21-269
Vector Analysis

Rohan Jain

Contents

Chapter 1

1.1 The Real Numbers

Definition 1.1.1: Partial Order

Let X be a set with a binary relation $\leqslant$. $\leqslant$ is a *partial order* if:

1. $x \leqslant x$ for all $x \in X$ (reflexivity)
2. $x \leqslant y$ and $y \leqslant z$ implies $x \leqslant z$ for all $x, y, z \in X$ (transitivity)
3. $x \leqslant y$ and $y \leqslant x$ implies $x = y$ for all $x, y \in X$ (antisymmetry)

Definition 1.1.2: Partially Ordered Set (poset)

A set X with a partial order $\leqslant$ is called a *partially ordered set* or *poset*. It is notated as $(X, \leqslant)$.

Definition 1.1.3: Total Order

A partial order $\leqslant$ is a *total order* if for all $x, y \in X$, we have $x \leqslant y$ or $y \leqslant x$.

Example 1.1.1 (poset)
Let Y be a set. Define $X = \{\text{all subsets of } Y\} = \mathcal{P}(Y)$. Let $E, F \in Y$, we say that $E \leqslant F$ if $E \subseteq F$. Then $(X, \leqslant)$ is a poset. This is not a total order.

Definition 1.1.4: Upper Bound, Bounded Above, Supremum, Maximum

Let $(X, \leqslant)$ be a poset. Let $E \subseteq X$.

1. $y \in X$ is an *upper bound* of E if $x \leqslant y$ for all $x \in E$.
2. E is *bounded above* if it has at least one upper bound.
3. If E is nonempty and bounded above, then the *supremum*, if it exists, of E, denoted $\sup E$, is the least upper bound of E.
4. E has a *maximum* if there is $y \in E$ such that $x \leqslant y$ for all $x \in E$.

Properties worth mentioning:

1. If E has a maximum, then $\sup E$ exists and is equal to the maximum.

   Proof. Let y be the maximum of E. If $z \in X$, is an upper bound of E, then $z \geqslant y$ because $y \in E$. Since z was arbitrary, this is true for any upper bound. Thus, y is the least upper bound of E. 😄

Example 1.1.2

Let Y be a nonempty set, $(\mathcal{P}(Y), \leqslant)$ poset.
Fix nonempty $Z \subseteq Y$.

$$E = \{W \subseteq Y : W \subset Z\}$$

Trivially, Z is an upper bound of E. Realize that any superset of Z is an upper bound as well. We can postulate that the supremum of E is Z. We will now prove it:

Proof. Need to show that if F is an upper bound of E, then $F \supseteq Z$. If $x \in Z$, then $\{x\} \in E$ by definition of E, so $F \supseteq x$ for all $x \in Z$. Thus, $F \supseteq Z$. 😊

Note that there is no maximum of E.

Definition 1.1.5: Lower Bound, Bounded Below, Infimum, Minimum

Let $(X, \leqslant)$ be a poset. Let $E \subseteq X$.

1. $y \in X$ is a *lower bound* of E if $y \leqslant x$ for all $x \in E$.
2. E is *bounded below* if it has at least one lower bound.
3. If E is nonempty and bounded below, then the *infimum*, if it exists, of E, denoted $\inf E$, is the greatest lower bound of E.
4. E has a *minimum* if there is $y \in E$ such that $y \leqslant x$ for all $x \in E$.

Going back to example 1.1.2, we can see that E is bounded below by $\emptyset$. The infimum of E is $\emptyset$. The minimum of E is also $\emptyset$.

Definition 1.1.6: Complete

Let $(X, \leqslant)$ poset. X is *complete* if every nonempty subset of X that is bounded above has a supremum.

Example 1.1.3 ($\mathbb{Q}$)

$(\mathbb{Q}, \leqslant)$ is not complete.

Claim 1.1.1 $\mathbb{R}$

There is a complete ordered field $(\mathbb{R}, +, \cdot, \leqslant)$. Its elements are called real numbers.

1.2 First Recitation, 1/18

Exercise 1.2.1 Function Example

Let X be the set of all functions $f : D_f \to Z$ with $D_f \subseteq Y$. We say that $f \leqslant g$ if $D_f \subseteq D_g$ and $f(x) = g(x)$ for all $x \in D_f$. Is $(X, \leqslant)$ a poset? Is it complete?

Proof. To show that $(X, \leqslant)$ is complete, we need to show that every nonempty subset of X that is bounded above has a supremum. Let $E \subseteq X$ be nonempty and bounded above. Let $G = \bigcup_{f \in E} D_f$. G is the union of all the domains of the functions in E. G is bounded above by the union of the upper bounds of the domains of the functions in E. Let $H = \bigcup_{f \in E} f(D_f)$. H is bounded above by the union of the upper bounds of the ranges of the functions in E. Let $F : G \to H$ be defined as $F(x) = f(x)$ for all $x \in D_f$. F is the supremum of E. 😊

1.3 Natural Numbers

Exercise 1.3.1

Take $(X, +, \cdot, \leqslant)$ ordered field. Prove:

1. If $0 \leqslant x$, then $-x \leqslant 0$.
2. If $x \leqslant y$, and $0 \leqslant z \neq 0$, then $xz \leqslant yz$.
3. For all $x \in X$, $0 \leqslant x^2$.
4. Prove $0 < 1$.

Proof. Fields have the following important properties:

- If $a \leqslant b$, then $a + c \leqslant b + c$.
- If $a, b \geqslant 0$, then $ab \geqslant 0$.

1. Take the first property with $a = 0$, $b = x$, and $c = -x$. Then $0 \leqslant x \implies 0 + (-x) \leqslant x + (-x) \implies -x \leqslant 0$.
2. If $x \leqslant y$, then $0 \leqslant y + (-x)$. By the second property, $0 \leqslant z \cdot (y + (-x)) = zy + (-zx)$. Then $0 \leqslant zy + (-zx) \implies zx \leqslant zy$.
3. We split into the three trichotomy cases:
   - If $x = 0$, then $0 \leqslant 0^2$.
   - If $x < 0$ with $x \neq 0$, then $0 \leqslant -x$. By the second property, $0 \leqslant (-x)^2 = (-x)(-x) = x^2$.
   - If $x > 0$, then $0 \leqslant x$. By the second property, $0 \leqslant x^2$.
4. FSOC, assume $0 > 1$ and multiply both sides by 1. Then we get $0 \cdot 1 > 1 \cdot 1 \Rightarrow 0 > (1)^2$, which is a contradiction to the third property we proved.

😄

Definition 1.3.1: Inductive

Take $E \subseteq \mathbb{R}$. E is *inductive* if $1 \in E$ and $x \in E$ implies $x + 1 \in E$.

Example 1.3.1 (Inductive Sets)

- $\mathbb{R}$ is inductive.
- $\{x \in \mathbb{R} : 0 \leqslant x\}$

  Proof. $1 \in E$ because $1 \geqslant 0$. If $x \in E$, then $x + 1 \geqslant 0$, so $x + 1 \in E$. 😄

Definition 1.3.2: Natural Numbers

The intersection of all inductive sets is denoted $\mathbb{N}$. The elements of $\mathbb{N}$ are called *natural numbers*.

Properties of $\mathbb{N}$:

- $\mathbb{N} \neq \emptyset$. Since $1 \in$ every inductive set, $1 \in \mathbb{N}$.
- $\mathbb{N}$ is an inductive set.

Theorem 1.3.1 Induction

For every $n \in \mathbb{N}$, let $P(n)$ be a proposition such that:

1. $P(1)$ is true.
2. If $P(n)$, then $P(n+1)$.

Then $P(n)$ is true for every $n \in \mathbb{N}$

Proof. $E = \{n \in \mathbb{N} \ : \ P(n)\}$ is inductive by 1. and 2. So, $\mathbb{N} \subseteq E$, but $E \subseteq \mathbb{N}$ by definition of $\mathbb{N}$. Thus, $E = \mathbb{N}$. ☺

Theorem 1.3.2 Archimedean Property

Let $a, b \in \mathbb{R}$ with $a > 0$. Then there is $n \in \mathbb{N}$ such that $na > b$.

Proof. If $b \leqslant 0$, then we take $n = 1$. Assume $b > 0$. For sake of contradiction, assume there does not exist n such that $na > b$. Then $E = \{na : n \in \mathbb{N}\}$ is bounded above by b. Let $c = \sup E$. $c - a \leqslant c$, so $c - a$ is not an upper bound of E. Thus, there is $n \in \mathbb{N}$ such that $c - a \leqslant na$. Then $c \leqslant (n+1)a$. But c is an upper bound of E, so $c \geqslant (n+1)a$. Thus, $c = (n+1)a$. But $c \in E$, so $c = na$ for some $n \in \mathbb{N}$. Thus, $na = (n+1)a$, so $n = n+1$, which is a contradiction. ☺

Definition 1.3.3: Integers

$\mathbb{Z} := \mathbb{N} \cup \{0\} \cup \{-n : n \in \mathbb{N}\}$

Theorem 1.3.3 Integer Part

For every $x \in \mathbb{R}$, there is a unique $k \in \mathbb{Z}$ such that $k \leqslant x < k+1$.

Definition 1.3.4: Integer Part

The k that satisfies the above theorem is called the *integer part* of x, denoted $\lfloor x \rfloor$.

Proof. Let $E = \{k \in \mathbb{Z} : k \leqslant x\}$. First we show that E is nonempty.

- If $x \geqslant 0$, then $0 \in E$, so E is nonempty.
- If $x < 0$, then $-x > 0$. By the Archimedean property, there is $n \in \mathbb{N}$ such that $n > -x$. Thus, $-n < x$. So, $-n \in E$, so E is nonempty.

Now we show that E is bounded from above. Very clearly, x is an upper bound. By supremum property, there is $L = \sup(E)$ and $L \in \mathbb{R}$. $L - 1$ is not an upper bound, which means that there is an element $k \in E$ such that $L - 1 < k$. But since L is the supremum, $L \geqslant k$. Thus, $L - 1 < k \leqslant L$. So, $L < k + 1$ so $k + 1 \notin E$. Now, $k \leqslant x$ since $k \in E$. Now we show that k is unique. Assume there is $m \in \mathbb{Z}$ such that $m \leqslant x < m+1$. Then $m \in E$, so $m \leqslant L$. But L is the supremum, so $L \geqslant m$. Thus, $L = m$. So, $k = m$. ☺

Definition 1.3.5: $\mathbb{Q}$

If $p \in \mathbb{Z}$ with $p \neq 0$, then $\exists p^{-1} \in \mathbb{R}$. Define $\mathbb{Q} = \{pq^{-1} : p, q \in \mathbb{Z}, p \neq 0\}$.

1.4 Density of Rationals

Theorem 1.4.1 Density of the Rationals

Let $a, b \in \mathbb{R}$ with $a < b$. Then there is $r \in \mathbb{Q}$ such that $a < r < b$.

Proof. We have $a < b \implies 0 = a + (-a) < b - a \implies 0 < \frac{1}{b-a}$. By the integer part theorem, there is $q \in \mathbb{Z}$ such that $\frac{1}{b-a} < q$. So now, $\frac{1}{q} < b - a \implies a < a + \frac{1}{q} < b$. Multiply both sides by $q > 0$ to get $aq < a + 1 < bq$. By the integer part theorem, there is $p \in \mathbb{Z}$ such that $p \leqslant qa < p + 1$ (i.e. $p = \lfloor qa \rfloor$). Since $qa < p + 1 \leqslant qa + 1 < qb$. Getting rid of unnecessary stuff, we have $qa < p + 1 < qb$. Thus, $a < \frac{p+1}{q} < b$. Let $r = \frac{p+1}{q}$. Then $r \in \mathbb{Q}$ and $a < r < b$. 😊

Definition 1.4.1: Irrational Numbers

$\mathbb{R} \setminus \mathbb{Q}$ is the set of *irrational numbers*.

Exercise 1.4.1 TODO in Recitation 1/23

- Prove that there is no $r \in \mathbb{Q}$ such that $r^2 = 2$.
- Prove that "$\sqrt{2}$" exists in $\mathbb{R}$. (prove that there is at least one irrational number)
  - Have to play with the set $E = \{x \in \mathbb{R} : x > 0, x^2 < 2\}$.

Theorem 1.4.2 Density of Irrationals

Let $a, b \in \mathbb{R}$ with $a < b$. Then there is $x \in \mathbb{R} \setminus \mathbb{Q}$ such that $a < x < b$.

Proof. $a < b \implies a\sqrt{2} < b\sqrt{2}$. By the density of rationals, there is $r \in \mathbb{Q}$ such that $a\sqrt{2} < r < b\sqrt{2}$. Then $a < \frac{r}{\sqrt{2}} < b$. Let $x = \frac{r}{\sqrt{2}}$. If $r = 0$, then $a\sqrt{2} < 0 < b\sqrt{2}$. By previous theorem, we can find $q \in \mathbb{Q}$ such that $a\sqrt{2} < q < 0 < b\sqrt{2}$. Then $a < \frac{q}{\sqrt{2}} < b$. Let $x = \frac{q}{\sqrt{2}}$. Then $x \in \mathbb{R} \setminus \mathbb{Q}$ and $a < x < b$. 😊

Note:

Take $x \in \mathbb{R}$, $E = \{r \in \mathbb{Q} : r < x\}$. x is the upper bound of E. This set is nonempty because we can take $x - 1 < r < x$. Now we prove that $x = \sup E$.

Proof. Assume $\exists L$ upper bound of E such that $L < x$. Then $L < x \implies$ there exists some $r \in \mathbb{Q}$ such that $L < r < x$, but $r \in E$, so L is not an upper bound of E. Thus, L cannot be an upper bound of E and x is the least upper bound of E. 😊

Since now we know that $\sqrt{2} = \sup\{r \in \mathbb{Q} : r < \sqrt{2}\}$, we can also define $3^{\sqrt{2}} = \sup\{3^r : r \in \mathbb{Q}, r < \sqrt{2}\}$.

Definition 1.4.2: x^0

Let $0 \neq x \in \mathbb{R}$. We define $x^0 = 1$.

Definition 1.4.3: x^n

Let $x \in \mathbb{R}$, $n \in \mathbb{N}$. We start with $x^1 := x$. Then assume x^m has been defined. Then we say $x^{m+1} := x^m \cdot x$.

Definition 1.4.4: $x^{p/m}$

Let $x \in \mathbb{R}$, $p \in \mathbb{Z}$, $m \in \mathbb{N}$. We say $x^{p/m} = \sqrt[m]{x^p}$.

Exercise 1.4.2 Properties of Exponenets

Let $x \in \mathbb{R}$, $r, q \in \mathbb{Q}$, and $x, r, q > 0$. Prove the following:

- $x^r \cdot x^q = x^{r+q}$
- $(x^r)^q = (x^q)^r = x^{rq}$

Proof. ☺

Definition 1.4.5: Negative Exponent

Take $x > 0$, $r = -\frac{p}{m}$ for $p, m \in \mathbb{N}$. First, we have that $x^{-r} := (x^{-1})^{p/m}$.

Exercise 1.4.3 More Properties of Exponents

Take $x \in \mathbb{R}, x > 0, r, q \in \mathbb{Q}$. Prove the following:

- If $r > 0$, prove that $x^r > 1$.
- If $r < q$, prove that $x^r < x^q$.

1.5 1/23 - Recitation - Proving Irrationality of $\sqrt{2}$

Existence of $\sqrt{2}$:

1. Let $E = \{x \in \mathbb{R} : x > 0, x^2 < 2\}$. Prove that E is non-empty and that E is bounded above.

   Proof. We know that $0 < 1$ and from that we get $1^2 = 1 < 2$, which can be checked by subtracting 1 from both sides. As such E is nonempty.

   Now we show that E is bounded above. We know that $2^2 = 4 > 2 > a^2 \in E$, so $2^2 > a^2 \Rightarrow 2 > a$, so 2 is an upper bound of E. ☺

2. By the completeness of $(\mathbb{R}, \leqslant)$, E has a supremum, L. Prove that $L > 0$ and that $L^2 = 2$.

   Proof. Since L is the least upper bound, it has to be greater than 1 which is in the set E. Therefore, $L > 1 > 0 \implies L > 0$.

   Now we show that $L^2 \geqslant 2$. For sake of contradiction, assume $L^2 < 2$. Since $L > 0$, this means that $L \in E$. By the density of rationals, there exists $r \in \mathbb{Q}$ such that $L < r < \sqrt{2}$. Since L is an upper bound of E, $r \notin E$. But $r \in \mathbb{Q}$, so $r^2 \neq 2$. Thus, $r^2 > 2$. Since $r > 0$, $r^2 > 2 \implies r > \sqrt{2}$. But $r < \sqrt{2}$, so we have a contradiction. Thus, $L^2 \geqslant 2$. ☺

3. Prove that if $y \in \mathbb{R} \setminus E$ and $y > 0$, then y is an upper bound of E.

   Proof. Assume $y \in \mathbb{R} \setminus E$ and $y > 0$. We need to show that y is an upper bound of E. Assume for sake of contradiction that y is not an upper bound of E. Then there exists $x \in E$ such that $x > y$. But $x \in E \implies x^2 < 2$. Since $y > 0$, $x^2 < 2 \implies y^2 < 2$. But $y \notin E$, so $y^2 \geqslant 2$. But this would mean that $y \in E$. Contradiction. Thus, y is an upper bound of E. ☺

4. Prove that $L^2 = 2$.

Proof. We know that $L^2 \geqslant 2$ from part 2. Now we show that $L^2 \leqslant 2$. Assume for sake of contradiction that $L^2 > 2$.

How small does $\epsilon > 0$ need to be such that $(L - \epsilon)^2 > 2$ as well.

Start with $(L-\epsilon)^2 = L^2 - 2L\epsilon + \epsilon^2$, which is greater than $L^2 - 2L\epsilon$ since $\epsilon > 0$. So now, how small does ϵ need to be such that $L^2 > 2 \implies L^2 - 2L\epsilon > 2$ too.

$$
\begin{aligned}
2L\epsilon &< 2 - L^2 \\
\epsilon &< \frac{2 - L^2}{2L}
\end{aligned}
$$

Since $L^2 > 2$, this means that an ϵ can be found. This means that L is not the least upper bound. Contradiction. Thus, $L^2 \leqslant 2$. 😄

1.6 Exponents

Definition 1.6.1: $\sqrt{2}$

$$\sqrt{2} := \sup\{x \in \mathbb{R} : x > 0, x^2 < 2\}$$

Exercise 1.6.1

For $n \in \mathbb{N}, n \geqslant 2$. Fix $x > 0$.

$$E = \{y \in \mathbb{R} : y > 0, y^n < x\}.$$

Prove that $l = \sup E$ satisfies $l^n = x$.

Proof. We first need to show that $\sup E$ exists. Let $y = x/(1+x)$. Then, $0 \leqslant y < 1$, so $y^n \leqslant y < x$. Thus, $y \in E$. So, E is nonempty. E is also bounded from above because x is an upper bound of E. Thus, $\sup E$ exists by the completeness of $\mathbb{R}$. Let $l = \sup E$. We now show that $l^n = x$.

First we show that $l^n \leqslant x$. FSOC, assume $l^n > x$. If you choose an $\epsilon > 0$ that is small enough, then $(l-\epsilon)^n > x$ as well. We can't do this because $y > l - \epsilon$ for some $y \in E$ since l is the supremum of E. As such, we arrive at a contradiction which means that $l^n \leqslant x$.

To show that $l^n \geqslant x$, assume FSOC that $l^n < x$. Then we can choose an ϵ such that $(l+\epsilon)^n < x$, meaning we have an element $(l+\epsilon)$ which is in E but bigger than the supremum, which is a contradiction.

Thus, $l^n \geqslant x$. 😆

Definition 1.6.2: $\sqrt[m]{x}$

$$\sqrt[m]{x} := \sup\{y \in \mathbb{R} : y > 0, y^m < x\}$$

Definition 1.6.3: $x^{p/q}$

$$x^{p/q} := \left(\sqrt[q]{x}\right)^p$$

Definition 1.6.4: x^q

For $q \in \mathbb{R}$, $q > 0$, and $x > 1$.

$$x^q := \sup\{x^r : r \in \mathbb{Q}, 0 < r < q\}$$

Example 1.6.1

$$\sqrt{2} = \sup\{r \in \mathbb{Q} : r > 0, r < \sqrt{2}\}$$

Theorem 1.6.1

Take $a, b \in \mathbb{R}$, $a, b > 0$ and $x \in \mathbb{R} > 1$. Then $x^a \cdot x^b = x^{a+b}$.

Proof. Let $E_i = \{x^r : r \in \mathbb{Q}, r > 0, r < i\}$. Consider E_a, E_b, E_{a+b}. Then let $l_i = \sup(E_i)$. Consider l_a, l_b, l_{a+b}. We want to show that $l_a \cdot l_b = l_{a+b}$ by showing that both $l_a \cdot l_b \leqslant l_{a+b}$ and $l_a \cdot l_b \geqslant l_{a+b}$.

Let $r \in \mathbb{Q}$ with $0 < r < a$. Let $s \in \mathbb{Q}$ with $0 < s < b$. Then we have that $x^r \cdot x^s = x^{r+s}$ (from the exercise two days ago and since $r, s \in \mathbb{Q}$.) we know that $0 < r + s < a + b$ and is rational. Thus, $x^{r+s} \in E_{a+b}$. Thus, $x^r \cdot x^s \leqslant l_{a+b}$.

We want to divide both sides by x^s while fixing r. So, we have that $x^r \leqslant \dfrac{l_{a+b}}{x^s}$, which is true for all $r \in \mathbb{Q}$, such that $0 < r < a$. Thus, $\dfrac{l_{a+b}}{x^s}$ is an upper bound for E_a. Thus, $l_a \leqslant \dfrac{l_{a+b}}{x^s}$. Thus, $x^s \leqslant \dfrac{l_{a+b}}{l_a}$, meaning that $\frac{l_{a+b}}{l_a}$ is an upper bound for E_b. Thus, $l_b \leqslant \dfrac{l_{a+b}}{l_a}$. Thus, $l_a \cdot l_b \leqslant l_{a+b}$.

Now we show that $l_a \cdot l_b \geqslant l_{a+b}$. Let $t \in \mathbb{Q}$ with $0 < t < a+b$. We need $0 < r \in \mathbb{Q} < a$ and $0 < s \in \mathbb{Q} < b$ with $t = r + s$. We start by looking at $t - a < b$. By the density of $\mathbb{Q}$, find $s \in \mathbb{Q}$ such that $t - a < s < b$. Take $s > 0$ beacuse $b > 0$. So $t - s < a$. By the density of $\mathbb{Q}$, find $0 < p \in \mathbb{Q}$ such that $t - s < p < a$. So $t < s + p$. So, $x^t < x^{s+p} = x^s x^p \leqslant l_a l_b$ since $x^s \in E_b$ and $x^p \in E_a$. We know that $l_a l_b$ is an upper bound of E_{a+b}, so $l_{a+b} \leqslant l_a l_b$.

Therefore $l_a \cdot l_b = l_{a+b}$. 😊

Definition 1.6.5: Negative Exponents

Let $x > 1$, $a < 0$. Then:

$$x^a := (x^{-a})^{-1}$$

Definition 1.6.6: Exponents between 0 and 1

Let $x \in \mathbb{R}$ with $0 < x < 1$ and $a > 0$. Then:

$$x^a := \left(\frac{1}{x}\right)^{-a}$$

An important note is that if we have $E \subseteq (0, \infty)$ with a bounded E. Then if we define $F = \{\dfrac{1}{x} : x \in E\}$, then we have the following:

$$\sup E = \frac{1}{\inf F}$$
$$\inf E = \frac{1}{\sup F}$$

1.7 1/25 - Recitation - Sequences of Set

Definition 1.7.1: Sequence of a Set

Given a set X, a sequence on X is a function $f : \mathbb{N} \to X$. We denote $f(n)$ as x_n. We can also denote the sequence as $\{x_n\}_{n=1}^{\infty}$.

Definition 1.7.2

Let $(X, \leqslant)$ be a poset and $\{x_n\}_{n=1}^{\infty}$ be a sequence on X. Then $E = \{x_n : n \in \mathbb{N}\}$ is a subset of X. We say that $\{x_n\}_{n=1}^{\infty}$ is bounded from above is the set E is bounded from above. We say that $\{x_n\}_{n=1}^{\infty}$ is bounded from below is the set E is bounded from below. We say that $\{x_n\}_{n=1}^{\infty}$ is bounded if it is bounded from above and below.

Definition 1.7.3: Limit Superior

Let $(X, \leqslant)$ be a poset. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence on X. Suppose $\{x_n\}_n$ is bounded from above. Then, we define the *limit superior* of x_n as $n \to \infty$ as:

$$\limsup_{n \to \infty} x_n = \inf_{n \in \mathbb{N}} \sup_{k \geqslant n} x_k$$

Definition 1.7.4: Limit Inferior

Let $(X, \leqslant)$ be a poset. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence on X. Suppose $\{x_n\}_n$ is bounded from below. Then, we define the *limit inferior* of x_n as $n \to \infty$ as:

$$\liminf_{n\to\infty} x_n = \sup_{n\in\mathbb{N}} \inf_{k\geqslant n} x_k$$

Exercise 1.7.1

1. Let $\{x_n\}_{n=1}^\infty$ be a sequence on $\mathbb{R}$ bounded above. Prove that $L \in \mathbb{R}$ is the $\limsup$ of $\{x_n\}_{n=1}^\infty$ iff for every $\epsilon > 0$, there exists $n_\epsilon \in \mathbb{N}$ such that:

   (a) $x_n < L + \epsilon$ for all $n \geqslant n_\epsilon$.

   (b) $L - \epsilon < x_n$ for infinitely many n.

Proof. Let $L \in \mathbb{R}$ be the $\limsup$ of $\{x_n\}_{n=1}^\infty$. Let $\epsilon > 0$. L being the lim sup means that $L = \inf_{n\in\mathbb{N}} \sup_{k\geqslant n} x_k$. Thus, $L \leqslant \sup_{k\geqslant n} x_k$ for all $n \in \mathbb{N}$. Thus, $L - \epsilon < \sup_{k\geqslant n} x_k$ for all $n \in \mathbb{N}$. Then $L - \epsilon$ is not an upper bound of $\{x_n\}_{n=1}^\infty$. Thus, there is $n_\epsilon \in \mathbb{N}$ such that $L - \epsilon < x_{n_\epsilon}$. Thus, $L - \epsilon < x_n$ for infinitely many n. Now we show that $x_n < L + \epsilon$ for all $n \geqslant n_\epsilon$. Assume for sake of contradiction that there is $n \geqslant n_\epsilon$ such that $x_n \geqslant L + \epsilon$. Then $L + \epsilon$ is an upper bound of $\{x_n\}_{n=1}^\infty$. But L is the $\limsup$, so $L \geqslant L + \epsilon$. Contradiction. Thus, $x_n < L + \epsilon$ for all $n \geqslant n_\epsilon$.

Now we show the other direction. Assume that for every $\epsilon > 0$, there exists $n_\epsilon \in \mathbb{N}$ such that $x_n < L + \epsilon$ for all $n \geqslant n_\epsilon$ and $L - \epsilon < x_n$ for infinitely many n. We want to show that L is the $\limsup$ of $\{x_n\}_{n=1}^\infty$. We know that L is an upper bound of $\{x_n\}_{n=1}^\infty$. We need to show that L is the least upper bound. Assume for sake of contradiction that L is not the least upper bound. Then there is $L' < L$ such that L' is an upper bound of $\{x_n\}_{n=1}^\infty$. Let $\epsilon = L - L'$. Then $L' < L - \epsilon$. But $L - \epsilon < x_n$ for infinitely many n. But $L' < L - \epsilon$, so L' is not an upper bound of $\{x_n\}_{n=1}^\infty$. Contradiction. 😊

1.8 Vector Spaces

Example 1.8.1 (Vector Spaces)

- Euclidean Space $\subseteq \mathbb{R}^n$. $x \in \mathbb{R}^n$ is a vector. $x = (x_1, \ldots, x_n)$.
- Polynomial Space from $\mathbb{R} \to \mathbb{R}$. $x \in \mathbb{R}^x$. $x = a_0 + a_1 x + \cdots + a_n x^n$.
- $f : [a, b] \to \mathbb{R}$ continuous functions.

Definition 1.8.1: Boundedness of Functions

Let E be a set and $f : E \to \mathbb{R}$.

1. f is bounded from above if the set $f(E) = \{y \in \mathbb{R} : y = f(x), x \in E\}$ is bounded from above.
2. f is bounded from below if the set $f(E) = \{y \in \mathbb{R} : y = f(x), x \in E\}$ is bounded from below.
3. f is bounded if $f(E)$ is bounded.

Definition 1.8.2: Inner Product

A function $(\cdot, \cdot) : V \times V \to \mathbb{R}$ is an *inner product* if it satisfies the following properties:

- $(x, x) \geqslant 0$ for all $x \in X$.
- $(x, x) = 0$ iff $x = 0$.
- $(x, y) = (y, x)$ for all $x, y \in X$.
- $(sx + ty, z) = s(x, z) + t(y, z)$ for all $x, y, z \in X$ and $s, t \in \mathbb{R}$.

Example 1.8.2 (Examples of Inner Products)

- $\mathbb{R}^n$ with dot products.
- $f : [a,b] \to \mathbb{R}$ with $(f,g) = \int_a^b f(x)g(x)dx$. This is is not an inner product because we can define:

$$f = \begin{cases} 1 & x = 0.5 \\ 0 & \text{otherwise} \end{cases}$$

which has an integral of 0. But $f \neq 0$. If we add that f is continuous, then it is an inner product.

Definition 1.8.3: Norm

Let V be a vector space with an inner product $(\cdot,\cdot)$. Then the *norm* of $x \in X$ is defined as $||\cdot|| : X \to [0,\infty)$ such that:

1. $||x|| = 0 \iff x = 0$
2. $||tx|| = |t|||x||$ for all $x \in X$
3. $||x+y|| \leqslant ||x|| + ||y||$ for all $x, y \in X$

Example 1.8.3 (Examples of Norms)

- $||x|| = \sqrt{(x,x)}$ for $x \in \mathbb{R}^n$
- $X = \{f : E \to \mathbb{R}, f \text{ bounded}\}$. $||f|| = \sup_{x\in E}|f(x)|$.
  - First property is obviously true.
  - For the second property, we use the fact that

$$\sup(tF) = \begin{cases} t\sup(F) & \text{if } t \geqslant 0 \\ t\inf(F) & \text{if } t < 0 \end{cases}$$

  - For the third property, we use the triangle inequality:

$$\sup|f+g| \leqslant \sup|f| + \sup|g|$$
$$|f(x)+g(x)| \leqslant |f(x)| + |g(x)| \leqslant \sup|f| + \sup|g|$$

Note:
Space of bounded functions denoted as $\ell^\infty(E) = \{f : E \to \mathbb{R} : f \text{ bounded}\}$.

Theorem 1.8.1 Cauchy Schwarz Inequality

Let X be a vector space with an inner product $(\cdot,\cdot)$. Then for all $x, y \in X$, we have that $|(x,y)| \leqslant \sqrt{(x,x)} \cdot \sqrt{(y,y)}$.

Proof. Let $y \neq 0$. Consider $(x+ty, x+ty) = (x, x+ty) + t(y, x+ty) = (x,x) + t(x,y) + t(y,x) + t^2(y,y)$. We can combine the middle terms to get $t^2(y,y)+2(x,y)+(x,x)$, which is quadratic in t. Take $t=-\dfrac{(x,y)}{(y,y)}$.

$$\begin{aligned}
0 &\leqslant (x,x) - 2\frac{(x,x)^2}{(y,y)} + \frac{(x,y)^2}{(y,y)} \\
0 &\leqslant (x,x)(y,y) - 2(x,y)^2 + (x,y)^2 \\
0 &\leqslant (x,x)(y,y) - (x,y)^2 \\
(x,y)^2 &\leqslant (x,x)(y,y) \\
|(x,y)| &\leqslant \sqrt{(x,x)}\cdot\sqrt{(y,y)}
\end{aligned}$$

☺

1.9 Inner Products, Norms, and Metric Spaces

Theorem 1.9.1

Let X be a vector space with an inner product $(\cdot,\cdot)$. Then $||x|| := \sqrt{(x,x)}$ is a norm.

Proof. We check the properties of norms:

1. $||x|| = 0 \iff \sqrt{(x,x)} = 0 \iff (x,x) = 0 \iff x = 0$.
2. $||tx|| = \sqrt{(tx,tx)} = \sqrt{t^2(x,x)} = |t|\sqrt{(x,x)} = |t|||x||$.
3. $||x+y||^2 = (x+y,x+y) = (x,x)+2(x,y)+(y,y) = ||x||^2+2(x,y)+||y||^2 \leqslant ||x||^2+2|(x,y)|+||y||^2 \leqslant ||x||^2+2||x||\cdot||y||+||y||^2 = (||x||+||y||)^2$.

☺

Corollary 1.9.1 Parallelogram Identity

Let X be a vector space with inner product $(\cdot,\cdot)$. Then for all $x,y\in X$, we have that

$$||x+y||^2+||x-y||^2 = 2||x||^2+2||y||^2$$

Proof.

$$\begin{aligned}
||x+y||^2+||x-y||^2 &= (x+y,x+y)+(x-y,x-y) \\
&= (x,x)+2(x,y)+(y,y)+(x,x)-2(x,y)+(y,y) \\
&= 2(x,x)+2(y,y) \\
&= 2||x||^2+2||y||^2
\end{aligned}$$

☺

If we subtract them instead, we get

$$\frac{||x+y||^2-||x-y||^2}{4} = (x,y) \tag{*}$$

So, if $||\cdot||$ is a norm, then if i want to define an inner product, I can use *.

Exercise 1.9.1

Let $||\cdot||$ be a norm. Then $(x,y) := \frac{1}{4}(||x+y||^2-||x-y||^2)$ is an inner product iif the parallelogram identity holds.

Linearity of inner products is the hard part to prove because we have to consider:

- $t \in \mathbb{N}$
- $t = \frac{1}{2}$
- $t \in \mathbb{Q}$
- $t \in \mathbb{R}$ (density of $\mathbb{Q}$)

Note:

For recitation:

1. $X = \{f : E \to \mathbb{R} \text{ bounded}\}, ||f|| = \sup_E |f|$, does not satisfy the parallelogram identity.
2. $x \in \mathbb{R}^N$, $||x||_1 = |x_1| + |x_2| + \cdots + |x_N|$ does not satisfy the parallelogram identity.

Definition 1.9.1: Metric

Let X be a set. A *metric* on X is a function $d : X \times X \to [0, \infty)$ such that:

1. $d(x, y) = 0 \iff x = y$
2. $d(x, y) = d(y, x)$ for all $x, y \in X$
3. $d(x, z) \leqslant d(x, y) + d(y, z)$ for all $x, y, z \in X$

Definition 1.9.2: Metric Space

A set X with a metric d is called a *metric space* and is denoted as (X, d).

Example 1.9.1 (Metrics)

Let X be a set. Then the following is a metric on X:

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

Theorem 1.9.2 If X is a vector space with $||\cdot||$ as a norm. Then

$$d(x, y) := ||x - y||$$

is a metric on X.

Proof. We check all the properties of metrics.

- $d(x, y) = 0 = ||x - y|| \Rightarrow 0 = x - y \iff x = y$.
- $d(x, y) = ||x - y|| = ||y - x|| = d(y, x)$.
- $d(x, y) = ||x - y|| = ||x - z + z - y|| \leqslant ||x - z|| + ||z - y|| = d(x, z) + d(z, y)$.

😄

Example 1.9.2

Let's define

$$d(x, y) = \left| \frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right|$$

as a metric on $\mathbb{R}$. However, this is not a norm because $d(tx, ty) \neq td(x, y)$.

Definition 1.9.3: Ball

Let (X, d) be a metric space. Let $x \in X$ and $r > 0$. Then the *ball* of radius r centered at x is defined as $B_r(x) = \{y \in X : d(x, y) < r\}$.

Example 1.9.3

- Take $X = \mathbb{R}^2$ with $(x, y) \in \mathbb{R}$. Then defin $||(x, y)||_\infty = \max(|x|, |y|)$ is a norm. Take $B((0, 0), 1) = \{(x, y) \in \mathbb{R}^2 : ||(x, y) - (0, 0)||_\infty < 1\}$. This is a square with vertices $(1, 1), (-1, 1), (-1, -1), (1, -1)$.
- If we have $||(x, y)||_1 = |x| + |y|$, then $B((0, 0), 1) = \{(x, y) \in \mathbb{R}^2 : ||(x, y) - (0, 0)||_1 < 1\}$. This is a square with vertices $(1, 0), (0, 1), (-1, 0), (0, -1)$.

Definition 1.9.4: Interior

Let (X, d) be a metric space and $E \subseteq X$. $x \in E$ is called an *interior point* of E if there is $B(x, r) \subseteq E$. The set of all interior points of E is called the *interior* of E and is denoted as E°.

Definition 1.9.5: Open Set

E is *open* if $E = E^\circ$.

1.10 Open Sets

Example 1.10.1 (Balls)

$B(x, r)$ is open.

Proof. Let $y \in B(x, r)$ and take $B(y, r - d(x, y))$. Let $z \in B(y, r - d(x, y))$. Then $d(x, z) \leqslant d(x, y) + d(y, z) < d(x, y) + r - d(x, y) = r$. Thus, $z \in B(x, r)$. Thus, $B(y, r - d(x, y)) \subseteq B(x, r)$. Thus, $B(x, r)$ is open. 😊

Example 1.10.2 ($\mathbb{R}$)

1. $E = (0, 1) \cap \mathbb{Q}$ is not open. Because the irrationals are dense, we can always find a rational number in any ball. Thus, $E^\circ = \emptyset$.
2. $E = (3, 4)$ is open. Let $x \in E$. Take $B(x, \min(x - 3, 4 - x))$. Then $B(x, \min(x - 3, 4 - x)) \subseteq E$. Thus, E is open.
3. $E = [3, 4)$ is not open. $E^\circ = (3, 4)$.
4. $E = \{x \in \mathbb{R} : x^3 - 3x + 4 > 0\}$. This is open and we'll be able to use continuity to prove this easily later.
5. $l^\infty([0, 1]) = \{f : [0, 1] \to \mathbb{R} \text{ bounded}\}$. $||f||_\infty = \sup_{[0,1]} |f|$. $d(f, g) = ||f - g||_\infty$. $E = \{f \in l^\infty([0, 1]) : f(x) > 0 \ \forall x \in [0, 1]\}$ is open? (finish in recitation)

Properties of open sets (X, d):

- $\emptyset$ is open. X is open.
- Infinite intersections of open sets are not necessarily open. For example, we have $\bigcap_{n=1}^{\infty}(-1/n, 1/n) = \{0\}$, which is not open.
- Finite intersections of open sets are open. Consider $U_1, \dots U_n$. Let $x \in \bigcap_{i=1}^{n} U_i$. Then $x \in U_i$ for all i. Since U_i is open, there exists $r_i > 0$ such that $B(x, r_i) \subseteq U_i$. Let $r = \min(r_1, \dots, r_n)$. Then $B(x, r) \subseteq U_i$ for all i. Thus, $B(x, r) \subseteq \bigcap_{i=1}^{n} U_i$.
- Unions of open sets are open because if a point in the union is contained in one of the open sets, then there is a ball in that set that is contained in the union.

Definition 1.10.1: Topological Space

Let X be a set. A *topology* on X is a collection $\mathcal{T}$ of subsets of X such that:

1. $\emptyset, X \in \mathcal{T}$.
2. If $U_1, \dots, U_n \in \mathcal{T}$, then $\bigcap_{i=1}^{n} U_i \in \mathcal{T}$. (finite intersections)
3. If $U_\alpha \in \mathcal{T}$ for all $\alpha \in A$, then $\bigcup_{\alpha \in A} U_\alpha \in \mathcal{T}$. (arbitrary unions)

Elements of $\mathcal{T}$ are called open sets.

Definition 1.10.2: Closed

Let (X, d) be a metric space. We say $C \subseteq X$ is *closed* if $X \setminus C$ is open.

Note that X and $\emptyset$ are both open and closed.

Example 1.10.3 (Open and Closed Sets)

- $[0, 1)$ is not open or closed.
- $[0, 1]$ is closed.

Properties of closed sets:

- $\emptyset$ and X are closed.
- Infinite intersections of closed sets are closed. (De Morgan's Law)
- Finite unions of closed sets are closed. For example, if we have $\bigcup_{m=1}^{\infty}(-\infty, -\frac{1}{m}) = (-\infty, 0)$ which is closed.

1.11 2/1 - Rectitation

Recall:

1. Let $\{x_n\}$ be a sequence bounded above in $\mathbb{R}$. Then $L \in \mathbb{R}$ is the limit superior of $\{x_n\}$ if for every $\epsilon > 0$, there exists $n_\epsilon \in \mathbb{N}$ such that:
   (a) $x_n < L + \epsilon$ for all $n \geqslant n_\epsilon$.
   (b) $x_n > L - \epsilon$ for infinitely many n.
2. Let $\{x_n\}$ be a sequence bounded below in $\mathbb{R}$. Then $L \in \mathbb{R}$ is the limit inferior of $\{x_n\}$ if for every $\epsilon > 0$, there exists $n_\epsilon \in \mathbb{N}$ such that:
   (a) $x_n < L + \epsilon$ for infinitely many n.
   (b) $x_n > L - \epsilon$ for all $n \geqslant n_\epsilon$.

Now consider the following sequence:

$$x_n = (-1)^n \frac{2n}{n+1} \in \mathbb{R}$$

Prove that $\limsup_{n\to\infty} x_n = 2$.

Proof. We need to show that for every $\epsilon > 0$, there exists $n_\epsilon \in \mathbb{N}$ such that:

1. $x_n < 2 + \epsilon$ for all $n \geqslant n_\epsilon$.
2. $2 - \epsilon < x_n$ for infinitely many n.

Let $\epsilon > 0$. We need to find $n_\epsilon \in \mathbb{N}$ such that $x_n < 2 + \epsilon$ for all $n \geqslant n_\epsilon$ and $2 - \epsilon < x_n$ for infinitely many n. We can find $n_\epsilon \in \mathbb{N}$ such that $2 - \epsilon < x_n$ for all $n \geqslant n_\epsilon$. Then $x_n < 2 + \epsilon$ for all $n \geqslant n_\epsilon$. Thus, $\limsup_{n\to\infty} x_n = 2$. 😄

Now prove that for any $\{x_n\}$ in $\mathbb{R}$, prove that $\liminf_{n\to\infty} x_n \leqslant \limsup_{n\to\infty} x_n$.

Proof. Comes quickly from properties of limits and that the inf is less than the sup. 😄

Now prove that $\liminf_{n\to\infty} -x_n = -\limsup_{n\to\infty} x_n$ and that $\limsup_{n\to\infty} -x_n = -\liminf_{n\to\infty} x_n$.

Proof. We start by using the property that $\inf(-E) = -\sup(E)$. Then we use the property that $\sup(-E) = -\inf(E)$. So,

$$\begin{aligned}
\liminf_{n\to\infty} -x_n &= \sup_{n\in\mathbb{N}} \inf_{k\geqslant n} -x_k \\
&= \sup_{n\in\mathbb{N}} -\sup_{k\geqslant n} x_k \\
&= -\inf_{n\in\mathbb{N}} \sup_{k\geqslant n} x_k \\
&= -\limsup_{n\to\infty} x_n
\end{aligned}$$

😄

1.12 Closure

Definition 1.12.1: Closure

Let (X, d) be a metric space with $A \subset X$. Then the *closure* of A is defined as $\overline{A}$, the intersection of all sets that contain E.

Definition 1.12.2: Boundary Point

Let (X, d) be a metric space with $E \subseteq X$. Then $x \in X$ is a *boundary point* of E if for every $r > 0$, $B(x, r) \cap E \neq \emptyset$ and $B(x, r) \cap (X \setminus E) \neq \emptyset$. The set of all boundary points is denoted as ∂E.

Theorem 1.12.1

Let (X, d) be a metric space and $E \subseteq X$. Then $\overline{E} = E \cup \partial E$.

Proof. Let $x \in \overline{E}$. FSOC, assume $x \notin E \cup \partial E$. Since $x \notin \partial E$, there exists $r > 0$ such that $B(x, r)$ that doesn't intersect with either E or complement of E. But since $x \notin E$, only the second option can occur. So there exists r such that $B(x, r) \cap E = \emptyset$. Because of that and the fact that $B(x, r)$ is open, it follows that $X \setminus B(x, r)$ is closed and contains E. By the definition of $\overline{E}$, we have that $\overline{E} \subseteq X \setminus B(x, r)$. But this is a contradiction because $x \in \overline{E}$.

Conversely, let $x \in E \cup \partial E$ and assume $x \notin \overline{E}$. Since $\overline{E}$ is closed, $X \setminus \overline{E}$ is open. Using the fact that $x \in E \cup \partial E$, we have that we can find a $B(x, r) \in X \setminus \overline{E}$. But this is a contradiction because $B(x, r)$ is open and contains E. Thus, $E \cup \partial E \subseteq \overline{E}$. 😄

Definition 1.12.3: Accumulation Point

Let (X, d) be a metric space with $E \subseteq X$. Then $x \in X$ is an *accumulation point* of E if for every $r > 0$, there exists $y \in E$ such that $y \neq x$ and $d(x, y) < r$.

Definition 1.12.4: Interval

$I \subseteq \mathbb{R}$ is an *interval* if we have that $z \in I$ for all $x < z < y$.

Definition 1.12.5: Rectangle

$R \subseteq \mathbb{R}^N$ is a *rectangle* if $R = I_1 \times \cdots \times I_N$ where $I_1, \ldots, I_N$ are intervals in $\mathbb{R}$.

Definition 1.12.6: Sequence

Let X be a set. A *sequence* is a function $f : \mathbb{N} \to X$. We denote $f(n)$ as x_n.

Definition 1.12.7: Convergent Sequence

Let (X, d) be a metric space. A sequence $\{x_n\}_{n=1}^{\infty}$ is *convergent* if there exists $x \in X$ such that for every $\epsilon > 0$, there exists $n_\epsilon \in \mathbb{N}$ such that $d(x, x_n) < \epsilon$ for all $n \geqslant n_\epsilon$. We write $x_n \to x$ as $n \to \infty$ or $\lim_{n \to \infty} x_n = x$.

1.13 Bolzano-Weierstrass

Theorem 1.13.1 Bolzano-Weierstrauss

If $E \subset \mathbb{R}^N$ is bounded and contains infinitely many distinct points, then E has an accumulation point

Proof.

Lenma 1.13.1 1

If $[a_n, b_n] \supseteq [a_{n+1}, b_{n+1}]$ for all n, then $\bigcap_{n=1}^{\infty}[a_n, b_n] \neq \emptyset$.

Proof. For all a_n and b_n, we have:

$$a_1 \leqslant a_2 \leqslant \cdots$$
$$b_1 \geqslant b_2 \geqslant \cdots$$

Let

$$A := \{a_1, a_2, \ldots\}.$$

We have that $a_n \leqslant b_n \leqslant b_1$ for all n. So A is bounded above, so by the supremum property, there exists $x = \sup A \in \mathbb{R}$ and $a_n \leqslant x$ for all $n \in \mathbb{N}$. We claim that $x \leqslant b_n$ as well. If not, then there exists $m \in \mathbb{N}$ such that $b_m < x$. Since x is an upper bound of A, we'll have that there's an $n \in \mathbb{N}$ such that $b_m < a_n \leqslant x$. Find $k \geqslant m, n$, then we have $b_m < a_n \leqslant a_k \leqslant b_k \leqslant b_m$, which is a contradiction. This proves the claim. Hence, $x \in [a_n, b_n]$ for all n. Thus, $x \in \bigcap_{n=1}^{\infty}[a_n, b_n]$. 😊

Lenma 1.13.2 2

Let R_n be a closed and bounded rectangle. Assume that $R_1 \supseteq R_2 \supseteq \cdots$. Then $\bigcap_{n=1}^{\infty} R_n \neq \emptyset$.

Proof. We know that

$$R_n = [a_{1,n}, b_{1,n}] \times \cdots \times [a_{N,n}, b_{N,n}]$$
$$R_{n+1} = [a_{1,n+1}, b_{1,n+1}] \times \cdots \times [a_{N,n+1}, b_{N,n+1}]$$

We can apply lemma 1 N times (for each of the components of R_n) to find that $x_1, x_2, \ldots, x_N \in \mathbb{R}$ such that $a_{i,n} \leqslant x_i \leqslant b_{i,n}$ for all $1 \leqslant i \leqslant N$. Then, if you take $x = (x_1, \ldots, x_N)$, then $x \in R_m$ for all n. Thus, $x \in \bigcap_{n=1}^{\infty} R_n$. 😊

Lenma 1.13.3 3

Let (X, d) be a metric space with $E \subseteq X$. Then $x \in X$ is an accumulation point of E if and only if there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in E such that $x_n \to x$ as $n \to \infty$.

Proof. Let $x \in X$ be an accumulation point of E. Take $r = \dfrac{1}{n}$. Find $x_n \in B\left(x, \dfrac{1}{n}\right) \cap E$ with $x_n \neq x$. We claim $x_n \to x$. Given $\epsilon > 0$, find $n_\epsilon \geqslant \dfrac{1}{\epsilon}$. Then $d(x, x_n) < \dfrac{1}{n} \leqslant \dfrac{1}{n_\epsilon}$ for all $n \geqslant n_\epsilon$. Thus, $x_n \to x$ as $n \to \infty$.

Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in E such that $x_n \to x$ as $n \to \infty$. We claim that $x \in \operatorname{acc}(E)$. Let $r > 0$ and take $\epsilon = r$. Then there exists $n_\epsilon \in \mathbb{N}$ such that $d(x, x_n) < \epsilon = r$ for all $n \geqslant n_\epsilon$. Thus, $x_n \in B(x, r) \cap E$ for all $n \geqslant n_\epsilon$. Thus, $x \in \operatorname{acc}(E)$. 😊

Now we prove the actual theorem. Let $E \subseteq \mathbb{R}^N$ be bounded. $E \subseteq B(0, r)$ for some r. Let Q_1 be the closed cube centered at 0 with sidelength $2r$. Pick some point $x_1 \in E \subseteq Q_1$. Subdivide Q_1 into 2^N closed cubes of sidelength $\dfrac{2r}{2}$. Let Q_2 be the closed cube containing x_1. Pick some point $x_2 \in E \cap Q_2$ with $x_2 \neq x_1$. Inductively, assume $Q_1 \supseteq Q_2 \supseteq \cdots \supseteq Q_n$ have been chosen. Then Q_n is a closed cube of sidelength $\dfrac{2r}{2^{n-1}}$ containing x_n. Each

Q_n contains infinitely many elements of E. Assume also that $x_1, x_2, \cdots x_n \in E$ have been chosen with $x_i \in Q_i$ and $x_i \neq x_j$ for $i \neq j$.

Now we can subdivide Q_n to get Q_{n+1} and continue this process infinitely.

By Lemma 2, we know that $\bigcap_{n=1}^{\infty} Q_n \neq \emptyset$. Let $x \in \bigcap_{n=1}^{\infty} Q_n$. Now we need to show there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in E such that $x_n \to x$ as $n \to \infty$ but $x_i \neq x$ for any i because then the rest of the points won't converge to x. If $x = x_i$ for some i, we can just pick another point.

So WLOG, assume $x_n \neq x$ for any n. So we claim $x_n \to x$ as $n \to \infty$. We know that in Q_n, the difference between any two points in this cube is given by:

$$||x_n - x|| = \sqrt{(x_{n,1} - x_1)^2 + (x_{n,2} - x_2)^2 + \cdots + (x_{n,N} - x_N)^2} \leqslant \sqrt{\frac{2r}{2^{n-1}} + \frac{2r}{2^{n-1}} + \cdots + \frac{2r}{2^{n-1}}} = \sqrt{N}\frac{2r}{2^{n-1}}$$

This value is less than ϵ for all large n, so this concludes the proof. 😊

1.14 2/6 - Recitation - Spaces

Let $X = \{f : [0,1] \to \mathbb{R} \text{ bounded}\}$. Define $\|f\| = \sup_{x\in[0,1]} |f(x)|$. Prove that $(X, \|\cdot\|)$ does not suffice parallelogram identity. That is, show a counterexample to the parallelogram identity, which is

$$\|f+g\|^2 + \|f-g\|^2 = 2\|f\|^2 + 2\|g\|^2$$

Proof. Counterexample: Let $f(x) = x$ and $g(x) = 1$. 😊

Now given a normed space which satisifies the parallelogram identity, can we define an inner product?

Proof. Yes. We can define $(f,g) = \frac{1}{4}(\|f+g\|^2 - \|f-g\|^2)$. We can prove that this is an inner product.

Linearity of products because the other properties are easy to prove. We need to show that $(x+y,z) = (x,z) + (y,z)$. I'm so lazy so I won't tbh.

We now show that $(tx,y) = t(x,y) \forall t \in \mathbb{Z}$. We proceed with induction for $t \in \mathbb{Z}^+$

Our two base cases are $t = 0, 1$. For $t = 0$, we have that $(0x, y) = (0, y) = 0 = 0(0, y)$. For $t = 1$, we have that $(x, y) = (x, y) = 1(x, y)$.

Now we assume that $(tx, y) = t(x, y)$ for some $t \in \mathbb{Z}^+$. Then we have that $(t+1)x = tx + x$. Then we have that $(t+1)x, y = (tx + x, y) = (tx, y) + (x, y) = t(x, y) + (x, y) = (t+1)(x, y)$. Thus, we have that $(tx, y) = t(x, y)$ for all $t \in \mathbb{Z}^+$.

Now we have to deal with $t \in \mathbb{Z}^-$. We have that $(tx, y) = -t(-x, y) = -t(x, y) = t(x, y)$. Thus, we have that $(tx, y) = t(x, y)$ for all $t \in \mathbb{Z}$.

To proceed, we deal with $t \in \mathbb{Q}$. We have that $t = \frac{m}{n}$ for some $m, n \in \mathbb{Z}$. Then we have that $n(tx, y) = (ntx, y) = (mx, y) = m(x, y) = t(mx, y) = t(n(x, y))$. Thus, we have that $n(tx, y) = t(n(x, y))$. Thus, we have that $(tx, y) = t(x, y)$ for all $t \in \mathbb{Q}$. 😊

1.15 Compactness

Definition 1.15.1: Subsequence

Let X be a set and $f : \mathbb{N} \to X$ a sequence. Let $g : \mathbb{N} \to \mathbb{N}$ be strictly increasing. Then $f \circ g : \mathbb{N} \to X$ is a *subsequence* of f. We denote m_k as $g(k)$, so $f(g(k)) = f(m_k) = x_{m_k}$. So we denote the whole sequence as $\{x_{m_k}\}_k$.

Definition 1.15.2: Sequentially Compact

Let (X, d) be a metric space. $K \subseteq X$ is *sequentially compact* if every sequence $\{x_n\}_n$ in K and there exists a subsequence $\{x_{n_k}\}_k$ such that $x_{n_k} \to x$ as $k \to \infty$ for some $x \in K$.

Example 1.15.1 ($\mathbb{R}$)

1. $(0, 1]$ is not sequentially compact. Consider the sequence $x_n = \dfrac{1}{n}$. This sequence has no convergent subsequence that tends to 0 since 0 is not in the set. The issue is that it's not closed.
2. $[0, \infty)$ is not sequentially compact. Consider the sequence $x_n = n$. This sequence has no convergent subsequence that tends to ∞ since ∞ is not in the set. So, $[0, \infty)$ is not sequentially compact. The issue is that it's not bounded.

Theorem 1.15.1
Let (X, d) be a metric space. If $K \subseteq X$ is sequentially compact, then K is closed and bounded.

Proof. Claim: K is closed. We want $X \setminus K$ to be open. Let $x \in X \setminus K$. We want $B(x, r) \in X \setminus K$ for some $r > 0$. By contradiction, for all $r > 0$, assume $\exists y \in B(x, r) \cap K$. Take $r = \frac{1}{m} \Rightarrow y_m \in B(x, \frac{1}{m}) \cap K$. $d(y_m, x) < \frac{1}{m} \to 0$, so $y_m \to x$. But $x \notin K$ even though $y_m \in K$. This is a contradiction, so K is closed.

Claim: K is bounded. By contradiction, assume K is not bounded. Let $x_0 \in X$. Then $K \nsubseteq B(x_0, r)$ for any $r > 0$. Take $r = n$. Then $\exists x_n \in K$ such that $d(x_n, x_0) \geqslant n$. So $\{x_n\}_n \in K$. K is sequentially compact, so there exists a subsequence $\{x_{n_k}\}_k$ such that $x_{n_k} \to x$ as $k \to \infty$ for some $x \in K$. But $n_k \leqslant d(x_{n_k}, x_0) \leqslant d(x_{n_k}, x) + d(x, x_0)$. But $d(x_{n_k}, x) \to 0$ as $k \to \infty$, so $n_k \to \infty < d(x_{m_k}, x_0) \leqslant d(x, x_0)$ which is a fixed number, so we have a contradiction. As such, K is bounded. 😊

Theorem 1.15.2
Let $K \subseteq \mathbb{R}^N$. Then K is sequentially compact if and only if K is closed and bounded.

Proof. We just showed the first direction. So, we need to show that if K is closed and bounded, then K is sequentially compact.

So now, assume K is closed and bounded. Let $\{x_n\}_n$ be a sequence in K. We want to show that there exists a subsequence $\{x_{n_k}\}_k$ such that $x_{n_k} \to x$ as $k \to \infty$ for some $x \in K$.

Consider the set $E = \{x_n : n \in \mathbb{N}\} \subseteq \mathbb{R}_N$. We now case on whether E has infinitely many distinct points or not.

If E doesn't have infinitely many distinct points, there exists $x \in K$ such that $x_n = x$ for einfinitely many n. Then $x_{n_k} = x$ for all k, so $x_{n_k} \to x$ as $k \to \infty$.

Now we consider the case where Bolzano-Weierstrass applies. By B-W, E has an accumulation point $x \in \mathbb{R}^N$. So we can find a subsequence $\{x_{n_k}\}_k$ such that $x_{n_k} \to x$ as $k \to \infty$. But $x \in K$ because K is closed. Thus, K is sequentially compact. 😊

Note:
Let $(X, \|\cdot\|)$ be a normed space. If every closed and bounded set is sequentially compact, then X has finite dimension.

Exercise 1.15.1
Recall $l^\infty([0,1]) = \{f : [0,1] \to \mathbb{R} \text{ bounded}\}$. Define $\|f\|_\infty = \sup_{x \in [0,1]} |f(x)|$. $B(0,1) = \{g \in l^\infty([0,1]) : \|g\|_\infty < 1\}$. Prove that $B(0,1) = \{g \in l^\infty([0,1]) : |g(x)| < 1 \ \forall x \in [0,1]\}$. Also prove that this not sequentially compact.

1.16 2/8 - Recitation

Let $n \in \mathbb{N}$, $x, y \in \mathbb{R}$.

1. Prove that $x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1})$.

   Proof. Base case: $n = 1$ is trivial.

   Now assume that for any $n \in \mathbb{N}$, $x^n - y^n = (x-y)(x^{n-1}+x^{n-2}y+\cdots+xy^{n-2}+y^{n-1})$. We want to show that this is true for $n+1$. We have that $x^{n+1}-y^{n+1} = x(x^n-y^n)+y^n(x-y) = x(x-y)(x^{n-1}+x^{n-2}y+\cdots+xy^{n-2}+y^{n-1})+y^n(x-y)$. Then we get $(x-y)(x^n+x^{n-1}y+\cdots+xy^{n-1}+y^n) = (x-y)(x^n+x^{n-1}y+\cdots+xy^{n-1}+y^n)$. ☺

2. Prove that when $|x-y| \leqslant 1$, then $|x^n - y^n| \leqslant n(1+|x|)^{n-1}|x-y|$.

   Proof. Let $|x-y| \leqslant 1$. Then we have that $|x^n - y^n| = |(x-y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1})| \leqslant |x-y|(|x^{n-1}| + |x^{n-2}y| + \cdots + |xy^{n-2}| + |y^{n-1}|) \leqslant |x-y|(|x^{n-1}| + |x^{n-2}||y| + \cdots + |x||y^{n-2}| + |y^{n-1}|) \leqslant |x-y|(|x^{n-1}|+|x^{n-2}||y|+\cdots+|x||y^{n-2}|+|y^{n-1}|) \leqslant |x-y|(|x^{n-1}|+|x^{n-2}|+\cdots+|x|+1) \leqslant n(1+|x|)^{n-1}|x-y|$. ☺

3. Let $E = \{x \in \mathbb{R} : x^n > 3\}$ for a fixed n. Prove that E is open.

   Proof. Let $x \in E$. We want to show that there is an $r > 0$ such that $B(x,r) \subseteq E$. Take $r = \dfrac{x^n - 3}{n(1+|x|)^{n-1}}$ and take $y \in B(x,r)$. Then $|x-y| < r \Rightarrow |x^n| - |y^n| \leqslant |x^n - y^n| \leqslant n(1+|x|)^{n-1}|x-y| < n(1+|x|)^{n-1}r < x^n - 3$. Then $y^n \geqslant x^n - n(1+|x|)^{n-1}r > 3$. Thus, $y \in E$. Thus, $B(x,r) \subseteq E$. Thus, E is open. ☺

4. Consider the space $l^\infty([0,1]) = \{f : [0,1] \to \mathbb{R} \text{ bounded}\}$. Define $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$.
   Let $E = \{f \in l^\infty([0,1]) : f(x) > 0 \ \forall x \in [0,1]\}$. Prove that E is not open.

   Proof. Consider
   $$f(x) = \begin{cases} x & x \in [0,1) \\ 1 & x = 1 \end{cases}$$
   Then let $r > 0$ and consider $g(x) = f(x) \cdot \dfrac{r}{2}$. Then $g(x) \in B(f,r)$. But $g(x) \notin E$ because $g(1) = \dfrac{r}{2}$. Thus, $B(f,r) \nsubseteq E$. Thus, E is not open. ☺

1.17 Limits

Definition 1.17.1: Limits

Let (X, d_X) and (Y, d_Y) be metric spaces, $E \subseteq X$, $f : E \to Y$. Let $x_0 \in \operatorname{acc} E$.
Take $l \in Y$. l is the *limit* of f as $x \to x_0$. We write $\lim_{x\to x_0} f(x) = l$ if for every $\epsilon > 0$, there exists $\delta > 0$ such that $0 < d_X(x,x_0) < \delta \Rightarrow d_Y(f(x), l) < \epsilon$. We can also write it as $f(x) \to l$ as $x \to x_0$.

Note:

Even if $x_0 \in E$, you don't take in the definition for the limit.

Theorem 1.17.1

Let (X, d_X) and (Y, d_Y) be metric spaces, $E \subseteq X$, $f : E \to Y$, and $x_0 \in \operatorname{acc} E$. If $\lim_{x\to x_0} f(x)$ exists, then it is unique.

Proof. Assume that $\lim_{x\to x_0} f(x) = l$ and $\lim_{x\to x_0} f(x) = m$. Take $\epsilon = \frac{d_Y(l,m)}{2} > 0$. Then there exists $\delta_1 > 0$ such that $0 < d_X(x,x_0) < \delta_1 \Rightarrow d_Y(f(x), l) < \epsilon$. There also exists $\delta_2 > 0$ such that $0 < d_X(x,x_0) < \delta_2 \Rightarrow d_Y(f(x), m) < \epsilon$. Take $\delta = \min(\delta_1, \delta_2)$. Then $0 < d_X(x,x_0) < \delta \Rightarrow d_Y(f(x), l) < \epsilon$ and $d_Y(f(x), m) < \epsilon$. Then $d_Y(l,m) \leqslant d_Y(l, f(x)) + d_Y(f(x), m) < 2\epsilon = d_Y(l,m)$. This is a contradiction, so $l = m$. ☺

Example 1.17.1 ($\mathbb{R}^2$)

Take $(x_0, y_0) \in \mathbb{R}^2$ and $y_0 \neq 0$. Compute

$$\lim_{(x,y)\to(x_0,y_0)} \frac{x}{y}$$

We want to show that this is $\dfrac{x_0}{y_0}$. We have the set $E = \{(x,y) \in \mathbb{R}^2 : y \neq 0\}$. We also know that $(x_0, y_0) \in \operatorname{acc} E$. What we know if that $(x,y) \to (x_0, y_0)$: $|x - x_0|$ and $|y - y_0|$ are going to be small. Then

$$\begin{aligned}
\left| f(x,y) - \frac{x_0}{y_0} \right| &= \left| \frac{x}{y} - \frac{x_0}{y_0} \right| \\
&= \left| \frac{xy_0 - x_0y_0}{yy_0} \right| \\
&= \left| \frac{xy_0 - x_0y_0 + x_0y_0 - x_0y}{yy_0} \right| \\
&= \left| \frac{y_0(x - x_0) + x_0(y_0 - y)}{yy_0} \right| \\
&\leqslant \frac{|y_0||x - x_0| + |x_0||y_0 - y|}{|y||y_0|} \\
&= \frac{|x - x_0|}{|y|} + \frac{|x_0||y_0 - y|}{|y||y_0|}
\end{aligned}$$

Then we have $\delta < \frac{|y_0|}{2}$. If $|y - y_0| < \delta < \frac{y_0}{2}$, then we get $|y| \geqslant \frac{|y_0|}{2} \Rightarrow \frac{1}{|y|} \leqslant \frac{2}{|y_0|}$.

$$\frac{|x - x_0|}{|y|} + \frac{|x_0||y_0 - y|}{|y||y_0|} \leqslant \frac{2|x - x_0|}{|y_0|} + \frac{2|x_0||y_0 - y|}{|y_0|^2}$$

Take $\delta = \min\left\{\epsilon, \frac{|y_0|}{2}\right\} > 0$. Then $0 < \|(x,y) - (x_0,y_0)\| < \delta$.

$$\begin{aligned}
&|x - x_0| = \sqrt{(x - x_0)^2} \leqslant \sqrt{(x - x_0)^2 + (y - y_0)^2} \\
&|y - y_0| \leqslant \delta
\end{aligned}$$

So,

$$\left| f(x,y) - \frac{x_0}{y_0} \right| < \epsilon \left(\frac{2}{|y_0} + \frac{2}{|y_0|^2} \right)$$

Say you can prove that for every $\epsilon > 0$, $\exists \delta > 0$ such that

$$d(f(x), l) < \epsilon|\log(\epsilon)| \text{ for all } x \in E \text{ such that } 0 < d(x, x_0) < \delta$$

For every $\eta > 0$ ("my epsilon"), since $\lim_{\epsilon\to0^+} \epsilon|\log(\epsilon)| = 0$, $\exists \delta_1 > 0$ such that $\epsilon|\log(\epsilon)| < \eta$ for all $0 < \epsilon < \delta_1$.

So given $\eta > 0$, take $0 < \epsilon < \delta_1$. Find η from $d(f(x), l) < \epsilon|\log(\epsilon)| < \eta$ for all $x \in E$ such that $0 < d(x, x_0) < \delta$. This means that

$$d_Y(f(x), l) < \epsilon|\log(\epsilon)| < \eta$$

for all $x \in E$, $0 < d(x, x_0) < \delta$. Thus, $\lim_{x\to x_0} f(x) = l$.

1.18 Limits Continued

Definition 1.18.1: Restriction

Assume that $\lim_{x\to x_0} f(x) = l$ exists. Let $F \subseteq E$ such that $x_0 \in \operatorname{acc} F$. The function $f : F \to Y$ is called the *restriction* of f to F. It is denoted as $f|_F$.

Note:

If $\lim_{x\to x_0} f(x) = l$, then $\lim_{x\to x_0} f|_F(x) = l$.

So to prove that the limit does not exist, you can conjure up two restrictions of the function and show that the limits are different.

Example 1.18.1 (Limits that don't exist)

Consider $\lim_{x\to 0} \sin\left(\frac{1}{x}\right)$. We can find two restrictions of the function and show that the limits are different.

- $\frac{1}{x} = 2\pi n + \frac{\pi}{2}$
- $x_n = \frac{1}{2\pi n + \frac{\pi}{2}}$

So we have:

- $\sin x_n = \sin\left(\frac{\pi}{2} + 2\pi n\right) = 1$
- $\sin x_n = \sin\left(\frac{1}{2\pi n}\right) = 0$

Thus, the limit does not exist.

Exercise 1.18.1 TODO in Recitation

- $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ (no)
- $\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2}$ (yes, 0)
- $\lim_{(x,y)\to(0,0)} \frac{x^{10000000000}y}{y-\sin(x)}$ (no)

Now we talk about the composition of limits.

Example 1.18.2

Consider

$$g(y) = \begin{cases} 1 & y \neq 0 \\ 2 & y = 0 \end{cases}.$$

The limit of $g(y)$ as $y \to 0$ is 1. Now consider $f(x) = 0$. The limit of $f(x)$ as $x \to x_0$ is 0. Now consider $g(f(x))$. The limit of $g(f(x))$ as $x \to x_0$ is 2.

Theorem 1.18.1 Composition of Limits

Let (X, d_X), (Y, d_Y), and (Z, d_Z) be metric spaces, $E \subseteq X$, $F \subseteq Y$, $f : E \to F$, $g : F \to Z$, and $x_0 \in \operatorname{acc} E$. Assume there exists $\lim_{x\to x_0} f(x) = l \in Y$. Assume $l \in \operatorname{acc} F$ and that there is $\lim_{y\to l} g(y) = L \in Z$. Assume that either $f(x) \neq l$ for all $x \in E$ or $l \in F$ and $g(l) = L$. Then there is $\lim_{x\to x_0} g(f(x)) = L$.

Proof. Since $\lim_{y\to l} g(y) = L$, there for every $\epsilon > 0$, there exists $\delta > 0$ such that $d_Z(g(y), L) < \epsilon$ for all $y \in F$ with $0 < d_Y(y, l) < \delta$. We would like to take $y = f(x)$. Use δ as "my epsilon" for the definition of the limit of $f(x)$. Then to find $\eta > 0$ such that $d_X(f(x), l) < \delta$ for all $x \in E$ with $0 < d_X(x, x_0) < \eta$. Now we split into cases:

- Assume $f(x) \neq l$ for all $x \in E$. Then $0 < d_Y(f(x), l)$ so we can take $y = f(x)$ to get $d_z(g(f(x)), L) < \epsilon$ for all $x \in E$ with $0 < d_X(x, x_0) < \eta$. This means that there exists $\lim_{x \to x_0} g(f(x)) = L$.
- Assume $l \in F$ and $g(l) = L$. If $f(x) = l$, then $d_Z(g(f(x)), L) = d_Z(g(l), L) = 0$ for all $x \in E$ with $0 < d_X(x, x_0) < \eta$. If $f(x) \neq l$, then take $y = f(x)$ to get $0 < d_Y(f(x), l)$ so we can take $y = f(x)$ to get $d_z(g(f(x)), L) < \epsilon$ for all $x \in E$ with $0 < d_X(x, x_0) < \eta$. This means that there exists $\lim_{x \to x_0} g(f(x)) = L$. This means that there exists $\lim_{x \to x_0} g(f(x)) = L$.

😄

Corollary 1.18.1 Limits of the Sum/Products/Quotients

Let (X, d) be a metric space and $E \in X$. Then take $f : E \to \mathbb{R}$ and $g : E \to \mathbb{R}$. Let $x_0 \in X$ and $x_0 \in \operatorname{acc} E$. Assume $\lim_{x \to x_0} f(x) = l$ and $\lim_{x \to x_0} g(x) = m$. Then we have the following results:

- $\lim_{x \to x_0} (f + g)(x) = l + m$.
- $\lim_{x \to x_0} (f \cdot g)(x) = l \cdot m$.
- $\lim_{x \to x_0} \frac{f}{g}(x) = \frac{l}{m}$.

Proof. We can use the composition of limits to prove this. We'll just proceed with the quotient case. Consider $x \to (f(x), g(x))$. Then consider the function that takes $(s, t) \to \frac{s}{t}$ and call it h. THen we have $\frac{f(x)}{g(x)} = h(f(x), g(x))$. We then have $\lim_{x \to x_0} (f(x), g(x)) = (l, m)$ and $\lim_{(s,t) \to (l,m)} h(s, t) = \frac{l}{m} = h(l, m)$. So now we can use the composition of limits to get $\lim_{x \to x_0} \frac{f}{g}(x) = \frac{l}{m}$.

The other two cases are similar. For products, you need to show that the limit as $(x, y) \to (x_0, y_0)$ of xy is $x_0 y_0$ and similarly for sum. 😄

1.19 Squeeze Theorem

Theorem 1.19.1 Squeeze Theorem

Let (X, d_X) be a metric space, $E \subseteq X$, $f : E \to \mathbb{R}$, $g : E \to \mathbb{R}$, and $h : E \to \mathbb{R}$. Let $x_0 \in \operatorname{acc} E$ and have $f \leqslant g \leqslant h$. Assume that $\lim_{x \to x_0} f(x) = l = \lim_{x \to x_0} h(x)$. Then $\lim_{x \to x_0} g(x) = l$.

Proof. Assume $\lim_{x \to x_0} f(x) = l = \lim_{x \to x_0} h(x)$. Then for every $\epsilon > 0$, there exists $\delta_1 > 0$ such that $0 < d_X(x, x_0) < \delta_1 \implies |f(x) - l| < \epsilon$ and $0 < d_X(x, x_0) < \delta_2 \implies |h(x) - l| < \epsilon$. Take $\delta = \min(\delta_1, \delta_2)$ and $x \in E$ with $0 < d(x, x_0) < \delta$. Then $l - \epsilon < f(x) < g(x) < h(x) < l + \epsilon$. Then $|g(x) - l| < \epsilon$ for all $x \in E$ with $0 < d_X(x, x_0) < \delta$. Thus, $\lim_{x \to x_0} g(x) = l$. 😄

Example 1.19.1

$$\lim_{x \to 0} |x|^a \sin \frac{1}{x} = 0$$

for all $a > 0$.

Let $Q > 0$. Then $O \leqslant ||x|^Q \sin \frac{1}{x}| \leqslant |x|^Q$. Since both sides tend to 0 as $x \to 0$, then the middle does as well.

Definition 1.19.1: Increasing

$f : E \to \mathbb{R}$ is *increasing* if $f(x) \leqslant f(y)$ for all $x \leqslant y$. It is strictly increasing if $f(x) < f(y)$ for all $x < y$.

Definition 1.19.2: Decreasing

$f : E \to \mathbb{R}$ is *decreasing* if $f(x) \geqslant f(y)$ for all $x \leqslant y$. It is strictly decreasing if $f(x) > f(y)$ for all $x < y$.

Definition 1.19.3: Divergent

Let (X, d_x) be a metric space with $E \subseteq X$, $x_0 \in \operatorname{acc} E$, and $f : E \to \mathbb{R}$. We say that f diverges to $+\infty$ as $x \to x_0$ if for every $M > 0 \in \mathbb{R}$, there exists $\delta > 0$ such that $f(x) > M$ for all $x \in E$ with $0 < d_X(x, x_0) < \delta$. We say that f diverges to $-\infty$ as $x \to x_0$ if for every $M < 0 \in \mathbb{R}$, there exists $\delta > 0$ such that $f(x) < M$ for all $x \in E$ with $0 < d_X(x, x_0) < \delta$.

Theorem 1.19.2

Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$ be increasing. Let $x_0 \in \mathbb{R}$. Assume x_0 is an accumulation point of $E \cap (-\infty, x_0)$. Then there is

$$\lim_{x \to {x_0}^-} f(x) = \sup_{E \cap (-\infty, x_0)} f(x)$$

Now if x_0 is an accumulation point of $E \cap (x_0, \infty)$, then there is

$$\lim_{x \to {x_0}^+} f(x) = \inf_{E \cap (x_0, \infty)} f(x)$$

Proof.

- Case 1: Assume f is bounded form above on $E \cap (-\infty, x_0)$. Let $l = \sup_{E \cap (-\infty, x_0)} f(x)$. Then for every $\epsilon > 0$, there exists $x_1 \in E \cap (-\infty, x_0)$ such that $l - \epsilon < f(x_1) \leqslant l$. Then for every $\epsilon > 0$, there exists $\delta > 0$ such that $l - \epsilon < f(x_1) \leqslant l$. Take $\delta = x_0 - x_1 > 0$. Let $x \in E$ with $x_0 - \delta < x < x_0$. Since f is increasing, we have $l - \epsilon < f(x_1) \leqslant f(x) \leqslant l < l + \epsilon$. Thus, $\lim_{x \to {x_0}^-} f(x) = l$.
- Case 2: If f is not bounded from above, then for every $M > 0$, there exists $x_1 \in E \cap (-\infty, x_0)$ such that $f(x_1) > M$. Let $\delta = x_0 - x_1 > 0$. Then for every $x \in E$ with $x_1 = x_0 - \delta < x < x_0$, we have $f(x) \geqslant f(x_1) > M$. Thus, $\lim_{x \to {x_0}^-} f(x) = +\infty$.

The other case is similar. 😊

Definition 1.19.4: Infinite Sum

Let X be a set and take $f : X \to [0, \infty]$. The *infinite sum* is defined as:

$$\sum_{x \in X} f(x) = \sup \left\{ \sum_{x \in F} f(x) : F \subseteq X \text{ finite} \right\}.$$

Lenma 1.19.1

Let X be nonempty with $f : X \to [0, \infty]$. Assume that $\sum_{x \in X} f(x) < \infty$. Then $\{x \in X : f(x) > 0\}$ is countable.

Proof. Take $n \in \mathbb{N}$ and define $X_n = \{x \in X : f(x) \geqslant \frac{1}{n}\}$. Let $E \subseteq X_n$ be finite. Then $\frac{1}{n}|E| < \sum_{x \in E} f(x) \leqslant M$. Then $|E| < nM$. Thus, X_n is countable. Then $\bigcup_{n \in \mathbb{N}} X_n = \{x \in X : f(x) > 0\}$ is countable. 😊

1.20 2/15 - LIMIT !!!!!!!!!

Example 1.20.1

- Let $f(x) = \frac{xy}{x^2y^2}$. Find $\lim_{(x,y)\to(0,0)} f(x,y)$. If we take the restriction $y = mx$, we see that the limit depends on m which is a contradiction.
- Let $f(x,y) = \frac{x^2y}{x^2+y^2}$. Find $\lim_{(x,y)\to(0,0)} f(x,y)$. We can use polar coordinates to show that this is 0.

1.21 This Theorem

Theorem 1.21.1

Take $I \subseteq \mathbb{R}$ to be an interval with $f : I \to \mathbb{R}$ increasing. Then for all but countably many $x_0 \in I$, there is $\lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^+} f(x) = f(x_0)$.

Proof. Let $I = [a,b]$. For every $x \in (a,b)$, there exists

$$\lim_{y\to x^+} =: f_+(x), \quad \lim_{y\to x^-} =: f_-(x).$$

Let $S(x) = f_+(x) - f_-(x) \geqslant 0$, which is the jump of f at x. Then we have that $\lim_{y\to x} f(y) = f(x) \iff S(x) = 0$. Let $J \in [a,b]$ be any finite subset, and write

$$J = \{x_1, \dots, x_k\}, \quad \text{where} x_1 < \cdots < x_k.$$

Since f is increasing, we have that

$$f(a) \leqslant f_-(x_1) \leqslant f_+(x_1) \leqslant f_-(x_2) \leqslant f_+(x_2) \leqslant \cdots \leqslant f_-(x_k) \leqslant f_+(x_k) \leqslant f(b).$$

So,

$$\sum_{x\in J} S(x) = \sum_{x\in J} f_+(x) - f_-(x) \leqslant f(b) - f(a),$$

which implies that $\sum_{x\in(a,b)} S(x) \leqslant f(b) - f(a)$. It follows that the amount of discontinuities is countable. 😃

Chapter 2

Definition 2.0.1: Series

Given a normed space X and a sequence $\{x_n\}_n$, of vectors in X, we call the nth-partial sum the vector $s_n = \sum_{k=1}^{n} x_k$. The sequence $\{s_n\}_n$ of partial sums is called infinite series or *series* and is denoted $\sum_{n=1}^{\infty} x_n$. If there exists $\lim_{n\to\infty} s_n = s \in X$, then we say that the series $\sum_{n=1}^{\infty} x_n$ *converges* to s and s is called the *sum* of the series. If the limit does not converge, we say the series *oscillates*.

2.1 More Series

Theorem 2.1.1

Let X be a normed space. Consdier the series $\sum_{n=1}^{\infty} x_n$. If the series converges, then $\lim_{n\to\infty} x_n = 0$.

Proof. We know by the hypothesis that $s_n = x_1 + \cdots + x_n$. Also $s_n \to s$ as $n \to \infty$. As such, we can write $x_n = s_n - s_{n-1}$ where both values on the RHS tend to s, meaning that x_n tends to 0 as $n \to \infty$. 😄

Note:

The above theorem is often useful to negate. In the exercises, we can also use teh fact that i $\lim_{n\to\infty}$ does not exists or does not equal 0, then $\sum_{n=1}^{\infty} x_n$ cannot converge.
Also important to note that this series is very much one directional. For example, consider the following sums:

$$\sum_{n=1}^{\infty} \frac{1}{n} \text{ diverges}$$
$$\sum_{n=1}^{\infty} \frac{1}{n^2} \text{ converges}$$

However, both values here tend to 0 as $n \to \infty$.

Example 2.1.1 (Geometric Series)

Consider $\sum_{n=1}^{\infty} x^n$. We know that $\lim_{n\to\infty} x^n = 0$ iff $|x| < 1$. So if $|x| \geqslant 1$, then the series does not converge. The theorem above does not help us for the $|x| < 1$ case. So let's compute the partial sum:

$$\begin{aligned} s_n &= \sum_{k=1}^{n} x^k \\ &= \frac{x^{n+1} - x}{x - 1} \end{aligned}$$

So we have that $\lim_{n\to\infty} s_n = \frac{x}{1-x}$ for $|x| < 1$.

Example 2.1.2

Consider $X = \ell^\infty(E) = \{f : E \to \mathbb{R} \text{ bounded}\}$ for $E \subseteq \mathbb{R}$. The norm here is the supremum norm. Consider the series $\sum_{n=1}^\infty f_n(x)$ of random functions in X. We need to check that

$$\sup_{x \in E} |f_n(x)| \to 0 \text{ as } n \to \infty$$

for the series to converge. If the limit $\neq 0$, then the series does not converge.

Example 2.1.3 (Combining the above)

Let our space be $\ell^\infty((-1,1))$ and consider the series $\sum_{n=1}^\infty f_n(x)$ where $f_n(x) = x^n$. We know that $\sup_{x\in(-1,1)} |x^n| = 1$ for all n. Thus, the series does not converge.

Theorem 2.1.2

Consider a series of nonnegative terms $\sum_{n=1}^\infty x_n$ in $\mathbb{R}$. Either the series converges or diverges to $+\infty$.

Proof. We know that $s_{m+1} \geqslant s_m$ for all m and that these values are increasing, so $\lim_{m\to\infty} s_n = \sup_n s_n \in [0, \infty]$. 😃

Theorem 2.1.3 Comparison Test

Let $\sum_{n=1}^\infty x_n$ and $\sum_{n=1}^\infty y_n$ be series of nonnegative terms in $\mathbb{R}$. Assume that $0 \leqslant x_n \leqslant y_n$ for all $n \geqslant N$ for some N. If $\sum_{n=1}^\infty y_n$ converges, then $\sum_{n=1}^\infty x_n$ converges. If $\sum_{n=1}^\infty x_n$ diverges, then $\sum_{n=1}^\infty y_n$ diverges.

Proof. Consider the first case. So let $s_n = \sum x_n$ and $t_n = \sum y_n$. Since y_n converges, $\lim t_n = T$ exists, so t_n is bounded by T. Then for all $n \geqslant N$:

$$\begin{aligned} s_n &:= x_1 + \cdots + x_{N-1} + x_N + \cdots + x_n \\ &\leqslant x_1 + \cdots x_{N-1} + y_N + \cdots + y_n \\ &\leqslant x_1 + \cdots + x_{N-1} + T \end{aligned}$$

Hence, $\{s_n\}$ is bounded and increasing, so it converges.

For the second case, we have that $s_n \to \infty$. So since

$$s_n \leqslant (x_1 + \cdots + x_{N-1}) + t_n$$

we have that $t_n \to \infty$ as $n \to \infty$. 😃

Example 2.1.4 (Examples)

1. $\sum_{n=1}^\infty \left(\frac{1+\cos n}{3}\right)^n$. We know that $\lim_{n\to\infty} \left(\frac{1+\cos n}{3}\right)^n = 0$ so the series isn't divergent.

   We will compare it to $0 \leqslant \left(\frac{1+\cos n}{3}\right)^n \leqslant \left(\frac{2}{3}\right)^n$. We know that $\sum_{n=1}^\infty \left(\frac{2}{3}\right)^n$ converges, so the series $\sum_{n=1}^\infty \left(\frac{1+\cos n}{3}\right)^n$ converges by the comparison test.

2. $\sum_{n=1}^\infty 1 - \cos\frac{1}{3^n}$. We know that $\lim_{n\to\infty} 1 - \cos\frac{1}{3^n} = 0$ so the series isn't divergent.

   We have that $\lim_{t\to 0} \frac{1-\cos t}{t} = 0$. Take $\epsilon = 1$ and find $\delta > 0$ such that $\left|\frac{1-\cos t}{t} - 0\right| < 1$ for all $0 < |t| < \delta$. We know that $-1 < \frac{1-\cos t}{t} < 1$. Now take $1 - \cos\frac{1}{3^n} < \frac{1}{3^n}$ for all n such that $\frac{1}{3^n} < \delta$.

   So, $1 - \cos\frac{1}{3^n} < \frac{1}{3^n}$ for all $n > N$. THe RHS converges so by comparison test, the LHS converges.

3. $\sum_{n=1}^\infty \frac{\sin\frac{1}{n^3}}{\log\left(1+\frac{1}{n}\right)}\left(e^{1/m} - 1\right)$. We know that $\sin\frac{1}{n^3} \sim \frac{1}{n^3}$, $\log\left(1+\frac{1}{n}\right) \sim \frac{1}{n}$ and $e^{1/n} - 1 \sim \frac{1}{n}$. So we have

that $\frac{\sin\frac{1}{n^3}}{\log\left(1+\frac{1}{n}\right)}\left(e^{1/m}-1\right) \sim \frac{1}{n^3}\cdot\frac{1}{n}\cdot\frac{1}{n} = \frac{1}{n^3}$.

4. Prove by induction that $n! > 2^n$ when $n \geqslant 4$. This implies that $\frac{1}{n!} \leqslant \frac{1}{2^n}$ for $n \geqslant 4$. Since $\sum_{n=1}^{\infty}\frac{1}{2^n} < \infty$, comparison test tells us that $\sum_{n=0}^{\infty}\frac{1}{n!} < \infty$. The sum of the series is called

$$e := \sum_{n=0}^{\infty}\frac{1}{n!}.$$

Theorem 2.1.4 Root Test

Let $x_n \geqslant 0$.

1. If $\limsup_{n\to\infty}\sqrt[n]{x_n} < 1$, then $\sum_{n=1}^{\infty} x_n < \infty$.
2. If $\limsup_{n\to\infty}\sqrt[n]{x_n} > 1$, then $\sum_{n=1}^{\infty} x_n = \infty$.
3. If $\limsup_{n\to\infty}\sqrt[n]{x_n} = 1$, then the test is inconclusive.

Proof.

1. Let $\ell = \limsup_{n\to\infty}\sqrt[n]{x^n}$. Assume $\ell < 1$. Find $\epsilon > 0$ such that $\ell + \epsilon < 1$. Then there exists N such that $\sqrt[n]{x_n} < \ell + \epsilon$ for all $n \geqslant N$. Then $\sqrt[n]{x_n} > \ell - \epsilon$ for infinitely many n. Taking the first inequality, we have that $x_n < (\ell+\epsilon)^n$ for all $n \geqslant N$. By the comparison test, we have that $\sum_{n=1}^{\infty}(\ell+\epsilon)^n$ converges, so $\sum_{n=1}^{\infty} x_n$ converges.
2. Assume $\ell > 1$. For $\epsilon > 0$ small, $(l-\epsilon) > 1$. So $x_n \geqslant (l-\epsilon)^n$ for infinitely many n. Since the RHS goes to infinity, a subsequence also goes to ∞, so $\lim_{n\to\infty} x_n \neq 0$ so the series cannot converge.

😄

Example 2.1.5 (Inconclusive Root Test)

Consider the series

$$\sum_{n=1}^{\infty}\frac{1}{n}.$$

Then, we have:

$$\sqrt[n]{\frac{1}{n}} = \left(\frac{1}{n}\right)^{\frac{1}{n}} = e^{\log\left(\frac{1}{n}\right)^{\frac{1}{n}}} = e^{\frac{\log\left(\frac{1}{n}\right)}{n}}$$

The exponent goes to 0, so the limit is 1.
Now consider the series

$$\sum_{n=1}^{\infty}\frac{1}{n^2}.$$

Then, we have:

$$\sqrt[n]{\frac{1}{n^2}} = \left(\frac{1}{n^2}\right)^{\frac{1}{n}} = e^{\log\left(\frac{1}{n^2}\right)^{\frac{1}{n}}} = e^{\frac{\log\left(\frac{1}{n^2}\right)}{n}}$$

The exponent goes to 0, so the limit is 1.
We see that the first series diverges and the second series converges, so the root test is inconclusive when the lim sup is 1.

Example 2.1.6 (More Root Test)

Consider the series

$$\sum_{n=1}^{\infty} \frac{n^2+1}{2^n}.$$

We have that $\lim_{n\to\infty} \frac{n^2+1}{2^n} = 0$, so the series isn't divergent. Consider the root test now. We have that

$$\sqrt[n]{\frac{n^2+1}{2^n}} = \frac{\sqrt[n]{n^2+1}}{2} = \frac{1}{2}e^{\frac{1}{n}\log(n^2+1)} \longrightarrow \frac{1}{2}e^0 = \frac{1}{2} < 1.$$

So the series converges.

Exercise 2.1.1

Let $x_n \geqslant 0$. Prove that

$$\liminf_{n\to\infty} \frac{x_{n+1}}{x_n} \leqslant \liminf_{n\to\infty} \sqrt[n]{x_n} \leqslant \limsup_{n\to\infty} \sqrt[n]{x_n} \leqslant \limsup_{n\to\infty} \frac{x_{n+1}}{x_n}.$$

Find an example where the last inequality is strict:

$$\limsup_{n\to\infty} \sqrt[n]{x_n} < \limsup_{n\to\infty} \frac{x_{n+1}}{x_n}.$$

An example is

$$x_n = \begin{cases} 1 & n \text{ odd} \\ 2 & n \text{ even} \end{cases}.$$

Definition 2.1.1: Ratio Test

Let $x_n > 0$.

1. If $\limsup_{n\to\infty} \frac{x_{n+1}}{x_n} < 1$, then the series converges.
2. If $\liminf_{n\to\infty} \frac{x_{n+1}}{x_n} > 1$, then the series diverges.

Proof. 1. This is a one-line proof. If the limit is less than 1, then the series converges by the root test by the exercise above. That is ,$\limsup_{n\to\infty} \frac{x_{n+1}}{x_n} < 1 \implies \limsup_{n\to\infty} \sqrt[n]{x_n} < 1$.

2. This is also a one-line proof. If the limit is greater than 1, then the series diverges by the root test by the exercise above. That is, $\liminf_{n\to\infty} \frac{x_{n+1}}{x_n} > 1 \implies \limsup_{n\to\infty} \sqrt[n]{x_n} > 1$.

😊

Example 2.1.7

Consider the sequence:

$$x_n = \begin{cases} \frac{1}{2^n} & n \text{ odd} \\ \frac{1}{3^n} & n \text{ even} \end{cases}.$$

We have that $\limsup_{n\to\infty} \frac{x_{n+1}}{x_n} = \infty$ and that $\liminf_{n\to\infty} \frac{x_{n+1}}{x_n} = 0$. As such, we cannot apply the ratio test in this case. We try the root test instead.
We have that $\limsup_{n\to\infty} \sqrt[n]{x_n} = \frac{1}{2}$ and that $\liminf_{n\to\infty} \sqrt[n]{x_n} = \frac{1}{3}$. Because of the first one, we know that the series converges.

Definition 2.1.2: Integral Test

Let $f : [1, \infty) \to [0, \infty)$ be a decreasing function in $[N, \infty)$. Then the series $\sum_{n=1}^{\infty} f(n)$ converges if and only if $\lim_{L\to\infty} \int_1^L f(x)\,dx$ converges.

Proof. We start with the forward direction. Consider $\int_N^\ell f(x)\,dx$ for integer ℓ. We have:

$$\int_N^\ell f(x)\,dx = \sum_{n=N}^{\ell-1} \int_n^{n+1} f(x)\,dx$$

If $n \leqslant x \leqslant n+1$ and f decreasing, we have that $f(n+1) \leqslant f(x) \leqslant f(n)$. For each n, we have:

$$\sum_{n=N}^{\ell-1} \int_n^{n+1} f(x)\,dx \leqslant \sum_{n=N}^{\ell-1} f(n)$$

If $\lim_{\ell\to\infty} \int_N^\ell f(x)\,dx$ diverges, then $\sum_{n=N}^{\infty} f(n)$ diverges. So,

$$\begin{aligned} \int_N^\ell f(x)\,dx &= \sum_{n=N}^{\ell-1} \int_n^{n+1} f(x)\,dx \\ &\geqslant \sum_{n=N}^{\ell-1} f(n+1) \end{aligned}$$

So if $\lim_{\ell\to\infty} \int_N^\ell f(x)\,dx$ converges, then $\sum_{n=N}^{\infty} f(n)$ converges since it is less than or equal to. 😊

2.2 More Series

Example 2.2.1 (Integral Test)

Consider:

- $\sum_{n=1}^{\infty} \frac{1}{n^a}$ for $a > 0$ First we check that $\lim_{n\to infty} \frac{1}{n^a} = 0$. This is indeed true. Now we define $f(x) = \frac{1}{x^a}$ for $x > 0$. This function is decreasing, so we can use the integral test. We have:

$$\begin{aligned} \int_1^\infty \frac{1}{x^a}\,dx &= \lim_{L\to\infty} \int_1^L \frac{1}{x^a}\,dx \\ &= \lim_{L\to\infty} \left[\frac{x^{1-a}}{1-a}\right]_1^L \\ &= \lim_{L\to\infty} \left[\frac{L^{1-a}}{1-a} - \frac{1}{1-a}\right] \\ &= \frac{-1}{a-1} \quad \text{if } a > 1 & &= \infty \quad \text{if } a < 1 \end{aligned}$$

If $a = 1$, then we have:

$$\begin{aligned} \int_1^\infty \frac{1}{x}\,dx &= \lim_{L\to\infty} \int_1^L \frac{1}{x}\,dx \\ &= \lim_{L\to\infty} [\log(x)]_1^L \\ &= \lim_{L\to\infty} [\log(L) - \log(1)] \\ &= \infty \end{aligned}$$

So, the series converges if $a > 1$ and diverges if $a \leqslant 1$.

- $\sum_{n=2}^{\infty} \frac{1}{n^a \log n}$. We have that $\lim_{n\to\infty} \frac{1}{n^a \log n} = 0$. We define $f(x) = \frac{1}{x^a \log x}$ for $x > 2$. This function is decreasing, so we can use the integral test. We have:

$$\begin{aligned}
\int_2^{\infty} \frac{1}{x^a \log x}\,dx &= \lim_{L\to\infty} \int_2^{L} \frac{1}{x^a \log x}\,dx \\
&= \lim_{L\to\infty} \left[\frac{\log(\log(x))}{1-a}\right]_2^L \\
&= \lim_{L\to\infty} \left[\frac{\log(\log(L))}{1-a} - \frac{\log(\log(2))}{1-a}\right] \\
&= \infty \quad \text{if } a > 1
\end{aligned}$$

Definition 2.2.1: Alternating Series

Let $\{a_n\}_n$ be a sequence of positive numbers. The series $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ is called *Alternating*

Theorem 2.2.1 Leibniz Test

Consider $\sum_{n=1}^{\infty}(-1)^n a_n$ where $a_n \geqslant 0$. If $\{a_n\}_n$ is decreasing and $\lim_{n\to\infty} a_n = 0$, then the series converges and $|S - s_n| \leqslant a_{n+1}$ for all n.

Proof. Write

$$\begin{aligned}
s_{2n+1} &= -a_1 + (a_2 - a_3) + (a_4 - a_5) + \cdots + (a_{2n} - a_{2n+1}) \\
&= -(a_1 - a_2) - (a_3 - a_4) - \cdots - (a_{2n-1} - a_{2n}) - a_{2n+1}
\end{aligned}$$

Since a_n is decreasing, we have that $a_i - a_{i-1} \geqslant 0$. And from the first equality, we get that $s_{2n+1} \leqslant s_{2n+3}$, meaning that s_{2n+1} is an increasing sequence. But from the second equality, we get that $s_{2n+1} \leqslant -a_{2n+1} \leqslant 0$. So, there exists:

$$\lim_{n\to\infty} s_{2n+1} = \sup_n s_{2n+1} = S \in (-\infty, 0]$$

Since $s_{2n+1} = s_{2n} + a_{2n+1}$ and $\lim_{n\to\infty} a_n = 0$, we have that $\lim_{n\to\infty} s_{2n} = S$. So, the series converges.

Moreover, we have that:

$$s_{2n} = -(a_1 - a_2) - (a_3 - a_4) - \cdots - (a_{2n-1} - a_{2n}),$$

which implies that $s_{2n} \geqslant s_{2n+2}$, meaning that s_{2n} is a decreasing sequence. So $\inf_n s_{2n} = S \in (-\infty, 0]$. Therefore, $s_{2n+1} \leqslant S \leqslant s_{2n}$. It follows that

$$\begin{aligned}
|S - s_{2n}| &= s_{2n} - S \leqslant s_{2n} - s_{2n+1} = a_{2n+1} \\
|S - s_{2n+1}| &= s_{2n+1} - S \leqslant s_{2n+2} - s_{2n+1} = a_{2n+1}
\end{aligned}$$

as desired. 😄

Corollary 2.2.1

Also if an alternating series converges, then the remainder $R_n = |S - S_n|$ satisfies $0 \leqslant R_n \leqslant a_{2n+1}$.

Proof. We have that $S_{2n+1} \leqslant S$ and that S_{2n} is decreasing. So $S = \inf_{n\in\mathbb{N}} S_{2n}$, so $S \leqslant S_{2n}$. This yields $|S - S_{2n}| = S_{2n} - S \leqslant S_{2n} - S_{2n+1} = a_{2n+1}$. For the other case, we have $|S - S_{2n+1}| + S - S_{2n+1} \leqslant S_{2n+2} - S_{2n+1} \leqslant a_{2n+2}$. 😄

Example 2.2.2

Consider the sequence

$$\sum_{n=1}^{\infty}(-1)^n \frac{n\log n}{1+n^2}.$$

We consider $\lim_{n\to\infty} \frac{n\log n}{1+n^2}$. This is similar to the limit of $\frac{\log n}{n}$, which diverges. So by comparison test, our limit diverges. So we have:

$$\begin{aligned} f(x) &= \frac{x\log x}{1+x^2} \\ f'(x) &= \frac{\log x + 1}{x^2+1} - \frac{2x^2\log x}{(x^2+1)^2} \end{aligned}$$

We can somehow show that $f'(x) < 0$ for all $x \geqslant N$ for some N

Theorem 2.2.2

Let $E, \ell^\infty(E) = \{f : E \longrightarrow \mathbb{R} \text{ bounded}\}$. Let $\{f_n\}_n \subset \ell^\infty(E)$ and $f \in \ell^\infty(E)$.

1. If $\sum_{n=1}^\infty \sup_{x\in E} |f_n(x)| < \infty$, then $\sum_{n=1}^\infty f_n(x)$ converges uniformly in E.
2. If $\sum_{n=1}^\infty f_n(x)$ converges uniformly to f, then $\lim_{n\to\infty} \sup_{x\in E} |f_n(x)| = 0$.

Example 2.2.3

1. Consider the series $\sum_{n=1}^\infty \frac{e^{nx}}{n}$, for $x \in \mathbb{R}$.
   $\frac{e^{nx}}{n} > 0 \, \forall x \in \mathbb{R}$.
   - For $x = 0$, we have $\sum_{n=1}^\infty \frac{1}{n} = \infty$.
   - For $x > 0$, we have $\lim_{n\to\infty} \frac{e^{nx}}{n} = \infty$.
   - For $x < 0$, $\left(\frac{e^{nx}}{n}^{1/n}\right) = \frac{e^x}{n^{1/n}} \longrightarrow e^x$ as $n \longrightarrow infty$.

   This shows that there is pointwise convergence when $x < 0$. So if we want to determine a subset were there is uniform convergence, then we have to consider only $x \in (-\infty, 0)$.

   So consider $E = (-\infty, -\epsilon)$ for some $\epsilon > 0$. Consider the sequence of functions defined as

   $$f_n(x) = \frac{e^{nx}}{n}$$

   for $x \in E$. Then we have:

   $$f_n'(x) = e^{nx} > 0 \implies \sup_{x\in E} |f_n(x)| = \frac{e^{-n\epsilon}}{n}$$

   So, by our theorem, we have that $\sum_{n=1}^\infty \frac{e^{nx}}{n}$ converges uniformly in E.

2. Consider $\sum_{n=1}^\infty \frac{x^{2n}}{\sqrt[3]{n}} \log\left(1 + \frac{x^2}{\sqrt[3]{n}}\right)$, for $x \in \mathbb{R}$.
   - for $x = 0$, it converges to 0.
   - for $|x| > 1$, we have that

   $$\lim_{n\to\infty} \frac{x^{2n}}{\sqrt[3]{n}} \log\left(1 + \frac{x^2}{\sqrt[3]{n}}\right) = \lim_{n\to\infty} \frac{x^{2n+2}}{n^{2/3}} \frac{\log\left(1 + \frac{x^2}{\sqrt[3]{n}}\right)}{x^2/\sqrt[3]{n}} = \infty$$

   - for $|x| < 1$, $\lim_{n\to\infty} \frac{x^{2n}}{\sqrt[3]{n}} \log\left(1 + \frac{x^2}{\sqrt[3]{n}}\right) = 0$

3. Consider $\sum_{n=1}^\infty \frac{x^n}{n}$ for $x \geqslant 0$.
   - for $x = 0$, there is pointwise convergence .
   - for $x \geqslant 1$, there is no pointwise convergence.
   - For $x \in (0, 1)$, we have that $\lim_{n\to\infty} \frac{x^n}{n} = 0$.

Theorem 2.2.3

Take some $x_n \in \mathbb{R}$ and consider the series $\sum_{n=1}^{\infty}$. If $\sum_{n\to\infty} |x_n|$ converges, then $\sum_{n=1}^{\infty} x_n$ converges.

Note:

The converse isn't true. Consider the alternating version of the harmonic series.

Definition 2.2.2

Let $t \in \mathbb{R}$. We define:

$$t^+ = \max(t, 0)$$
$$t^- = \max(-t, 0)$$

From these, we derive:

$$|t| = t^+ + t^- \quad \text{and} \quad t = t^+ - t^-$$

Proof. We have $0 \leqslant x_n^+ \leqslant |x_n|$. By comparison test, we have that $\sum_{n=1}^{\infty} x_n^+$ converges. We also have $0 \leqslant x_n^- \leqslant |x_n|$. By comparison test, we have that $\sum_{n=1}^{\infty} x_n^-$ converges. Remember that by the limit of the sum,

$$\sum_{n=1}^{\infty} x_n^+ = \lim_{\ell\to\infty} \sum_{n=1}^{\ell} x_n^+$$
$$\sum_{n=1}^{\infty} x_n^- = \lim_{\ell\to\infty} \sum_{n=1}^{\ell} x_n^-$$

So, we have that

$$\sum_{n=1}^{\infty} x_n = \lim_{\ell\to\infty} \sum_{n=1}^{\ell} x_n^+ - x_n^- = \lim_{\ell\to\infty} \sum_{n=1}^{\ell} x_n.$$

This implies that x_n converges as desired. ☺

Theorem 2.2.4

Let E be a set and $\ell^\infty(E) = \{f : E \to \mathbb{R} \text{ bounded}\}$. Let $\{f_n\}_n \subset \ell^\infty(E)$ and $f \in \ell^\infty(E)$. Then,

1. If $\sum_{n=1}^{\infty} \sup_{x\in E} |f_n(x)| < \infty$, then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly in E.
2. If $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly to f, then $\lim_{n\to\infty} \sup_{x\in E} |f_n(x)| = 0$.

Proof. Let $a_n = \sup_{x\in E} |f_n(x)|$. We know that the sum of the a_n converges in $\mathbb{R}$. Fix an $x \in E$. We have that $0 \leqslant |f_n(x)| \leqslant a_n$. By the comparison test, we have that $\sum_{n=1}^{\infty} |f_n(x)|$ converges pointwise. So by the previous theorem, $\sum_{n=1}^{\infty} f_n(x)$ converges pointwise in $\mathbb{R}$. This isn't good enough; we want uniform convergence. That is, we want:

$$\|f - \sum_{n=1}^{\infty} f_n\|_\infty \to 0$$

FINSIH THIS LATER ☺

2.3 Continuity

Definition 2.3.1: Continuity

Let (X, d_x) and (Y, d_Y) be metric spaces. Let $E \subseteq X$ and $f : E \to Y$. Let $x_0 \in E$ and assume $x_0 \in \operatorname{acc} E$. We say that f is continuous at x_0 if there is $\lim_{x \to x_0} f(x) = f(x_0)$. We say that f is continuous on E if f is continuous at all $x_0 \in E$.
We denote $C(E)$ as the continuous functions on E.

Example 2.3.1

1. Consider sequences. That is, $f\mathbb{N} \to \mathbb{R}$. This is continous because $\mathbb{N} \cap \operatorname{acc} \mathbb{N} = \emptyset$.
2. If we have $f : [0, 1] \cup \{3\} \to \mathbb{R}$, we only check continuity at $x_0 \in [0, 1]$. f is continuous at 3.
3. The sum, product, quotient (denominator nonzero), and composition of two continuous functions is continuous.

Exercise 2.3.1 Continuity

- x^n continuous
- $\sin(x)$ continuous
- $\cos(x)$ continuous

Definition 2.3.2: Relatively Open

Let (X, d_X) be a metric space and $E \subseteq X$. We say that $F \subseteq E$ is *relatively open* in E if $F = E \cap U$ with U open.

Theorem 2.3.1

Let (X, d_X) and (Y, d_Y) be metric spaces. Let $E \subseteq X$ and $f : E \to Y$. Then f is continuous on E if and only if for all open sets $V \subseteq Y$, $f^{-1}(V)$ is relatively open in E.

Example 2.3.2

Let $F = \{(x, y) \in \mathbb{R}^2 : x + \sin y > 4\}$.Then $f(x, y) = x + \sin y$ is continuous because $F = f^{-1}((4, \infty))$.

Proof. We start with the forward direction. Assume that f is continuous on E. Let $V \subseteq Y$ be open. Consider $f^{-1}(V)$. If $V \neq \emptyset$, let $x_0 \in f^{-1}(V)$. Then $f(x_0) \in V$. Find $B_Y(f(x_0), \epsilon) \subseteq V$. If $x_0 \in \operatorname{acc} E$, then we can find $\delta > 0$ such that if $x \in E$ and $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \epsilon$. So if $x \in B(x_0, \delta) \cap E$, then $f(x) \in B_Y(f(x_0), \epsilon) \subseteq V$. So $B(x_0, \delta) \cap E \subseteq f^{-1}(V)$.

If x_0 isn't an accumulation point, then there exists $\delta > 0$ such that $B(x_0, \delta) \cap E = \{x_0\}$. So $f^{-1}(V) = \{x_0\}$.

So we just have $f^{-1}(V) = E \cap \bigcup_{x \in E} B(x, \delta_x)$. So $f^{-1}(V)$ is relatively open in E.

For the backward direction, assume that $f^{-1}(V)$ is relatively open for all $V \subseteq Y$ that are open. We want to show that f is continuous. Let $x_0 \in E \cap \operatorname{acc} E$. We want to show that the limit as $x \to x_0$ of $f(x)$ is $f(x_0)$. Take $V = B_Y(f(x_0), \epsilon)$. Since $f^{-1}(V)$ is relatively open, $f^{-1}(B(f(x_0), \epsilon)) = E \cap U$ for some open U. So $x_0 \in U$, so there exists $\delta > 0$ such that $B(x_0, \delta) \cap E \subseteq U$. So if $x \in B(x_0, \delta) \cap E$, then $f(x) \in B_Y(f(x_0), \epsilon)$. So $\lim_{x \to x_0} f(x) = f(x_0)$ because $d(f(x), f(x_0)) < \epsilon$. ☺

Theorem 2.3.2

Let (X, d_X) and (Y, d_Y) be metric spaces. Take $K \subseteq X$ that is sequentially compact and $f : K \longrightarrow Y$ that is continuous. Then $f(K)$ is sequentially compact.

Proof. Let $y_n \in f(K)$. Then there exists $x_n \in K$ such that $f(x_n) = y_n$. Since K is sequentially compact, there exists a subsequence $\{x_{n_k}\}_k$ that converges to $x \in K$. Since f is continuous, we have that $f(x_{n_k}) \longrightarrow f(x)$ if $x \in \operatorname{acc} E$. If $x \notin \operatorname{acc} E$, then it is a constant sequence and we are done. So $f(x) \in f(K)$. 😃

Theorem 2.3.3 Weierstrass Theorem

Let (X, d_X) be a metric space and $K \subseteq X$ that is sequentially compact. Let $f : K \longrightarrow \mathbb{R}$ be continuous. Then f is bounded and attains its bounds.

Proof. We know that $f(K)$ closed and bounded by the previous theorem. As such, there exists $\sup f(K) = \ell \in \mathbb{R}$. We want to show that this is the maximum. Consider $\ell - \frac{1}{n}$. This is not an upper bound for $f(K)$, so there exists $x_n \in K$ such that $f(x_n) > \ell - \frac{1}{n}$. Since K is sequentially compact, there exists a subsequence $\{x_{n_k}\}_k$ that converges to $x \in K$. So,

$$\ell - \frac{1}{n_k} < f(x_{n_k}) \leqslant \ell$$

As $k \longrightarrow \infty$, we have that $f(x_{n_k}) \longrightarrow \ell$. So $f(x) = \ell$. So f attains its maximum. 😃

Theorem 2.3.4 Let $I \subseteq \mathbb{R}$ be an interval and $f : I \longrightarrow \mathbb{R}$ continuous and assume there exist x_1 and x_2 such that $f(x_1) < 0 < f(x_2)$. Then there exists $x_0 \in I$ such that $f(x_0) = 0$.

Proof. Assume $x_1 < x_2$. So $lim_{x \to x_1} = f(x_1) < 0$. Let $\epsilon = -f(x_1)/2$ to find $\delta_1 > 0$ such that $f(x) < 0$ in $[x_1, x_1 = \delta_1]$.

We also have that $\lim_{x \to x_2} f(x) = f(x_2) > 0$. Let $\epsilon = f(x_2)/2 > 0$. So we can find $\delta_2 > 0$ such that $f(x) > 0$ in $[x_2 - \delta_2, x_2]$.

Consider the set $E = \{x \in [x_1, x_2] : f(x) < 0\}$, which is bounded above by $x_2 - \delta_2$. So $\sup E = \ell \in [x_1 + \delta_2, x_2 - \delta - 2]$ exists.

We claim that $f(\ell) = 0$. If $f(\ell) < 0$, then by dfn of continuity, $f(\ell) = \lim_{x \to \ell} f(x)$. Take $\epsilon = -\ell/2$ and find $\delta_3 > 0$ such that $f(x) < 0$ in $[\ell - \delta_3, \ell + \delta_3]$. So $\ell + \delta_3 \in E$, which is a contradiction because ℓ is a maximum. If $f(\ell) > 0$, then take $\epsilon = \ell/2$ and find $\delta_4 > 0$ such that $f(x) > 0$ in $[\ell - \delta_4, \ell + \delta_4]$. So $\ell - \delta_4 \in E$, which is a contradiction because it is then a better lower bound than ℓ. So $f(\ell) = 0$ by trichotomy. 😃

Corollary 2.3.1

A polynomial of odd degree has at least one zero.

Proof. Consider $p(x)$ that has odd degree. WLOG assume the first coefficient is positive. So we have:

$$\lim_{x \to \infty} p(x) = \infty \qquad \lim_{x \to -\infty} p(x) = -\infty$$

By the previous theorem, we have that there exists $x_0 \in \mathbb{R}$ such that $p(x_0) = 0$. 😃

Corollary 2.3.2

Consider the interval $I \subseteq \mathbb{R}$ with $f : I \longrightarrow \mathbb{R}$ continuous. Then $f(I)$ is an interval with endpoints $\inf f(I)$ and $\sup f(I)$.

Proof. Let $y_1, y_2 \in f(I)$ with $y_1 < y_2$ and let $y_1 < t < y_2$. Then we wnant to show that $t \in f(I)$. Consider $g(x) = f(x) - t$. There exists x_1 and x_2 such that $f(x_1) = y_1$ and $f(x_2) = y_2$. So $g(x_1) < 0 < g(x_2)$. So by the previous theorem, there exists $x_0 \in I$ such that $g(x_0) = 0$. So $f(x_0) = t$. 😄

Corollary 2.3.3

Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ that is continuous and injective. Then $f^{-1} : f(I) \to \mathbb{R}$ is continuous.

Example 2.3.3 (bad)

Consider $E = [0,1] \cup (2,3]$. Then let:

$$f(x) = \begin{cases} x & x \in [0,1] \\ x - 1 & x \in (2,3] \end{cases}$$

This does not have a continuous inverse.

Proof. First we show that f is monotone. Assume FSOC that there are $a < b$ such that $f(a) < f(b)$. Then f is strictly increasing. 😄

Theorem 2.3.5

Let (X, d_X) and (Y, d_Y) be metric spaces. Let $K \subseteq X$ be sequentially compact. Let $f : K \to Y$ be continuous and injective. Then $f^{-1} : f(K) \to X$ is continuous.

Proof. Let $y_0 \in f(K)$ and let $y_0 \in \operatorname{acc} f(K)$. We claim that $\lim_{y \to y_0} f^{-1}(y) = f^{-1}(y_0)$. So BWOC, there exists $\epsilon > 0$ such that for every δ, we can find $y \in B_Y(y_0, \delta)$ such that $d_X(f^{-1}(y), f^{-1}(y_0)) \geqslant \epsilon$. Let $\delta = 1/n$. Then we can find $y_n \in B_Y(y_0, 1/n)$ such that $d_X(f^{-1}(y_n), f^{-1}(y_0)) \geqslant \epsilon$. $y_n \in f(K)$, so there is $x_n \in K$ such that $f(x_n) = y_n$. Since K is sequentially compact, there exists a subsequence $\{x_{n_k}\}_k$ that converges to $x_0 \in K$. So since f is continuous, we have that $x_{n_k} \to x_0 \implies f(x_{n_k}) = y_{n_k} \to f(x_0) = y_0$. $d_X(f^{-1}(y_{n_k}), f^{-1}(y_0)) \geqslant \epsilon > 0$. But we have that $d_X(x_{n_k}, x_0) \to 0$, contradiction. As such, we have that $\lim_{y \to y_0} f^{-1}(y) = f^{-1}(y_0)$ and so f^{-1} is continuous. 😄

2.4 Differentiability

Definition 2.4.1: Directional Derivatives

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Let $f : E \to Y$, $E \subseteq X$, $x_0 \in E$, $v \in X$, $\|v\|_X = 1$. $L = \{x \in E : x = x_0 + tv \text{ for some } t \in \mathbb{R}\}$. Assume $x_0 \in \operatorname{acc} L$.
The directional derivative of f at x_0 in the direction of v is:

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t},$$

whenever the limit exists. It is denoted as $\frac{\partial f}{\partial v}(x_0)$.

Note:

If $X = \mathbb{R}^N$ and ℓ_i is the ith vector in the canonical basis, then $\frac{\partial f}{\partial \ell_i}(x_0)$ is the ith partial derivative of f at x_0. Also, if we have $f(x, y, z)$, then $\frac{\partial f}{\partial x}(x_0, y_0, z_0)$ is the same as $\lim_{t \to 0} \frac{f(x_0, y_0 + t, z_0) - f(x_0, y_0, z_0)}{t}$.

Definition 2.4.2: One-dimensional Derivative

Let $X = \mathbb{R}$ and $v = 1$. Then:

$$\frac{\partial f}{\partial v}(x_0) = \lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t} = f'(x_0)$$

This is the one-dimensional derivative of f at x_0.

If $f'(x_0)$ exists, then f is continuous at x_0. This is not true for $N \geqslant 2$.

Definition 2.4.3: Differentiability

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Let $E \subseteq X$ and $f : E \to Y$. We say that f is differentiable at $x_0 \in E \cap \operatorname{acc} E$ if there exists a linear function $L : X \to Y$ and continuous such that $\lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|_X} = 0$. We denote L as $df(x_0)$. L is called the differential of f at x_0.

Theorem 2.4.1 (Useful to negate)

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Let $E \subseteq X$ and $f : E \to Y$. Let $x_0 \in E \cap \operatorname{acc} E$. If f is differentiable at x_0, then f is continuous at x_0.

Proof. Let $L = df(x_0)$. Then we have:

$$\begin{aligned} f(x) - f(x_0) &= f(x) - f(x_0) - L(x - x_0) + L(x - x_0) \\ &= \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|_X} \|x - x_0\|_X + L(x - x_0) \\ &= 0 \cdot 0 + L(0) = 0 \end{aligned}$$

So f is continuous at x_0. ☺

Theorem 2.4.2 (Also useful to negate)

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Let $E \subseteq X$ and $f : E \to Y$. Let $x_0 \in E \cap \operatorname{acc} E$. Assume f is differentiable at x_0. Let $v \in X$ with $\|v\|_X = 1$. Assume $x_0 \in \operatorname{acc} L$, $L = \{x \in E : x = x_0 + tv \text{ for some } t \in \mathbb{R}\}$. Then $\frac{\partial f}{\partial v}(x_0) = T(v)$ where T is the differential of f at x_0.

Proof. Want to show:

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} = T(v).$$

By the definition of differentiability, we know that:

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

where $T : X \to Y$ is linear and continuous. Take $x = x_0 + tv$ (restriction). Then we have:

$$\begin{aligned}
\frac{f(x) - f(x_0) - T(x - x_0)}{\|x - x_0\|_X} &= \frac{f(x_0 + tv) - f(x_0) - T(tv)}{\|tv\|_X} \\
&= \frac{f(x_0 + tv) - f(x_0) - tT(v)}{|t|\|v\|_X} \\
&= \frac{f(x_0 + tv) - f(x_0) - tT(v)}{|t|}
\end{aligned}$$

If we take the limit from the right, we get:

$$\begin{aligned}
\lim_{t \to 0^+} \frac{f(x) - f(x_0) - T(x - x_0)}{\|x - x_0\|_X} &= 0 \\
&= \lim_{t \to 0^+} \frac{f(x_0 + tv) - f(x_0) - tT(v)}{|t|} \\
&= \lim_{t \to 0^+} \frac{f(x_0 + tv) - f(x_0)}{t} - T(v) \\
&\implies \lim_{t \to 0^+} \frac{f(x_0 + tv) - f(x_0)}{t} = T(v)
\end{aligned}$$

If we take the limit from the left, we get:

$$\begin{aligned}
\lim_{t \to 0^-} \frac{f(x) - f(x_0) - T(x - x_0)}{\|x - x_0\|_X} &= 0 \\
&= \lim_{t \to 0^-} \frac{f(x_0 + tv) - f(x_0) - tT(v)}{|t|} \\
&= \lim_{t \to 0^-} \frac{f(x_0 + tv) - f(x_0)}{t} - T(v) \\
&\implies \lim_{t \to 0^-} \frac{f(x_0 + tv) - f(x_0)}{t} = T(v)
\end{aligned}$$

😄

Note:

Let $X = \mathbb{R}^N$ and $x_0 \in E^\circ$. Then let $f : E \to \mathbb{R}$. Assume f is differentiable at x_0. Then for all $v \in \mathbb{R}^N$ with $\|v\| = 1$, we have that $\frac{\partial f}{\partial v}(x_0) = T(v)$. But we have that:

$$v = (v_1, v_2, \dots, v_N)$$

which means that

$$v = \sum_{i=1}^{N} v_i e_i.$$

Since T is linear, we have that:

$$\frac{\partial f}{\partial v}(x_0) = T(v) = \sum_{i=1}^{N} v_i T(e_i)$$
$$= \sum_{i=1}^{N} v_i \frac{\partial f}{\partial x_i}(x_0)$$

Definition 2.4.4: Gradient

Let $X = \mathbb{R}^N$ and $x_0 \in E^\circ$. Let $f : E \to \mathbb{R}$. Assume f is differentiable at x_0. Then the gradient of f at x_0 is:

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \dots, \frac{\partial f}{\partial x_N}(x_0)\right)$$

Note:

So to check that f is differentiable at x_0, we need to check:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - \nabla f(x_0) \cdot (x - x_0)}{\|x - x_0\|} = 0$$

Theorem 2.4.3 (useful to prove differentiability at most points)

Let $E \subseteq \mathbb{R}^N$ and $x_0 \in E^\circ$. Let $f : E \to \mathbb{R}$. Assume

1. $\frac{\partial f}{\partial x_i}$ exists in $B(x_0, r) \subseteq E$ for all $i = 1, 2, \dots, N$.
2. $\frac{\partial f}{\partial x_i}$ is continuous at x_0 for all $i = 1, 2, \dots, N$.

Then f is differentiable at x_0.

Lenma 2.4.1 Rolle

Let $f : [a, b] \in \mathbb{R}$ with $a < b$ be continuous in $[a, b]$ and differentiable on (a, b). Also assume $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.

Proof. $[a, b]$ is sequentially compact and f is continuous, so there are max $f(x) = M$ and min $f(x) = m$. If $M = m$, then f is constant and we are done since the derivative is 0 for all x. Assume $M > m$. Since $f(a) = f(b)$, we have that either M or m is in the interior. Say $m = f(c)$ for $c \in (a, b)$. So $f(x) - f(c) \geqslant 0$ for all $x \in [a, b]$. If $x - c > 0$, then we have

$$\frac{f(x) - f(c)}{x - c} \geqslant 0$$

So,

$$\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} = f'(c) \geqslant 0.$$

If $x - c < 0$, then we have

$$\frac{f(x) - f(c)}{x - c} \leqslant 0$$

So,

$$\lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} = f'(c) \leqslant 0.$$

So $f'(c) = 0$. 😊

Lemma 2.4.2 MVT

Let $f : [a, b] \to \mathbb{R}$ be continuous in $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that:

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

Proof. Want g such that $g(b) = g(a)$. Take $g(x) = f(x) - (x - a)\frac{f(b)-f(a)}{b-a}$. Let's make sure this works:

$$g(b) = f(b) - (b - a)\frac{f(b) - f(a)}{b - a} = f(b) - f(b) + f(a) = f(a)$$
$$g(a) = f(a) - (a - a)\frac{f(b) - f(a)}{b - a} = f(a)$$

So $g(a) = g(b)$. So by Rolle's theorem, there exists $c \in (a, b)$ such that $g'(c) = 0$. So $g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} = 0$. So $f'(c) = \frac{f(b)-f(a)}{b-a}$ as desired. 😊

Now we actually prove Theorem 2.2.12.

Proof. Take $x \in B(x_0, r)$ and consider $f(x) - f(x_0)$. We need to define these vectors in terms of their components:

$$x = (x_1, x_2, \dots, x_N) \quad x_0 = (x_{01}, x_{02}, \dots, x_{0N})$$

So, we alternate the definition of $f(x) - f(x_0)$ component-wise:

$$f(x) - f(x_0) = f(x_1, x_2, \dots, x_N) - f(x_{01}, x_2, \dots, x_N) + f(x_{01}, x_2, \dots, x_N) - f(x_{01}, x_{02}, x_3, \dots, x_N) + \dots$$

Consider $g(x_1) = f(x_1, x_2, \dots, x_N)$ where the last $N - 1$ components are fixed. By MVT,

$$g(x_1) - g(x_{01}) = g'(c_1)(x_1 - x_{01})$$

where c_1 is between x_1 and x_{01}. That is,

$$f(x_1, x_2, \dots, x_N) - f(x_{01}, x_2, \dots, x_N) = \frac{\partial f}{\partial x_1}(c_1, x_2, \dots, x_N)(x_1 - x_{01})$$

So

$$\begin{aligned} f(x) - f(x_0) = {} & \frac{\partial f}{\partial x_1}(c_1, x_2, \dots, x_N)(x_1 - x_{01}) + \frac{\partial f}{\partial x_2}(x_{01}, c_2, x_3, \dots, x_N)(x_2 - x_{02}) + \dots \\ & + \frac{\partial f}{\partial x_{N-1}}(x_{01}, x_2, \dots, c_{N-1}, x_n)(x_{N-1} - x_{0N-1}) + f(x_{01}, \dots, x_{0N-1}, x_N) - f(x_{01}, \dots, x_{0N}) \end{aligned}$$

First let's define $z_m = (x_{01}, \dots, c_m, \dots, x_N)$. We can now consider:

$$\begin{aligned} f(x) - f(x_0) - \nabla f(x_0) \cdot (x - x_0) &= f(x) - f(x_0) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0i}) \\ &= \left(\frac{\partial f}{\partial x_1}(z_1) - \frac{\partial f}{\partial x_1}(x_0)\right)(x_1 - x_{01}) + \dots \\ &+ f(x_{01}, \dots, x_{0N-1}, x_N) - f(x_{01}, \dots, x_{0N}) \end{aligned}$$

Now we divide by $\|x - x_0\|$ to get:

$$
\begin{aligned}
&\frac{|f(x) - f(x_0) - \nabla f(x_0) \cdot (x - x_0)|}{\|x - x_0\|} \leqslant \\
&\left|\frac{\partial f}{\partial x_1}(z_1) - \frac{\partial f}{\partial x_1}(x_0)\right| \frac{|x_1 - x_{01}|}{\|x - x_0\|} + \ldots + \\
&\left|\frac{\partial f}{\partial x_{N-1}}(z_{N-1}) - \frac{\partial f}{\partial x_{N-1}}(x_0)\right| \frac{|x_{N-1} - x_{0N-1}|}{\|x - x_0\|} + \\
&\frac{|f(x_{01}, \ldots, x_{0N-1}, x_N) - f(x_0) - \frac{\partial f}{\partial x_N}(x_0)(x_N - x_{0N})|}{\|x - x_0\|}
\end{aligned}
$$

All the terms that aren't the last one end up tending to 0 as $x \longrightarrow x_0$. Consider the last term and call it A. So,

$$
\begin{aligned}
A &= \left|\frac{f(x_0, \ldots, x_{0N-1}, x_N) - f(x_0)}{x_N - x_{0N}} - \frac{\partial f}{\partial x_N}(x_0)\right| \\
&= \left|\frac{f(x_0 + t\ell_n) - f(x_0)}{t} - \frac{\partial f}{\partial x_N}(x_0)\right| \to 0
\end{aligned}
$$

if we set $t = x_N - x_{0N}$. So f is differentiable at x_0. 😄

2.5 Differentiation Rules

Theorem 2.5.1 Chain Rule

Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, and $(Z, \|\cdot\|_Z)$ be normed spaces. Let $E \subseteq X$, $F \subseteq Y$, $f : E \longrightarrow Y$, $g : F \longrightarrow Z$. Assume $f(E) \subseteq F$ and f is differentiable at $x_0 \in E$ and g is differentiable at $f(x_0)$. Then $g \circ f$ is differentiable at x_0 and:

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0).$$

Moreover, if $x_0 \in \operatorname{acc}\{x_0 + tv \in E : t \in \mathbb{R}\}$, then

$$\frac{\partial(g \circ f)}{\partial v}(x_0) = d(g(f(x_0)))\left(\frac{\partial f}{\partial v}(x_0)\right).$$

Lenma 2.5.1

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Take $T : X \longrightarrow Y$ linear and continuous. Then there exists a constant $M > 0$ such that $\|T(x)\|_Y \leqslant M\|x\|_X$ for all $x \in X$.

Proof. Let $\epsilon = 1$, T continuous at $x = 0$. By the definition of continuity, we can find $\delta > 0$ such that $\|T(x)\|_Y < \epsilon = 1$ if $\|x\|_X < \delta$. Take any $x \in X$ where $x \neq 0$. Then let $x_1 = \frac{\delta}{2}\frac{x}{\|x\|_X}$. So, $\|x_1\|_X = \frac{\delta}{2}$. So $\|T(x_1)\|_Y < 1$. So,

$$\begin{aligned}
\|T(x_1)\|_Y &= \|T\left(\frac{\delta}{2}\frac{x}{\|x\|_X}\right)\|_Y \\
&= \frac{\delta}{2}\frac{\|T(x)\|_Y}{\|x\|_X} < 1 \\
\implies \|T(x)\|_Y &< \frac{2}{\delta}\|x\|_X
\end{aligned}$$

So we can take $M = \frac{2}{\delta}$ to complete the proof. 😄

Lenma 2.5.2

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Take $f : E \longrightarrow Y$ and assume f differentiable at some point $x_0 \in E$. Then there exists a constant $C > 0$ and $\delta > 0$ such that $\|f(x) - f(x_0)\|_Y \leqslant C\|x - x_0\|_X$ for all $x \in B(x_0, \delta)$.

Proof. f is differentiable, so we know that $\lim_{x \to x_0} \frac{f(x) - f(x_0) - T(x - x_0)}{\|x - x_0\|_X} = 0$ for some $T : X \longrightarrow Y$ linear and continuous. Take $\epsilon = 1$. Find $\delta > 0$ such that

$$\|f(x) - f(x_0) - T(x - x_0)\|_Y < \epsilon\|x - x_0\|_X$$

if $\|x - x_0\|_X < \delta$. We have:

$$\begin{aligned}
\|f(x) - f(x_0)\|_Y &\leqslant \|f(x) - f(x_0) - T(x - x_0)\|_Y + \|T(x - x_0)\|_Y \\
&\leqslant \epsilon\|x - x_0\|_X + M\|x - x_0\|_X \\
&= (M + 1)\|x - x_0\|_X
\end{aligned}$$

as desired. 😄

Now we actually prove the chain rule.

Proof. We know there exists $T : X \to Y$ linear and continuous such that $\lim_{x \to x_0} \frac{f(x)-f(x_0)-T(x-x_0)}{\|x-x_0\|_X} = 0$. We also know there exists $L : Y \to Z$ linear and continuous such that $\lim_{y \to f(x_0)} \frac{g(y)-g(f(x_0))-L(y-f(x_0))}{\|y-f(x_0)\|_Y} = 0$. We want to show that $\lim_{x \to x_0} \frac{g(f(x))-g(f(x_0))-(L \circ T)(x-x_0)}{\|x-x_0\|_X} = 0$. We start by analyzing the numerator:

$$\begin{aligned}
&\|g(f(x)) - g(f(x_0)) - L(T(x - x_0))\|_Z \\
&\leqslant \|g(f(x)) - g(f(x_0)) - L(f(x) - f(x_0))\|_Z + \|L(f(x) - f(x_0)) - L(T(x - x_0))\|_Z \\
&= \|g(f(x)) - g(f(x_0)) - L(f(x) - f(x_0))\|_Z + \|L(f(x) - f(x_0) - T(x - x_0))\|_Z
\end{aligned}$$

We will study these quantities separately.

We start with the second quantity. By Lemma 1, there exists $M > 0$ such that $\|L(f(x) - f(x_0) - T(x - x_0))\|_Z \leqslant M\|f(x) - f(x_0) - T(x - x_0)\|_Y$. Given $\epsilon > 0$, there exists $\delta_1 > 0$ such that $\|f(x) - f(x_0) - T(x - x_0)\|_Y < \epsilon\|x - x_0\|_X$ if $0 < \|x - x_0\|_X < \delta_1$. So,

$$\begin{aligned}
\|L(f(x) - f(x_0) - T(x - x_0))\|_Z &\leqslant M\|f(x) - f(x_0) - T(x - x_0)\|_Y \\
&\leqslant M\epsilon\|x - x_0\|_X
\end{aligned}$$

for all $x \in E$ with $0 < \|x - x_0\|_X < \delta_1$.

Now we study the first quantity. Given the same $\epsilon > 0$, there exists $\delta_2 > 0$ such that $\|g(y) - g(f(x_0)) - L(y - f(x_0))\|_Z < \epsilon\|y - f(x_0)\|_Y$ if $0 < \|y - f(x_0)\|_Y < \delta_2$. We want to take $y = f(x)$ so we need to check that $0 < \|f(x) - f(x_0)\|_Y < \delta_2$. By Lemma 2, there exists $C > 0$ and $\delta_3 > 0$ such that $\|f(x) - f(x_0)\|_Y \leqslant C\|x - x_0\|_X$ for all $x \in E$ such that $0 < \|x - x_0\|_X < \delta_3$. So we can take $\delta = \min\{\delta_1, \delta_3, \delta_2/C\}$. So we have that we can take $y = f(x)$ because $\|f(x) - f(x_0)\|_Y < \delta_2$. So we have that:

$$\begin{aligned}
\|g(f(x)) - g(f(x_0)) - L(f(x) - f(x_0))\|_Z &< \epsilon\|f(x) - f(x_0)\|_Y \\
&\leqslant \epsilon C\|x - x_0\|_X
\end{aligned}$$

for all $x \in E$ such that $0 < \|x - x_0\|_X < \delta$. So we have that:

$$\|g(f(x)) - g(f(x_0)) - L(f(x) - f(x_0))\|_Z + \|L(f(x) - f(x_0) - T(x - x_0))\|_Z < \epsilon(M + C)\|x - x_0\|_X$$

Dividing by $\|x - x_0\|_X$, we get that:

$$\frac{\|g(f(x)) - g(f(x_0)) - L(f(x) - f(x_0))\|_Z + \|L(f(x) - f(x_0) - T(x - x_0))\|_Z}{\|x - x_0\|_X} < \epsilon(M + C)$$

Now we prove the second part of the theorem. f differentiable at x_0 implies that $\dfrac{\partial f}{\partial v}(x_0) = T(v)$. $g \circ f$ being differentiable at x_0 implies that $\dfrac{\partial (g \circ f)}{\partial v}(x_0) = L(T(v))$. So we have that:

$$L(T(v)) = L\left(\frac{\partial f}{\partial v}(x_0)\right)$$

as desired. 😃

Theorem 2.5.2 Quotient Rule

Let $(X, \|\cdot\|)$ be a normed space and $f : E \longrightarrow \mathbb{R}^2$ with f differentiable at $x_0 \in E$, with $f_2(x) \neq 0$ for all $x \in E$. Then $\frac{f_1}{f_2}$ is differentiable at x_0 and:

$$\left(\frac{f_1}{f_2}\right)'(x_0) = \frac{f_2(x_0)f_1'(x_0) - f_1(x_0)f_2'(x_0)}{(f_2(x_0))^2}$$

Proof. We can use the fact that if $g : \mathbb{R}^2 \longrightarrow \mathbb{R}\ \{0\} \longrightarrow \mathbb{R}$ with $g(y_1, y_2) = \frac{y_1}{y_2}$, then:

$$\frac{\partial g}{\partial y_1}(y_1, y_2) = \frac{1}{y_2}$$
$$\frac{\partial g}{\partial y_2}(y_1, y_2) = -\frac{y_1}{y_2^2}$$

We apply chain rule.

$$\begin{aligned}
\frac{\partial}{\partial v}\left(\frac{f_1}{f_2}\right)(x_0) &= \frac{\partial}{\partial v}(g \circ f)(x_0) \\
&= dg(f(x_0))\left(\frac{\partial f}{\partial v}(x_0)\right) \\
&= \nabla g(f(x_0)) \cdot \frac{\partial f}{\partial v}(x_0) \\
&= \frac{\partial g}{\partial y_1}(f_1(x_0), f_2(x_0))\frac{\partial f_1}{\partial v}(x_0) + \frac{\partial g}{\partial y_2}(f_1(x_0), f_2(x_0))\frac{\partial f_2}{\partial v}(x_0) \\
&= \frac{1}{f_2(x_0)}f_1'(x_0) - \frac{f_1(x_0)}{(f_2(x_0))^2}f_2'(x_0) \\
&= \frac{f_2(x_0)f_1'(x_0) - f_1(x_0)f_2'(x_0)}{(f_2(x_0))^2}
\end{aligned}$$

as desired. 😄

Definition 2.5.1: Jacobian Matrix

Let $E \subseteq \mathbb{R}^N$ and $f : E \longrightarrow \mathbb{R}^M$ and $H \subseteq \mathbb{R}^M$ with $g : H\ to \mathbb{R}^L$. Assume f is differentiable at x_0 and g is differentiable at $f(x_0)$. The Jacobian matrix of f at $x_0 \in E$ is the $M \times N$ matrix:

$$\begin{bmatrix} \nabla f_1(x_0) \\ \nabla f_2(x_0) \\ \vdots \\ \nabla f_M(x_0) \end{bmatrix}$$

and is denoted as $Jf(x_0)$.

2.6 Higher Order Derivatives

Definition 2.6.1: Higher Order Derivative

Take $E \subseteq \mathbb{R}^N$ and $f : E \to \mathbb{R}$. Assume there is $\dfrac{\partial f}{\partial x_i} : E \to \mathbb{R}$. Take $j \in [1, N]$. Then,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).$$

In general, we have that:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \neq \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Theorem 2.6.1 Symmetry of Higher Order Derivatives

Let $E \subseteq \mathbb{R}^N$ and $f : E \to \mathbb{R}$. Assume there exists $\dfrac{\partial f}{\partial x_i}, \dfrac{\partial f}{\partial x_j}, \dfrac{\partial^2 f}{\partial x_i \partial x_j}$ in $B(x_0, r) \subseteq E$. And assume that $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ is continuous at x_0. Then $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{\partial^2 f}{\partial x_j \partial x_i}$.

Lenma 2.6.1

Let $A : ((-r, r) \setminus \{0\})^2 \to \mathbb{R}$. Assume that there exists:

1. $\lim_{(s,t)\to(0,0)} A(s, t) = \ell \in \mathbb{R}$.
2. For every $s \in (-r, r) \setminus \{0\}$,

$$\exists \lim_{t \to 0} A(s, t) \in \mathbb{R}.$$

Then, there is

$$\lim_{s \to 0}\left(\lim_{t \to 0} A(s, t)\right) = \lim_{(s,t)\to(0,0)} A(s, t) = \ell.$$

Proof. For every $\epsilon > 0$, there exists $0 < \delta < r$ such that

$$|A(s, t) - \ell| < \epsilon$$

for all s, t with $0 < \sqrt{s^2 + t^2} < \delta$. So, we have:

$$\ell - \epsilon \leqslant A(s, t) \leqslant \ell + \epsilon.$$

Fix $s \in (-\delta, \delta) \setminus \{0\}$. By the second condition, we have that:

$$\lim_{t \to 0} A(s, t) = \ell_s$$

for some ℓ_s. So take $t \to 0$ in the inequality above to get:

$$\ell - \epsilon \leqslant \ell_s \leqslant \ell + \epsilon$$

for the same choice of δ. So,

$$|\ell_s - \ell| < \epsilon$$

for all $0 < |s| < \delta$. So $\lim_{s \to 0} \ell_s = \ell$, as desired. 😊

Now we turn to the proof of the symmetry of higher order derivatives.

Proof. Step 1: Let $N = 2$. Let $\vec{x}_0 = (x_0, y_0)$. Consider the rectangle R determined by the corners $(x_0, y_0), (x_0 + s, y_0), (x_0, y_0 + t), (x_0 + s, y_0 + t)$ and assume $R \subseteq B(\vec{x}_0, r)$. Let

$$A(s,t) = \frac{f(x_0+s, y_0+t) - f(x_0+s, y_0) - f(x_0, y_0+t) + f(x_0, y_0)}{st}.$$

Fix t, move s. Consider the following functions

$$g(x) = f(x, y_0 + t) - f(x, y_0)$$

We attempt to write $A(s,t)$ in terms of g:

$$\begin{aligned} A(s,t) &= \frac{f(x_0+s, y_0+t) - f(x_0+s, y_0) - f(x_0, y_0+t) + f(x_0, y_0)}{st} \\ &= \frac{g(x_0+s) - g(x_0)}{st} \\ &= \frac{g(x_0+s) - g(x_0)}{s} \cdot \frac{1}{t} \end{aligned}$$

By MVT, there exists c_s between x_0 and $x_0 + s$ such that:

$$g(x_0+s) - g(x_0) = s g'(c_s)$$

So we have that:

$$A(s,t) = \frac{g'(c_s)}{t}$$

Consider $g'(x) = \frac{\partial f}{\partial x}(x, y_0 + t) - \frac{\partial f}{\partial x}(x, y_0)$. So we have that:

$$A(s,t) = \frac{\frac{\partial f}{\partial x}(c_s, y_0+t) - \frac{\partial f}{\partial x} c_s, y_0)}{t}$$

Consider the function:

$$h(y) = \frac{\partial f}{\partial x}(c_s, y)$$

So,

$$A(s,t) = \frac{h(y_0+t) - h(y_0)}{t} = h'(c_t) = \frac{\partial^2 f}{\partial y \partial x}(c_s, c_t)$$

where c_t is between y_0 and $y_0 + t$. As $(s,t) \to (0,0)$, we have that $c_s \to x_0$ and $c_t \to y_0$. So we have that:

$$\lim_{(s,t)\to(0,0)} A(s,t) = \lim_{(s,t)\to(0,0)} \frac{\partial^2 f}{\partial y \partial x}(c_s, c_t) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

Now we consider

$$\begin{aligned} \lim_{t\to 0} A(s,t) &= \frac{1}{s} \lim_{t \to 0} \frac{f(x_0+s, y_0+t) - f(x_0+s, y_0)}{t} - \frac{f(x_0, y_0+t) - f(x_0, y_0)}{t} \\ &= \frac{1}{s}\left(\frac{\partial f}{\partial y}(x_0+s, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)\right) \end{aligned}$$

So now the preconditions for the lemma are satisfied. So we have that:

$$\begin{aligned} \lim_{(s,t)\to(0,0)} A(s,t) = \lim_{s\to 0}\left(\lim_{t\to 0} A(s,t)\right) &= \lim_{s\to 0} \frac{\frac{\partial f}{\partial y}(x_0+s, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)}{s} \\ &= \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \end{aligned}$$

But we also have that:

$$\lim_{(s,t)\to(0,0)} A(s,t) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

So we have that $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$ as desired.

Now consider the case where $N \geqslant 2$. Say $1 < i < j < N$. Then realize our function looks like:

$$F(x_i, x_j) := f(x_1, x_2, \cdots, x_i, \cdots, x_j, \cdots, x_N)$$

which is a function of two variables. So we can apply the previous case to get that:

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i}$$

as desired. 😄

2.7 Taylor Series

Theorem 2.7.1 Taylor's Formula

Let $f : (a, b) \to \mathbb{R}$ be k times differentiable for $k \in \mathbb{N}$ with $x_0 \in (a, b)$. Then:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_k(x)$$

where $R_k(x) = \frac{f^{(k-1)}(c)}{(k+1)!}(x - x_0)^{k+1}$ is the error term for some c between x and x_0. Moreover, if there exists $M > 0$ such that $|f^{(k+1)}(x)| \leqslant M$ for every $x \in (a, b)$, then

$$|R_k(x)| \leqslant \frac{M}{(k+1)!}|x - x_0|^{k+1},$$

which tends to zero faster than $(x - x_0)^{k+1}$ as $x \to x_0$. That is,

$$\lim_{x \to x_0} \frac{R_k(x)}{(x - x_0)^{k+1}} = 0.$$

Or,

$$R_k(x) = o((x - x_0)^{k+1}).$$

Example 2.7.1 ($\log(x + 1)$)

We have that $f(x) = \log(x+1)$, $f'(x) = \dfrac{1}{x+1}$, $f''(x) = -\dfrac{1}{(x+1)^2}$. Consider $x_0 = 0$ and $(a, b) = (-1/2, 1/2)$. Then, we have that the $k = 1$ exapansion is:

$$\begin{aligned} f(x) &= f(0) + f'(0)(x - 0) + o(x - 0) \\ &= 0 + 1x + o(x) \end{aligned}$$

Dividing both sides by x, we get that:

$$\begin{aligned} \frac{\log(1 + x)}{x} &= \frac{x + o(x)}{x} \\ &= 1 + \cancel{o(1)}^{\,0} \end{aligned}$$

Now we prove Taylor's Formula.

Proof. Fix $x \in (a, b)$. Consider:

$$g(t) = f(x) - \left[f(t) + f'(t)(x - t) + \cdots + \frac{f^{(k)}(t)}{k!}(x - t)^k\right] - R_k(x)\frac{(x - t)^{k+1}}{(x - x_0)^{k+1}}$$

Let $t \in (x_0, x)$. Then,

$$\begin{aligned} g(x) &= f(x) - f(x) + 0 + \cdots + 0 = 0 \\ g(x_0) &= R_k(x) - R_k(x) = 0 \end{aligned}$$

So, we can apply Rolle's Theorem to get that there exists c between x and x_0 such that $g'(c) = 0$. So we have that:

$$\begin{aligned}
0 &= \cancelto{0}{f'(c)} - f'(c) - f''(c)(x-c) - \cdots - \frac{f^{(k)}(c)}{k!}k(x-c)^{k-1} - R_k(x)\frac{(k+1)(x-c)^k}{(x-x_0)^{k+1}} \\
&= -f'(c) - \frac{f^{(k+1)}(c)}{k!}(x-c)^k + f'(c) + R_k(x)\frac{(k+1)(x-c)^k}{(x-x_0)^{k+1}} \\
&= -\frac{f^{(k+1)}(c)}{k!}(x-c)^k + R_k(x)\frac{(k+1)(x-c)^k}{(x-x_0)^{k+1}}
\end{aligned}$$

So,

$$\frac{f^{(k+1)}(c)}{k!} = R_k(x)\frac{k+1}{(x-x_0)^{k+1}} \implies R_k(x) = \frac{f^{(k+1)}(c)}{(k+1)!}(x-x_0)^{k+1}$$

as desired. 😊

Corollary 2.7.1

Let $f : (a,b) \to \mathbb{R}$ be k times differentiable with $f', \cdots, f^{(k)}$ continuous. Then, for every $x, x_0 \in (a,b)$,

$$f(x) = f(x_0) + \sum_{n=1}^{k} \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n + o((x-x_0)^k).$$

Proof. Apply Taylor's Formula with k to get:

$$\begin{aligned}
f(x) &= f(x_0) + \sum_{n=1}^{k-1} \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n + \frac{f^{(k)}(c)}{k!}(x-x_0)^k \\
&= f(x_0) + \sum_{n=1}^{k} \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n + [f^{(k)}(c) - f^{(k)}(x_0)]\frac{(x-x_0)^k}{k!}
\end{aligned}$$

😊

Definition 2.7.1: Multi-Index

A multi-index is a vector $\alpha \in \mathbb{N}_0$, $\alpha = (\alpha_1, \cdots, \alpha_N)$. The length of α is $|\alpha| = \alpha_1 + \cdots + \alpha_N$. We define $\alpha! = \alpha_1! \cdots \alpha_N!$ and $x^\alpha = x_1^{\alpha_1} \cdots x_N^{\alpha_N}$. We define the derivative of f with respect to α as:

$$\frac{\partial^\alpha}{\partial x^\alpha} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}}.$$

Example 2.7.2

Let $N = 3$ and $\alpha = (2,0,3)$. Then, $(x,y,z)^\alpha = x^2 z^3$. So, $\frac{\partial^\alpha}{\partial x^\alpha} = \frac{\partial^2}{\partial x^2}\frac{\partial^3}{\partial z^3}$.

Definition 2.7.2: Class

Let $U \subseteq \mathbb{R}^N$ open, $k \in \mathbb{N}$, and $f : U \to \mathbb{R}$. We say f is of class C^k if all partial derivatives of order less than or equal to k exist and are continuous.

If $f \in C^k(V)$, $k \geqslant 2$, you can apply theorem about symmetry of derivatives.

Example 2.7.3

Let $f \in C^5(k)$ and consider the the domain as $\mathbb{R}^3$. Then, we have that:

$$\frac{\partial^2}{\partial y^2}\left(\frac{\partial}{\partial x}\left(\frac{\partial}{\partial z}f\right)\right) = \frac{\partial}{\partial x}\left(\frac{\partial^2}{\partial y^2}\left(\frac{\partial}{\partial z}f\right)\right)$$

Theorem 2.7.2 Taylor's Formula

Let $V \in \mathbb{R}^N$ open, $k \in \mathbb{N}$. $f : V \to \mathbb{R}$ of class C^k. Let $x_0 \in V$. Then:

$$f(x) = f(x_0) + \sum_{|\alpha|=1}^{k} \frac{\partial^\alpha f}{\partial x^\alpha}(x_0)\frac{(x-x_0)^\alpha}{\alpha!} + o(\|x-x_0\|^k)$$

for x near x_0.

Note:

If $k = 1$,

$$\begin{aligned} f(x) &= f(x_0) + \sum_{|\alpha|=1} \frac{1}{\alpha!}\frac{\partial^\alpha f(x_0)}{\partial x^\alpha}(x-x_0)^\alpha + o(\|x-x_0\|) \\ &= f(x_0) + \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0i}) + o(\|x-x_0\|) \\ &= f(x_0) + \nabla f(x_0)\cdot(x-x_0) + o(\|x-x_0\|). \end{aligned}$$

So,

$$\frac{f(x) - f(x_0) - \nabla f(x_0)\cdot(x-x_0)}{\|x-x_0\|} = \frac{o(\|x-x_0\|)}{\|x-x_0\|} \longrightarrow 0.$$

This is differentiability at x_0.

Example 2.7.4 (Multinomial Theorem)

Let $x \in \mathbb{R}^N$, $n \in \mathbb{N}$. Prove that:

$$(x_1 + x_2 + \cdots + x_N)^n = \sum_{|\alpha|=n} \frac{x^\alpha n!}{\alpha!}$$

Example 2.7.5

$V \in \mathbb{R}^N$ open, $f : V \to \mathbb{R}$ of class C^k. $v \in \mathbb{R}^N$. Define:

$$\begin{aligned} v\cdot\nabla &= v_1\frac{\partial}{\partial x_1} + \cdots + v_N\frac{\partial}{\partial x_N} \\ (v\cdot\nabla)^n &= (v\cdot\nabla)(v\cdot\nabla)^n \\ &= \sum_{|\alpha|=n} v^\alpha \frac{n!}{\alpha!}\frac{\partial^\alpha}{\partial x^\alpha} \end{aligned}$$

Now we prove the theorem.

Proof. Let $\|v\| = 1$, and consider $x_0 + tv$ for $t \in \mathbb{R}$. Define $g(t) = f(x_0 + tv)$. We have that

$$g'(t) = \nabla f(x_0 + tv)\cdot v = (v\cdot\nabla)f(x_0+tv)$$

$$g''(t) = (v \cdot \nabla)^2 f(x_0 + tv)$$

and so on. So we have:

$$\begin{aligned} g^{(n)}(t) &= (v \cdot \nabla)^n f(x_0 + tv) \\ &= (v_1 \frac{\partial}{\partial x_1} + \cdots + v_N \frac{\partial}{\partial x_N})^n f(x_0 + tv) \\ &= \sum_{|\alpha|=n} v^\alpha \frac{n!}{\alpha!} \frac{\partial^\alpha f}{\partial x^\alpha}(x_0 + tv) \end{aligned}$$

Apply Taylor's formula to g. We have that:

$$g(t) = g(0) + g'(0)t + \cdots + \frac{g^{(n)}(0)}{n!} t^n + \cdots + \frac{g^{(k)}(0)}{k!} t^k + R_k(t)$$

where $R_k(t) = \left(\frac{g^{(k)}(c)}{k!} - \frac{g^{(k)}(0)}{k!} \right) t^k$. So then,

$$f(x_0 + tv) = f(x_0) + (v \cdot \nabla) f(x_0 + tv) t + \cdots + \sum_{|\alpha|=k} \frac{k!}{\alpha!} v^\alpha \frac{\partial^\alpha f}{\partial x^\alpha}(x_0) \frac{1}{k!} t^k + \sum_{|\alpha|=k} \frac{k!}{\alpha!} v^\alpha \left(\frac{\partial^\alpha f}{\partial x^\alpha}(x_0 + cv) - \frac{\partial^\alpha f}{\partial x^\alpha}(x_0) \right) \frac{1}{k!} t^k$$

The $k!$ cancels out, and we are now ready to choose t and v. We choose $v = \frac{x - x_0}{\|x - x_0\|}$ and $t = \|x - x_0\|$. So we have that:

$$x_0 + tv = x$$

So,

$$f(x) = f(x_0) + \cdots + \sum_{|\alpha|=k} \frac{1}{\alpha!} \frac{\partial^\alpha f}{\partial x^\alpha}(x_0) \frac{(x - x_0)^\alpha}{\|x - x_0\|^k} \|x - x_0\|^k + \sum_{|\alpha|=k} \frac{1}{\alpha!} \left(\frac{\partial^\alpha f}{\partial x^\alpha}(x_0 + cv) - \frac{\partial^\alpha f(x_0)}{\partial x^\alpha} \right) (x - x_0)^\alpha$$

Cancelling out the terms, we get the desired result. 😊

2.8 Local Min and Local Max

Definition 2.8.1: Local Min and Local Max

Let (X, d) be a metric space, $E \subseteq X$, and $f : X \to \mathbb{R}$.

1. f has a *local maximum* at $x_0 \in E$ if $f(x) \leqslant f(x_0)$ for all $x \in E \cap B(x_0, r)$ for some $r > 0$.
2. f has a *local minimum* at $x_0 \in E$ if $f(x) \geqslant f(x_0)$ for all $x \in E \cap B(x_0, r)$ for some $r > 0$.

Theorem 2.8.1

Let $(X, \|\cdot\|)$ be a normed space, $f : E \to \mathbb{R}$. Assume f has a local maximum at $x_0 \in E^\circ$ and f is differentiable at x_0. Then $df(x_0) = 0$.

Proof. Let $\|v\| = 1$. Take $g(t) = f(x_0 + tv)$. We know that $g(t) \leqslant g(0)$ for all $|t| < r$. (Proof of Rolle's Theorem) $\implies g'(0) = 0$. So, $g'(t) = df(x_0 + tv)(v)$, and $g'(0) = df(x_0)(v) = \dfrac{\partial f}{\partial v}(x_0) \implies df(x_0) = 0$. 😄

Definition 2.8.2: Hessian

Let $E \subseteq \mathbb{R}^N$ and $f : E \to \mathbb{R}$. Define $Hf(x_0)$ as the Hessian matrix of f and x_0 as:

$$Hf(x_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \cdots & \frac{\partial^2}{\partial x_N \partial x_1}(x_0) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \cdots & \frac{\partial^2}{\partial x_N \partial x_2} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_N}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_N^2}(x_0) \end{pmatrix}$$

This is an $N \times N$ matrix.

If $f \in C^2$, then $Hf(x_0)$ is symmetric. A $N \times N$ symmetric matrix implies that the eigenvalues are real.

Exercise 2.8.1

Take $p(t) = t^N + a_{N-1}t^{N-1} + \cdots + a_1 t + a_0$.

1. All the roots are positive $\iff$ the coefficients switch signs.
2. All the roots are negative $\iff$ all the coefficients are positive.

Theorem 2.8.2

Let $U \subseteq \mathbb{R}^N$ open and $f : U \longrightarrow \mathbb{R}$ of class C^2. Assume $x_0 \in U$ and that $\nabla f(x_0) = 0$. Let $\lambda_1, \lambda_2, \ldots, \lambda_N$ be the eigenvalues of $Hf(x_0)$. Then:

1. If $\lambda_1, \lambda_2, \ldots, \lambda_N > 0$, then f has a local minimum at x_0.
2. If f has a local minimum at x_0, then $\lambda_1, \lambda_2, \ldots, \lambda_N \geqslant 0$.
3. If $\lambda_1, \lambda_2, \ldots, \lambda_N < 0$, then f has a local maximum at x_0.
4. If f has a local maximum at x_0, then $\lambda_1, \lambda_2, \ldots, \lambda_N \leqslant 0$.

Lenma 2.8.1

If $H = (h_{i,j})_{i,j=1}^N$ is an $N \times N$ symmetric matrix and all the eigenvalues are positive, then $\sum_{i,j=1}^N h_{i,j} x_i x_j \geqslant c\|x\|^2$ for all $x \in \mathbb{R}^N$ and some $c > 0$.

Proof. Let $x \in \mathbb{R}^N$. Treat x as a $1 \times N$ matrix. We can show:

$$xHx^T = \sum_{i,j=1}^N h_{i,j} x_i x_j.$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{R}$ be the eigenvalues of H. Let $e_1, e_2, \ldots, e_N$ be the corresponding eigenvectors. We can assume that they form an orthonormal basis for $\mathbb{R}^N$. So $x \in \mathbb{R}^N$ can be written as a linear combination of the eigenvectors:

$$x = \sum_{i=1}^N c_i e_i.$$

So,

$$\begin{aligned} Hx^T &= \sum_{i=1}^N c_i H e_i^T \\ &= \sum_{i=1}^N c_i \lambda_i e_i^T \end{aligned}$$

So, we have that:

$$\begin{aligned} xHx^T &= \sum_{j=1}^n c_j e_j \sum_{i=1}^N c_i \lambda_i e_i^T \\ &= \sum_{i,j=1}^N c_i c_j \lambda_i e_j^T e_i \\ &= \sum_{i=1}^N c_i^2 \lambda_i \\ &\geqslant c \sum_{i=1}^N c_i^2 && (c = \min\{\lambda_1, \lambda_2, \ldots, \lambda_N\}) \\ &= c\|x\|^2 && \text{(by orthonormality)} \end{aligned}$$

😄

We move to the proof of the actual theorem.

Proof. We start by proving 1.
Assume $\lambda_1, \lambda_2, \ldots, \lambda_N > 0$. We know that $Hf(x_0)$ is symmetric. So, we can apply the lemma to get that:

$$yHf(x_0)y^T = \sum_{i,j=1}^{N} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) y_i y_j \geqslant c\|y\|^2$$

for all $y \in \mathbb{R}^N$ and some $c > 0$. Apply Taylor's formula to f:

$$f(x) = f(x_0) + \sum_{|\alpha|=1}^{2} \frac{1}{\alpha!}\frac{\partial^\alpha f}{\partial x^\alpha}(x_0)(x - x_0)^\alpha + o(\|x - x_0\|^2)$$

Since the gradient is zero, we know that:

$$\frac{\partial f}{\partial x_i}(x_0) = 0$$

for all i. So, we have that:

$$\begin{aligned}
f(x) &= f(x_0) + \sum_{|\alpha|=2}^{2} \frac{1}{\alpha!}\frac{\partial^\alpha f}{\partial x^\alpha}(x_0)(x - x_0)^\alpha + o(\|x - x_0\|^2) \\
&= f(x_0) + \frac{1}{2}\sum_{i,j=1}^{N} \frac{\partial^2 f}{\partial x_i x_j}(x_0)(x_i - x_{0i})(x_j - x_{0j}) + o(\|x - x_0\|^2) \\
&\geqslant f(x_0) + \frac{c}{2}\|x - x_0\|^2 + o(\|x - x_0\|^2) && \text{(by result from lemma)} \\
&= f(x_0) + \|x - x_0\|^2 \left(\frac{c}{2} + \frac{o(\|x - x_0\|^2)}{\|x - x_0\|^2}\right) \\
&\geqslant f(x_0) + \|x - x_0\|^2 \left(\frac{c}{2} - \frac{c}{4}\right) && (\text{take } \epsilon = \tfrac{c}{4}) \\
&\implies f(x) > f(x_0)
\end{aligned}$$

for $\|x - x_0\| < \delta$. So, f has a local minimum at x_0.
Now we prove 2. Assume f has a local minimum at x_0. That is $f(x) \geqslant f(x_0)$ for all $x \in B(x_0, r) \subseteq U$ for some $r > 0$. We want to show that $\lambda_1, \lambda_2, \ldots, \lambda_N \geqslant 0$. So BWOC, assume $\lambda_j < 0$. Take $x = x_0 + te_j$ for $t > 0$ where e_j is λ_j's eigenvector. Apply Taylor's formula:

$$\begin{aligned}
f(x_0 + te_j) &= f(x_0) + \frac{1}{2}(x_0 + te_j - x_0)Hf(x_0)(x_0 + te_j - x_0)^T + o(t^2) \\
&= f(x_0) + \frac{1}{2}\lambda_j t^2 \cancel{\|e_j\|^2}^{1} + o(t^2) \\
&= f(x_0)t^2 \left(\frac{\lambda_j}{2} + \frac{o(t^2)}{t^2}\right) \\
&\leqslant f(x_0) + t^2 \left(\frac{\lambda_j}{2} + \frac{\lambda_j}{4}\right) && (\epsilon = -\tfrac{\lambda_j}{4}) \\
&\implies f(x_0 + te_j) < f(x_0)
\end{aligned}$$

for t small enough. So, f does not have a local minimum at x_0. So, $\lambda_1, \lambda_2, \ldots, \lambda_N \geqslant 0$ by contradiction. ☺

2.9 Implicit Functions

Let $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^M$ and $E \subseteq \mathbb{R}^N \times \mathbb{R}^M$. $f : E \to \mathbb{R}^M$, $f = (f_1, f_2, \ldots, f_M)$.

$$\frac{\partial f}{\partial x}(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_M} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_M}{\partial x_1} & \cdots & \cdots & \frac{\partial f_M}{\partial x_N} \end{pmatrix},$$

which is an $M \times N$ matrix. Similarly for y,

$$\frac{\partial f}{\partial y}(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_M} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_M}{\partial y_1} & \cdots & \cdots & \frac{\partial f_M}{\partial y_M} \end{pmatrix},$$

which is an $M \times M$ matrix (can find the determinant). Now we state the theorem.

Theorem 2.9.1 Implicit Function Theorem
Let $U \subseteq \mathbb{R}^N \times \mathbb{R}^M$ be an open set and $f : U \to \mathbb{R}^M$ of class C^k. Let $(a,b) \in U$ such that $f(a,b) = 0$. Assume that $\det \frac{\partial f}{\partial y}(a,b) \neq 0$. Then, $r_1 > 0, r_2 > 0$ and a unique function $g : B(a, r_1) \to B(b, r_2)$ of class C^k such that $g(a) = b$, $f(x, g(x)) = 0$ for all $x \in B(a, r_1)$ and $B(a, r_1) \times B(b, r_2) \subseteq U$.

Proof. We start with $M = 1$. This means that $y \in \mathbb{R}$. So, $f : U \to \mathbb{R}$. Assume $\frac{\partial f}{\partial y}(a,b) > 0$, $f \in C^k$. So we know that $\frac{\partial f}{\partial y}$ is continuous at the point (a,b). As such, we can find a ball centered at (a,b) such that $\frac{\partial f}{\partial y} > 0$ for all points in the ball. Consider $B(a,r) \times [b-r, b+r] \subseteq$ our ball. Consider $h(y) = f(a,y)$. $h'(y) = \frac{\partial f}{\partial y}(a,y) > 0$, meaning that h is strictly increasing and $h(b) = 0$. So, $h(b+r) = f(a, b-r) > 0$, and $h(b-r) = f(a, b-r) < 0$. f is also continuous at $(a, b+r)$ and $(a, b-r)$.

As such, we can find $r_3 > 0$ such that $f > 0$ in a ball $B((a, b+r), r_3)$. Additionally, we can find $r_4 > 0$ such that $f < 0$ in a ball $B((a, b-r), r_4)$. Let $r_5 = \min(r_3, r_4)$. Let $x \in B(a, r_5)$. Let $k(y) = f(x,y)$. Then $k'(y) = \frac{\partial f}{\partial y}(x,y) > 0$. $k(b+r) > 0$ and $k(b-r) < 0$. So, $\exists y$ such that $k(y) = 0$ that is unique since k is strictly increasing.

So far, we have proved that for every $x \in B(a, r_5)$, there exists a unique $y \in [r-b, r+b]$ such that $f(x,y) = 0$. Define $g(x) = y \implies f(x, g(x)) = 0$.

Step 2: We claim that g is continuous. Let $x_0 \in B(a, r_5)$. Fix small $\epsilon > 0$. Call $p(y) = f(x_0, y)$ which is strictly increasing. $p(g(x_0)) = f(x_0, g(x_0)) = 0$. So, $p(g(x_0) + \epsilon) > 0$ and $p(g(x_0) - \epsilon) < 0$. f continuous at $(x_0, g(x_0) + \epsilon)$, so $f > 0$ at a ball centered at $x(, g(x_0) + \epsilon)$. Similarly, $f < 0$ at a ball centered at $(x_0, g(x_0) - \epsilon)$. So, $f(x, g(x)) = 0$ for all x in a ball centered at x_0. Define r_6 for the former and r_7 for the latter. Let $r_8 = \min(r_6, r_7)$. Let $x \in B(x_0, r_8)$.

We know that $f(x, g(x_0) + \epsilon) > 0$ and $f(x, g(x_0) - \epsilon) < 0$. Also, $f(x, g(x)) = 0 \implies g(x_0) - \epsilon < g(x) < g(x_0) + \epsilon$ as desired.

Step 3: Prove that g is differentiable. So, let $x, x_0 \in B(a,r)$, $y_0, y \in (b-r, b+r)$. Consider the fubnction $q(s) = f(s, (x,y) + (1-s)(x_0, y_0))$ for $s \in [0,1]$. By the mean value theorem, $q(1) - q(0) = g'(c)(1)$ for some $c \in (0,1)$. Meaning,

$$f(x,y) - f(x_0, y_0) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(c(x,y) + (1-c)(x_0,y_0))(x_i - x_{0i}) + \frac{\partial f}{\partial y}(c(x,y) + (1-c)(x_0,y_0))(y - y_0).$$

Take $(x_0, y_0) = (x_0, g(x_0))$ and $(x, y) = (x_0 + te_k, g(x_0 + te_k))$. So, we have that:

$$\begin{aligned}
f(x_0 + te_k, g(x_0 + te_k)) - f(x_0, g(x_0)) = 0 & \\
&= 0 + 0 + \cdots + \frac{\partial f}{\partial x_k}(\overline{x}, \overline{y})(x_{0k} + t - x_{0k}) + \frac{\partial f}{\partial y}(\overline{x}, \overline{y})(g(x_0 + te_k) - g(x_0)) \\
&= \frac{\partial f}{\partial x_k}(\overline{x}, \overline{y})t + \frac{\partial f}{\partial y}(\overline{x}, \overline{y})(g(x_0 + te_k) - g(x_0)) \\
\implies \quad \frac{g(x_0 + te_k) - g(x_0)}{t} &= -\frac{\frac{\partial f}{\partial x_k}(\overline{x}, \overline{y})}{\frac{\partial f}{\partial y}(\overline{x}, \overline{y})}
\end{aligned}$$

Now take $t \longrightarrow 0$ to get:

$$\frac{\partial g}{\partial x_k}(x_0) = -\frac{\frac{\partial f}{\partial x_k}(x_0, g(x_0))}{\frac{\partial f}{\partial y}(x_0, g(x_0))}$$

for every $x_0 \in B(a, r)$. So, since the partials of f are continuous and g is continuous, the partials of g are continuous and therefore g is differentiable. 😄

We now use this result to prove the inverse function theorem.

Theorem 2.9.2 Inverse Function Theorem

Let $U \subseteq \mathbb{R}^N$ be open and $f : U \longrightarrow \mathbb{R}^N$ of class C^k. Assume $a \in U$ and that $\det Jf(a) \neq 0$. Then there is $r > 0$ such that $f : B(a, r) \longrightarrow \mathbb{R}^N$ is one-to-one, $f^{-1} : f(B(a, r)) \longrightarrow \mathbb{R}^N$ is of class C^k and f^{-1}.
Also, $f(B(a, r))$ is open.

Proof. Define $h(x, y) = f(x) - y$. Let $b = f(a)$. So, $h(a, b) = 0$. Then,

$$\begin{aligned}
\frac{\partial h}{\partial x}(x, y) &= \frac{\partial f}{\partial x}(x) \\
\det \frac{\partial h}{\partial x}(x, y) &= \det \frac{\partial f}{\partial x}(x) \neq 0 && \text{(by hypothesis)}
\end{aligned}$$

So, we can apply the implicit function theorem to find a function $g : B(b, r_1) \longrightarrow B(a, r_2)$, $g(b) = a$, of class C^k such that $h(g(y), y) = 0$. So, $f(g(y)) = y$.

Let $V = B(a, r_2) \cap f^{-1}(B(b, r_1))$ (pre-image). So $f(V) = B(b, r_1)$. So $f : V \longrightarrow B(b, r_1)$ is onto. We want to show that it is injective as well. Assume $\exists x_1, x_2 \in V$ such that $f(x_1) = f(x_2) = y \in B(b, r_1)$. But the implicit function theorem tells us that there is a unique $x \in B(a, r_2)$ such that $h(x, y) = 0$. So, $x_1 = x_2$. So, $g = f^{-1}$ is of class C^k. 😄

2.10 Lagrange Multipliers

Definition 2.10.1: Constrained Extrema

Let $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}^N$. Let $F \subseteq E$ and $x_0 \in F$. We say that:

- f attains a *constrained local minimum* at x_0 if there exists $r > 0$ such that $f(x) \geqslant f(x_0)$ for all $x \in B(x_0, r) \cap F$.
- f attains a *constrained local maximum* at x_0 if there exists $r > 0$ such that $f(x) \leqslant f(x_0)$ for all $x \in B(x_0, r) \cap F$.

The set F is called our *constraint*.

Theorem 2.10.1 Lagrange Multipliers

Let $U \subseteq \mathbb{R}^N$ be an open set, $f : U \to \mathbb{R}$ a function of class C^1 and let $g : U \to \mathbb{R}^M$ a function of class C^1 where $M \leqslant N$. Then let

$$F := \{x \in U : g(x) = 0\}.$$

Let $x_0 \in F$ and assume that f attains a constrained local minimum (or maximum) at x_0. If the vectors $\nabla g_i(x_0)$ for $i = 1, 2, \ldots, M$ are linearly independent, then there exist $\lambda_1, \lambda_2, \ldots, \lambda_M \in \mathbb{R}$ such that:

$$\nabla f(x_0) = \sum_{i=1}^{M} \lambda_i \nabla g_i(x_0).$$

Proof. write later ☺

2.11 Lebesgue Measure

Definition 2.11.1: Length, Rectangle, Volume

$I \subseteq \mathbb{R}$ interval. The *length* of I is defined as the $\text{length}(I) = \sup I - \inf I$. A *rectangle* in $\mathbb{R}^N$ is a set of the form $I_1 \times I_2 \times \cdots \times I_N$ where I_i are intervals in $\mathbb{R}$. The *volume* of a rectangle is defined as $\text{vol}(I_1 \times I_2 \times \cdots \times I_N) = \text{length}(I_1)\,\text{length}(I_2) \cdots \text{length}(I_N)$.

Definition 2.11.2: Lebesgue Outer Measure

Let $E \subseteq \mathbb{R}^N$. The *Lebesgue outer measure* of E is defined as:

$$\mathcal{L}_o^N(E) := \inf \left\{ \sum_{n=1}^{\infty} \text{vol}(R_m) : E \subseteq \bigcup_{n=1}^{\infty} R_m, R_m \text{ rectangles} \right\}$$

Exercise 2.11.1

R rectangle, R closed. Let R_n be open rectangles such that $R \subseteq \bigcup_{m=1}^{\infty} R_m$. Prove that there is $\ell \in \mathbb{N}$ such that $R \subseteq \bigcup_{m=1}^{\ell} R_m$.

Example 2.11.1

R rectangle, $\mathcal{L}_o^N(\partial R) = 0$.

Theorem 2.11.1

R rectangle. Then $\mathcal{L}_o^N(R) = \mathrm{vol}(R)$.

Proof. Let $R_1 = R$ and $R_2 = R_3 = \cdots = \emptyset$. Then we have that $\mathcal{L}_o^N(R) \leqslant \sum_{n=1}^\infty \mathrm{vol}(R_n) = \mathrm{vol}(R)$.

Now assume R is closed. Let R_n be rectangles $\bigcup_{n=1}^\infty R_n \supseteq R$. Fix $\epsilon > 0$, find $S_n \supseteq R_n$ open with $\mathrm{vol}(S_n) \leqslant \mathrm{vol}(R_n) + \epsilon 2^{-n}$.

Now we can apply the exercise to find some $\ell \in \mathbb{N}$ such that $R \subseteq \bigcup_{n=1}^\infty S_n$. So by other example,

$$\mathrm{vol}(R) \leqslant \sum_{n=1}^{\infty} \mathrm{vol}(S_n) \leqslant \sum_{n=1}^{\ell} \mathrm{vol}(S_n) \leqslant \sum_{n=1}^{\ell} \mathrm{vol}(R_n) + \epsilon$$

That is,

$$\mathrm{vol}(R) \leqslant \sum_{n=1}^{\infty} \mathrm{vol}(R_n)$$

as $\epsilon \longrightarrow 0$. So,

$$\mathrm{vol}(R) \leqslant \mathcal{L}_o^N(R) \leqslant \mathrm{vol}(R)$$

and we are done.

If R is not closed, $\mathrm{vol}(\bar{R}) = \mathrm{vol}(R) = \mathcal{L}_o^N(\bar{R}) \leqslant \mathcal{L}_o^N(R) + \mathcal{L}_o^N(\partial R) = \mathcal{L}_o^N(R)$. 😄

Properties of Lebesgue outer measure:

1. $E \subseteq F$, then $\mathcal{L}_o^N(E) \leqslant \mathcal{L}_o^N(F)$.
2. $E \subseteq \bigcup_{k=1}^\infty E_k$, then $\mathcal{L}_o^N(E) \leqslant \sum_{k=1}^\infty \mathcal{L}_o^N(E_k)$.

   Proof. If $\sum_{k=1}^\infty \mathcal{L}_o^N(E_k) = \infty$, then the inequality is trivial. So assume that the sum is finite. Fix $\epsilon > 0$. $\mathcal{L}_o^N(E_k)$ find rectangles $R_{n,k}$ such that $E_k \subseteq \bigcup_{n=1}^\infty R_{n,k}$ and $\sum_{n=1}^\infty \mathrm{vol}(R_{n,k}) \leqslant \mathcal{L}_o^N(E_k) + \epsilon 2^{-k}$. So,

   $$E \subseteq \bigcup_{k=1}^{\infty} E_k \subseteq \bigcup_{k=1}^{\infty}\bigcup_{n=1}^{\infty} R_{n,k} = \bigcup_{j=1} S_j$$

   So, $\mathcal{L}_o^N(E) \leqslant \sum_{j=1}^\infty \mathrm{vol}(S_j) = \sum_{k=1}^\infty \sum_{n=1}^\infty \mathrm{vol}(R_{n,k}) \leqslant \sum_{k=1}^\infty \mathcal{L}_o^N(E_k) + \epsilon$. Letting $\epsilon \longrightarrow 0^+$ gives the desired result. 😄

3. $E, F \subseteq \mathbb{R}^N$ with $\mathrm{dist}(E,F) > 0$, where dist is the infimum of the collection $\|x - y\|$ for $x \in E$, $y \in F$. Then $\mathcal{L}_o^N(E \cup F) = \mathcal{L}_o^N(E) + \mathcal{L}_o^N(F)$.

   Proof. Let R_n be rectangles, $E \cup F \subseteq \bigcup_{n=1}^\infty R_n$. Can we assume $\operatorname{diam} R_n < \mathrm{dist}(E,F)$ (partition R_n into smaller rectangles and use previous exercise). So,

   $$\bigcup_{n=1}^{\infty} R_n = \bigcup_{R_n \cap E \neq \emptyset} R_n \cup \bigcup_{R_n \cap F \neq \emptyset} R_n$$

   So, $\mathcal{L}_o^N(E) + \mathcal{L}_o^N(F) \leqslant \sum_{R_n \cap E \neq \emptyset} \mathrm{vol}(R_n) + \sum_{R_n \cap F \neq \emptyset} \mathrm{vol}(R_n) = \sum_{n=1}^\infty \mathrm{vol}(R_n) = \mathcal{L}_o^N(E \cup F)$. 😄

Note:

You can construct $E, F \subseteq \mathbb{R}^N$, $E \cap F = \emptyset$ such that

$$\mathcal{L}_o^N(E \cup F) < \mathcal{L}_o^N(E) + \mathcal{L}_o^N(F).$$

(Relies on axiom of choice)

Definition 2.11.3: Lebesgue Measure

A set $E \subseteq \mathbb{R}^N$ is *Lebesgue measurable* if for every $\epsilon > 0$, there exists an open set $U \supseteq E$ such that $\mathcal{L}_o^N(U \setminus E) < \epsilon$.
We define the *Lebesgue measure* of E to be $\mathcal{L}^N(E) = \mathcal{L}_o^N(E)$.

Properties of Lebesgue measure:

1. Open sets are Lebesgue measurable. (If E is open, take $U = E$, so $U \setminus E = \emptyset$.)
2. If $\mathcal{L}_o^N(E) = 0$, then E is Lebesgue measurable.

   Proof. Let $\epsilon > 0$. Find R_n rectangles such that $\bigcup R_n \supseteq E$ and $\sum \text{vol}(R_n) < \epsilon$. Find S_n open rectangle with $S_n \supseteq R_n$ and $\text{vol}(S_n) < \text{vol}(R_n) + \epsilon 2^{-n}$. So, let $U = \bigcup S_n \supseteq E$, so U is open. Then,

   $$\mathcal{L}_o^N(U \setminus E) \leqslant \mathcal{L}_o^N(U) \leqslant \sum \text{vol}(S_n) \leqslant \sum \text{vol}(R_n) + \epsilon < \epsilon + \epsilon = 2\epsilon$$

   and we are done. 😊

3. If E_n is Lebesgue measurable, $n \in \mathbb{N}$, then $\bigcup E_n$ is Lebesgue measureable.

   Proof. Fix $\epsilon > 0$, find $U_n \supseteq E_n$ open, $\mathcal{L}_o^N(U_n \setminus E_n) < \epsilon 2^{-n}$. So, $U = \bigcup U_n \supseteq \bigcup E_n$. So,

   $$\mathcal{L}_o^N(U \setminus \bigcup E_n) \leqslant \sum \mathcal{L}_o^N(U_n \setminus E_n) < \epsilon$$

   and we are done. 😊

Theorem 2.11.2

Sequentially compact sets are Lebesgue measurable.

Lenma 2.11.1

$C \subset \mathbb{R}^N$ closed, $K \subset \mathbb{R}^N$ sequentially compact such that $C \cap K = \emptyset$. Then $\operatorname{dist}(C, K) > 0$.

Proof. Assume $\operatorname{dist}(C, K) = 0$. Take $\epsilon = \frac{1}{n}$, find $x_n \in G$, $y_n \in K$ such that $\|x_n - y_n\| < \frac{1}{n}$. Since K is sequentially compact, $\exists y_{n_k} \to y \in K$. So,

$$\|x_{n_k} y\| \leqslant \|x_{n_k} - y_{n_k}\| + \|y_{n_k} - y\|$$

Both values on the RHS tend to 0, so $x_{n_k} \to y$, $x_{n_k} \in C$ as well. This implies $y \in C$, but $C \cap K$ is empty, so contradiction. ☺

Lenma 2.11.2

$U \subseteq \mathbb{R}^N$ open. Then, $U = \bigcap Q_n$ where Q_n are closed cubes with disjoint interiors.

Proof. Subdivide $\mathbb{R}^N$ into cubes of sidelength one. If $Q_n \subseteq U$, keep it. If $Q_n \subseteq \mathbb{R}^n \setminus U$, throw it away. If $Q_n \cap U \neq \emptyset$, $\overline{Q}_n \cap (\mathbb{R}^n \setminus U) \neq \emptyset$, subdivide Q_n into cubes of sidelength $\frac{1}{2}$. Keep going. Find a sequence of cubes Q_n (the ones you kept). We have that $\overline{Q}_n \subseteq U$ and we want $\bigcup \overline{Q}_n = U$.

Let $x \in U$, then $\operatorname{dist}(x, \partial U) > 0$. x was in one of the cubes of side one T_0. Either $\overline{T}_0 \subseteq U$ or T_0 was subdivided. In the first case, we are done. Otherwise $x \in T_1$, one of the subcubes $\frac{\sqrt{N}}{2^n} < \frac{r}{2}$. ☺

We turn to the proof of the theorem.

Proof. Take K sequentially compact. Fix $\epsilon > 0$. Find open rectangles R_n such that $\bigcup R_n \supseteq K$ and $\sum \operatorname{vol} R_n \leqslant \mathcal{L}_o^N(K) + \epsilon$.

Let $U = \bigcup R_n$.

$$\mathcal{L}_o^N(U) \leqslant \sum \operatorname{vol} R_n \leqslant \mathcal{L}_o^N(K) + \epsilon$$

So, $U \setminus K$ is open, and by a previous lemma, we have:

$$U \setminus K = \bigcup Q_n$$

where Q_n are closed cubes with disjoint interiors. Let $C_\ell = \bigcup_{n=1}^{\ell} Q_n$. We know that C_ℓ is closed and $C\ell \cap K = \emptyset$. By the other lemma, $\operatorname{dist}(C_\ell, K) > 0$, so

$$\begin{aligned}
\mathcal{L}_o^N(C_\ell \cup K) &= \mathcal{L}_o^N(C_\ell) + \mathcal{L}_o^N(K) \\
&= \sum_{n=1}^{\ell} \operatorname{vol} Q_n + \mathcal{L}_o^N(K) \\
\implies \sum_{n=1}^{\ell} \operatorname{vol} Q_n &\leqslant \mathcal{L}_o^N(C_\ell \cup K) - \mathcal{L}_o^N(K) \\
&\leqslant \mathcal{L}_o^N(U) - \mathcal{L}_o^N(K) \leqslant \epsilon
\end{aligned}$$

Taking $\ell \to \infty$ completes the proof;

$$\mathcal{L}_o^N(U \setminus K) \leqslant \sum_{n=1}^{\ell \to \infty} \operatorname{vol} Q_n \leqslant \epsilon$$

as desired. ☺

Theorem 2.11.3
Closed sets are Lebesgue measureable.

Proof. Let $C \in \mathbb{R}^N$ be closed. If C is bounded, then we're done. Otherwise, $C = \bigcup_{n=1}^{\infty} C \cap \overline{B(0,n)}$. Each $C \cap B(0,n)$ is bounded, so Lebesgue measurable. So, C is Lebesgue measurable, as the union of Lebesgue measurable sets is Lebesgue measurable. 😃

Theorem 2.11.4
If $E \subseteq \mathbb{R}^N$ is Lebesgue measureable, then $\mathbb{R}^N \setminus E$ is Lebesgue measureable.

Proof. Let E be Lebesgue measureable. Let $\epsilon = \frac{1}{n}$. Let U_n be open, $U_n \supseteq E$, and $\mathcal{L}_o^N(U_n \setminus E) \leqslant \frac{1}{n}$. Then we have $\mathbb{R}^N \setminus U_n \subseteq \mathbb{R}^N \setminus E$. Let $C_n := \mathbb{R}^N \setminus U_n$. Then C_n is closed, so Lebesgue measureable.

We have that $F := \bigcup C_n \subseteq \mathbb{R}^N \setminus E$, where the LHS is LM. We want to show that $\mathbb{R}^n \setminus E = F \cup S$, where S is a set with measure zero. That is, we claim $\mathcal{L}_o^N((\mathbb{R}^n \setminus E) \setminus F) = 0$.

So,

$$x \in (\mathbb{R}^n \setminus E) \setminus F \subseteq U_n \setminus E.$$

if $x \notin F$, then $x \notin C_n \implies x \in U_n$. So,

$$0 \leqslant \mathcal{L}_o^N((\mathbb{R}^N \setminus E) \setminus F) \leqslant \mathcal{L}_o^N(U_n \setminus E) \leqslant \frac{1}{n}.$$

So as we send $n \to \infty$, we have that $\mathcal{L}_o^N((\mathbb{R}^N \setminus E) \setminus F) = 0$, meaning that $\mathbb{R}^N \setminus E$ is Lebesgue measureable. 😃

Theorem 2.11.5
E_n Lebesgue measureable, $n \in \mathbb{N}$. Then $\bigcap E_n$ is Lebesgue measureable.

Proof. We know that the union of Lebesgue measureable sets is Lebesgue measureable. So, $\mathbb{R}^N \setminus \bigcup \mathbb{R}^N \setminus E_n$ is Lebesgue measureable. By the previous theorem, $\bigcap E_n$ is Lebesgue measureable. 😃

Theorem 2.11.6 Important

Let E_n be Lebesgue measureable such that $E_i \cap E_j = \emptyset$ for $i \neq j$. Then $\mathcal{L}^N\left(\bigcup E_n\right) = \sum \mathcal{L}^N(E_n)$.

Proof. **Step 1:** Assume that the E_n are bounded. Consider $\mathbb{R}^N \setminus E_n$, which we know is LM. By the definition of LM, given an $\epsilon > 0$, we can find some U_n open that contains $\mathbb{R}^N \setminus E_n$ and $\mathcal{L}_o^N(U_n \setminus (\mathbb{R}^N \setminus E_n)) \leqslant \frac{\epsilon}{2^n}$.

Let $K_n := \mathbb{R}^n \setminus U_n \subseteq E_n$. Then K_n is compact, so LM. Then for $i \neq j$, $K_j \cap K_i = \emptyset$, meaning $\operatorname{dist}(K_i, K_j) > 0$. So,

$$\mathcal{L}_o^N\left(\bigcup_{n=1}^{\ell} K_n\right) = \sum_{n=1}^{\ell} \mathcal{L}_o^N(K_n).$$

Then,

$$\mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty} E_n\right) \geqslant \mathcal{L}_o^N\left(\bigcup_{n=1}^{\ell} E_n\right) \geqslant \mathcal{L}_o^N\left(\bigcup_{n=1}^{\ell} K_n\right) = \sum_{n=1}^{\ell} \mathcal{L}_o^N(K_n) \geqslant \sum_{n=1}^{\ell} \mathcal{L}_o^N(E_n) - \mathcal{L}_o^N(E_n \setminus K_n) \geqslant \sum_{n=1}^{\ell} \mathcal{L}_o^N(E_n) - \epsilon.$$

Now let $\ell \to \infty$ to get:

$$\mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty} E_n\right) \geqslant \sum_{n=1}^{\infty} \mathcal{L}_o^N(E_n) - \epsilon.$$

Send $\epsilon \to 0$ to get the desired result.

Step 2: Now we consider the case where the E_n are not necessarily bounded. Let

$$\begin{aligned} F_1 &= B(0,1) \\ F_2 &= B(0,2) \setminus B(0,1) \\ &\vdots \\ F_i &= B(0,i) \setminus B(0,i-1) \\ &\vdots \end{aligned}$$

Then we have $E_n = E_n \cap \bigcup F_i = \bigcup_i (E_n \cap F_i)$. So,

$$\begin{aligned} \mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty} E_n\right) &= \mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty}\bigcup_{i=1}^{\infty}(E_n \cap F_i)\right) \\ &= \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \mathcal{L}_o^N(E_n \cap F_i) \\ &= \sum_{n=1}^{\infty} \mathcal{L}_o^N\left(\bigcup_{i=1}^{\infty} E_n \cap F_i\right) \\ &= \sum_{n=1}^{\infty} \mathcal{L}_o^N(E_n). \end{aligned}$$

😃

Corollary 2.11.1

If we have $E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n$, basically do the same as above with $F_1 = E_1, F_2 = E_2 \setminus E_1, \ldots$ to get:

$$\mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty} E_n\right) = \sup_n \mathcal{L}_o^N(E_n) = \lim_{n\to\infty} \mathcal{L}_o^N(E_n).$$

And,

$$\mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mathcal{L}_o^N(F_n) = \lim_{\ell\to\infty} \sum_{n=1}^{\ell} \mathcal{L}_o^N(F_n) = \lim_{\ell\to\infty} \mathcal{L}_o^N\left(\bigcup_{n=1}^{\ell} F_n\right) = E_\ell.$$

Note that $\mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty} F_n\right) = \mathcal{L}_o^N\left(\bigcup_{n=1}^{\infty} E_n\right)$.

2.12 Lebesgue Integration

Definition 2.12.1: Characteristic Function

Let E be a set. The characteristic function of the set E is defined as:

$$\chi_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}.$$

Definition 2.12.2: Simple Function

Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measureable set. A function $s : E \longrightarrow \mathbb{R}$ is called a *simple function* if it is of the form:

$$s(x) = \sum_{i=1}^{n} a_i \chi_{E_i}(x)$$

where $a_i \in \mathbb{R}$ and E_i are Lebesgue measureable sets.

Definition 2.12.3: Lebesgue Integral (simple)

Let $s : E \longrightarrow [0, \infty)$ be a simple function. That is:

$$s = \sum_{i=1}^{n} c_i \chi_{E_i},$$

with disjoint E_j. Then the Lebesgue integral of s over E is defined as:

$$\int_E s\, dx := \sum_{i=1}^{n} c_i \mathcal{L}^N(E_i).$$

If $c_i = 0$, $c_i \mathcal{L}^N(E_i) = 0$.

Properties of the Lebesgue integral:

1. $\int_E cs\,dx = c\int_E s\,dx$.
2. If s, t simple, then $\int_E (s+t)\,dx = \int_E s\,dx + \int_E t\,dx$.

Definition 2.12.4: Lebesgue Measureable Function

Let $E \subseteq \mathbb{R}^N$ be Lebesgue measureable. A function $f : E \longrightarrow [0,\infty)$ is called Lebesgue measureable if there exists a sequence of simple functions s_n with $0 \leqslant s_n \leqslant f$ such that $s_n \longrightarrow f$ pointwise in E.

Properties. Let $E \subseteq \mathbb{R}^N$ LM, $f, g : E \longrightarrow [0,\infty)$ LM:

1. $f + g$ is LM.
2. fg is LM.
3. $\max(f, g)$ is LM.
4. $\min(f, g)$ is LM.

Definition 2.12.5: Lebesgue Integral

Let $f : E \longrightarrow [0,\infty)$ be a Lebesgue measureable function. Then the Lebesgue integral of f over E is defined as:

$$\int_E f\,dx = \sup\left\{\int_E s\,dx : 0 \leqslant s \leqslant f, s \text{ simple}\right\}.$$

Note:

$f : E \longrightarrow [0,\infty)$ LM. WLOG, can assume $0 \leqslant s_n \leqslant s_{n+1} \leqslant f$ for all n. By def, $\exists s_n$ simple $0 \leqslant s_n \leqslant f$ such that $s_n \longrightarrow f$ pointwise. So redefine s_n to be increasing.
This still converges to f pointwise. If we have $x \in E$, then:

$$f(x) > \epsilon < s_n(x) \leqslant f(x),$$

but

$$f(x) - \epsilon < s_n(x) \leqslant \bar{s}_n(x) \leqslant f(x).$$

So, $\bar{s}_n \longrightarrow f$ pointwise.

Theorem 2.12.1

Let $f : E \longrightarrow [0, \infty)$ be a Lebesgue measureable function. Then the set:

$$E_a = \{x \in E : f(x) > a \geqslant 0\}$$

is Lebesgue measureable.