdot |\det A|$. $\square$

Now we understand $|\det A|$, but $\det A$ has a sign. Let's discuss the meaning of the sign.

Definition. (Frame)
Let V be a vector space of dimension n. An n-tuple $(a_1, \ldots, a_n)$ of linearly independent vectors in V is called an **n-frame** in V.

Why is this interesting: In physics, a rotating object has centripetal force $\vec{a}$. The velocity vector $\vec{v}$ is tangential to motion. Then $(\vec{v}(t), \vec{a}(t))$ is a frame in $\mathbb{R}^2$. The frame can tell us about the direction of rotation.

Definition. (Orientation of Frame)
An n-frame $(a_1, \ldots, a_n)$ in $\mathbb{R}^n$ is called **right-handed** if $\det[a_1 \cdots a_n] > 0$ and **left-handed** if $\det[a_1 \cdots a_n] < 0$.
The **positive orientation** in $\mathbb{R}^n$ is the collection of right-handed n-frames in $\mathbb{R}^n$. The negative orientation in $\mathbb{R}^n$ is the collection of left-handed n-frames in $\mathbb{R}^n$.

Example:
In dimension:

(1) the Positive Orientation is $\mathbb{R}_+$ and the Negative Orientation is $\mathbb{R}_-$.

(2) $(a, b) \in$ PO if rotating a to be parallel with b occurs faster from a counterclockwise rotation. Similarly, $(a, b) \in$ NO if rotating a to be parallel with b occurs faster from a clockwise rotation.

Proposition. (Orientation Under Linear Map)
Let M be an $n \times n$ nonsingular matrix. Let $(a_1, \ldots, a_n)$ be a frame in $\mathbb{R}^n$.
(a) If $\det M > 0$, then $(Ma_1, \ldots, Ma_n)$ and $(a_1, \ldots, a_n)$ have the same orientation.
(b) If $\det M < 0$, then $(Ma_1, \ldots, Ma_n)$ and $(a_1, \ldots, a_n)$ have opposite orientations.
(c) $(a, b, c) \in$ PO provided that the right hand rule holds. Position your hand parallel to a, curl your finger towards b, if your thumb points towards c, then the orientation is positive. If your thumb points away from c, then the orientation is negative.

Proof.(for $n = 2$)
Let $A = [a_1 \cdots a_n] \implies [Ma_1 \cdots MMa_n] = MA$. $\det MA = (\det M)(\det A)$. The proof follows. $\square$

Isometries on $\mathbb{R}^n$

Definition. (Isometry)
A map $h : \mathbb{R}^n \to \mathbb{R}^n$ is called an **isometry** if it preserves distance.

$$d(h(x), h(y)) = d(x, y)$$

Theorem. (Preserving Distance ≡ Preserving Angles)
Let $h : \mathbb{R}^n \to \mathbb{R}^n$ and $h(0) = 0$. Then, h is an isometry if and only if h preserves the inner product.

Note that the isometry preserves the distance induced by the innerproduct.

Proof.

$$\|u - v\| = (u - v) \cdot (u - v) = u \cdot u - 2u \cdot v + v \cdot v = \|u\|^2 - 2u \cdot v + \|v\|^2$$

($\Longleftarrow$) We want to express distance purely in terms of dot product.

$$\|h(x) - h(y)\|^2 = h(x) \cdot h(x) - 2h(x) \cdot h(y) + h(y) \cdot h(y) = x \cdot x - 2x \cdot y + y \cdot y = \|x - y\|^2$$

because dot products are preserved.

($\Longrightarrow$) We want to express dot product in terms of distances.

We have $u \cdot v = \frac{1}{2}(\|u\|^2 + \|v\|^2 - \|u - v\|^2)$, this is the **polarization identity**.

$$h(x) \cdot h(y) = \frac{1}{2}(\|h(x)\|^2 + \|h(y)\|^2 - \|h(x) - h(y)\|^2) = \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|x - y\|^2) = x \cdot y$$

because distance is preserved.

□

<u>Recall</u>: A matrix ix called **orthogonal** if $A^T A = I$.

Theorem. (Isometries Are Affine Orthogonal Maps)
Let $h : \mathbb{R}^n \to \mathbb{R}^n$, $h(0) = 0$. Then, h preserves the dot product if and only if h is a linear transformation with orthogonal matrix representation.

Corollary. $h : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry $\iff$ $\exists A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ and $b \in \mathbb{R}^n$ such that $h(x) = Ax + b$.

Proof.

($\Longleftarrow$)

$$h(x) \cdot h(y) = hh(x)^T h(y) = (Ax)^T Ay = x^T A^T Ay = x^T Iy = x \cdot y$$

($\Longrightarrow$) Let

$$A = \begin{bmatrix} | & & | \\ h(e_1) & \cdots & h(e_n) \\ | & & | \end{bmatrix}$$

$e_1, \ldots, e_n$ standard basis of $\mathbb{R}^n$.

First we check that $A^T A = I$.

$$A^T A = \begin{bmatrix} | & & | \\ h(e_1) & \cdots & h(e_n) \\ | & & | \end{bmatrix} = \sum_{i,j\in[n]} h(e_i)^T h(e_j) E_{ij} = \sum_{i,j\in[n]} e_i^T e_j E_{ij} = I$$

> **Corollary**. (Isometries Preserve Volume)
> If $h : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry and $S \subseteq \mathbb{R}^n$ is rectifiable, then $h(S)$ is rectifiable and $v(h(S)) = v(S)$.

Proof.
Let $h(x) = Ax + b$. Then $Dh = A$. So $\det(A)^2 = \det(A^T)\det(A) = \det(A^T A) = \det(I) = 1$. So $\det(A)$ is ± 1.

$$v(h(S)) = \int_{h(S)} 1 = \int_S 1 \cdot |\det Dh| = v(S).$$

$\square$

CONCLUDING MATERIAL

What Comes Next?

There is a lot more analysis after 395.
In 396

(a) We will compute integrals (mass, "volume" or "total charge") of scalar functions on **manifolds** (k-dimensional sets) in $\mathbb{R}^n$.

(b) We will compute integrals ("work" or "flux") of vector valued functions on manifolds.

(c) We will discuss scalar, vector, and tensor fields. We will discuss multilinear algebra including tensors, differential forms.

(d) We will extend to Fundamental Theorem of Calculus to **Stokes Theorem**.

In math there are several branches: analysis, algebra, differential geometry, algebraic geometry, topology, logic, combinatorics, and more.

Within analysis there are several branches: (1) real analysis, (2) complex analysis, (3) Fourier analysis, (4) functional analysis, and (5) differential equations.

"Enjoy your ... what should you enjoy?"

Final Exam Logistics

The topics that will be excluded are from before Midterm 1:

(a) IBL

(b) min/max

(c) Lagrange Multipliers.

Limits, continuity, partial differentiation, chain rule, inverse, and implicit function theorem could be covered.

Questions to Know

Answer the following questions with specificity:

1. What is a diffeomorphism? A diffeomorphism of Euclidean space is a differentiable function whose inverse exists and is differentiable.
2. If f is continuous, then f maps compact sets to compact sets. What can we say if f is C^1? If f is a diffeomorphism?
3. What are contented (rectifiable) sets?
4. What is the relation between rectifiable sets and measure zero sets?
5. How do we compute an integral?
6. What are some interesting integral computations?

7. What is an integral?
8. Under what conditions is an integral defined?
9. When should we use Change of Variables?
10. What is a partition of unity?
11. How does the behavior of a continuous function differ from a continuous function with compact support.
12. How are rectifiable sets related to integrals?
13. What difficulties does integration in $\mathbb{R}^n$ pose compared to integration in $\mathbb{R}$.
14. What is an exhaustion of a set?