richardwu.ca

MATH 247 Final Exam Guide

Calculus 3 (Advanced)

Spiro Kargiannis • Winter 2018 • University of Waterloo

Last Revision: April 10, 2018

Table of Contents

Abstract

These notes are intended as a resource for myself; past, present, or future students of this course, and anyone interested in the material. The goal is to provide an end-to-end resource that covers all material discussed in the course displayed in an organized manner. These notes are my interpretation and transcription of the content covered in lectures. The instructor has not verified or confirmed the accuracy of these notes, and any discrepancies, misunderstandings, typos, etc. as these notes relate to course's content is not the responsibility of the instructor. If you spot any errors or would like to contribute, please contact me directly.

1 Topology

Theorem 1.1 (Cauchy-Schwarz inequality)**.**

$$\|x \cdot y\| \leq \|x\|\|y\|$$

Definition 1.1 (Open ball)**.** Let $x \in \mathbb{R}^n$ and $r > 0$. The **open ball** at radius r centred at x is denoted

$$B_r(x) = \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$$

Definition 1.2 (Closed ball)**.** Let $x \in \mathbb{R}^n$, $r > 0$. The **closed ball** of radius $r > 0$ centred at x is denoted

$$\overline{B_r(x)} = \{y \in \mathbb{R}^n \mid \|x - y\| \leq r\}$$

Definition 1.3 (Open sets)**.** A subset $U \subseteq \mathbb{R}^n$ is called an **open set** (or open) *iff* $\forall x \in U$, $\exists r > 0$ (r depends on x) such that $B_r(x) \subseteq U$.

1. Let $U_\alpha \subseteq \mathbb{R}^n$ be open $\forall \alpha \in A$ (countably or uncountably many), then $\bigcup_{\alpha \in A} U_\alpha$ is open.
2. Let $U_1, \ldots, U_k$ be open (**must be finite** number of sets). Then $\bigcap_{j=1}^k U_j$ is open.

Definition 1.4 (Closed sets)**.** A subset $F \subseteq \mathbb{R}^n$ is called **closed** if $F^c = \mathbb{R} \setminus F$ is open.

1. If $F_1, \ldots, F_k$ is closed, then $\bigcup_{j=1}^k F_j$ is closed.
2. If F_α is closed $\forall \alpha \in A$, then $\bigcap_{\alpha \in A} F_\alpha$ is closed.

Definition 1.5 (Interior)**.** Let $A \subseteq \mathbb{R}^n$ (could be $\varnothing$). The interior of A^0 or $\text{int}(A)$ is

$$\bigcup_{\substack{V \subseteq A \\ V \text{open in } \mathbb{R}^n}} V$$

It is the union of **all** open subsets of $\mathbb{R}^n$ that are contained in A.

1. A^o is open.
2. A is open *iff* $A^o = A$.

Definition 1.6 (Closure)**.** Let $A \subseteq \mathbb{R}^n$ (could be $\varnothing$). The interior of $\overline{A}$ or $\text{cl}(A)$ is

$$\bigcap_{\substack{F \supseteq A \\ F \text{closed in } \mathbb{R}^n}} F$$

It is the intersection of **all** closed subsets of $\mathbb{R}^n$ that contains A.

1. $\overline{A}$ is closed.
2. A is closed *iff* $\overline{A} = A$.

The closure of the open ball $B_\epsilon(x)$ is the closed ball $\overline{B_\epsilon(x)}$.

Definition 1.7 (Boundary)**.** Let $A \subseteq \mathbb{R}^n$. We define the **boundary** of A denoted $\partial A = \mathrm{bd}(A)$ to be

$$\{x \in \mathbb{R}^n \mid B_\epsilon(x) \cap A \neq \varnothing, B_\epsilon(x) \cap A^c \neq \varnothing \quad \forall \epsilon > 0\}$$

That is, $x \in \partial A$ *iff* every open ball centred at x contains a point in A **and** a point in A^c.
Note that

$$\partial B_\epsilon(x) = \{y \in \mathbb{R}^n \mid \|y - x\| = \epsilon\} = \partial(\overline{B_\epsilon(x)})$$

(this is **not** true in general for all sets).

Proposition 1.1 (Characterization of boundary)**.** Let $A \subseteq \mathbb{R}^n$, then

$$\partial A = \overline{A} \setminus A^o$$

This follows from the two claims:

1. $$x \in \overline{A} \iff B_\epsilon(x) \cap A \neq \varnothing \quad \forall \epsilon > 0$$
2. $$x \notin A^o \iff B_\epsilon(x) \cap A^c \neq \varnothing \quad \forall \epsilon > 0$$

Definition 1.8 (Sequential characterization of limits)**.** Let (x_k) be a sequence of points in $\mathbb{R}^n, k \in \mathbb{N}$. We say (x_k) **converges** to a point $x \in \mathbb{R}^n$ *iff* for any $\epsilon > 0$, $\exists N \in \mathbb{N}$ (N depends on ϵ in general)

$$k \geq N \Rightarrow \|x_k - x\| < \epsilon$$

If (x_k) converges to x, we denote

$$\lim_{k \to \infty} x_k = x$$

where x is **the limit** of x_k.
The limit of a convergent sequence is **unique**.

Definition 1.9 (Neighbourhood)**.** Let $x \in \mathbb{R}^n$. A subset $U \in \mathbb{R}^n$ is called a **neighbourhood** of x if $\exists \epsilon_0 > 0$ such that $B_{\epsilon_0}(x) \subseteq U$.

Proposition 1.2 (Convergent sequences and closed sets)**.** $x \in \overline{A}$ *iff* $\exists (x_k) \in A$ such that $\lim_{k\to\infty} x_k = x$.

Definition 1.10 (Bounded sequences)**.** A sequence (x_k) in $\mathbb{R}^n$ is called **bounded** if $\exists M > 0$ such that

$$\|x_k\| \leq M \quad \forall k \in \mathbb{N}$$

Definition 1.11 (Cauchy sequences)**.** A sequence (x_k) is called **Cauchy** if for any $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$k, l \geq N \Rightarrow \|x_k - x_l\| < \epsilon$$

Proposition 1.3 (Convergent is Cauchy)**.** (x_k) is a convergent sequence *iff* it is Cauchy.

Lemma 1.1 (Convergence implies bounded)**.** Every convergent sequence is bounded.

Definition 1.12 (Subsequences)**.** Let (x_k) be a sequence in $\mathbb{R}^n$. Let $1 \leq k_1 < k_2 < \ldots < k_e < k_{e+1} < \ldots$ be a sequence of $1, 2, 3, 4, \ldots$. Then the corresponding sequence (y_l) (or (x_{k_l})) in $\mathbb{R}^n$ given by $y_l = x_{k_l}$ is called a **subsequence** of (x_k).

Proposition 1.4 (Subsequences converges to same limit)**.** Suppose $(x_k) \to x$. Then any subsequence (x_{k_l}) of (x_k) also converges to the same limit x.

Theorem 1.2 (Bolzano-Weierstrass)**.** In $\mathbb{R}^n$, every **bounded** sequence has a **convergent subsequence**.
This convergent subsequence is **not** unique.

Definition 1.13 (Connected sets)**.** Let E be a non-empty subset of $\mathbb{R}^n$.
We say E is **disconnected** if there exists a **separation** for E. A separation of E is a pair U, V open sets in $\mathbb{R}^n$ such that

1. $E \cap U \neq \varnothing$
2. $E \cap V \neq \varnothing$
3. $E \cap U \cap V = \varnothing$
4. $E \subseteq U \cup V$

E is **connected** if $\nexists$ any separation of E.

Theorem 1.3 (0,1 closed interval is connected)**.** Let $E = [0, 1] \subseteq \mathbb{R}$. Then E is connected.

Definition 1.14 (Convex sets)**.** A **non-empty** subset E of $\mathbb{R}^n$ is called **convex** if for any $x, y \in E$ then

$$tx + (1 - t)y \in E \quad \forall t \in [0, 1]$$

i.e. the line segment between any 2 points in E is contained in E.

Corollary 1.1 (Convex implies connected)**.** Any convex subset E of $\mathbb{R}^n$ is connected.
This implies that $\mathbb{R}^n$ is connected.

Definition 1.15 (Open cover)**.** Let E be a subset of $\mathbb{R}^n$. An **open cover** of E is a collection of open subsets U_α, $\alpha \in A$, of $\mathbb{R}^n$ such that

$$E \subseteq \bigcup_{\alpha \in A} U_\alpha$$

(finite or infinite union of open subsets).

Definition 1.16 (Compact sets)**.** The subset E is called **compact** *iff* every open cover of E admits a **finite subcover**.
That is, if $\bigcup U_\alpha$, $\alpha \in A$, is an open cover of E, then $\exists$ a finite subset A_0 of A such that

$$E \subseteq \bigcup_{\alpha \in A_0} U_\alpha$$

Theorem 1.4 (Heine-Borel)**.** Let E be a subset of $\mathbb{R}^n$. E is **compact** *iff* E is both **closed and bounded**.

2 Limits and continuity

Definition 2.1 (Limits of functions)**.** Let $V \subseteq \mathbb{R}^n$ be an *open set* with $x_0 \in V$. Let $f : V \setminus \{x_0\} \to \mathbb{R}^m$ for some m (i.e. f is defined at all points of V except *possibly* at x_0).
We say $\lim_{x\to x_0} f(x)$ exists and equals $L \in \mathbb{R}^m$ **iff** $\forall \epsilon > 0$, $\exists \delta > 0$ such that

$$0 < \|x - x_0\| < \delta \Rightarrow \|f(x) - L\| < \epsilon$$

(note that $B_\delta(x_0) \subseteq V$ must hold).

Example 2.1 (Showing limit does not exist)**. Key idea:** find some path (towards x) that does not have a constant limit.
Suppose we wish to find

$$\lim_{(x,y)\to(2,3)} \frac{(x-2)^2}{(x-2)^2 + (y-3)^2}$$

where $f(x, y)$ defined everywhere except $(2, 3)$.
Suppose we have paths/lines with slope m where $(y - 3) = m(x - 2)$. Along this line we have

$$\begin{aligned} f(x,y) &= \frac{(x-2)^2}{(x-2)^2 + (y-3)^2} \\ &= \frac{1}{1+m^2} \end{aligned}$$

So f is a constant function which depends on the slope of the line/path.

Example 2.2 (Showing limit does exist)**. Key idea:** use the definition and reduce $\|f(x) - L\|$ to $\|x - a\| < \delta$.
Suppose we wish to find

$$\lim_{(x,y)\to(0,0)} \frac{x^4}{x^2 + y^2}$$

We *expect the limit to converge* since the degree of the numerator is > degree of denominator, thus numerator $\to 0$ "much faster" than the denominator so the quotient should go to zero.
Observe that

$$\frac{x^2}{x^2+y^2} \leq 1 \quad (x,y) \neq (0,0)$$

Thus

$$\begin{aligned} |\frac{x^4}{x^2+y^2}| = \frac{x^4}{x^2+y^2} &= x^2\left(\frac{x^2}{x^2+y^2}\right) \\ &\leq x^2 && \frac{x^2}{x^2+y^2} \leq 1 \\ &\leq x^2 + y^2 \\ &< \delta^2 = \epsilon \end{aligned}$$

Thus we can take $\delta = \sqrt{\epsilon}$.

Proposition 2.1 (Sequential characterization of limits of functions)**.** For $f : V\setminus\{x_0\} \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\lim_{x\to x_0} f(x) = L$ *iff* the sequence $f(x_k)$ converges to L *for every* sequence (x_k) in $V \setminus \{x_0\}$ converging to x_0.

Example 2.3 (Showing limits DNE with sequential characterization)**.** Suppose we want to solve

$$\lim_{(x,y)\to(0,0)} \frac{x}{\sqrt{x^2+y^2}} \cos(\frac{1}{\sqrt{x^2+y^2}})$$

By sequential characterization of limits

$$\lim_{(x,y)\to(0,0)} h(x,y) = 0 \iff \lim_{k\to\infty} h(x_k, y_k) = 0$$

for all sequences $(x_k, y_k) \in \mathbb{R}^2$ converging to $(0,0)$.
Thus consider $(x_k, y_k) = (\frac{(-1)^k}{k\pi}, 0)$, so we have

$$\begin{aligned} h(x_k, y_k) &= \frac{(-1)^k \frac{1}{k\pi}}{\sqrt{\frac{1}{k^2\pi^2}}} \cos(\frac{1}{\sqrt{\frac{1}{k^2\pi^2}}}) \\ &= (-1)^k \cos(k\pi) \\ &= 1 \quad \forall k \end{aligned}$$

Similarly when $(x_k, y_k) = (\frac{(-1)^{k+1}}{k\pi}, 0)$, we have the limit approaching to -1. Since they have different limits, then the limit DNE so f_x is not continuous at $(0,0)$.

Proposition 2.2 (Properties of limits)**.** Let $f, g : V \setminus \{x_0\} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and suppose

$$\lim_{x\to x_0} f(x) = L \qquad \lim_{x\to x_0} g(x) = M$$

then

$$\begin{aligned} &\lim_{x\to x_0} (f(x) + g(x)) = L + M && \text{(additive)} \\ &\lim_{x\to x_0} cf(x) = cL && \text{(scalar multiplicative)} \\ &\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{L}{M} && \text{if } m = 1, M \neq 0 \\ &\lim_{x\to x_0} (f(x)g(x)) = LM && \text{if } m = 1 \end{aligned}$$

Definition 2.2 (Component functions)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, U is open. Then for $x \in U$

$$f(x) = (f_1(x), \dots, f_m(x)) \in \mathbb{R}^m$$

$f_i : U \to \mathbb{R}$, $1 \leq i \leq m$ are the **component functions** of f (real-valued).

Lemma 2.1 (Convergence of components)**.** $x_0 \in V$ open in $\mathbb{R}^n$. Let $f : V \setminus \{x_0\} \to \mathbb{R}^m$. Then $\lim_{x\to x_0} f(x) = L = (L_1, \dots, L_m)$ **iff** $\lim_{x\to x_0} f_i(x) = L_i \; \forall i = 1, 2, \dots, m$.

Theorem 2.1 (Squeeze theorem)**.** Suppose $f, g, h : V \setminus \{x_0\} \to \mathbb{R}$ $(m = 1!)$. If $f(x) \leq g(x) \leq h(x) \; \forall x \in V \setminus \{x_0\}$ (this only needs to hold in a neighbourhood of x_0) and $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = L \in \mathbb{R}$, then

$$\lim_{x\to x_0} g(x) = L$$

Proposition 2.3 (Norm of limits)**.** Suppose $f : V \setminus \{x_0\} \to \mathbb{R}^m$ and $\lim_{x \to x_0} f(x) = L$ then

$$\lim_{x \to x_0} \|f(x)\| = \|\lim_{x \to x_0} f(x)\| = \|L\|$$

Definition 2.3 (Continuity (at a point))**.** f is **continuous** at x_0 *iff* $\forall \epsilon > 0$, $\exists \delta > 0$ such that

$$\|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \epsilon$$

i.e. $\lim_{x \to x_0}$ exists and equals $f(x_0)$.

Definition 2.4 (Sequential characterization of continuity)**.** By the sequential characterization of limits, f is continuous at x_0 *iff* whenever (x_k) is a sequence in U converging to x_0, then $f(x_k)$ is a sequence in $\mathbb{R}^m$ converging to $f(x_0)$.

Definition 2.5 (Continuity (on a set))**.** f is **continuous on** U (an open set) if it is continuous at every $x \in U$.

Proposition 2.4 (Continuity of components)**.** If $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, f is continuous at $x_0 \in U$ *iff* $f_i : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is continuous at x_0 for all $i = 1, \ldots, n$.

Proposition 2.5 (Composition is continuous)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be continuous on U. Let $g : V \subseteq \mathbb{R}^m \to \mathbb{R}^p$ be continuous on V. Suppose $f(U) = \{f(x) \mid x \in U\} \subseteq V$ so the composition

$$h = g \circ f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$$

is defined $g(f(x))$. Then $h = g \circ f$ is continuous on U.

Proposition 2.6 (Dot product is continuous)**.** Suppose $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Define $f \cdot g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ by

$$(f \cdot g)(x) = f(x) \cdot g(x) = f_1(x)g_1(x) + f_2(x)g_2(x) + \ldots + f_m(x)g_m(x)$$

If f, g continuous at x_0, then $f \cdot g$ is continuous at x_0.

Definition 2.6 (Inverse image)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, U is open. Let $A \subseteq \mathbb{R}^m$.
The **inverse image** of A under f is denoted $f^{-1}(A)$ and is defined to be

$$f^{-1}(A) = \{x \in U \mid f(x) \in A\}$$

Proposition 2.7 (Continuous iff inverse image of open/closed is open/closed)**.** $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, U is open. Then f is continuous on U *iff* $f^{-1}(V)$ is **open** in $\mathbb{R}^n$ whenever V is **open** in $\mathbb{R}^m$.
Similarly, f is continuous *iff* $f^{-1}(V)$ is closed whenever V is closed.

Remark 2.1 (Continuity and open/closed domain)**.** From above, note it is **not true** that if U is open, then $f(U)$ is open for a continuous f on U. Consider $f(x) = x^2$ and $U = (-1, 1) \Rightarrow f(U) = [0, 1]$.
Similarly for closed.

Proposition 2.8 (Continuity carries compact into compact)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ continuous on U which is open. Let $K \subseteq U$ be **compact**. Then $f(K) = \{f(x) \mid x \in K\}$ is **compact** in $\mathbb{R}^m$.

Proposition 2.9 (Continuity carries connected into connected)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ continuous on U which is open. Let $E \subseteq U$ be **connected** on $\mathbb{R}^n$. Then $f(E)$ is **connected**.

Theorem 2.2 (Extreme value theorem)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, U is open $(m = 1!)$ and f is **continuous** on U.

Let $K \subseteq U$ be **compact**. Then $\exists x_1, x_2$ in K

$$f(x_1) \leq f(x) \leq f(x_2) \quad \forall x \in K$$

and x_1, x_2 need not be unique.

Theorem 2.3 (Intermediate value theorem)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, where U open ($m = 1!$).
Suppose f is **continuous** on U and let $E \subseteq U$ be connected. Let $x, y \in E$ such that $f(x) < f(y)$. Then for **each** $w \in (f(x), f(y))$, $\exists z \in E$ such that $f(z) = w$.

Definition 2.7 (Uniform continuity)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, U is open, and let $D \subseteq U$.
We say that f is **uniformly continuous on** D **iff** $\forall \epsilon > 0 \; \exists \delta > 0$ such that

$$\forall x, y \in D \quad \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon$$

Theorem 2.4 (Uniform continuity and compact sets)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be continuous on U open. Let $K \subseteq U$ be **compact**.
Then f is uniformly continuous on K.

3 Differentiation

Definition 3.1 (Single variable differentiability)**.** Let $f : U \subseteq \mathbb{R} \to \mathbb{R}$, U open, and $a \in U$. We say f is **differentiable** at a **iff**

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

exists. If so, we call the limit the **derivative** of f at a and we denote it

$$f'(a) = \frac{df(a)}{dx} = (Df)_a$$

Remark 3.1 (Single variable differentiability implies continuity)**.** If $f : \mathbb{R} \to \mathbb{R}$ is differentiable at a then f is continuous at a.

Definition 3.2 (Partial derivative)**.** Let $i \in \{1, \ldots, n\}$. The **partial derivative** of f *in the x_i-direction* at the point a is defined to be

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h}$$

if it exists.

Definition 3.3 (Directional derivative)**.** Consider the rate of change of f at a in the direction of *any* unit vector u (i.e. in between the standard vectors e_i).
This is called the **directional derivative** of f at a in the u-direction and is denoted

$$(D_u f)_a = \lim_{h \to 0} \frac{f(a + hu) - f(a)}{h}$$

(for $f : \mathbb{R} \to \mathbb{R}$).

Definition 3.4 (Class of continuous functions)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, U open. We say f is in $C^0(U)$ if f is continuous on U.
In general, for $k \in \mathbb{N}$, f is in $C^k(U)$ if f is in $C^{k-1}(U)$ and all $\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}$ exist and are continuous on U.

Theorem 3.1 (Mean value theorem)**.** Let $f : U \subseteq \mathbb{R} \to \mathbb{R}$ ($m = n = 1!$), U open, be continuous on $[a, b] \in U$ and differentiable on (a, b). There $\exists c \in (a, b)$ such that

$$f(b) - f(a) = f'(c)(b - a)$$

Theorem 3.2 (Commutativity of mixed partial derivatives)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, U open. Let $a \in U$. Suppose $\frac{\partial f}{\partial x_j}$, $\frac{\partial f}{\partial x_k}$ **exist and are continuous** ($j \neq k, j, k \in \{1, \ldots, n\}$) on a *neighbourhood of* a.
Furthermore, suppose that $\frac{\partial^2 f}{\partial x_j \partial x_k}$ exists in a *neighbourhood of* a and is continuous on a.
Then $\frac{\partial^2 f}{\partial x_k \partial x_j}$ exists at a and

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_k}(a)$$

Definition 3.5 (Differentiability)**.** For $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, U open, let $x_0 \in U$.
We say f is **differentiable** at x_0 if $\exists$ a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - T(h)\|}{\|h\|} = 0$$

Proposition 3.1 (Differentiability implies continuity)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, U open, and $a \in U$. Suppose f is differentiable at a. Then f is **continuous** at a.

Theorem 3.3 (Differentiable map is matrix of partial derivatives)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $a \in U$. Suppose f is differentiable at a.
We have

$$f(x) \in \mathbb{R}^m = (f_1(x), \ldots, f_m(x))$$

where $f_j : U \subseteq \mathbb{R}^n \to \mathbb{R}$ are the component functions of f, $1 \leq j \leq m$.
Then all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exists at a for $1 \leq i \leq m$, $1 \leq j \leq n$. Moreover,

$$T = (Df)_a = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ldots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix}$$

is the $m \times n$ matrix whose (i, j)-entry is $\frac{\partial f_i}{\partial x_j}(a)$. This shows $(Df)_a$ is unique if it exists.

Definition 3.6 (Gradient)**.** For $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ (note $m = 1!$), $a \in U$, and f differentiable at a, then $(Df)_a$ is a $1 \times n$ matrix also called the **gradient** denoted

$$(\nabla f)(a) = (Df)_a = \begin{bmatrix} \frac{\partial f}{\partial x_1}(a) & \ldots & \frac{\partial f}{\partial x_n}(a) \end{bmatrix}$$

Lemma 3.1 (Differentiability of components)**.** Let $f : U : \mathbb{R}^n \to \mathbb{R}^m$, $a \in U$. Then f is differentiable at a *iff* each component function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at a $\forall i = 1, \ldots, m$.

Proposition 3.2 (Linear combinations are differentiable)**.** Let $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Suppose f, g both differentiable at $a \in U$. Let $\lambda, \mu \in \mathbb{R}$. Then $\lambda f + \mu g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ or

$$(\lambda f + \mu g)(x) = \lambda f(x) + \mu g(x)$$

is differentiable at a and

$$(D(\lambda f + \mu g))_a = \lambda (Df)_a + \mu (Dg)_a$$

Theorem 3.4 (Partials exist and continuous implies differentiability)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $a \in U$. Suppose all $\frac{\partial f_i}{\partial x_j}$ **exists on a neighbourhood of** a and are **continuous at** a.
Then f is **differentiable** at a.
(The premises are sufficient but not necessary).

Remark 3.2 (Checking for differentiability)**.** To check if $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in U$

1. If f is **not continuous at** a, then f is **not differentiable** at a
2. If any of $\frac{\partial f_i}{\partial x_j}$ do not exist at a, f is **not differentiable** at a
3. Let $(Df)_a$ be the $m \times n$ matrix whose i, j entry is $\frac{\partial f_i}{\partial x_j}(a)$. Then f is differentiable at $a \iff$
$$\lim_{h\to 0} \frac{\|f(x_0 + h) - f(x_0) - T(h)\|}{\|h\|} = 0$$
4. We can avoid step 3 if we know all $\frac{\partial f_i}{\partial x_j}$ exist on a n'h'd of a and are continuous at a (this implies f is differentiable at a by theorem 3.4).

Proposition 3.3 (Product rule for differentiability)**.** Let $U \subseteq \mathbb{R}^n$, $f, g : U \to \mathbb{R}^m$, $a \in U$.
Suppose f, g are both differentiable at a. Then we claim $f \cdot g : U \to \mathbb{R}$, where $(f \cdot g)(x) = f(x) \cdot g(x)$ is differentiable at a and
$$D(f \cdot g)_a = f(a)^T (Dg)_a + g(a)^T (Df)_a \tag{3.1}$$

Theorem 3.5 (Chain rule)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $a \in U$. Let $g : V \subseteq \mathbb{R}^n \to \mathbb{R}^p$ be differentiable at $b = f(a) \in V$. Assume $f(U) \subseteq V$.
Then $g \circ f : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at a and
$$D(g \circ f)_a = (Dg)_{f(a)}(Df)_a$$

Proposition 3.4 (Linearization using derivative)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$. Then
$$f(x) - f(x_0) = (Df)_{x_0}(x - x_0) + R_{x_0}(h)$$
where $h = x - x_0$ for some remainder term $R_{x_0}(h)$.
We say f is **differentiable at** x_0 *iff* $\lim_{h\to 0} \frac{R_{x_0}(h)}{\|h\|} = 0$.

Definition 3.7 (Graph of function)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. The **the graph of** f is
$$\begin{aligned} \Gamma_f &= \{(x_1, \ldots, x_n, y) \in \mathbb{R}^{n+1} \mid y = f(x_1, \ldots, x_n)\} \\ &= \{(x_1, \ldots, x_n, f(x_1, \ldots, x_n) \mid (x_1, \ldots, x_n) \in U\} \end{aligned}$$

Theorem 3.6 (Rolle's theorem)**.** Let $f : \mathbb{R} \to \mathbb{R}$, f is continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists at least one $c \in (a, b)$ such that $f'(c) = 0$.

Theorem 3.7 (Single variable Taylor's theorem)**.** Let $I \subseteq \mathbb{R}$ be an *interval*, let p be a non-negative integer. Let $h : I \to \mathbb{R}$ be $(p + 1)$-times differentiable on I. Let $t_0 \neq t \in I$. Then $\exists\theta$ between t_0 and t (exclusively) such that
$$h(t) = \sum_{k=0}^{p} \frac{h^{(k)}(t_0)}{k!}(t - t_0)^k + \frac{h^{(p+1)}(\theta)}{(p + 1)!}(t - t_0)^{p+1}$$
where θ may not be unique.

Theorem 3.8 (Taylor's theorem)**.** We denote

$$(D^{(k)}f)_a(\xi) = \sum_{j_1,\dots,j_k=1}^{n} \frac{\partial^k f}{\partial x_{j_1}\dots\partial x_{j_k}}(a)\xi_1\dots\xi_k$$

for $k \geq 1$ and $(D^{(0)}f)_a = f(a)$.
Let $U \subseteq \mathbb{R}^n$ open, $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be in $C^{p+1}(U)$. Let $a \in U$, $\xi \in \mathbb{R}^n$ such that $\{a + t\xi \mid t \in [0,1]\} \subseteq U$.
Then $\exists\theta \in (0,1)$ such that

$$f(a+\xi) = \sum_{k=0}^{p} \frac{(D^{(k)}f)_a(\xi)}{k!} + \frac{1}{(p+1)!}(D^{(p+1)}f)_{a+\theta\xi}(\xi)$$

Proposition 3.5 (Lipschitz function)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. Suppose $f \in C^1(U)$. Let K be a **compact** subset of $\mathbb{R}^n$ with $K \subseteq U$. If $E \subseteq K$ is **convex**, $\exists$ a constant $M > 0$ (depending on f and on K but not on E) such that

$$\|f(x) - f(y)\| \leq M\|x - y\| \quad \forall x, y \in E$$

This says the *restriction* of f on E is **Lipschitz**: in particular any Lipshitz function on a set E is **uniformly continuous** on E (for any ϵ, choose $\delta = \frac{\epsilon}{M}$). Note however that uniform continuity *does not imply* Lipschitz.

Theorem 3.9 (More general Taylor's theorem)**.** $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ (U open, as always). Suppose $f \in C^p(U)$ (previously had $C^{p+1}(U)$). Let $a \in U$, $\xi \in \mathbb{R}^n$ such that $\{a + t\xi, t \in [0,1]\} \subseteq U$. Then

$$f(x) = \sum_{k=0}^{p} \frac{D^{(k)}f)_a(\xi)}{k!} + R_{a,p}(x)$$

where $x = a + \xi$ and where

$$\lim_{x\to a} \frac{R_{a,p}(x)}{\|x-a\|^p} = 0$$

3.1 Optimization

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ *real-valued* be differentiable on U.

Definition 3.8 (Local minimum)**.** Let $a \in U$. We say f has a **local minimum** at a if $\exists\epsilon > 0$ such that

$$f(x) \geq f(a) \quad \forall x \in B_\epsilon(a)$$

Definition 3.9 (Local maximum)**.** We say f has a **local maximum** at a if $\exists\epsilon > 0$ such that

$$f(x) \leq f(a) \quad \forall x \in B_\epsilon(a)$$

Definition 3.10 (Critical points)**.** A point $a \in U$ such that $(\nabla f)(\vec{a}) = 0$ is called a **critical point** of f.

Definition 3.11 (Saddle point)**.** A critical point $a \in U$ of f is called a **saddle point** if $\exists\epsilon > 0$ such that $\forall\epsilon' \in (0,\epsilon)$, $\exists x, yB_{\epsilon'}(a)$

$$f(x) < f(a) < f(y)$$

Definition 3.12 (Bilinear symmetric forms)**.** H is **bilinear** on $\mathbb{R}^n$ i.e. $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

$$H(av + bw, u) = aH(v,u) + bH(w,u)$$
$$H(v, aw + bu) = aH(v,w) + bH(v,u)$$

where $a, b \in \mathbb{R}$ and $u, v, w \in \mathbb{R}^n$.
We have for $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{j=1}^n y_j e_j$

$$H(x,y) = \sum_{i,j=1}^n H(e_i, e_j) x_i y_j$$

Denote $H_{ij} = H(e_i, e_j)$ where H is an $n \times n$ matrix. H is **symmetric** if $H(x,y) = H(y,x)$ for all $x, y \in \mathbb{R}^n$ i.e. *iff* $H_{ij} = H_{ji}$.

Definition 3.13 (Quadratic form)**.** We define the **quadratic form** Q associated to the symmetric bilinear form H to be the map $Q : \mathbb{R}^n \to \mathbb{R}$ given by

$$Q(x) = H(x,x) = \sum_{i,j=1}^n H_{ij} x_i x_j$$

Notice $Q(0) = 0$ **always**.

1. We say Q is **positive definite** if $Q(x) > 0 \; \forall x \neq \vec{0}$.
2. We say Q is **positive semi-definite** if $Q(x) \geq 0 \; \forall x \in \mathbb{R}^n$.
3. We say Q is **negative definite** if $Q(x) < 0 \; \forall x \neq \vec{0}$.
4. We say Q is **negative semi-definite** if $Q(x) \leq 0 \; \forall x \in \mathbb{R}^n$.
5. We say Q is **indefinite** if $\exists x, y \in \mathbb{R}^n$ such that $Q(x) > 0$, $Q(y) < 0$.

   For indefinite, *non-degenerate* means no $z \neq \vec{0} \Rightarrow Q(z) = 0$. *Degenerate* if there is such a z.

Lemma 3.2 (Bounds on quadratic forms)**.** Let Q be a quadratic form associated to symmetric bilinear form of H.

1. If Q is positive definite, $\exists M > 0$ such that $Q(x) \geq M\|x\|^2 \; \forall x \in \mathbb{R}^n$.
2. If Q is negative definite, $\exists M > 0$ such that $Q(x) \leq -M\|x\|^2 \; \forall x \in \mathbb{R}^n$.

Definition 3.14 (Hessian)**.** The **Hessian** of f at $a \in U$ is the $n \times n$ **symmetric** matrix $(\text{Hess } f)_a$ whose i, j entry is

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

Theorem 3.10 (Second derivative test)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, $f \in C^2(U)$. Let a be a critical point for f ($(\nabla f)(a) = \vec{0}$).
Let $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$ and H be the Hessian of f at a with quadratic form Q.

1. If Q is **positive definite**, then f has a **local min** at a.
2. If Q is **negative definite**, then f has a **local max** at a.
3. If Q is **indefinite**, then a is a **saddle point** of f.

(otherwise test fails and *any of the 3* can happen).

Example 3.1 (Second derivative test fails)**.** Consider

$$f(x,y) = x^4 + y^2 \quad g(x,y) = -x^4 - y^2 \quad h(x,y) = x^3 + y^2$$

which all have one critical point at $(0,0)$. Note their Hessians at $(0,0$ are

$$(\text{Hess } f)_{(0,0)} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \quad (\text{Hess } g)_{(0,0)} = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} \quad (\text{Hess } h)_{(0,0)} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

Note that $\exists x \neq \vec{0}$ where $x^T H x = 0$ (so not definite). Furthermore, they all map to either positive or negative values so they are not indefinite.

Definition 3.15 (Matrix norm)**.** Define the **norm** on $\mathbb{R}^{n \times n}$ by taking the usual *Euclidean norm* on $\mathbb{R}^{n^2}$

$$\|A\|^2 = \sum_{i,j=1}^{n} A_{ij}^2$$

Note that

$$\|Ax\| \leq \|A\| \|x\| \quad \forall x \in \mathbb{R}^n$$

Theorem 3.11 (Inverse function theorem)**.** Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be in $C^k(U)$ for some $k \geq 1$.
Let $V = f(U)$, let $a \in U$ such that $(Df)_a$ is *invertible* (note that $n = m$ since we require square matrices for invertibility).
Then $\exists$ open set $\tilde{U} \subseteq U$ containing a, an open set $\tilde{V} \subseteq V$ contain $f(a)$, and a map $g : \tilde{V} \to \tilde{U}$ (with $g(\tilde{V}) = \tilde{U}$) such that $g(f(x)) = x \ \forall x \in \tilde{U}$ and $f(g(y)) = y \ \forall y \in \tilde{V}$.
Moreover, $g \in C^k(\tilde{V})$ for the *same* k **and if** $b \in \tilde{V}$ then

$$(Dg)_b = [(Df)_{f^{-1}(b)}]^{-1}$$

Also

$$f\Big|_{\tilde{U}} : \tilde{U} \to \tilde{V}$$

is a bijection.

Example 3.2 (Applying inverse function theorem)**.** Let $(x,y) = f(u,v) = (uv, u^2 + v^2)$ where $f : \mathbb{R}^2 \to \mathbb{R}^2$. Note that $f \in C^\infty(\mathbb{R}^2)$ since f_i are polynomials.
We want to prove f^{-1} exists and is C^∞ in some nonempty open set containing $(2,5)$.
For $f(a,b) = (2,5)$, find all points $(u,v) \in \mathbb{R}^2$ such that $f(u,v) = (2,5)$.

$$\begin{aligned} uv = 2 &\Rightarrow v = \frac{2}{u} \\ u^2 + v^2 = 5 &\Rightarrow u^2 + \frac{4}{u^2} = 5 \\ &\Rightarrow u^4 - 5u^2 + 4 = 0 \\ &\Rightarrow (u^2 - 1)(u^2 - 4) = 0 \end{aligned}$$

So $(u,v) = \{(1,2), (-1,-2), (2,1), (-2,-1)\}$.
Note that

$$(Df)_{(u,v)} = \begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix} = \begin{bmatrix} v & u \\ 2u & 2v \end{bmatrix}$$

Thus $det((Df)_{(u,v)} = 2v^2 - 2u^2 = 2(v^2 - u^2) \neq 0$ for any of our points.
So by the inverse function theorem, for any of these 4 points (a, b) there is an open n'h'd $\tilde{U}$ of (a, b) and an open n'h'd of $\tilde{V}$ of $(2, 5)$ such that $f : \tilde{U} \to \tilde{V}$ is invertible and $f^{-1} \in C^\infty(\tilde{V})$.

Theorem 3.12 (Implicit function theorem)**.** Let $f : W \subseteq \mathbb{R}^{n+m} \to \mathbb{R}^n$ be in $C^k(W)$ for $k \geq 1$. Suppose $f(y_0, x_0) = 0$ for some $(y_0, x_0) \in W$.
Let A be the $n \times n$ matrix where $A_{ij} = \frac{\partial f_i}{\partial y_j}(y_0, x_0)$.
If $det(A) \neq 0$ (i.e. A invertible) then $\exists W' \subseteq W$ open n'h'd of (y_0, x_0) and an open n'h'd U of x_0 in $\mathbb{R}^m$ and a function $h : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$, $h \in C^k(U)$ for the same k such that

$$\{(y, x) \in W' \mid f(y, x) = 0\} = \{(h(x), x), x \in U\}$$

i.e. on W', the points where $f = 0$ can be expressed as y as a function of x.

Example 3.3 (Applying implicit function theorem)**.** Given $x_0, y_0, u_0, v_0, s_0, t_0$ **nonzero** real numbers that satisfy the simultaneous equations

$$u^2 + sx + ty = 0 \quad v^2 + tx + sy = 0 \quad 2s^2x + 2t^2y - 1 = 0 \quad s^2x - t^2y = 0$$

(this is almost impossible to solve explicitly: we may only want to know it exists).
Show that $\exists$ smooth (C^∞) functions $u(x, y), v(x, y), s(x, y), t(x, y)$ defined on an open n'h'd of (x_0, y_0) such that u, v, s, t satisfy the equations and

$$u(x_0, y_0) = u_0 \quad v(x_0, y_0) = v_0 \quad s(x_0, y_0) = s_0 \quad t(x_0, y_0) = t_0$$

We'll apply the implicit function theorem. Define $f : \mathbb{R}^6 = \mathbb{R}^{4+2} \to \mathbb{R}^4$ where

$$f(u, v, s, t, x, y) = \begin{bmatrix} u^2 + sx + ty \\ v^2 + tx^sy \\ 2s^2x + 2t^2y - 1 \\ s^2x - t^2y \end{bmatrix} \in \mathbb{R}^4$$

By hypothesis, $f(u_0, v_0, s_0, t_0, x_0, y_0) = 0$. Also

$$Df = \begin{bmatrix} 2u & 0 & x & y & \dots \\ 0 & 2v & y & x & \dots \\ 0 & 0 & 4sx & 4ty & \dots \\ 0 & 0 & 2sx & -2ty & \dots \end{bmatrix}$$

So we have

$$A = \begin{bmatrix} 2u_0 & 0 & x_0 & y_0 \\ 0 & 2v_0 & y_0 & x_0 \\ 0 & 0 & 4s_0x_0 & 4t_0y_0 \\ 0 & 0 & 2s_0x_0 & -2t_0y_0 \end{bmatrix}$$

where $det(A) = (2u_0)(2v_0)(-8s_0x_0t_0y_0 - 8s_0x_0t_0y_0) = 64u_0v_0s_0t_0x_0y_0 \neq 0$ since they're all non-zero.
So u, v, s, t exist by be the implicit function theorem in a n'h'd of (x_0, y_0) and are in C^∞ (since f is in C^∞, polynomials).

Theorem 3.13 (Methods of Lagrange multipliers)**.** Let $1 \leq k \leq n$. Let $W \subseteq \mathbb{R}^n$ (open). Let $f : W \subseteq \mathbb{R}^n \to \mathbb{R}$ and $g : W \subseteq \mathbb{R}^n \to \mathbb{R}^k$ (component functions $g_1, \dots, g_k$ are the constraint functions).
Let $S = \{w \in W \mid g(x) = 0\}$ (the "constraint" set). Let $a \in S$.

Suppose

1. f has a local extrema at a subject to the constraints $g(x) = 0$ (i.e. f restricted to S has a local extrema at a).
2. $rank((Dg)_a) = k$ (where $(Dg)_a$ is $k \times n$ thus maximal rank).

Then $\exists \lambda \in \mathbb{R}^k$ such that

$$(Df)_a + \lambda(Dg)_a = \vec{0}$$

Example 3.4 (Applying Lagrange multipliers)**.** Find all extrema of $f(x, y, z) = x^2 + y^2 + z^2$ subject to the 2 constraints: $x - y = 1$ and $y^2 - z^2 = 1$.
There exists points on constraint set with arbitrary large distance from origin (no global max).
We know there **will exist** a global min (which will also be a local min). We expect 2 local minima since $y^2 - z^2 = 1$ cuts twice into the other constraint plane.
We have

$$g_1(x, y, z) = x - y - 1 = 0 \quad g_2(x, y, z) = y^2 - z^2 - 1 = 0$$

and from Lagrange multipliers we know $\nabla f + \lambda \nabla g_1 + \mu \nabla g_2 = 0$, thus

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

From the last constraint, either $\mu = 1$ or $z = 0$:

$\mu = 1$ Then the second equation becomes $4y = \lambda$ and the first equation becomes $2x + 4y = 0$ so $x = -2y$.

From our original constraint equations, we have from g_1 $-3y = 1 \Rightarrow y = \frac{-1}{3}$ and from g_2 $\frac{1}{9} - z^2 = 1 \Rightarrow z^2 = \frac{-8}{9}$ which is a **contradiction** since squares are always positive.

$z = 0$ From g_2 we have $y = \pm 1$ and from g_1 we have $x = y + 1$.

Thus we have two solutions $(2, 1, 0)$ and $(0, -1, 0)$ (which satisfy all the other equations too).

Thus we have $f(2, 1, 0) = 5$ (some local min) and $f(0, -1, 0 = 1$ (global min).

4 Integration

We almost always consider only sets D that are **bounded**. We sometimes require $f : D \to \mathbb{R}^m$ to be bounded: this requires further effort.

Definition 4.1 (Box)**.** Let $I = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subseteq \mathbb{R}^n$. I is the Cartesian product of closed bounded intervals, i.e. $x \in I \iff a_i \leq x \leq b_i$ for all $i = 1, \ldots, n$. We'll call I a **box** in $\mathbb{R}^n$.
It is clear I is **compact** because it's closed and bounded.

Definition 4.2 (Size of box)**.** Define the **size of a box** I $\mu(I) \in \mathbb{R}$ to be

$$\mu(I) = (b_1 - a_1)(b_2 - a_2) \ldots (b_n - a_n) = \prod_{k=1}^{n} (b_k - a_k)$$

Definition 4.3 (Zero size)**.** Let $E \subseteq \mathbb{R}^n$. We say E has **zero size** (and write $\mu(E) = 0$) *iff* $\forall \epsilon > 0$, $\exists$ boxes $I_1, \ldots, I_N$ with $E \subseteq \bigcup_{k=1}^N I_k$ and $\sum_{k=1}^N \mu(I_k) < \epsilon$ (i.e. we can cover E by finitely many boxes whose sizes sum to as small as we want). Note $I_1, \ldots, I_k$ need not be disjoint.

Proposition 4.1 (Continuous graphs of compact sets have zero size)**.** Let $K \subseteq \mathbb{R}^n$ be **compact**. Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be continuous on U with $K \subseteq U$. Define

$$\Gamma_{f,K} = \{(x, f(x)) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n\}$$

or the "graph of f over the set K". Then $\Gamma_{f,k}$ has size zero.

Corollary 4.1 (Boundary of box have zero size)**.** Let I be a box in $\mathbb{R}^n$ then ∂I has **zero size**.

Definition 4.4 (Non-zero size)**.** $E \subseteq \mathbb{R}^n$ **does not have zero size** *iff* $\exists \epsilon_0 > 0$ such that $\forall$ finite collections of boxes $I_1, \ldots, I_n$ with $E \subseteq \bigcup_{j=1}^N I_j$, we have $\sum_{j=1}^n \mu(I_j) \geq \epsilon_0$.
Non-zero size *does not imply* it's sizeable (consider unbounded set).

Lemma 4.1 (Technical lemma for non-zero size)**.** A subset $E \subseteq \mathbb{R}^n$ **does not have zero size** *iff* $\exists \tilde{\epsilon_0} > 0$ such that $\forall$ finite collections of boxes $I_1, \ldots, I_n$ with $E \subseteq \bigcup_{j=1}^N I_j$ where $int(I_j) \cap E \neq \varnothing$, we have $\sum_{j=1}^n \mu(I_j) \geq \tilde{\epsilon_0}$.

Definition 4.5 (Partitions of box)**.** Let $I = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$ be a box in $\mathbb{R}^n$.
For $j \in \{1, \ldots, n\}$, choose $t_{j,0}, t_{j,1}, \ldots, t_{j,N_j}$ $N_j \geq 1$ such that $a_j = t_{j,0} < t_{j,1} < t_{j,2} < \ldots < t_{j,N_j-1} < t_{j,N_j} = b_j$.
Let $P_j = \{t_{j,l} \mid l = 0, 1, \ldots, N_j\}$ and $P = P_1 \times P_2 \times \ldots \times P_n$.
Then $x \in P \iff x_j \in P_j$ for all $j = 1, \ldots, n$.
Such a P is a called a **partition** of the box.
P has $N_1 \times N_2 \times \ldots \times N_n$ elements.

Definition 4.6 (Subdivision of box)**.** A partition P of I determines a **subdivision** of I into $N_1 \times \ldots \times N_n$ boxes of the form

$$I_{k_1,\ldots,k_n} = [t_{1,k_1}, t_{1,k_1+1}] \times \ldots \times [t_{n,k_n}, t_{n,k_n+1}]$$

where $k_j \in \{0, 1, \ldots, N_j - 1\}$.

Definition 4.7 (Riemann sum)**.** Let $f : I \to \mathbb{R}^m$, $I \subseteq \mathbb{R}^n$ be a box and P be a partition of I. Let I_α where $\alpha \in P$ be the corresponding subdivision of I. Then $I = \bigcup_{\alpha \in P} I_\alpha$.
For each I_α choose $x_\alpha \in I_\alpha$. Then the **Riemann sum** for f with respect to partition P is

$$S(f, P) = \sum_{\alpha \in P} f(x_\alpha)\mu(I_\alpha)$$

Definition 4.8 (Refinement of partition)**.** Let P and Q be two partitions of the same box I. We say that Q is a **refinement** of P if $P_j \subseteq Q_j$ for all $j = 1, \ldots, n$.
That is, all the subboxes J_B of I corresponding to Q are themselves subboxes of one of the subboxes I_α of I corresponding to P.
Note that

$$I = \bigcup_\alpha I_\alpha = \bigcup_\beta J_\beta$$

each I_α is a union of some J_β's.

Remark 4.1 (Common refinement)**.** Let P and Q be any two partitions of I. There always exists a partition R of I where $R_j = P_j \cup Q_j$ for all j that is a **common refinement**.

Definition 4.9 (Riemann integral of box)**.** Let $I \subseteq \mathbb{R}^n$ be a box. Let $f : I \to \mathbb{R}^m$. f is **Riemann integrable** on I **iff** $\exists y \in \mathbb{R}^m$ such that $forall \epsilon > 0$, there exists a partition P_ϵ of I such that, for **any** refinement P of P_ϵ and **any** Riemann sum $S(f, P)$ corresponding to P, we have

$$\|S(f, P) - y\| < \epsilon$$

Moreover, the **Riemann integral** of f over I is

$$y = \int_I f$$

Theorem 4.1 (Cauchy criterion for Riemann integrable)**.** Let $I \subseteq \mathbb{R}^n$ be a box, $f : I \to \mathbb{R}^m$. f is Riemann integrable on I *iff* $\forall \epsilon > 0$, $\exists$ partition P_ϵ of I such that for any refinements P, Q of P_ϵ and any Riemann sums $S(f, P), S(f, Q)$ we have

$$\|S(f, P) - S(f, Q)\| < \epsilon$$

Lemma 4.2 (Simplified Cauchy criterion)**.** $f : I \to \mathbb{R}^m$ is integrable on I **iff** $\forall \epsilon > 0$, $\exists$ partition P_ϵ of I such that if $S_1(f, P_\epsilon)$, $S_2(f, P_\epsilon)$ are **any two Riemann sums** for f with respect to P_ϵ, then

$$\|S_1(f, P_\epsilon) - S_2(f, P_\epsilon)\| < \epsilon$$

(note: we no longer require that this holds for all Riemann sums of all refinements).

Theorem 4.2 (Integrable on box when bounded size zero discontinuity)**.** Let $I \subseteq \mathbb{R}^n$ be a box. Let $f : I \to \mathbb{R}^m$ be **bounded**. Let $S \subseteq I$ be the points in I where f **fails to be continuous**. If S has **size zero**, then f is Riemann integrable on I.

Definition 4.10 (Riemann integral)**.** Let $D \subseteq \mathbb{R}^n$ be **bounded**. Let $f : D \to \mathbb{R}^m$. Choose any box I in $\mathbb{R}^n$ such that $D \subseteq I$ (possible since D is bounded).
Define $\tilde{f} : I \to \mathbb{R}^m$ by $\tilde{f}(x) = f(x)$ if $x \in D$ and $\tilde{f}(x) = 0$ if $x \notin D$ (i.e. $\tilde{f}$ is the extension by zero of f from D to I).
We say f is **Riemann integrable** on D *iff* $\tilde{f}$ is Riemann integrable on I and we write

$$\int_D f = \int_I \tilde{f}$$

Theorem 4.3 (Integrable on bounded general set with size zero boundary)**.** Let $D \neq \varnothing$ and bounded subset of $\mathbb{R}^n$ with $\mu(\partial D) = 0$. Suppose $f : D \to \mathbb{R}^m$ is **bounded** and **continuous**, then f is **integrable** on D.

Definition 4.11 (Indicator function)**.** Let $D \subseteq \mathbb{R}^n$ be **any subset**. We define the **indicator function** $X_D : \mathbb{R}^n \to \mathbb{R}$ of D to be

$$X_D(x) \begin{cases} 1 & \text{, if } x \in D \\ 0 & \text{, if } x \notin D \end{cases}$$

Definition 4.12 (Size of general sets)**.** Let $D \subseteq \mathbb{R}^n$ be **bounded**. We say that D is **sizeable** *iff* X_D is integrable on D.
If D is sizeable, then the **size** of D is

$$\mu(D) = \int_D 1 = \int_D X_D = \int_I X_D$$

for any box $I \supseteq D$.

Theorem 4.4 (Characterization of sizeability)**.** Let $D \neq \varnothing$ and bounded. Then D is **sizeable** *iff* ∂D **has size zero**.

Proposition 4.2 (Properties of Riemann integrals)**.** Let $D \subseteq \mathbb{R}^n$ be bounded.

1. Let $f, g : D \to \mathbb{R}^m$ be integrable on D. Let $\lambda, \mu \in \mathbb{R}$. Then $\lambda f + \mu g$ is integrable on D and
$$\int_D (\lambda f + \mu g) = \lambda \int_D f + \mu \int_D g$$
2. Let $f : D \to \mathbb{R}$ ($m = 1!$) be integrable on D and **non-negative**. Then $\int_D f \geq 0$.
3. Let D be sizeable and $f : D \to \mathbb{R}$ ($m = 1!$) be integrable on D and $\exists M_1, M_2 \in \mathbb{R}$ such that $M_1 \leq f(x) \leq M_2$ $\forall x \in D$. Then $M_1\mu(D) \leq \int_D f \leq M_2\mu(D)$.

Theorem 4.5 (Mean value theorem for integration)**.** Let $D \subseteq \mathbb{R}^n$ be **compact, connected, and sizeable**. Let $f : D \to \mathbb{R}$ be **continuous** on D. Then $\exists$ at least one point $x_0 \in D$ such that
$$\int_D f = f(x_0)\mu(D) \tag{4.1}$$

Theorem 4.6 (Fubini's theorem)**.** Let I be a box in $\mathbb{R}^n$. Let J be a box in $\mathbb{R}^m$. Then $I \times J$ is a box in $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$.
Let $f : I \times J \to \mathbb{R}^p$. Suppose that f is integrable on $I \times J$ **and** for each $x \in I$ the function where $y \mapsto f(x, y)$ i.e. $f(x, \cdot) : J \to \mathbb{R}^p$ is integrable on J, that is
$$\int_J f(x, \cdot) = F(x) \in \mathbb{R}^m$$
exists for all $x \in I$ where $F : I \to \mathbb{R}^m$.
Then F is integrable on I and
$$\int_I F = \int_{I \times J} f$$
or $\int_I \left(\int_J f(x, \cdot)\right) = \int_{I \times J} f$.

Corollary 4.2 (1-D Fubini)**.** Let $f(x, y) : \mathbb{R}^{1+1} \to \mathbb{R}$. If $\int_a^b f(x, y)\,\mathrm{d}x$ exists $\forall y \in [c, d]$ and $\int_{[a,b]\times[c,d]} f(x, y)$ exists, then $\int_{[a,b]\times[c,d]} f(x, y) = \int_c^d \left(\int_a^b f(x, y)\,\mathrm{d}x\right) \mathrm{d}y$.

Corollary 4.3 (1-D Fubini with bounding functions)**.** Let $\phi, \psi : [a, b] \to \mathbb{R}$ **continuous** with $\phi(x) \leq \psi(x)$ for all $x \in [a, b]$.
Let $D = \{(x, y) \in \mathbb{R}^2, x \in [a, b], \phi(x) \leq y \leq \psi(x)\}$.
Let $c \leq d$ such that $c \leq \phi(x) \leq \psi(x) \leq d$ for all $x \in [a, b]$.
Let $f : D \to \mathbb{R}$ be bounded such that the set $D_0 \in \mathbb{R}^2$ of its points of discontinuity has size zero, **and**, $\{y \in [c, d], (x, y) \in D_0\}$ also has size zero as a subset of $\mathbb{R}^1$, for each $x \in [a, b]$.
Then f is integrable on D and
$$\int_D f = \int_a^b \left(\int_{\phi(x)}^{\psi(x)} f(x, y)\,\mathrm{d}y\right) \mathrm{d}x$$

Example 4.1 (Finding volume using Fubini's)**.** Find the volume of the region D lying inside the "elliptic" cylinder $x^2 + 4y^2 = 4$ above the x-y plane and below the plane $z = 2 + x$.
We want to find $Vol(D) = \int_D 1$.
Note that the base is an ellipse since $x^2 + 4y^2 = 4 \Rightarrow \left(\frac{x}{2}\right)^2 + y^2 + 1$. We then extend this ellipse along the z axis. Note that the plane is a function of only x. It is also not parallel to the x-y plane so our cylinder has a slanted top.

From our ellipse we see that

$$
\begin{aligned}
-2 \leq x \leq 2 \\
-\sqrt{1-\frac{x^2}{4}} \leq y \leq \sqrt{1-\frac{x^2}{4}}
\end{aligned}
$$

And of course from our plane and the x-y plane we have

$$0 \leq z \leq 2+x$$

Thus we have

$$
\begin{aligned}
Vol(D) &= \int_{-2}^{2} \int_{-\sqrt{1-\frac{x^2}{4}}}^{\sqrt{1-\frac{x^2}{4}}} \int_{0}^{2+x} 1 \quad \mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{-2}^{2} \int_{-\sqrt{1-\frac{x^2}{4}}}^{\sqrt{1-\frac{x^2}{4}}} (2+x)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{-2}^{2} 2(2+x)\sqrt{1-\frac{x^2}{4}}\,\mathrm{d}x \\
&= 4\int_{0}^{2} 2\sqrt{1-\frac{x^2}{4}}\,\mathrm{d}x \\
&= 8\int_{0}^{2} \sqrt{1-\frac{x^2}{4}}\,\mathrm{d}x \\
&= 16\int_{0}^{1} \sqrt{1-u^2}du && x=2u, \mathrm{d}x = 2du \quad 0 \leq x \leq 2, 0 \leq u \leq 1 \\
&= 4\pi
\end{aligned}
$$

Example 4.2 (Re-ordering order of integration using Fubini's)**.** Suppose we wanted to find the integral

$$
\begin{aligned}
\int_0^1 \int_z^1 \int_0^x e^{x^3}\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}z &= \int_0^1 \int_z^1 e^{x^3} y\Big|_{y=0}^{y=x} \mathrm{d}x\,\mathrm{d}z \\
&= \int_0^1 \int_z^1 xe^{x^3}\,\mathrm{d}x\,\mathrm{d}z
\end{aligned}
$$

Fubini's theorem says we can change the order of integration. Note that we had $0 \leq z \leq 1$ and $z \leq x \leq 1$ from our integral bounds. Let us express x first then find the bounds of z in terms of x, i.e. $0 \leq x \leq 1$ and $0 \leq z \leq x$. Thus we have

$$
\begin{aligned}
\int_0^1 \int_0^x xe^{x^3}\,\mathrm{d}z\,\mathrm{d}x &= \int_0^1 \left(xe^{x^3} z\Big|_{z=0}^{z=x} \right) \mathrm{d}x \\
&= \int_0^1 x^2 e^{x^3}\,\mathrm{d}x \\
&= \frac{1}{3}e^{x^3}\Big|_0^1 \\
&= \frac{1}{3}(e-1)
\end{aligned}
$$

Theorem 4.7 (Change of variables theorem)**.** Let $U \subseteq \mathbb{R}^n$ be open and non-empty. Let $K \subseteq U$ be **compact**, non-empty and **sizeable**. Suppose $\psi : U \to \mathbb{R}^n$ is in $C^1(U)$. Suppose $\exists$ a subset $D \subseteq K$ with $\mu(D) = 0$ such that

1. $\psi\Big|_{K \setminus D}$ is injective
2. $det((D\psi)_x) \neq 0$ for all $x \in K \setminus D$.

Then $\psi(K)$ is **sizeable** and for any $f : \psi(K) \to \mathbb{R}^p$ which is *continuous*, then f is integrable on $\psi(K)$ and

$$\int_{\psi(K)} f = \int_K (f \circ \psi)|det(D\psi)|$$

where $f \circ \psi : K \to \mathbb{R}^p$ and $|det(D\psi)|$ is the "scaling factor".

Example 4.3 (Cylindrical coordinates)**.** We have three axes r, θ, z for cylindrical coordinates where $0 \leq r < \infty$ is the distance from the centre on the x-y plane ($r^2 = x^2 + y^2$), $0 \leq \theta \leq 2\pi$ is the angle from the positive x-axis, and $-\infty < z < \infty$ is the distance along the z-axis.

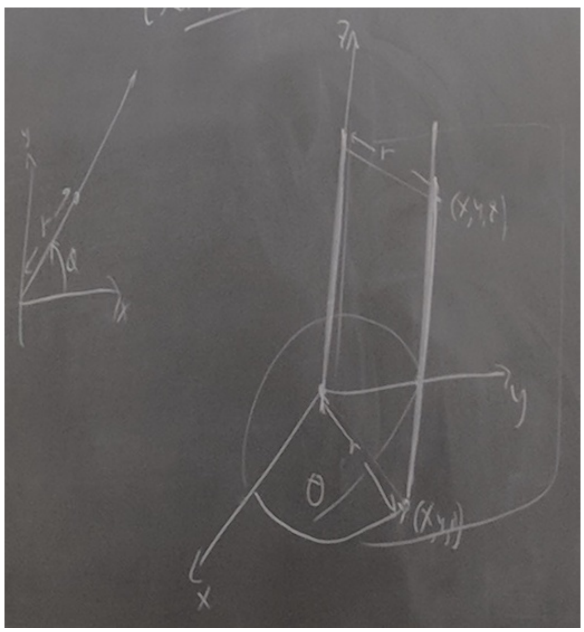

Figure 4.1: Cylindrical coordinates in the xyz $\mathbb{R}^3$ space.

We thus have $(x, y, z) = \psi(r, \theta, z) = (r\cos\theta, r\sin\theta, z)$ and

$$|det(D\psi)| = r$$

Find the volume of the region **above** the paraboloid $z = x^2 + y^2 = r^2$. and inside the sphere $x^2 + y^2 + z^2 = 12$. These two surfaces intersect when $z = x^2 + y^2 \geq 0$ and $x^2 + y^2 + z^2 = 12$. So

$$\begin{aligned} &z^2 + z - 12 = 0 \\ &(z - 3)(z + 4) = 0 \\ &z = 3 && z \geq 0 \end{aligned}$$

Intersect is the circle of radius $\sqrt{3}$ in the plane $z = 3$ (we only want to integrate r up to this maximum radius).
We want the volume above the paraboloid so when the z-axis is larger, thus we have the min bound for $z \geq x^2+y^2 = r^2$.
We want the volume inside the sphere so we want $x^2 + y^2 + z^2 \leq 12$ or $r^2 + z^2 \leq 12 \Rightarrow z \leq \sqrt{12 - r^2}$ (note the other inequality $z \geq -\sqrt{12 - r^2}$ is already accounted for by our min bound above).
Thus the region D has bounds

$$\begin{aligned}
&0 \leq \theta \leq 2\pi \\
&0 \leq r \leq \sqrt{3} \\
&r^2 \leq z \leq \sqrt{12 - r^2}
\end{aligned}$$

So the volume of D is

$$\begin{aligned}
Vol(D) &= \int_0^{2\pi} \int_0^{\sqrt{3}} \int_{r^2}^{\sqrt{12-r^2}} 1r \, \mathrm{d}z \, \mathrm{d}r \, \mathrm{d}\theta \\
&= \int_0^{2\pi} \int_0^{\sqrt{3}} r(\sqrt{12 - r^2} - r^2) \, \mathrm{d}r \, \mathrm{d}\theta \\
&= 2\pi \int_0^{\sqrt{3}} r(\sqrt{12 - r^2}) - r^3 \, \mathrm{d}r \\
&= 2\pi \left(\frac{-1}{3}(12 - r^2)^{\frac{3}{2}} - \frac{r^4}{4} \right) \Big|_0^{\sqrt{3}} \\
&= 2\pi \left(\frac{(12)^{\frac{3}{2}}}{3} - 9 - \frac{9}{4} \right)
\end{aligned}$$

Example 4.4 (Spherical coordinates)**.** We have three axes ρ, θ, ϕ in spherical coordinates where $0 \leq \rho < \infty$ is the distance from the origin, $0 \leq \theta \leq 2\pi$ is the distance from the postiive x-axis, and $0 \leq \phi \leq \pi$ is the angle from the positive z-axis.

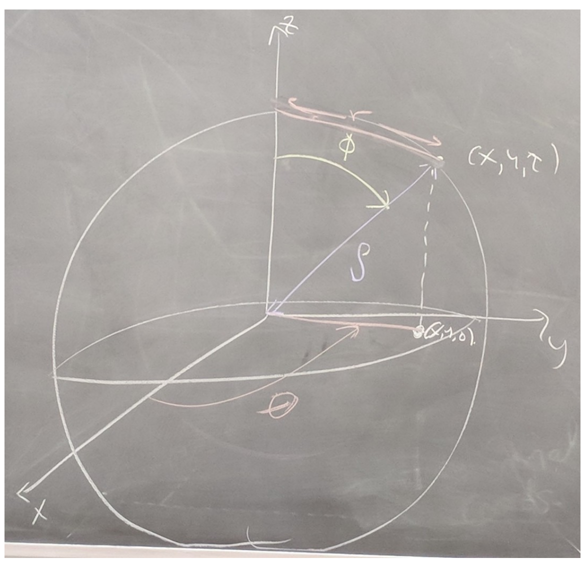

Figure 4.2: Spherical coordinates with ρ, θ, ϕ against the xyz axes.

If we look at z against ρ and r from spherical coordinates

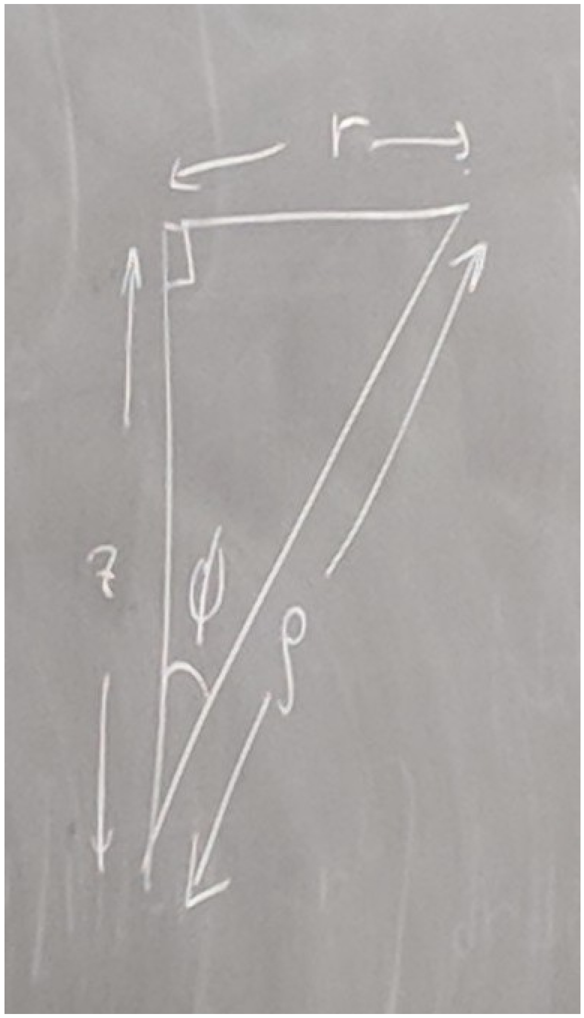

Figure 4.3: The z-axis with respect to ρ and r from cylindrical coordinates.

So $(x, y, z) = \psi(\rho \sin\phi \cos\theta, \rho \sin\phi \sin\theta, \rho \cos\phi)$ and

$$|det(D\psi)| = \rho^2 \sin\phi$$

Find $\int_D g$ where $g(x, y, z) = 1 - \sqrt{x^2 + y^2 + z^2} = 1 - \rho$ and D is the region above the cone $z = \frac{1}{\sqrt{3}}\sqrt{x^2 + y^2}$ and inside the sphere $x^2 + y^2 + z^2 = 1$.
Note that $z = \frac{1}{\sqrt{3}}r$ so $\cos\phi_0 = \frac{1}{2}$ thus $\phi_0 = \frac{\pi}{3}$ (the angle from the z-axis of the cone).
Thus we have for the bounds of D

$$0 \le \theta \le 2\pi$$
$$0 \le \phi \le \frac{\pi}{3}$$
$$0 \le \rho \le 1$$

Thus we have

$$\begin{aligned}
\int_D g &= \int_0^{2\pi}\int_0^{\frac{\pi}{3}}\int_0^1 (1-\rho)\rho^2 \sin\phi \,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta \\
&= 2\pi\int_0^{\frac{\pi}{3}}\int_0^1 (\rho^2-\rho^3)\sin\phi\,\mathrm{d}\rho\,\mathrm{d}\phi \\
&= 2\left(\frac{1}{3}-\frac{1}{4}\right)\int_0^{\frac{\pi}{3}}\sin\phi\,\mathrm{d}\phi \\
&= \frac{\pi}{6}(-\cos\phi)\Big|_0^{\frac{\pi}{3}} \\
&= \frac{\pi}{6}\left(1-\frac{1}{2}\right) \\
&= \frac{\pi}{12}
\end{aligned}$$