Honours Analysis I and II Class Notes

Fall 2024

Taught by: Dr. Axel Hundemer and Prof. Jérôme Vetois

Notes by: Nikhil Raman

Contents

1 Introduction

1.1 Sample Proofs

Method 1: Direct Proof

Start with a true statement and derive from this using a series of implications

> **Example 1.1**
> Show that $x \mapsto x^2$ is strictly increasing on $[0, \infty)$, i.e. show that $f : [0, \infty) \to \mathbb{R},\ f(x) = x^2$ is strictly increasing.

Proof. We know $\forall\ 0 \leq x < y : x^2 < y^2$. Let $0 \leq x < y$. Then:

$$y^2 - x^2 = (y - x)(y + x)$$

$(y - x) > 0$ since we defined $y > x$, and $(y + x) > 0$ since both y and x are positive. Therefore

$$\begin{aligned} &y^2 - x^2 > 0 \\ \implies &y^2 > x^2 \end{aligned}$$

□

Method 2: Proof by Contradiction

Assume the opposite of the result and lead this to a contradiction.

> **Example 1.2**
> Prove that $\sqrt{2}$ is irrational.

Proof. Assume that $\sqrt{2}$ is rational.

$$\begin{aligned} &\implies \exists\ a, b \in \mathbb{N} : \gcd(a, b) = 1 \text{ and } \sqrt{2} = \frac{a}{b} \\ &\implies (\sqrt{2})^2 = \frac{a^2}{b^2} \\ &\implies 2b^2 = a^2 \end{aligned}$$

$2b^2$ is even which means a^2 is even and a is even.

$$\begin{aligned} &\implies \exists\ c \in \mathbb{N} : a = 2c \\ &\implies 2b^2 = (2c)^2 \\ &\implies b^2 = 2c^2 \end{aligned}$$

This means that b is also even $\lightning$ since if both a and b are even, $\gcd(a, b) \neq 1$. Thus $\sqrt{2}$ is irrational. □

Method 3: Proof by Induction

Let P_n be a statement on the natural numbers. If we can prove:

(i) P_1 holds

(ii) $P_n \implies P_{n+1}$

then P_n holds $\forall\, n \in \mathbb{N}$.

Example 1.3
Prove that:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

Proof. Base case:

$$1 = \frac{1 \cdot 2}{2} = 1$$

$n \to n+1$: Assume that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ for some $n \in \mathbb{N}$. Then

$$1 + 2 + \cdots + n + n + 1 = \frac{n(n+1)}{2} + (n+1)$$
$$\implies \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2} = \frac{(n+1)[(n+1)+1]}{2}$$

□

Theorem 1.4 (Bernoulli's Inequality)
$\forall\, x \geq -1$ and $n \in \mathbb{N}$, $(1+x)^n \geq 1 + nx$.

Proof. By induction on $\mathbb{N}$.

Base case:

$$(1+x)^1 = (1+x) \geq 1 + x$$

$n \to n+1$: Assume $(1+x)^n \geq 1 + nx$. Then:

$$\implies (1+x)^{n+1} = (1+x)^n(1+x)$$
$$\geq (1+nx)(1+x) = 1 + x + nx + nx^2$$
$$= 1 + nx^2 + (n+1)x$$
$$\geq 1 + (n+1)x \text{ since } nx^2 \geq 0\ \forall\, n \in \mathbb{N}$$

□

Note
There is a special case of Bernoulli:

$$2^n \geq n + 1\ \forall\, n \in \mathbb{N}$$
$$\implies n < 2^n\ \forall\, n \in \mathbb{N}$$

1.2 The Well Ordering Principle of Natural Numbers

Theorem 1.5
Every non-empty subset $S \subset \mathbb{N}$ has a least (or minimal) element, i.e. $\exists\, s \in S : \forall\, t \in S, s \leq t$.

Note that most other sets of numbers d o not have this property. For example, any subset of $\mathbb{Z}$, e.g. $S = \{\ldots, -3, -2, -1, 0\}$ has no least elements. Consider $\mathbb{Q}$ with $S_1 = \{\frac{1}{n}, n \in \mathbb{N}\}$.

Claim 1.6. *S_1 does not have a least element*

Proof. Assume $s \in S_1$ is a least element. Then $s > 0$. Let $n \in \mathbb{N}$ such that $n \geq \frac{1}{s}$. This means $\frac{1}{n} < s$ ↯ Therefore s is not a least element. □

Proof of Well Ordered Principle. We can split this up into two cases.

(Case 1) S is a finite subset of NN. Let $n := |S|$.

Base case: $S = \{a_1\}$ for some $a_1 \in \mathbb{N}$. Then a_1 is the least element of S.

$n \to n+1$: Assume every subset S of $\mathbb{N}$ has a least element. Let $S = \{a_1, a_2, \ldots, a_n, a_{n+1}\}$. Further let $S' = \{a_1, a_2, \ldots, a_n\}$. Since S' has a least element s by our inductive hypothesis, either:

- $s \leq a_{n+1} \wedge s \leq a_i, \forall\, 1 \leq i \leq n$, which means s is the least element of S
- $s > a_{n+1} \wedge s \leq a_i, \forall\, 1 \leq i \leq n$ which means $\forall\, 1 \leq i \leq n+1,\ a_{n+1} \leq a_i$ so a_{n+1} is the least element.

(Case 2) S is infinite.

Let t_0 be any element of S. Then let $S' = S \cap \{1, 2, \ldots, t_0\} \implies S$ is a finite and non-empty subset of $\mathbb{N}$, meaning S' has a least element s. We will show that s is also the least element of S. Let $t \in S$ be arbitrary. Again there are two cases:

- $t \in S' \implies s \leq t$
- $t \notin S' \implies t > t_0 \implies s \leq t_0 < t$

Thus in either case, finite or infinite, a non-empty subset of the naturals has a least element. □

1.3 Introduction to Logic

Will do later

1.4 Set Theory

Definition 1.7 (Union and Intersection)
The Union of sets A and B is defined as $A \cup B := \{x : x \in A \vee x \in B\}$.

The Intersection of A and B is $A \cap B := \{x : x \in A \wedge x \in B\}$

We can write unions/intersections of more than 2 sets:

$$\begin{aligned}
&A_1 \cup A_2 \cup \cdots \cup A_n \text{ for } n \in \mathbb{N}\\
&:= \{x : x \in A_1 \lor x \in A_2 \lor \cdots \lor x \in A_n\}\\
&= \{x : \exists\, 1 \le i \le n : x \in A_i\}\\
&= \bigcup_{i=1}^{n} A_i
\end{aligned}$$

The same can be done for intersections:

$$\begin{aligned}
&A_1 \cap A_2 \cap \cdots \cap A_n \text{ for } n \in \mathbb{N}\\
&:= \{x : x \in A_1 \land x \in A_2 \land \cdots \land x \in A_n\}\\
&= \{x : \forall\, 1 \le i \le n,\ x \in A_i\}\\
&= \bigcap_{i=1}^{n} A_i
\end{aligned}$$

Infinitely Many Sets

Let $A_1, A_2, A_3, \ldots$ be sets. Then:

$$\begin{aligned}
&A_1 \cup A_2 \cup A_3 \cup \ldots\\
&= \bigcup_{i\in\mathbb{N}} A_i\\
&= \{x : \exists\, i \in \mathbb{N} : x \in A_i\}\\
&\bigcap_{i\in\mathbb{N}} A_i = \{x : \forall\, i \in N,\ x \in A_i\}
\end{aligned}$$

In fact we can define unions/intersections for any non-empty index set. Let S be a non-empty set and let A_i be a set $\forall\, i \in S$. Then $\bigcup_{i\in S} A_i = \{x : \exists\, i \in S : x \in A_i\}$.

Example 1.8
Show that $\bigcap_{x\in\mathbb{R}^+} [0, x] = \{0\}$

Proof. RHS $\subseteq$ LHS: $\forall\, x > 0, 0 \in [0, x] \implies 0 \subseteq \bigcap_{x\in\mathbb{R}^+} [0, x]$ 1 RHS $\subseteq$ LHS: Let $x_0 \in \bigcap_{x\in\mathbb{R}^+} [0, x]$. Then $x_0 \ge 0$. Assume $x_0 > 0$, and let $x := \frac{1}{2}x_0$. Then $0 \le x \le x_0 \implies x_0 \notin [0, x]$ $\lightning$. Thus $x_0 = 0$, which means any element of $\bigcap_{x\in\mathbb{R}^+} [0, x]$ must be 0. 2 With 1 + 2, we have $\bigcap_{x\in\mathbb{R}^+} [0, x] = \{0\}$. ◻

Complements

Definition 1.9 (Set Difference)
$A \setminus B := \{x : x \in A \land x \notin B\}$

Theorem 1.10 (De Morgan's Laws)
For all sets A, B, C, the following holds:

1. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
2. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

Proof. □

1.5 Functions

Definition 1.11
Let D, E be sets. A function f from D to E is a rule that assigns each $x \in D$ a uniquely determined $y \in E$, i.e. $y = f(x)$. D is called the domain and E is the codomain.

Some remarks about functions and sets:

- $f(D)$ can be a true subset of E, where $f(D) \subseteq E \land f(D) \neq E$. We can write this as $f(D) \subsetneq E$
- If $f(D) = E$, f is surjective, and if $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)\ \forall\, x_1, x_2 \in D$, f is injective.

1.6 Absolute Value

The absolute value function is defined as:

$$f : \mathbb{R} \to \mathbb{R}_0^+ , f(x) = |x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

Which can alternatively be formulated as:

$$|x| = \max\{x, -x\}$$

With the following properties, $\forall\, x, y \in \mathbb{R}$:

(1) $|x| = \sqrt{x^2}$

(2) $|xy| = |x||y|$

(3) $-|x| \leq x \leq |x|$

(4) $|x + y| \leq |x| + |y|$ and $|x - y| \geq |x| - |y|$ (The Triangle Inequality)

Proof. We will prove these by breaking down into cases where x is positive or negative:

(1) Let $x \geq 0$. Then $\sqrt{x^2} = x$ and $|x| = x$. Thus $\sqrt{x^2} = |x|$. Now let $x < 0$. $\sqrt{x^2} = -x = |x|$. In all cases, $\sqrt{x^2} = |x|$

(2) $|xy| \overset{(1)}{=} \sqrt{(xy)^2} = \sqrt{x^2}\sqrt{y^2} \overset{(1)}{=} |x||y|$

(3) Let $x \geq 0$. Then $|x| = x \implies x \leq |x|$, and therefore $-|x| \leq 0 \leq x \implies -|x| \leq x \leq |x|$. Now let $x < 0$. $|x| = -x > 0$, and $-|x| = -(-x) = x \leq x \implies -|x| \leq x \leq |x|$.

(4) $x + y \overset{(3)}{\leq} |x| + y \overset{(3)}{\leq} |x| + |y|$ and for the same reason $-x - y \leq |x| - y \leq |x| + |y|$. Now we have $-x - y \leq |x| + |y| \implies -(x+y) \leq |x| + |y|$. Thus $\max\{(x+y), -(x+y)\} \leq |x| + |y| \implies |x+y| \leq |x| + |y|$. To show the second part, note that $|x| = |x + y - y| \leq |x - y| + |y|$ by the triangle inequality, which implies $|x| - |y| \leq |x - y| \iff |x - y| \geq |x| - |y|$ which is what we wanted.

□

2 The Real Numbers

2.1 Bounded Sets

Definition 2.1
A nonempty subset $S \subseteq \mathbb{R}$ is said to be bounded from above if $\exists\, u \in \mathbb{R} : \forall\, s \in S, s \leq u$. Any such u is an upper bound of S.

Similarly, $S \subset \mathbb{R}$ is called bounded from below if $\exists\, v \in \mathbb{R} : \forall\, s \in S, v \leq s$. v is called a lower bound.

A nonempty subset S of $\mathbb{R}$ is <u>bounded</u> if it has an upper bound and a lower bound.

Example 2.2
Let $S = [0, 1]$. Any number $u \geq 1$ is an upper bound. On the other hand, show that $(0, \infty]$ is not bounded from above.

Proof. Assume $\exists$ an upper bound u. Then $u \geq 0$, however $u + 1 \in (0, \infty] \implies u \leq u + 1$ ↯. Thus this interval does not have an upper bound. □

Definition 2.3
Let S be a nonempty subset of $\mathbb{R}$ that is bounded from above. If there is an upper bound u such that $u \in S$, u is the maximum or greatest element.

Claim 2.4. *If S has a maximum, it is uniquely determined.*

Proof. Let u, u' be two maximums of S. Then u, u' are upper bounds of S, which means $\forall\, s \in S, s \leq u \wedge s \leq u'$. Since $u, u' \in S$, we have that $u \leq u' \wedge u' \leq u \implies u = u'$. □

Example 2.5
Consider the interval $S = [a, b]$. b is an upper bound contained in $S \implies b = \max S$. Now show that $S = [a, b)$ has no max.

Proof. Assume $m = \max S$. Then $m \in S \implies a \leq m < b$. Take the average of m and b to get $m' := \frac{m+b}{2}$, so that $a \leq m < m' < b$. Thus $\exists\, m' \in S : m < m'$, so is not the max. $\square$

Exercise 2.6
Show that $S = (a, b]$ has no min.

Proof. Let $s := \min S$. Then $a < s \leq b$ and s. Let $s' := \frac{a+s}{2}$ such that $a < s' < s \leq b$. Then s is no longer the min. $\square$

Definition 2.7 (Supremum and Infimum)
Let $S \neq \varnothing$ and $S \subseteq \mathbb{R}$.

1. Let S be bounded from above. An upper bound s is called a <u>supremum</u> ($s = \sup S$), or least upper bound if $s \leq u$ where u is any upper bound of S.
2. Let S be bounded from nelow. A lower bound t is called an <u>infimum</u> ($t = \inf S$) or greatest lower bound if $t \geq v$ where v is any lower bound.

Claim 2.8. *Both* sup *and* inf*, if they exist, are uniquely determined.*

Proof. Consider the subset S, and let s, s' be $\sup S$. Then since any sup is also an upper bound, $s \leq s' \wedge s' \leq s \implies s = s'$. Same argument holds for inf. $\square$

Example 2.9
$S = [a, b)$. Show that $a = \inf S$ and $b = \sup S$.

Proof. Clearly, a is a lower bound. Let v be any lower bound of S. For this to be true, $v \leq a$ which means a fits the requirements for an infimum. Secondly, b is an upper bound of S. Let u be any other upper bound; we must show that $b \leq u$. Assume $u < b$. Then $a \leq u \leq b$. Take the average of b and u, $\frac{b+u}{2}$. Then $a \leq u < \frac{b+u}{2} < b$. Therefore $\frac{b+u}{2} \in S$ and $u < \frac{b+u}{2}$ $\lightning$. Thus $b \leq u$. $\square$

Theorem 2.10
Let $S \neq \varnothing$ and $S \subseteq \mathbb{R}$.

1. Let S be bounded from above and have a maximum $\max S$. Then $\sup S = \max S$.
2. Let S be bounded from below and have a minimum $\min S$. Then $\inf S = \max S$.

Note that this implies the existence of the inf and sup.

Proof. We will show both sides, but they are very similar:

1. Let $m := \max S$, and u be any upper bound of S. Then $\forall s \in S$, $s \leq u$. Since $m \in S$, $m \leq u$. Thus m is the least upper bound and $m = \sup S$.
2. Let $t := \min S$ and v be any lower bound of S. Then $v \leq s\ \forall\, s \in S$. Since $t \in S$, $v \leq t \iff t \geq v$, which means t is the greatest lower bound and thus the inf of S.

□

The Fundamental Question

Given any nonempty subset of the real numbers that is bounded from above, is there always a supremum? Likewise, if its bounded from below, does the infimum always exist?

This is not really the right question, as we never made a proper definition of $\mathbb{R}$. In modern math, $\mathbb{R}$ is constructed from $\mathbb{Q}$ by the process of "completion." As a result, every non-empty subset of $\mathbb{R}$ that's bounded from above has a sup.

We will assume the axiom of completeness, i.e. $\mathbb{R}$ is complete.

2.2 Consequences of Completeness

Theorem 2.11 (Archimedean Property of $\mathbb{R}$)
Let $x \in \mathbb{R}$ be arbitrary. Then $\exists\, n \in \mathbb{N} : x < n$.

Proof. Assume $\exists x_0 \in \mathbb{R} : \forall\, n \in \mathbb{N}, n < x_0$, i.e. we assume that there is a real number bigger than every single natural number. Let S be the set $S := \{x \in \mathbb{R} : \forall\, n \in \mathbb{N}, n < x\}$. By assumption, S is nonempty, and we deduce that S is bounded from below by 0 (let $x \leq 0$, then $x < 1$ and $x \notin S$). By completeness of $\mathbb{R}$, let $t := \inf S$. t is then the greatest lower bound of S.

$$\begin{aligned} &\implies t + 1 \text{ is not a lower bound of } S \\ &\implies \exists x \in S : x < t + 1 \\ &\implies x - 1 < t \\ &\implies x - 1 \notin S \\ &\implies \exists\, n \in \mathbb{N} : x - 1 < n \\ &\quad\; x < n + 1 \; \lightning \end{aligned}$$

This is a contradiction since $x \in S$ was assumed to be larger than every natural, and $n + 1 \in \mathbb{N}$ defies this. Thus $S = \varnothing$ and $\forall\, x \in \mathbb{R},\ \exists\, n \in \mathbb{N} : x < n$. □

Corollary 2.12
Let $\varepsilon > 0$. Then $\exists\, n \in \mathbb{N} : 0 < \frac{1}{n} \leq \varepsilon$.

Proof. By the Archimedean Property, $\exists\, n \in \mathbb{N} : n > \frac{1}{\varepsilon} \implies \frac{1}{n} < \varepsilon$ □

2.3 Density

Definition 2.13
A subset $S \subset \mathbb{R}$ is called dense in $\mathbb{R}$ if $\forall\, a, b \in \mathbb{R}$ with $a < b$, $\exists\, x \in S : a < x < b$.

Theorem 2.14
$\mathbb{Q}$ is dense in $\mathbb{R}$

Proof. Let $a, b \in \mathbb{R}$, $a < b$. There are three cases:

Case 1: $0 < a < b$. Let $n \in \mathbb{N}$ and consider $\frac{1}{n}$. We can add this k times to get a rational number $\frac{k}{n}$ contained between a and b, provided that n is small enough. We need each $\frac{1}{n}$ to be smaller than $b - a$. By the Archimedean Property, let $n > b - a$. Then consider the set $\{\frac{1}{n}, \frac{2}{n}, \dots\}$. We know (again by the Archimedean Property) that $\exists\, k_0 \in \mathbb{N} : \frac{k_0}{n} > a$ $(k_0 > na)$. Thus we can define the nonempty set:

$$S := \{j \in \mathbb{N} : \frac{j}{n} > a\} \subseteq \mathbb{N}$$

S fits the criteria of the Well Ordering Principle, so it has a least element k, which means

$$\begin{aligned} \frac{k-1}{n} &= \frac{k}{n} - \frac{1}{n} < a \\ \implies \frac{k}{n} &< \frac{1}{n} + a = (b-a) + a = b \\ \implies \frac{k}{n} &< b. \implies a < \frac{k}{n} < b \end{aligned}$$

Case 2: $a < b \leq 0$. This implies $0 \leq -b < -a$, and by case 1:

$$\begin{aligned} &\exists\, x \in \mathbb{Q} : -b < x < -a \\ &\implies a < -x < b \end{aligned}$$

Case 3: $a < 0 < b$. Let $x = 0 \in \mathbb{Q}$. Then $a < x < b$. Thus, in all cases, we've found $x \in \mathbb{Q} : a < x < b$. □

Theorem 2.15 (Nested Interval Property)
Let $[a_1, b_1] \supseteq [a_2, b_2] \supseteq [a_3, b_3] \supseteq \dots$ with $a_1, b_1, a_2, b_2, a_3, b_3, \dots \in \mathbb{R} : \forall\, n \in \mathbb{N}, a_n \leq b_n$. Then $\bigcap_{n \in \mathbb{N}} [a_n, b_n] \neq \varnothing$.

Let $I_n := [a_n, b_n]\ \forall\, n \in \mathbb{N}$. We have $I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$ which means that $a_1 \leq a_2 \leq a_3 \dots$ (fact 1) and $b_1 \geq b_2 \geq b_3 \geq \dots$ (fact 2). Finally, $\forall\, n \in \mathbb{N}, a_n \leq b_n$ (fact 3). We will first prove some lemmas:

Lemma 2.16
$\forall\, n, k \in \mathbb{N}, a_n \leq b_k$.

Proof. Let $n \leq k$. Then $a_n \leq a_k \leq b_k$. Then let $n > k$. In this case $a_n \leq b_n \leq b_k$. In all cases, $a_n \leq b_k$. □

Let $A := \{a_1, a_2, \ldots\}$. By Lemma, b_1 constitutes an upper bound of A. Since $A \subseteq \mathbb{R}$ is nonempty and bounded, by completeness, let $a := \sup A$. Similarly, $B := \{b_1, b_2, \ldots\}$ is bounded from below. Let $b := \inf B$.

Lemma 2.17
$\forall k \in \mathbb{N}, a \leq b_k$.

Proof. By Lemma above, b_k is an upper bound of A $\forall k \in \mathbb{N}$. Since a is the least upper bound of A, $a \leq b_k$. □

Lemma 2.18
$a \leq b$.

Proof. By Lemma above, $a \leq b_k$ $\forall k \in \mathbb{N}$, so a is a lower bound of B. Since b is the greatest lower bound of B, $b \geq a$. □

Proof of Theorem. Therefore, with facts 1,2,3, and our lemmas,

$$\begin{aligned} & a_n \leq a \leq b \leq b_n \\ \implies & [a,b] \subseteq [a_n, b_n] \ \forall\, n \in \mathbb{N} \\ \implies & [a,b] \subseteq \bigcap_{n\in\mathbb{N}} [a_n, b_n] \\ \implies & \bigcap_{n\in\mathbb{N}} [a_n, b_n] \neq \varnothing \end{aligned}$$

□

2.4 Cardinality

Our goal is to measure the size of sets. This is simple in the finite case, where $|S|$:= the number of elements in S. Cantor's ides was to use functions to compare the sizes of sets. Consider the function $f : A \to B$ where $|A| = n$, $|B| = k$. If $n > k$, there exists a surjective map $f : A \to B$. f is defined by:

$$f(a_j) = \begin{cases} b_j & \forall\, 1 \leq j \leq k \\ b_k & \forall\, k+1 \leq j \leq n \end{cases}$$

Claim 2.19. *There is no injective map from $A \to B$.*

Proof. Assume there is an injective $f : A \to B$ where $A = \{a_1, \ldots, a_k, a_{k+1}, \ldots, a_n\}$ and $\forall\, 1 \leq j \leq k$, $b_j := f(a_j)$.

$$\begin{gathered} \implies f(A) = \{b_1, b_2, \ldots, b_n\}, \ |\{b_1, b_2, \ldots, b_n\}| = n \\ \{b_1, b_2, \ldots, b_n\} \subseteq B, \ |B| = k \implies |\{b_1, b_2, \ldots, b_n\}| \leq k < n \\ \implies n < n \ \lightning \end{gathered}$$

Thus no injective map can exist between A and B as long as $|A| = n < k = |B|$. □

Definition 2.20
A and B have the same cardinality, i.e. $|A| = |B|$ if there exists a bijection $f : A \to B$. Furthermore we say the cardinality of A is strictly less than B, i.e. $|A| < |B|$ if there does not exist any surjection between A and B

Something interesting arises from this definition of cardinality. Take the sets $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ and $\mathbb{N} = \{1, 2, \ldots\}$. Clearly $\mathbb{N} \subsetneq \mathbb{N}_0$. However, they do have the same number of elements. Why? Define $f : \mathbb{N} \to \mathbb{N}_0, f(n) := n - 1$. Note that $f(n) \in \mathbb{N}_0 \forall\, n \in \mathbb{N}$. f is bijective.

Injective: let $f(n) = f(k)$. Then $n - 1 = k - 1 \implies n = k$. Surjective: let $n \in \mathbb{N}_0$ be arbitrary. Then $\exists\, m \in \mathbb{N} : f(m) = n$, especially $m = n + 1$. Thus f is indeed bijective, and by definition $|\mathbb{N}| = |\mathbb{N}_0|$. Note that if A and B are finite, with $A \subsetneq B$, it is impossible for $|A| = |B|$, so this is a quirk of infinite sets. We can take this further:

Example 2.21
Show that $|\mathbb{N}| = |\mathbb{Z}|$.

Proof. The elements of $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ will be arranged in a way such that a bijective between it and $\mathbb{N}$ is clear. Consider the complete list of integers:

$$0, 1, -1, 2, -2, 3, -3, \ldots$$

Let

$$a_1 = 0, a_2 = 1, a_3 = -1, \ldots$$

and define the map $f : \mathbb{N} \to \mathbb{Z}, f(n) = a_n$. f is injective since each a_i corresponds to a unique natural number, and surjective since all integers are hit by the a_i's. Thus f is a bijection and $|\mathbb{N}| = |\mathbb{Z}|$. We can even explicitly construct f:

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ \frac{-n-1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

This is also bijective: let $f(n) = f(k)$. Then □

Definition 2.22
A set A is countably infinite if $|A| = |\mathbb{N}|$, i.e. $\exists\, f : A \to \mathbb{N}$, bijective. A set B is countable if it is finite or countably infinite.

Theorem 2.23
Let A be a set such that $\exists\, f : \mathbb{N} \to A$, surjective. Then A is countable.

Proof. Since f is surjective, □

Theorem 2.24
Let $A \subseteq \mathbb{N}$. Then A is countable.

Proof. Either A is a finite subset of $\mathbb{N}$, or it is infinte. The finite case is trivial. Now suppose Suppose $A = \{a_1, a_2, \dots\}$ is an infinite subset of $\mathbb{N}$. We can define the function $f : \mathbb{N} \to A$:

$$f(n) = \begin{cases} a_n & \text{if } n \in A \\ a_1 & \text{if } n \notin A \end{cases}$$

f is surjective: divide $\mathbb{N}$ into A and $(\mathbb{N} \setminus A)$. We know that $f(A \cup B) = f(A) \cup f(B)$ for any function f. Thus $f(\mathbb{N}) = f(A \cup (\mathbb{N} \setminus A)) = f(A) \cup f(\mathbb{N} \setminus A) = A \cup \{a_1\} = A$. Thus $f(\mathbb{N}) = A$ as required for surjectivity. By our theorem above, A is countable. □

Theorem 2.25
$\mathbb{Q}$ is countably infinite.

Proof. Restrict ourselves to $\mathbb{Q}^+$ for now. Consider the "matrix like" arrangement pictured below.

$$\begin{array}{cccccc}
\frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\
\frac{2}{1} & \frac{2}{2} & \frac{2}{3} & \frac{2}{4} & \frac{2}{5} & \cdots \\
\frac{3}{1} & \frac{3}{2} & \frac{3}{3} & \frac{3}{4} & \frac{3}{5} & \cdots \\
\frac{4}{1} & \frac{4}{2} & \frac{4}{3} & \frac{4}{4} & \frac{4}{5} & \cdots \\
\frac{5}{1} & \frac{5}{2} & \frac{5}{3} & \frac{5}{4} & \frac{5}{5} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots &
\end{array}$$

Then, we can create the following enumeration by counting along the diagonals:

$$a_1 = 1, a_2 = \frac{1}{2}, a_3 = \frac{1}{3}, a_4 = \frac{2}{2}, a_5 = \frac{3}{1}, \dots$$

So that we can define the surjection $f : \mathbb{N} \to \mathbb{Q}^+$, $f(n) = a_n$. This is a surjection and not a bijection because in our scheme every element of $\mathbb{Q}^+$ was counted infinitely often.

This shows that $\mathbb{Q}^+$ is countable, but since $\mathbb{N} \subset \mathbb{Q}^+$, and $\mathbb{N}$ is countably infinite, $\mathbb{Q}^+$ is countably infinite. To show that $\mathbb{Q}$ is countable, we alter our enumaration slightly:

$$0, a_1, -a_1, a_2, -a_2, \dots$$

This provides a complete enumeration of $\mathbb{Q}$, and thus $\mathbb{Q}$ is countably infinite. □

Despite all this, $\mathbb{R}$ is uncountable. We will prove a stronger result:

Theorem 2.26
Every interval $[a, b] \subseteq \mathbb{R}$, $a < b$ is uncountable.

Proof. Assume $[a, b]$ is countable. Then there exists an enumeration (without repetition) $s_1, s_2, \dots$ for every element in $[a, b]$. We will construct a nested sequence of closed and bounded intervals so that at least one does not contain s_1. The way to guarantee that the subinterval does not contain s_1 is to create 3 different ones. Let that interval be I so that $s_1 \notin I$. Further divide I into 3 subintervals, one of which is I_2 so that $s_2 \notin I_2$. We continue this so that we have a nested sequence of intervals:

$$I \supseteq I_2 \supseteq I_3 \supseteq \dots \text{ with } s_n \notin I_n \ \forall\, n \in \mathbb{N} \text{ (by construction)}$$

Then, by the Nested Interval Property of $\mathbb{R}$:

$$\bigcap_{n\in\mathbb{N}} I_n \neq \varnothing$$

Let x be in that intersection of intervals. Since we have a complete enumeration $[a, b]$, $x = s_j$ for some $j \in \mathbb{N}$. However, $s_j \notin \bigcap_{n\in\mathbb{N}} I_n$ $\lightning$.

Thus, $[a, b]$ is uncountable, and resultingly $\mathbb{R} \supseteq [a, b]$ is uncountable. □

3 Sequences

Definition 3.1 (Sequence)
A sequence of real numbers is a function $f : \mathbb{N} \to \mathbb{R}$ with $a_1 := f(1), a_2 := f(2), \dots$ We write $(a_n)_{n\in\mathbb{N}}$ or simply (a_n).

3.1 The Limit of a Sequence

Definition 3.2 (Limit of a Sequence)
Let (a_n) be a sequence. We say that the limit of $\lim(a_n) = L$, or (a_n) converges to L, if $\forall\, \varepsilon > 0,\ \exists\, N \in \mathbb{N} : \forall\, n \geq N, |a_n - L| < \varepsilon$.

Example 3.3
Show that $\lim(\frac{1}{n}) = 0$.

Proof. Let $\varepsilon > 0$. Then we need $|a_n - L| = |\frac{1}{n} - 0| = \frac{1}{n} < \varepsilon$. Let $N > \frac{1}{\varepsilon}$ (exists by Archimedean Property). Then $\forall\, n \geq N > \frac{1}{\varepsilon} \implies \frac{1}{n} < \varepsilon$, i.e. $\forall\, n \geq N, |a_n - L| < \varepsilon$. □

Example 3.4
Show that $\lim(\frac{n^2}{n^2+1}) = 1$.

Proof. Let $\varepsilon > 0$. We need:

$$\left|\frac{n^2}{n^2+1} - 1\right| = \left|\frac{-1}{n^2+1}\right| = \frac{1}{n^2+1} < \varepsilon$$
$$n^2 + 1 > n^2 \implies \frac{1}{n^2+1} < \frac{1}{n^2} < \varepsilon \iff n > \frac{1}{\sqrt{\varepsilon}}$$

Let $N > \frac{1}{\sqrt{\varepsilon}}$. Then $\forall\, n \geq N, |\frac{n^2}{n^2+1} - 1| = \frac{1}{n^2+1} < \frac{1}{n^2} < \varepsilon$. □

Theorem 3.5
The limit of a sequence, if it exists, is uniquely determined. In other words, if $\lim(a_n) = L_1$ and $\lim(a_n) = L_2$, $L_1 = L_2$.

Lemma 3.6
Let $x \in \mathbb{R}$ such that $\forall\, \varepsilon > 0, 0 \leq x < \varepsilon$. Then $x = 0$.

Proof. Assume that $x > 0$. Then $\forall\, \varepsilon > 0, 0 < x < \varepsilon$. Let $\varepsilon := \frac{x+0}{2}$ so that $0 < \frac{x}{2} < x < \varepsilon \implies 0 < \varepsilon < x < \varepsilon \lightning$. Thus $x = 0$. □

Proof of Theorem. Let $\varepsilon > 0$ and let $\lim(a_n) = L_1$, $\lim(a_n) = L_2$. Then $\exists\, N_1, N_2 \in \mathbb{N} : \forall\, n \geq N_1, |a_n - L_1| < \frac{\varepsilon}{2}$ and $\forall\, n \geq N_2, |a_n - L_2| < \frac{\varepsilon}{2}$. Then:

$$\begin{aligned} |L_1 - L_2| &= |a_n + L_1 - a_n - L_2| = |(-a_n + L_1) + (a_n - L_2)| \\ &\overset{\Delta \text{ Ineq.}}{\leq} |(-a_n + L_1)| + |(a_n - L_2)| = \underbrace{|(a_n) - L_1|}_{<\frac{\varepsilon}{2}} + \underbrace{|(a_n - L_2)|}_{<\frac{\varepsilon}{2}} \\ &\implies 0 \leq |L_1 - L_2| < \varepsilon \end{aligned}$$

By Lemma, $|L_1 - L_2| = 0 \implies L_1 = L_2$. □

Definition 3.7 (eps Neighborhood)
Let $a \in \mathbb{R}$, $\varepsilon > 0$. The ε-neighborhood about a is denoted $V_a(\varepsilon) := \{x \in \mathbb{R} : |x - a| < \varepsilon\} = (a - \varepsilon, a + \varepsilon)$.

Using this notion the definition of a limit can be rewritten as $\lim(a_n) = L \iff \forall\, \varepsilon > 0, \exists\, N : \forall\, n \geq N, a_n \in V_L(\varepsilon)$.

Definition 3.8
A sequence is divergent if it doesn't converge.

Example 3.9
Show that $((-1)^n)$ diverges.

Proof. Let $L \in \mathbb{R}$. Need to find ε so that limit definition fails.

Let $\varepsilon := 1$. Then $\exists\, N : \forall\, n \geq N, |(-1)^n - L| < 1$. First let $n \geq N$ be even. Then $|(-1)^n - L| = |1 - L| < 1 \implies -2 < L < 0$. Then consider an $n \geq N$ that is odd. We have $|-1 - L| < 1 \implies 0 < L < 2$. Under the two conditions, $0 < L < 0 \lightning$.

Thus $((-1)^n)$ diverges. □

Theorem 3.10
Let $0 < a < 1$. Then $\lim(a^n) = 0$.

Proof. Let $\varepsilon > 0$. Note that if $0 < a < 1$, $\frac{1}{a} > 1$. Let $b := \frac{1}{a} - 1 > 0$. Then:

$$\begin{aligned} a &= \frac{1}{b+1} \\ \implies a^n &= \frac{1}{(b+1)^n} \\ &\overset{\text{Bernoulli}}{\leq} \frac{1}{1+nb} < \frac{1}{nb} < \varepsilon \\ \iff n &> \frac{1}{b\varepsilon} \end{aligned}$$

Let $N > \frac{1}{b\varepsilon}$. Then $|a^n - 0| < \varepsilon$. □

Exercise 3.11
Prove that $\lim(a^n) = 0 \ \forall \ |a| < 1$.

3.2 Algebraic Limit Laws

Theorem 3.12
Let (a_n) be a convergent sequence. Then (a_n) is bounded, i.e. $\exists \, M > 0 : \forall \, n \in \mathbb{N}, |a_n| \leq M$.

Note that the converse is not true, the counterexample being $((-1)^n)$. It is bounded by 1 but diverges.

Proof. Let $\lim(a_n) = L$ and let $\varepsilon := 1$. Then, $\exists N : \forall \, n \geq N, |a_n - L| < 1$. Note that:

$$\begin{aligned} |a_n| = |a_n - L + L| &\overset{\Delta \text{ ineq.}}{\leq} |a_n - L| + |L| \\ &< 1 + |L| \end{aligned}$$

Let $M := \max\{a_1, a_2, \ldots, a_{N-1}, 1 + L\}$. Then, $\forall n \in \mathbb{N}$, $a_n \leq M$ meaning (a_n) is bounded. □

Exercise 3.13
Show that (n) diverges.

Proof. Suppose (n) is convergent. Then, (n) must be bounded, i.e. $\exists \, M > 0 : |n| \leq M, \forall \, n \in \mathbb{N}$. However, if $n > M$, we have a contradiction. Thus (n) cannot be convergent. □

Theorem 3.14 (Algebraic Limit Laws)
Let (a_n), (b_n) be convergent sequences, and $c \in \mathbb{R}$. Then the following holds:

1. $(a_n + b_n)$ converges with $\lim(a_n + b_n) = \lim(a_n) + \lim(b_n)$
2. $(a_n -_n N)$ converges with $\lim(a_n - b_n) = \lim(a_n) - \lim(b_n)$
3. (ca_n) converges with $\lim(ca_n) = c \lim(a_n)$
4. $(a_n b_n)$ converges with $\lim(a_n b_n) = \lim(a_n) \lim(b_n)$
5. If $b_n \neq 0 \; \forall \; n \in \mathbb{N}$ and $\lim(b_n) \neq 0$, then $(\frac{a_n}{b_n})$ converges with $\lim(\frac{a_n}{b_n}) = \frac{\lim(a_n)}{\lim(b_n)}$

Proof. (d) Let $a := \lim(a_n)$ and $b := \lim(b_n)$. Let $\varepsilon > 0$. Consider:

$$|a_n b_n - ab| = |a_n b_n - ab_n + ab_n - ab| \text{ (Symmetry Breaking)}$$
$$\implies |b_n(a_n - a) + a(b_n - b)| \leq \underbrace{|b_n||a_n - a|}_{(*)} + \underbrace{|a||b_n - b|}_{(**)}$$

Let us examine $(*)$ since (b_n) is convergent, it is bounded, so $\exists \, M > 0 : |b_n| \leq M \forall n \in \mathbb{N}$. Also, (a_n) converges to a, so $\exists N_1 : \forall n \geq N_1, |a_n - a| < \frac{\varepsilon}{2\tilde{M}} \implies |b_n||a_n - a| < \frac{\varepsilon}{2}$.

$(**)$ For the same reason, let $\tilde{M} > 0 : |a_n| \leq \tilde{M}$. Since (b_n) converges to b, $\exists N_2 : \forall n \geq N_2, |b_n - b| < \frac{\varepsilon}{2\tilde{M}} \implies |a||b_n - b| < \frac{\varepsilon}{2}$.

Let $N = \max\{N_1, N_2\}$. Then, $\forall n \geq N, (*) + (**) < \varepsilon \implies |a_n b_n - ab| < \varepsilon \implies \lim(a_n b_n) = ab$. □

Definition 3.15 (Null Sequence)
A sequence (a_n) with limit 0 is a null sequence.

Exercise 3.16
Let (a_n) be a null sequence and $c \in \mathbb{R}$. Show that (ca_n) is a null sequence.

The next result will be essential later on for constructing convergent sequences contained in sets.

Theorem 3.17 (Null Sequence Criterion)
Let (a_n) be a sequence, $L \in \mathbb{R}$ and (b_n) a null sequence, with both $a_n, b_n \geq 0 \; \forall n \in \mathbb{N}$. Then, if $\exists \, K \in \mathbb{N} : \forall \, n \geq K, |a_n - L| \leq b_n$, $\lim(a_n) = L$.

Proof. Let $\varepsilon > 0$. Since (b_n) is a null sequence, $\exists \tilde{N}$ such that $\forall n \geq \tilde{N}, |b_n| < \varepsilon$. Let $N := \max\{K, \tilde{N}\}$. Then $\forall n \geq N, |a_n - L| < b_n < \varepsilon$

$\implies \forall n \geq N, |a_n - L| < \varepsilon$

$\implies \lim(a_n) = L$ □

3.3 Limits and Order

Theorem 3.18
Let (a_n), (b_n) be convergent sequences such that $\exists k \in \mathbb{N} : \forall n \geq k$, $a_n \leq b_n$ or $a_n \geq b_n$. Then $\lim(a_n) \leq \lim(b_n)$ or $\lim(a_n) \geq \lim(b_n)$.

Note that this does not hold for strict inequalities: $\forall n > 1$, $\frac{1}{n^2} < \frac{1}{n}$, but $\lim(\frac{1}{n^2}) = \lim(\frac{1}{n}) = 0$.

Proof. □

Theorem 3.19 (Squeeze Theorem)
Let (a_n), (b_n), (c_n) be sequences such that $\exists k \in \mathbb{N} : \forall n \geq k$, $a_n \leq b_n \leq c_n$ and $\lim(a_n) = \lim(c_n) = L$. Then (b_n) converges with $\lim(b_n) = L$.

Proof. Let $\varepsilon > 0$. Since (a_n) is convergent, $\exists N_1 : \forall n \geq N_1$, $a_n \in V_\varepsilon(L)$, i.e. $L - \varepsilon < a_n < L + \varepsilon$. Also since (a_n) is convergent, $\exists N_2 : \forall n \geq N_2$, $L - \varepsilon < c_n < L + \varepsilon$..

Let $N = \max\{N_1, N_2, k\}$. Then, $\forall n \geq N$, the following holds:

$$\begin{aligned} L - \varepsilon < a_n \leq b_n \leq c_n < L + \varepsilon \\ \implies L - \varepsilon < b_n < L + \varepsilon \\ \implies b_n \in V_\varepsilon(L) \\ \implies \lim(b_n) = L \end{aligned}$$

□

Theorem 3.20 (Monotone Convergence)
A monotone sequence is a sequence that is either increasing or decreasing.

1. Let (a_n) be increasing and bounded from above. Then (a_n) is convergent.
2. Let (b_n) be decreasing and bounded from below. Then (b_n) is convergent.

Proof. Let $A := \{a_1, a_2, a_3, \ldots\}$. Since (a_n) is bounded from above, A is a bounded subset of $\mathbb{R}$, meaning completeness applies; let $a := \sup A$. Since a is the sup, $\forall\ \varepsilon > 0$, $a - \varepsilon$ is not an upper bound, meaning $\exists\ a_k : a - \varepsilon < a_k$. Also, since a is an upper bound, $a_k \leq a_{k+1} \leq a_{k+2} \leq \cdots \leq a < a + \varepsilon$. Combing these two facts, we get:

$$\begin{aligned} a - \varepsilon < a_k < a + \varepsilon \\ \implies a_k \in V_\varepsilon(a)\ \forall\ n \geq k \\ \implies \lim(a_n) = a \end{aligned}$$

□

3.4 Euler's Number e

$$\begin{aligned}
\frac{a_{n+1}}{a_n} &= \frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{\left(1+\frac{1}{n}\right)^n} = \frac{\left(\frac{n+2}{n+1}\right)^{n+1}}{\left(\frac{n+1}{n}\right)^n} = \frac{(n+2)^{n+1}\cdot n^n}{(n+1)^{n+1}\cdot(n+1)^n} = \frac{n+2}{n+1}\cdot\frac{(n+2)^n\cdot n^n}{(n+1)^{2n}} \\
&= \frac{n+2}{n+1}\cdot\left[\frac{n\cdot(n+2)}{(n+1)^2}\right]^n = \frac{n+2}{n+1}\cdot\left[\frac{n^2+2n+1-1}{n^2+2n+1}\right]^n = \frac{n+2}{n+1}\cdot\left[1-\frac{1}{n^2+2n+1}\right]^n \\
&= \frac{n+2}{n+1}\cdot\left[1+\left(-\frac{1}{(n+1)^2}\right)\right]^n \overset{\text{Bernoulli}}{\geq} \frac{n+2}{n+1}\cdot\left[1-n\frac{1}{(n+1)^2}\right] \\
&= \frac{n+2}{n+1}\cdot\frac{n^2+2n+1-n}{(n+1)^2} = \frac{(n+2)\cdot\left(n^2+n+1\right)}{(n+1)^3} \\
&= \frac{n^3+n^2+n+2n^2+2n+2}{n^3+3n^2+3n+1} = \frac{n^3+3n^2+3n+2}{n^3+3n^2+3n+1} > 1
\end{aligned}$$

3.5 Subsequences

Definition 3.21 (Subsequence)
Let (a_n) be a sequence and let $n_1, n_2, n_3, \dots \in \mathbb{N}$ such that $n_1 \leq n_2 \leq n_3 \leq \dots$. Then $(a_{n_k}) = (a_{n_1}, a_{n_2}, a_{n_3}, \dots)$ is a subsequence of (a_n).

Theorem 3.22
Let (a_n) be a convergent sequence and (a_{n_k}) a subsequence. Then (a_{n_k}) converges and $\lim(a_{n_k}) = \lim(a_n)$.

Lemma 3.23
$\forall\, k \in \mathbb{N}, n_k \geq k$

Proof. We will perform induction w.r.t. k.

Base case ($k = 1$): $n_1 \geq 1$ is true since $n_1 \in \mathbb{N}$.

Assume $n_k \geq k$ for some $k \in \mathbb{N}$. Then, $n_{k+1} > n_k \geq k$. Since we have a natural number strictly larger than another, they must differ by at least 1, meaning $n_{k+1} \geq n_k + 1 \geq k + 1 \implies n_{k+1} \geq k + 1$ □

Proof of Theorem. Let $\varepsilon > 0$, $\lim(a_n) := L$. Then $\exists\, N : \forall\ \geq N, |a_n - L| < \varepsilon$. Let $k \geq N$. By Lemma, we know $n_k \geq k \geq N$, so $|a_{n_k} - L| < \varepsilon\ \forall\, k \geq N$. □

Corollary 3.24
Let (a_n) be a sequence with two subsequences (a_{n_k}) and (a_{n_j}) that converge differently. Then (a_n) is divergent.

Proof. If (a_n) was convergent, any subsequence would have the same limit as (a_n). Thus (a_n) cannot be convergent. □

Theorem 3.25 (Bolzano-Weierstrass)
Any bounded sequence of real numbers has a (monotone) convergent subsequence.

Proof 1 (Excluding Monotonicity). Since (x_n) is bounded, let $M > 0$ be such that $\forall\, n \in \mathbb{N}$, $|x_n| \leq M$, i.e. $-M \leq x_n < M$. If we split $I_1 := [-M, M]$ into two subintervals, at least one contains infinitely many points. Call that interval I_1. Now split I_2 into two subintervals; again, at least one will contain infinitely many points; call that I_3, and so on. Then we have a sequence of nested intervals $I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$. □

3.6 Cauchy Sequences

Definition 3.26 (Cauchy Sequence)
A sequence (x_n) is a cauchy sequence if $\exists N \in \mathbb{N} : \forall\, n, m \geq N$, $|x_n - x_m| \leq \varepsilon$.

Lemma 3.27
Every cauchy sequence is bounded.

Theorem 3.28
A sequence of real numbers is convergent if and only if it is cauchy.

3.7 Applications: Contractive Sequences

Definition 3.29 (Contractive Sequence)
A sequence (x_n) is contractive if $\forall\, n \in \mathbb{N}$, $|x_{n+2} - x_{n+1}| \leq c|x_{n+1} - x_n|$.

Theorem 3.30
Every contractive sequence is convergent.

Lemma 3.31 (Sums of Finite Geometric Series)
Let $a \neq 1, n \in \mathbb{N}$. Then the series $1 + a + a^2 + \cdots + a^n$ sums to $\frac{1-a^{n+1}}{1-a}$.

Example 3.32
Let $x_1 := 2$, $x_{n+1} = 2 + \frac{1}{x_n}$. Find $\lim(x_n)$.

Solution.

3.8 Infinite Limits

Definition 3.33
Let (x_n) be a sequence such that $\forall\, M \in \mathbb{R}, \exists N : \forall n \geq N, x_n > M$. We say that (x_n) diverges to infinity, or $\lim(x_n) = +\infty$.

Example 3.34
Show that $\lim(n) = +\infty$

Solution. Let $M \in \mathbb{R}$. By the Archimedean Property, $\exists N \in \mathbb{N} : N > M$. Then, $\forall n \geq N, x_n = n \geq N > M$. Hence $\lim(x_n) = +\infty$.

Definition 3.35
Let (x_n) be a sequence such that $\forall\, M \in \mathbb{R}, \exists N : \forall n \geq 22, x_n < M$. We say that (x_n) diverges to infinity, or $\lim(x_n) = -\infty$.

Theorem 3.36
If $\lim(x_n) = +\infty$ or $-\infty$, then (x_n) diverges.

Theorem 3.37
Let (a_n), (b_n) be sequences such that $\lim(b_n) = +\infty$ and $\exists K : \forall n \geq K, a_n \geq b_n$. Then $\lim(a_n) = +\infty$.

Proof. Let $M \in \mathbb{R}$. Then, since b_n diverges, $\exists \tilde{N} : \forall n \geq \tilde{N}, b_n > M$. Let $N := \max\{K, \tilde{N}\}$. Then $\forall n \geq N, a_n \geq b_n > M$. Thus $\lim(a_n) = +\infty$. □

4 Topology

4.1 Open and Closed Sets

We will start by generalizing the notion of open and closed intervals.

Definition 4.1 (Open Set)
A subset $U \subseteq \mathbb{R}$ is called open if $\forall x \in U, \exists \varepsilon > 0$ such that $V_\varepsilon(x) \subseteq U$.

Theorem 4.2
Every open interval is open.

Proof. Open intervals are of the form (1): (a, b), $a < b$, (2): (a, ∞), (3): $(-\infty, a)$.

(1) Let $I := (a, \infty)$, $a \in \mathbb{R}$. Let $x \in I$ and $\varepsilon := x - a > 0$. Then $V_\varepsilon(x) = (x - \varepsilon, x + \varepsilon) = (x - (x - a), x + (x - a)) = (a, 2x - a) \subseteq (a, \infty) \implies V_\varepsilon(x) \subseteq I$.

(2) If $I = (-\infty, b)$, $x \in I$, then let $\varepsilon := b - x$. Then $V_\varepsilon(x) = (2x - b, b) \subset (-\infty, b)$

(3) $I = (a, b)$, $a < b$. Let $x \in I$, $\varepsilon := \min\{x - a, b - x\}$. Then $V_\varepsilon(x) = (x - \varepsilon, x + \varepsilon) \subseteq (x - \varepsilon, b)$ since $\varepsilon \leq b - x$

$\implies V_\varepsilon(x) \subset (x - (x - a), b) = (a, b)$ since $\varepsilon \leq b - a$

$\implies V_\varepsilon(x) \subseteq (a, b)$

Trivial case: $I = \mathbb{R}$. Thus in all cases, I is open. □

Example 4.3 (Open Sets)
Some examples of open sets:

1. $\varnothing$ is open by definition
2. $\mathbb{R}$ is open by last theorem
3. $\mathbb{R} \setminus \{0\}$

The fact that $\mathbb{R} \setminus \{0\}$ is open follows from any x in the set is non zero, and hence $\mathbb{R} \setminus \{0\} = (-\infty, x) \cup (x, \infty)$ which are both open by last theorem.

Theorem 4.4
Arbitrary unions of open sets are open.

Proof. The theorem can be stated as such: if J is an index set and U_j is open $\forall j \in J$, then

$$U = \bigcup_{j \in J} U_j$$

is open. Let $x \in U$. Then $\exists k \in J : x \in U_j$. Since U_j is open, $\exists \varepsilon > 0 : V_\varepsilon(x) \subseteq U_j$. But since U is the union of all U_i's, $U_j \subseteq U$.

$\implies V_\varepsilon(x) \subseteq U_j \subseteq U$

$\implies U$ is open. □

Theorem 4.5
Finite intersections of open sets are open.

Proof. Let U be the finite intersection of open sets such that

$$U = \bigcap_{j=1}^{n} U_j$$

for some $n \in \mathbb{N}$. Let $x \in U$. Then $\forall 1 \leq j \leq n, x \in U_j$, since each U_j is open, $\exists \varepsilon_1, \ldots, \varepsilon_n$ such that $V_{\varepsilon_1}(x) \subseteq U_1, \ldots, V_{\varepsilon_n}(x) \subseteq U_n$.

Let $\varepsilon := \min\{\varepsilon_1 \ldots, \varepsilon_k\}$. Then $\forall 1 \leq j \leq n, V_\varepsilon(x) \subseteq V_{\varepsilon_j}(x) \subseteq U_j$

$\implies V_\varepsilon(x) \subseteq \bigcap_{i=1}^{n} U_j = U$

$\implies U$ is open. □

Note 4.6
This does not generalize to infinite intersctions. Take $U_n = (-\frac{1}{n}, \frac{1}{n})$, $n \in \mathbb{N}$. Then:

$$U = \bigcap_{n \in \mathbb{N}} U_n = \{0\}$$

which is not open.

Definition 4.7 (Closed Set)
A subset $A \subseteq R$ is closed if its complement A^C is open.

Theorem 4.8
Every closed interval is closed.

Proof. (1) $A = \mathbb{R}$. Then $A^C = \varnothing$ which is open.

(2) $A = [a, \infty) \implies A^C = (-\infty, a)$ which is open.

(3) $A = (-\infty, b] \implies A^C = (b, \infty)$ which is open.

(4) $A = [a, b]$. Then $A^C = (-\infty, a) \cup (b\infty)$, which are both open, and by our theorem their union is open.

Thus for all closed intervals, we have shown their complement is open, so the intervals are closed. □

This proof is a testament to the power of a good definition: instead of reproving the same thing for closed sets, we simply defined them in a way that the result is easily adaptable from the proofs we already did for open sets. Further, notice that if $A = \varnothing$, then $A^C = \mathbb{R}$ which is open. So both $\varnothing$ and $\mathbb{R}$ satisfy the property of being open and closed at the same time.

Theorem 4.9
Finite unions of closed sets are closed, and arbitrary intersetions of closed sets are closed.

Proof. (1) Let $A_1, A_2, \ldots, A_n$ be closed sets, and let $A := \bigcup_{i=1}^n A_i$. Then:

$$A^C = \mathbb{R} \setminus \bigcup_{i=1}^n A_i = \bigcap_{i=1}^n (\mathbb{R} \setminus A_i)$$

Since each A_i is closed, $\mathbb{R} \setminus A_i$ is open, and we finite intersections of open sets are open. Hence A^C is open and A is closed.

(2) Let I be an index set so that $\forall i \in I$, A_i is closed. Let $A := \bigcap_{i=1}^n A_i$. Then:

$$A^C = \mathbb{R} \setminus \bigcap_{i=1}^n A_i = \bigcap_{i=1}^n (\mathbb{R} \setminus A_i)$$

Again, we know arbitrary unions of open sets are open, so A^C is open and A is closed. □

Definition 4.10 (Boundary)
Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is a boundary point of A if $\forall \varepsilon > 0$, $V_\varepsilon(x) \cap A \neq \varnothing$ and $V_\varepsilon(x) \cap A^C \neq \varnothing$. The set of all boundary points of A is called the boundary of A and is denoted ∂A.

Example 4.11
Let $I = [a, \infty)$. Show that $\partial I = \{a\}$.

Proof. Let $x > a$. We know (a, ∞) is open, so $\exists \varepsilon > 0$ such that $V_\varepsilon(x) \subset (a, \infty) \subset I$.

$\implies V_\varepsilon(x) \cap I^C = \varnothing$ and $x \notin \partial I$.

Let now $x < a$ and consider the open interval $(-\infty, a)$. $x \in (-\infty, a) \implies \exists \varepsilon > 0$ such that $V_\varepsilon(x) \subseteq (-\infty, a) \subset I^C$

$\implies V_\varepsilon(x) \cap I = \varnothing$ and $x \notin \partial I$.

However, $V_\varepsilon(a) \cap I \neq \varnothing$ since $a + \frac{\varepsilon}{2} \in V_\varepsilon(a) \cap I$ and $a - \frac{\varepsilon}{2} \in V_\varepsilon(a) \cap I^C$. Thus neither the intersection of the eps neighborhood of I and I^C is empty and $a \in \partial I$. No other real number satisfies this, so $\partial I = \{a\}$. $\square$

Example 4.12 (Boundaries of Intervals on $\mathbb{R}$)
$\partial[a, b] = \partial[a, b) = \partial(a, b] = \partial(a, b) = \{a, b\}$.

Theorem 4.13
Let $A \subseteq \mathbb{R}$. Then:

1. A is open if and only if it does not contain any of its boundary points, i.e. $A \cap \partial A = \varnothing$ or $\partial A \subset A^C$
2. A is closed iff A contains all of its boundary points, i.e. $\partial A \subseteq A$.

Proof. $\square$

It should be noted despite this discussion that most subsets of $\mathbb{R}$ are neither open nor closed. For example, $I = [0, 1)$ is neither open nor closed, since $\partial I = \{0, 1\}$, where $0 \in I$ but $1 \notin I$.

This perhaps motivates the following definitions, which allows us to create closed and open sets related to a set that may be neither open nor closed.

Definition 4.14 (Interior and Closure)
Let $S \subseteq \mathbb{R}$.

1. The interior of S, denoted $\mathring{S}$ or $\operatorname{int}(S)$, is defined as $S \setminus \partial S$
2. The closure of S, denoted $\overline{S}$, is defined as $S \cup \partial S$.

Since, $A \setminus B = A \cap B^C$, we can also express the interior as $\mathring{S} = S \cap (\partial S)^C$.

Theorem 4.15
Let $S \subseteq \mathbb{R}$. Then:

1. $\mathring{S} = \bigcup_{U_i \subseteq S} U_i$ such that all U_i are open
2. $\overline{S} = \bigcap_{A_i \supseteq S} A_i$ such that all A_i are closed.

To restate this theorem, the $\mathring{S}$ is the largest open subset of S, since it is comprised of the union of all the subsets of S that are open. Likewise, the $\overline{S}$ is the smallest closed superset of S, being the intersection of all closed sets containing S.

Proof. 1. ($\implies$) Let $x \in S \setminus \partial S$. Since $x \notin \partial S$, we have:

$\neg(\forall \varepsilon > 0 : V_\varepsilon(x) \cap S \neq \varnothing \wedge V_\varepsilon(x) \cap S^C \neq \varnothing)$

$\equiv \exists \varepsilon > 0 : V_\varepsilon(x) \cap S = \varnothing \vee V_\varepsilon(x) \cap S^C = \varnothing$

Since $x \in S, V_\varepsilon(x) \cap S \neq \varnothing$, so it must be that $V_\varepsilon(x) \cap S^C = \varnothing$, and thus $V_\varepsilon(x) \subseteq S$. Since $V_\varepsilon(x)$ is open, $x \in \bigcup_{U_i \subseteq S} U_i$

($\impliedby$) Let $x \in \bigcup_{U_i \subseteq S} U_i$. Then $x \in S$ and there exists some $U_i \subseteq S$ open such that $x \in U_i$. Let $\varepsilon > 0$ such that $V_\varepsilon(x) \subseteq U_i$. Since U_i is entirely contained in S, $V_\varepsilon(x) \cap S^C = \varnothing$, meaning $x \notin \partial S$. Hence $x \in S$ and $x \notin \partial S$, so $x \in S \setminus \partial S$.

2. We will prove this using the previous statement and taking complements.

Note that $(\mathring{S})^C = (S \cap (\partial S)^C) = S^C \cup (\partial S)^C = \overline{S^C}$. From above we know:

$$(\mathring{S})^C = \left(\bigcup_{U_i \subseteq S} U_i\right)^C = \bigcap_{U_i \subseteq S} U_i^C$$

by De Morgan's law. Hence:

$$\overline{S} = \bigcap_{U_i \subseteq S^C} U_i^C = \bigcap_{U_i^C \supseteq S} U_i^C = \bigcap_{A_i \supseteq S} A_i$$

With the latter condition following from U_i being open implying $A_i := U_i^C$ is closed. □

Theorem 4.16
Let $S \subseteq \mathbb{R}$. Then:

1. If S is bounded from above, $\sup S \in \partial S$
2. If S is bounded from below, $\inf S \in \partial S$

Proof. (1) Let $s := \sup S$. Then s is the least upper bound, and $existsx \in S$ such that $s - \varepsilon < x \leq s$

$\implies x \in (s - \varepsilon, s) \subseteq V_\varepsilon(s)$

$\implies V_\varepsilon(s) \cap S \neq \varnothing$

At the same time, $(s, s + \varepsilon) \subseteq S^C$ since s is an upper bound.

$\implies V_\varepsilon(s) \cap S^C \supset (s, s+\varepsilon) \neq \emptyset$

Both of these conditions imply $s \in \partial S$.

(2) Exercise. □

4.2 Sequences and Topology, Compactness

Definition 4.17
Let $S \subseteq \mathbb{R}$. A sequence (x_n) is in S if $\forall n \in \mathbb{N}$, $x_n \in S$.

Theorem 4.18
Let $A \subseteq \mathbb{R}$ be closed and x_n be a convergent sequence in A. Then $\lim(x_n) \in A$.

Proof. Let $x := \lim(x_n)$ and suppose $x \notin A$. Then $x \in A^C$, which is open, so $\exists \varepsilon > 0$ such that $V_\varepsilon(x) \subseteq A^C$.

Since (x_n) is in A, $\forall n \in \mathbb{N}$, $x_n \notin V_\varepsilon(x)$. But through the definition of a convergence sequence $\forall \varepsilon > 0$, $\exists N : \forall n \geq N$, $x_n \in V_\varepsilon(x)$, which contradicts the previous line. □

The converse of this theorem is also true, namely that if there is a convergence sequence in a set whose limit is in the set, then the set is closed.

We will now begin a discussion on compact sets. There are multiple definitions on what this could be, each with varying mileage depending on the topological space one is working with. For example, in $\mathbb{R}$, the regular definition suffices, but in $\mathbb{R}^n$, we will need sequential compactness. These are equivalent for our purposes, but that equivalence fails for general metric spaces, and a new definition altogether is needed for general topological spaces.

Definition 4.19 (Compact Set)
$A \subseteq \mathbb{R}$ is compact if it is closed and bounded.

Definition 4.20
$A \subseteq \mathbb{R}$ is sequentially compact if all sequences (x_n) in A have a convergent subsequence (x_{n_k}) such that $\lim(x_{n_k}) \in A$.

We will show that on $\mathbb{R}$, these are equivalent.

Theorem 4.21
$A \subseteq \mathbb{R}$ is compact if and only if it is sequentially compact.

Proof. ($\implies$) Let A be compact and (x_n) a sequence in A. Since A is then bounded, by Bolzano-Weierstrass, (x_n) has a convergent subsequence (x_{n_k}) that is also in A. Hence by our last theorem, $\lim(x_{n_k}) \in A$, so A is sequentially compact.

($\impliedby$) Let A be sequentially compact and assume A is not boounded. Then, $\forall n \in \mathbb{N}$, $\exists x_n > n \in A$. Now consider (x_n) in A, which possesses a convergent subsequence (x_{n_k}) since A is sequentially compact. However, by definition, all terms of the subsequence must satisfy $x_{n_k} \geq n_k \geq k$ which means that (x_{n_k}) is unbounded and thereby cannot be convergent. This leads to a contradiction and thus A must be bounded.

Further assume A is not closed. Then A^C is not open, so we have:

$\neg(\forall x \in A^C, \exists \varepsilon > 0 : V_\varepsilon(x) \subseteq A)$

$\equiv \exists x \in A^C : \forall \varepsilon > 0, V_\varepsilon(x) \not\subseteq A^C \iff V_\varepsilon(x) \cap A \neq \varnothing$

Since the intersection of $V_\varepsilon(x)$ and A is nonempty for all ε positive, select one such $x \in A^C$. Then define, $\forall n \in \mathbb{N}, x_n \in V_{1/n}(x) \cap A$. We now have a sequence (x_n) in A, satisfying $\forall n \in \mathbb{N}, |x_n - x| < \frac{1}{n}$, so by the Null Sequence Criterion, (x_n) converges to x, and thus any subsequence x_{n_k} has limit x which lies in A^C. This again provides a contradiction to sequential compactness, since convergent subsequences' limits must lie in A. Hence A is closed.

Since A is necessarily closed and bounded, A is compact. □

Theorem 4.22 (Nested Compact Set Property)
Let $A_1 \supseteq A_2 \supseteq \ldots$ be a nested sequence of nonempty compact subsets of $\mathbb{R}$. Then $\bigcap_{n \in \mathbb{N}} A_n \neq \varnothing$.

Proof. □

4.3 Accumulation Points

Definition 4.23 (Accumulation Point)
Let (x_n) be a sequence. A point $x \in \mathbb{R}$ is an accumulation point of (x_n) if (x_n) has a convergent subsequence (x_{n_k}) such that $\lim(x_{n_k}) = x$.

Accumulation points generalize the notion of limits to subsequences. Immediately, we see that a sequence is convergent, it must have only one accumulation point, which is its unique limit. An alternating series like $((-1)^n)$ has accumulation points 1 and -1, since its only possible subsequences are odd or even indices. Finally, the set of accumulation points for a sequence enumerating $\mathbb{Q}$ is $\mathbb{R}$, which follows from the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$, and thus convergent sequences of rationals can be constructed around any real number!

Theorem 4.24
Let (x_n) be a sequence. Then the set of accumulation points of (x_n) is closed.

Lemma 4.25
$x \in \mathbb{R}$ is an accumulation point of (x_n) if and only if $\forall \varepsilon > 0, V_\varepsilon(x)$ contains infinitely many terms of (x_n).

Proof. ($\implies$) Let x be an accumulation point. Then $\exists (x_{n_k})$ convergent whose limit is x. Hence by the definition of convergence, $\exists K : \forall k \geq K, x_{n_k} \in V_\varepsilon(x)$, so $V_\varepsilon(x)$ contains infinitely many terms of (x_n).

($\Longleftarrow$) Let x be a point such that there are infinitely terms of (x_n) in any ε neighborhood of x. Then, $\forall k \in \mathbb{N}, \exists x_{n_k} \in V_{1/n}(x)$. Therefore $\forall k \in \mathbb{N}, |x_{n_k} - x| < \frac{1}{k}$, and (x_{n_k}) characterizes a convergent subsequence whose limit is x, meaning x is an accumulation point. □

Corollary 4.26
Let (x_n) be a sequence. Then x is not an accumulation point if and only if $\exists \varepsilon > 0$ such that $V_\varepsilon(x)$ contains finitely many terms of (x_n)

Proof of Theorem. Let A denote the set of accumulation points, and let $x \in A^C$. By last corollary, $\exists \varepsilon > 0$ such that $V_\varepsilon(x)$ contains finitely many terms of (x_n) since x is not an accumulation point. Further let $y \in V_\varepsilon(x)$ be arbitrary. Since $V_\varepsilon(x)$ is open, $\exists \tilde{\varepsilon} > 0$ such that $V_{\tilde{\varepsilon}}(y) \subseteq V_\varepsilon(x)$. Then, since $V_\varepsilon(x)$ contains finitely many terms of (x_n), so does $V_{\tilde{\varepsilon}}(y)$, so y is not an accumulation point. Hence $y \in A^C$. Since y was arbitrary, $V_\varepsilon(x) \subseteq A^C$ meaning A^C is open, and resultingly A is closed. □

Example 4.27
Let $(x_n) = (1, 1, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{3}, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots)$ Find the set of accumulation points of (x_n).

Solution. Let A denote the set of accumulation points. Immediately, we recognize the following are subsequences of (x_n): $(1, 1, \dots)$, $(\frac{1}{2}, \frac{1}{2}, \dots)$, $(\frac{1}{3}, \frac{1}{3}, \dots)$, and so on. So the set $S := \{\frac{1}{k}, k \in \mathbb{N}\}$ is contained in A. But S is not closed, so $A \neq S$.

Nonetheless, we see that $\partial S = \{0\}$, and $\overline{S} = S \cup \{0\} \subseteq A$, so 0 is an accumulation point of (x_n). We will show that $\overline{S} = A$; having already shown the first inclusion, it remains to prove $A \subseteq \overline{S}$, i.e. if x is an accumulation point it must be $\frac{1}{k}$ for some $k \in \mathbb{N}$ or 0.

Let $x \in A$. Then

Theorem 4.28
Let (x_n) be a sequence and A the set of its accumulation points. Then:

1. if (x_n) is bounded from above, A is bounded from above
2. if (x_n) is bounded from below, A is bounded from below
3. if (x_n) is bounded, A is bounded

Proof. (a) Let $M > 0$ be an upper bound such that $\forall n \in \mathbb{N}, x_n \leq M$, and let $x \in A$. Then $\exists (x_{n_k})$ with limit x. Since $\forall k \in \mathbb{N}, x_{n_k} \leq M$, and limits preserve order, $\lim(x_{n_k}) \leq M$, and thus $x \leq M$. Hence A is bounded from above as well.

(b),(c) exercise. □

Corollary 4.29
Let (x_n) be a bounded sequence. Then A is compact.

Proof. By Theorem 4.24, we know A is closed, and by above theorem A is bounded since (x_n) is bounded. Since A is closed and bounded, it is compact. □

4.4 limsup and liminf

We already established that the sup and inf of a set are contained in its boundary. Further, the set of accumulation points of a is closed and hence contains its boundary. Putting these together with our last theorem, if (x_n) is bounded, its set of accumulation points is compact and contains its sup and inf (in this case, the maximum and minimum), so there is a maximum and minimum accumulation point for any bounded sequence, which merit a special designation.

Definition 4.30 (lim sup and lim inf)
Let (x_n) be a sequence, A the set of its accumulation points.

1. If (x_n) is bounded from above, we call the greatest accumulation point the limes superior of (x_n), denoted $x^* = \limsup(x_n)$
2. If (x_n) is bounded from below, we call the smallest accumulation point the limes inferior of (x_n), denoted $x_* = \liminf(x_n)$

Note that $x^* = \max A$ and $x_* = \min A$.

In this context, $\limsup((-1)^n) = 1$ and $\liminf((-1)^n) = 1$. Revisiting Example 4.27, $\limsup(x_n) = \max A = 1$, and $\liminf(x_n) = \min A = 0$. The two objects also play an important role in power series, as the radius of convergence of a power series $\sum_{n=1}^{\infty} a_n(x-x_0)^n$ is given by $R = \frac{1}{\limsup(\sqrt[n]{(a_n)})}$.

Theorem 4.31
Let (x_n) be a bounded sequence. Then $\forall \varepsilon > 0, \exists N : \forall n \geq N, x_n \in (x_* - \varepsilon, x^* + \varepsilon)$.

If (x_n) is convergent, then $x^* = x_* = x$, and the theorem reduces to the standard definition of convergence.

Proof. Let $\varepsilon > 0$ and assume the theorem does not hold. Then there are infinitely many terms of x_n outside $(x_* - \varepsilon, x^* + \varepsilon)$, so there are either infinitely many $x_n \geq x^* + \varepsilon$ or infinitely many $x_n \leq x_* - \varepsilon$.

Let us examine the first case. Collect all $x_n \geq x^* + \varepsilon$ into a subsequence (x_{n_k}). Since (x_n) is bounded, by Bolzano-Weierstrass, x_{n_k} has a convergent subsequence $(x_{n_{k_j}})$, whose limit we will call x. By definition, $\forall j \in \mathbb{N}, x_{n_{k_j}} \geq x^* + \varepsilon$ so its limit $x > x^*$.

However, since x is necessarily an accumulation point of (x_n), $x \leq x^*$, since x^* is the greatest accumulation point. This provides a contradiction. The same argument follows for the second case, and consequently, the theorem must hold. □

5 Functional Limits and Continuity

So far we have looked at limits of sequences, indexed by the natural numbers. We defined the limit with the ε definition, which proved to be intuitive. In a similar fashion, given a function

f defined on a domain $D \subseteq \mathbb{R}$, we want to make precise the notion of the limit of f as its input approaches some constant c, i.e. $\lim_{x \to c} f$. We will do this two ways.

Definition 5.1 ($\varepsilon - \delta$)
Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$. We say L is the limit of f as x approaches c if $\forall \varepsilon > 0, \; \exists \delta > 0 : \forall x \in D \setminus \{x\}, |x - c| < \delta \implies |f(x) - L| < \varepsilon$.

Importantly, c need not lie in the δ-neighborhood of c, since we are only interested in the behavior of f as it gets arbitrarily close to c. Its value at c, $f(c)$, is immaterial to the limit, and this concept is illustrated in the figure below. Therefore, it will be useful to define a new kind of neighborhood that is centered around a point but does not actually include it.

Definition 5.2 (Punctured Neighborhood)
Let $c \in \mathbb{R}$, $\varepsilon > 0$. Then $V_\varepsilon^*(c) := V_\varepsilon(c) \setminus \{c\}$ is the punctured ε-neighborhood about c.

Using this, we can rewrite the $\varepsilon - \delta$ definition in a number of logically equivalent ways:

$\forall \varepsilon > 0, \; \exists \delta > 0 : \forall x \in D, 0 < |x - c| < \delta \implies |f(x) - L| < \varepsilon$

$\equiv \forall \varepsilon > 0, \; \exists \delta > 0 : \forall x \in D, x \in V_\delta^*(c) \implies f(x) \in V_\varepsilon(L)$

$\equiv \forall \varepsilon > 0, \; \exists \delta > 0 : \forall x \in V_\delta^*(c) \cap D, f(x) \in V_\varepsilon(L)$

$\equiv \forall \varepsilon > 0, \; \exists \delta > 0 : f(V_\delta^*(c) \cap D) \subseteq V_\varepsilon(L)$

Example 5.3
Show the following limits:

1. Let $f : \mathbb{R} \to \mathbb{R}, x \longmapsto 2x$, and $c \in \mathbb{R}$. Show that $\lim_{x \to c} f = 2c$.
2. $f : \mathbb{R} \to \mathbb{R}, x \longmapsto x^2$, and $c \in \mathbb{R}$. Show that $\lim_{x \to c} f = c^2$.

Proof. (1) Let $\varepsilon > 0$, and $\delta > 0$ be arbitrary for now. Let $0 < |x - c| < \delta$. Then $|f(x) - L| = |2x - 2c| = 2|x - c| < 2\delta$, which is less than eps if $\delta < \frac{\varepsilon}{2}$.

Now let $\delta < \frac{\varepsilon}{2}$. Then $\forall x \in \mathbb{R}, 0 < |x - c| < \delta \implies |f(x) - 2c| < \varepsilon$.

(2) Let $\varepsilon > 0$, and $\delta > 0$ be arbitrary for now, and $0 < |x - c| < \delta$. Then $|x^2 - c^2| = |(x - c)(x + c)| = |x - c| \cdot |x + c| < \delta|x + c|$. Since $|x + c| = |x - c + 2c|$, by the triangle inequality, $\delta|x + c| \leq \delta(|x - c| + 2|c|) < \delta(\delta + 2|c|)$.

Suppose $\delta < 1$. Then $\delta(\delta + 2|c|) < \delta(1 + 2|c|) < \varepsilon \iff \delta < \frac{\varepsilon}{1+2|c|}$. Thus, let $\delta < \min\{1, \frac{\varepsilon}{1+2|c|}\}$. Whenever $0 < |x - c| < \delta$, it holds that $|x^2 - c^2| < \varepsilon$, so $\lim_{x \to c} x^2 = c^2$. □

Clearly, showing a limit through the $\varepsilon - \delta$ method can get quite cumbersome with more complex functions. Hence, we can also (equivalently) formulate an alternative definition of a limit involving sequences, which will allow us to capitalize on the algebraic limit laws we already proved.

Definition 5.4 (Sequential Definition of a Limit)
Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$. We say L is the limit of f as x approaches c if $\forall (x_n)$ in $D \setminus \{c\}$ that converge to c, it holds that $\lim(f(x_n)) = L$.

Revisiting our example of the limit of $f(x) = x^2$ as x approaches c, we can show it's c^2 by letting (x_n) be a sequence in $\mathbb{R} \setminus \{c\}$ such that $\lim(x_n) = c$. Then $\lim(f(x_n)) = \lim(x_n^2) = \lim(x_n)^2 = c^2$ by algebraic limit laws, so $\lim_{x \to c} x^2 = c^2$.

Theorem 5.5
The $\varepsilon - \delta$ and sequential definition of a limit are equivalent.

Proof. ($\implies$) Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$, $c, L \in \mathbb{R}$ such that $\lim_{x \to c} f = L$ by $\varepsilon - \delta$, i.e. $\forall \varepsilon > 0, \exists \delta > 0 : \forall x \in V_\delta^*(x) \cap D, |f(x) - L| < \varepsilon$. Let (x_n) be a sequence in $D \setminus \{c\}$ such that $\lim(x_n) = c$. Let $\varepsilon, \delta > 0$ satisfy the statement above. Since $(x_n) \to c$, $\exists N \in \mathbb{N} : \forall n \geq N, x_n \in V_\delta(c)$, whilst $x_n \in D \setminus \{c\}$.

$$\implies \forall n \geq N, x_n \in V_\delta(c) \cap D$$

$$\implies \forall n \geq N, |f(x) - L| < \varepsilon$$

$$\implies \lim(f(x_n)) = L$$

($\impliedby$) Let (x_n) be an arbitrary sequence in $D \setminus \{c\}$ converging to c, and such that $\lim(f(x_n)) = L$. Suppose that the $\varepsilon - \delta$ limit is not true, i.e.:

$$\neg(\forall \varepsilon > 0, \exists \delta > 0 : x \in V_\delta^*(c) \implies |f(x) - L| < \varepsilon)$$

$$\equiv \exists \varepsilon > 0 : \forall \delta > 0, \exists x \in V_\delta^*(c) : |f(x) - L| \geq \varepsilon \ (*)$$

Let ε satisfy the above, and choose $\forall n \in \mathbb{N}, \delta := \frac{1}{n}$. Then let $x_n \in V_{1/n}^*(c) \cap D$. By the Null Sequence Criterion, $\lim(x_n) = c$, and resultingly since (x_n) is in $D \setminus \{c\}$, $\lim(f(x_n)) = L$.

$\implies \exists N \in \mathbb{N}$ such that $|f(x) - L| < \varepsilon$ by the definition of convergence. But this contradicts $(*)$, so it must be that the $\varepsilon - \delta$ definition holds. □

We would also like to have a convienient way to show limits of functions do not exist without attempting to draw out contradictions every time. The following criterions present two methods:

Theorem 5.6 (Two Sequence Criterion for the Non-Existence of a Limit)
Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$, $c \in \mathbb{R}$. If there exists (x_n) and (u_n) in $D \setminus \{c\}$ such that $\lim(x_n) = \lim(u_n) = c$, but $\lim(f(x_n)) \neq \lim(f(u_n))$, then $\lim_{x \to c} f$ does not exist.

Proof. Follows immediately from the sequential definition of a limit. □

Example 5.7
Let $f : \mathbb{R} \to \mathbb{R}$ be the Dirichlet function, that is:

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

Show that $\lim_{x\to c} f$ does not exist $\forall c \in \mathbb{R}$.

Proof. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we construct a sequence of rationals converging to any real number, say (x_n) such that (x_n) is in $\mathbb{Q} \setminus \{c\}$ and $\lim(x_n) = c$. Likewise, the irrationals are also dense so we can construct a sequence of irrations, (u_n) such that (u_n) is in $(\mathbb{R} \setminus \mathbb{Q}) \setminus \{c\}$ and $\lim(u_n) = c$. However, $\forall n \in \mathbb{N}$, we have that $f(x_n) = 1$ and $f(u_n) = 0$, so $\lim(f(x_n)) = 1$ and $\lim(f(u_n)) = 0$. Hence by the two sequence criterion for the non-existence of a limit, $\lim_{x\to c} f$ does not exist. □

Theorem 5.8 (One Sequence Criterion for the Non-Existence of a Limit)
Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$, $c \in \mathbb{R}$. Let (x_n) be a sequence in $D \setminus \{c\}$ such that $\lim(x_n) = c$. If $\lim(f(x_n))$ does not exist, then $\lim_{x\to c} f$ does not exist.

Proof. Also immediate from the sequential definition of a limit. □

Example 5.9
Let $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $x \longmapsto \frac{1}{x}$. Show that $\lim_{x\to 0} f$ does not exist.

Proof. Let $(x_n) = \left(\frac{1}{n}\right)$. Then $\lim(x_n) = 0$, but $\lim(f(x_n)) = \left(\frac{1}{\frac{1}{n}}\right) = n = +\infty$. Thus $\lim(f(x_n))$ does not exist and $\lim_{x\to c} f$ does not exist. □

5.1 Uniqueness of Limits and Cluster Points

We saw that with the ε definition of a limit for sequences, it was easy to show that limits of convergent sequences are unique. Does the same hold for functions? We will need some additional requirements.

Suppose that $D := \{0\} \cup [1, 2]$, and $f : D \to \mathbb{R}$ is constantly equal to 0. What will determine $\lim_{x\to 0} f$? It turns out (in a tricky sense) that any real number $L \in \mathbb{R}$ will satisfy this:

Let (x_n) be any sequence in $D \setminus \{0\} = [1, 2]$ such that $\lim(x_n) = 0$. Immediately, we see that there are no such sequences. So $\lim(f(x_n)) = L$. Alternatively, let $\varepsilon, \delta > 0$ be arbitrary. $V_\delta^*(0) \cap D = \varnothing$, so $\forall x \in V_\delta^*(0)$ it must hold that $|f(x) - L| < \varepsilon$.

Perhaps this happens because 0 is an "isolated" point in D. It does not make sense to look at the limit of a function as it approaches some c if the function is not defined near c at all. This is precisely what the concept of a isolated and cluster points formalize.

Definition 5.10 (Cluster Point and Isolated Point)
Let $D \subseteq \mathbb{R}$:

1. A point $x \in \mathbb{R}$ is called a cluster point (or limit point) of $\exists(x_n)$ in $D \setminus \{x\}$ such that $\lim(x_n) = x$.
2. A point $x \in D$ is an isolated point if $\exists\delta > 0$ such that $V_\delta(x) \cap D = \varnothing$

Now we see that in our previous example, 0 was an isolated point, so there was no comprhensible (much less unique) limit of f near 0.

Theorem 5.11
Let $D \subset \mathbb{R}$:

1. $c \in D$ is isolated if and only if $\exists\delta > 0$ such that $V_\delta^*(c) \cap D = \varnothing$
2. $c \in \mathbb{R}$ is a cluster point if and only if $\forall\delta > 0, V_\delta^*(c) \cap D \neq \varnothing$

Proof. (a) Exercise.

(b) ($\implies$) Let c be a cluster point. Then $\exists(x_n)$ in $D \setminus \{c\}$ such that $\lim(x_n) = c$. Let $\delta > 0$:

$\implies \exists N \in \mathbb{N} : \forall n \geq N, x_n \in V_\delta(c)$

Since (x_n) is in $D \setminus \{c\}$, $\forall n \in \mathbb{N}$, $x_n \neq c$.

$\implies \forall n \geq N, x_n \in V_\delta(c) \setminus \{c\}$

$\implies V_\delta^*(c) \neq \varnothing$

($\impliedby$) Assume that $\forall\delta > 0, V_\delta^*(c) \neq \varnothing$. Let, $\forall n \in \mathbb{N}$, $\delta := \frac{1}{n}$ and consider $x_n \in V_\delta^*(c)$. Then (x_n) is in $D \setminus \{c\}$ and $\forall n \in \mathbb{N}$, $|x_n - c| < \frac{1}{n}$ so (x_n) converges to c by the Null Sequence Criterion. □

Theorem 5.12
Let $D \subseteq \mathbb{R}, f : D \to \mathbb{R}, c \in \mathbb{R}$ such that $\lim_{x\to c} f$ exists. Then:

1. If c is a cluster point of D, then $\lim_{x\to c} f$ is uniquely determined.
2. If $c \in D$ is an olsated point, then $\lim_{x\to c} f$ is arbitrary.

Proof. (1) Follows from the sequential definition of a limit, since limits of sequences are unique.

(2) Shown in the beginning of this subsection. □

For the remainder of our discussion of limits, we will assume that "c" is always a cluster point.

5.2 Limit Laws for Functions

Theorem 5.13 (Algebraic Limit Laws)
Let $f, g : D \subseteq \mathbb{R} \to \mathbb{R}$, c a cluster point of D. Assume that $\lim_{x\to c} f$ and $\lim_{x\to c} g$ exist. Then:

1. $\lim_{x\to c}(f+g) = \lim_{x\to c} f + \lim_{x\to c} g$
2. $\lim_{x\to c}(f-g) = \lim_{x\to c} f - \lim_{x\to c} g$
3. $\lim_{x\to c} kf = k \cdot \lim_{x\to c} f \ \forall k \in \mathbb{R}$
4. $\lim_{x\to c}(fg) = \lim_{x\to c} f \cdot \lim_{x\to c} g$
5. $\lim_{x\to c} \frac{f}{g} = \frac{\lim_{x\to c} f}{\lim_{x\to c} g}$ if $\forall x \in D$, $g(x) \neq 0$ and $\lim_{x\to c} g \neq 0$.

Proof. The proof for all of these will follow from Theorem 3.14.

(1) Let (x_n) in $D \setminus \{c\}$ such that $\lim(x_n) = c$. Then $\lim_{x\to c} f + \lim_{x\to c} g = \lim_{x\to c} f(x_n) + \lim_{x\to c} g(x_n)$ by the sequential definition, which then equals $\lim(f(x_n)+g(x_n)) = \lim((f+g)(x_n))$ by the algebraic limit laws for sequences. Again by the sequential definition, $\lim((f+g)(x_n)) = \lim_{x\to c}(f+g)$.

(2),(3),(4),(5) exercise.

□

Theorem 5.14 (Squeeze Theorem)
Let $f, g, h : D \subseteq \mathbb{R} \to \mathbb{R}$, c a cluster point of D. Assume that $\forall x \in D, f(x) \leq g(x) \leq h(x)$, with $\lim_{x\to c} f = \lim_{x\to c} h = L$. Then $\lim_{x\to c} g = L$.

Proof. Let (x_n) in $D \setminus \{c\}$ such that $\lim(x_n) = c$. By the sequential definition, $\lim(f(x_n)) = \lim(h(x_n)) = L$. Since $\forall n \in \mathbb{N}, f(x_n) \leq g(x_n) \leq h(x_n)$, we conclude by the squeeze theorem for sequences that $\lim(g(x_n))$ exists and equals L. So by the sequential definition again, $\lim_{x\to c} g = L$. □

Example 5.15
Let $f : \mathbb{R} \setminus \{0\}, x \longmapsto x\sin(\frac{1}{x})$. Show that $\lim_{x\to 0} f = 0$.

Lemma 5.16
$\lim_{x\to 0} |x| = 0$

Proof. Let $\varepsilon > 0$, $\delta := \varepsilon$. Then $0 < |x-0| < \delta \implies ||x| - 0| = |x| < \delta = \varepsilon$. □

Proof of Example. Note that $\forall x \in \mathbb{R} \setminus \{0\}, -1 \leq \sin(\frac{1}{x}) \leq 1$

$\implies -x \leq x\sin(\frac{1}{x}) \leq x$ as long as $x > 0$

$\implies -|x| \leq x\sin(\frac{1}{x}) \leq |x|$

The same holds if $x < 0$. By our lemma and limit laws, we see that $\lim_{x\to 0} -|x| = -0 = 0$, so by squeeze theorem $\lim_{x\to 0} x\sin(\frac{1}{x}) = 0$. □

5.3 Continuity

We will define continuity in several ways.

Definition 5.17 (Limit Definition of Continuity)
Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$, $c \in D$. We say f is continuous at c if $\lim_{x \to c} f = f(c)$, i.e. the limit of f at c exists and equals the function value. In any other case, f is discontinuous at c.

There are two ways in which f can fail to be continuous: either $\lim_{x \to c} f$ exists but does not equal $f(c)$, or $\lim_{x \to c}$ does not exist at all.

If f is continuous at all $x \in D$, we say f is continuous on D.

Example 5.18
$f : \mathbb{R} \to \mathbb{R}$, given by:

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

Show that $\lim_{x \to 0} f$ does not exist.

Proof. (Using the two sequence criterion) Let $(x_n) := \frac{1}{n}$. Then $\forall n \in \mathbb{N}, x_n > 0$, so $f(x_n) = 1$. Let $(u_n) := -\frac{1}{n}$. Similarly, $f(u_n) = 0$. Both sequences converge to 0, but their images converge to 1 and 0. Hence $\lim_{x \to 0} f$ does not exist, and f is discontinuous at 0. □

Definition 5.19 (Sequential Definition of Continuity)
Let $f : D \to \mathbb{R}$, $c \in D$. We say f is continuous at c if $\forall (x_n)$ in D converging to c, $\lim(f(x_n)) = f(c)$.

Definition 5.20 ($\varepsilon - \delta$ Definition of Continuity)
Let $f : D \to \mathbb{R}$, $c \in D$. We say f is continuous at c if $\forall \varepsilon > 0, \exists \delta > 0 : \forall x \in V_\delta(c) \cap D, |f(x) - f(c)| < \varepsilon$.

It would be nonsensical to consider $D \setminus \{c\}$ or $V_\delta^*(c)$, since we are especially interested in the behavior of f at c to determine continuity.

Theorem 5.21
All three definitions of continuity are equivalent.

Proof. (Limit $\iff$ $\varepsilon - \delta$) Let f be continuous at c by the $\varepsilon - \delta$ definition. Then $\forall \varepsilon > 0, \exists \delta > 0 :$ $\forall x \in V_\delta(c) \cap D, |f(x) - f(c)| < \varepsilon$. Since $V_\delta(c) \supset V_\delta^*(c)$, the same holds $\forall x \in V_\delta^*(c) \cap D$, meaning $\lim_{x \to c} f$ exists and equals $f(c)$.

Now suppose that f is continuous at c by the limit definition. Then $\lim_{x \to c} f = f(c)$, so by the definition of the limit of a function, $\forall \varepsilon > 0, \exists \delta > 0 : x \in V_\delta^*(c) \cap D \implies |f(x) - f(c)| < \varepsilon$. If $x = c$, $x \in V_\delta(c)$, and $|f(x) - f(c)| = 0 < \varepsilon$, so the $\varepsilon - \delta$ definition is satisfied.

(Limit $\iff$ Sequential) exercise. □

Theorem 5.22 (Algebraic Continuity Theorems)
Let $f, g : D \subseteq \mathbb{R} \to \mathbb{R}$, $c \in D$ such that f, g are continuous at c. Then:

1. $(f + g)$ is continuous at c
2. $(f - g)$ is continuous at c
3. kf is continuous at c
4. (fg) is continuous at c
5. $\frac{f}{g}$ is continuous at c if $\forall x \in D, g(x) \neq 0$

Proof. Follows from algebraic limit laws for functions. □

5.4 Applications of Continuity

Theorem 5.23
All polynomials are continuous on $\mathbb{R}$, and rational functions $\frac{P(x)}{Q(x)}$ are continuous everywhere except roots of Q.

Proof. A polynomial function is of the form $P(x) = a_0 + a_1 x + \cdots + a_n x^n$. Since the identity function $x \longmapsto x$ is continuous, by (4), x^k is continuous $\forall k \in \mathbb{N}$, and by (3), $a_k x^k$ is continuous. Finally by (1), the sum of all of them is continuous.

A rational function $R(x) = \frac{P(x)}{Q(x)}$ is a ratio of two polynomials, which we just saw were continuous, and a polynomial is different from zero everywhere except for its roots. Thus by (5), $R(x)$ is continuous everywhere but the roots of $Q(x)$. □

Theorem 5.24
Compositions of continuous functions are continuous; that is, if $f : A \subseteq \mathbb{R} \to \mathbb{R}$, $c \in A$, $g : B \subseteq \mathbb{R} \to \mathbb{R}$, $d \in B$, such that f, g are continuous at c, d respectively, and $f(A) \subseteq B$ with $f(a) = b$, then $g \circ f$ is continuous at c.

Proof. (Using Sequential Definition) Let $(x_n) \to c$ be a sequence in A. Since f is continuous at c, $\lim(f(x_n)) = f(c) = d$. From this, we see $f(x_n)$ is a sequence in B that converges to d. Since g is continuous at D, $lim(g(f(x_n))) = g(d)$

$\implies \lim(g \circ f(x_n)) = (g \circ f)(c)$ so the function $g \circ f$ is continuous at c by the sequential definition of continuity.

(Using $\varepsilon - \delta$) Let $\varepsilon > 0$. $y = g(u)$ is continuous at d, so $\exists \gamma > 0 : \forall u \in V_\gamma(d) \cap B, |g(u) - g(d)| < \varepsilon$ $(*)$

Likewise, $u = f(x)$ is continuous at c, so $\exists \delta > 0 : \forall x \in V_\delta(c) \cap A, |f(x) - f(c)| < \gamma$

$\implies |u - d| < \gamma$

$\implies u \in V_\gamma(d) \cap B$

By $(*)$, $|g(f(x)) - g(f(c))| < \varepsilon$, so $\forall x \in V_\delta(c) \cap A$, $|(g \circ f)(x) - (g \circ f)(c)| < \varepsilon$.

$\implies g \circ f$ is continuous at a. □

Now, for example, since $\sqrt{\cdot} : \mathbb{R}_0^+ \to \mathbb{R}$ is continuous, the function $\sqrt{x^2+1}$ is continuous as it's the composition of a polynomial and the square root.

5.5 *Continuity and Topology*

Since continuous functions have nice properties, how do images these functions preserve characteristics of the domain we input. For example, if D is open, is $f(D)$ open? Are the images of closed or bounded domains also closed or bounded?

The answer to these questions is no in general. See the following examples: if $f(x) = x^2$, the open set $(-1, 1)$ gets mapped to $f(-1, 1) = [0, 1)$ which is not open. Likewise, if $f(x) = \frac{1}{x}$, the closed set $[1, \infty)$ is mapped to $(0, 1]$ which is not closed. The same function maps the bounded set $(0, 1]$ to $[1, \infty)$ which is not bounded.

Despite this, it turns out if a domain is compact, the properties of comapctness are actually preserved under continuous functions.

Theorem 5.25 (Preservation of Compactness)
Let A be compact, and $f : A \to \mathbb{R}$ be continuous. Then $f(A)$ is compact.

Proof. Since A is compact, every (x_n) in A has a convergent subsequence (x_{n_k}) whose limit lies in A. Let (y_n) be an arbitrary sequence in $f(A)$. Since the function $f : A \to f(A)$ is necessarily surjective, $\forall y_n \in f(A), \exists x_n \in A$ such that $f(x_n) = y_n, \forall n \in \mathbb{N}$. Let $x \in A$ be the limit of (x_{n_k}), and $y = f(x)$. f is continuous on A, and thus epsecially continuous at x, so by the sequential definition of continuity, $\lim(f(x_{n_k})) = f(x) = y$

$\implies \lim(y_{n_k}) = y$

$\implies (y_n)$ has a convergent subsequence, and since it was arbitary, $f(A)$ is sequentially compact and hence compact. □

Definition 5.26 (Global Maximum and Minimum)
Let $f : D \to \mathbb{R}$.

1. $x_0 \in D$ is a global/absolute maximum of f if $\forall x \in D, f(x) \leq f(x_0)$
2. $x_1 \in D$ is a global/absolute minimum of f if $forall x \in D, f(x_1) \leq f(x)$

Theorem 5.27 (Extreme Value Theorem)
Let $f : D \to \mathbb{R}$ be continuous, D compact and non-empty. Then f obtains both an absolute max and an absolute min on D.

Proof. Since D is compact and non-empty and f is continuous, $f(D)$ is compact and nonempty. Especially, it is bounded, so $f(D)$ has a sup and inf, both of which belong to its boundary. Since $f(D)$ is closed, it follows that it contains all of its boundary, so $\sup f(D) \in f(D)$ and $\inf f(D) \in f(D)$

$\implies \sup f(D) = \max f(D), \inf f(D) = \min f(D)$

$\implies \forall x \in D, \min f(D) \leq f(x) \leq \max f(D)$

$\implies \exists x_0, x_1 \in D$ such that $f(x_0) = \min f(D), f(x_1) = \max f(D)$ and $f(x_0) \leq f(x) \leq f(x_1)$.

So x_0 is an absolute minimum, and x_1 is a absolute maximum. □

Theorem 5.28 (Localization of Roots)
Let $f : [a, b] \to \mathbb{R}$, f continuous such that f has opposite signs at a and b ($f(a) > 0, f(b) < 0$ or $f(a) < 0, f(b) > 0$). Then $\exists c \in (a, b)$ such that $f(c) = 0$.

Proof. Let $I_0 = [a, b]$, $a_0 = a$, $b_0 = b$. Divide I_0 into two subintervals of euqla width. If $f(\frac{a+b}{2} = 0)$, we've found the root of if f. If not, there are two possibilites:

- if $f(\frac{a+b}{2}) > 0$, let $a_1 = \frac{a_0+b_0}{2}$, $b_1 = b_0$, and $I_1 = [a_1, b_1]$
- if $f(\frac{a+b}{2}) < 0$, let $a_1 = a_0$, $b_1 = \frac{a_1+b_0}{2}$, and $I_1 = [a_1, b_1]$

Then proceed in the same fashion: if $f(\frac{a_1+b_1}{2}) = 0$, we've found the root; if not, there are two cases as above. There are two possibilities for the algorithm: either it terminates after finitely many steps with a root, or we obtain a nested sequence $I_0 \supseteq I_1 \supseteq I_2 \supseteq \ldots$ of closed and bounded intervals. By the Nested Interval Property of $\mathbb{R}$, $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$. Let $c \in \bigcap_{n\in\mathbb{N}} I_n$.

$\implies \forall n \in \mathbb{N}_0, c \in [a_n, b_n]$

$\implies |a_n - c| \leq b_n - a_n$ and $|b_n - c| \leq b_n - a_n = \frac{1}{2^n}(b - a)$

$\implies \lim(a_n) = \lim(b_n) = c$ by the Null Sequence Criterion

Then, by the sequential definition of continuity we can say $\lim(f(a_n)) = f(c)$ and $\lim(f(b_n)) = f(c)$ Since by construction, $f(a_n) > 0$ and $f(b_n) < 0$ $\forall n \in \mathbb{N}$, and limits preserve weak inequalities, $\lim(f(a_n)) \geq 0$ and $\lim(f(b_n)) \leq 0$

$\implies f(c) \leq 0$ and $f(c) \geq 0$ so $f(c)$ must equal 0. □

Localization of roots is a special case of a stronger result we will prove:

Theorem 5.29 (Intermediate Value Theorem)
Let $f : [a, b] \to \mathbb{R}$, f continuous. Let $y_0 \in \mathbb{R}$ be between $f(a)$ and $f(b)$, i.e. $f(a) < y_0 < f(b)$ or $f(b) < y_0 < f(a)$. Then $\exists x_0 \in (a, b)$ such that $y_0 = f(x_0)$.

Proof. Define $g : [a, b] \to \mathbb{R}$, $g(x) = f(x) - y_0$. g is continuous on $[a, b]$; $g(a) = f(a) - y_0$, and $g(b) = f(b) - y_0$. From the two cases, either $g(a) < 0, g(b) > 0$ or $g(a) > 0, g(b) < 0$. In either case, by the Localization of Roots Theorem, $\exists x_0 \in (a, b)$ such that $g(x_0) = 0$.

$\implies f(x_0) - y_0 = 0$

$\implies f(x_0) = y_0$. □

Theorem 5.30 (Preservation of Intervals)
Let I be an interval on $\mathbb{R}$ and $f : I \to \mathbb{R}$ be continuous. Then $f(I)$ is also an interval.

The result will depend on the so-called "characterization of intervals," i.e. if I is an interval if and only if $\forall a, b \in I, a < b, [a, b] \subseteq I$.

Proof. Let $y_1, y_2 \in f(I)$ with $y_1 < y_2$. Obviously, $\exists x_1, x_2 \in I$ such that $f(x_1) = y_1, f(x_2) = y_2$. Let $y_1 < y < y_2$ be arbitrary. Then by the Intermediate Value Theorem, $\exists x$ between x_1 and x_2 so that $f(x) = y$

$$\implies y \in f(I)$$

$$\implies [y_1, y_2] \subseteq f(I)$$

Hence by the characterization of intervals, $f(I)$ is an interval. □

5.6 *Uniform and Lipschitz Continuity*

Definition 5.31
Let $f : D \to \mathbb{R}$. f is said to be uniformly continuous on D if $\forall \varepsilon > 0, \exists \delta > 0 : \forall x, u \in D, |x - u| < \delta \implies |f(x) - f(u)| < \varepsilon$.

Intuitively speaking, it means that the value of δ depends only on the given value of ε, with no regard to *where* on the domain we are examining continuity. From the definition, it is clear that every uniformly continuous is continuous, but does the converse hold? In general no.

Theorem 5.32 (Sequential Criterion for the Absence of Uniform Continuity)
Let $f : D \to \mathbb{R}$. f is not uniformly continuous if and only if $\forall \varepsilon > 0, \exists (x_n), (u_n)$ in D such that:

- $\lim(x_n - u_n) = 0$
- $\forall n \in \mathbb{N}, |f(x_n) - f(u_n)| \geq \varepsilon$.

Lemma 5.33
$\neg(P \implies Q) \equiv P \wedge (\neg Q)$

Proof. $P \implies Q \equiv \neg P \vee Q$

$$\implies \neg(P \implies Q) \equiv \neg(\neg P \vee Q) \equiv P \wedge \neg Q$$ □

Proof of Theorem. ($\implies$) Let f not be uniformly continuous on D, i.e. $\neg(\forall \varepsilon > 0, \exists \delta > 0 : \forall x, u \in D, |x - u| < \delta \implies |f(x) - f(u)| < \varepsilon)$

$$\implies \exists \varepsilon > 0 : \forall \delta > 0, \exists x, u \in D \text{ such that } |x - u| < \delta \wedge |f(x) - f(u)| \geq \varepsilon.$$

Choose such an ε, and let $\forall n \in \mathbb{N}, \delta := \frac{1}{n}, x_n, u_n$ such that $|x_n - u_n| < \delta$ and $|f(x_n) - f(u_n)| \geq \varepsilon$. Then we have $\forall n \in \mathbb{N}, |x_n - u_n| < \frac{1}{n}$. By the Null Sequence Criterion, the sequence $(x_n - u_n)$ converges to 0 and $|f(x_n) - f(u_n)| \geq \varepsilon$.

($\Longleftarrow$) Let $\varepsilon > 0$, (x_n), (u_n) as in the theorem, and assume f is uniformly continuous.

$\implies \exists\delta > 0 : \forall x, u \in D, |x-u| < \delta \implies |f(x) - f(u)| < \varepsilon.$

Since $\lim(x_n - u_n) = 0$, $\exists N \in \mathbb{N} : \forall n \geq N, |x_n - u_n| < \delta$. By the definition of uniform continuity, $|f(x_n) - f(u_n)| < \varepsilon$ since (x_n), (u_n) are in D. But this contradicts the assumption of the theorem, so f cannot be uniformly continuous. □

Example 5.34
Show that $f(x) = \frac{1}{x}$ is uniformly continuous on $[a, \infty)$ for any $a > 0$ and f is not uniformly continuous on $(0, \infty)$.

Proof. Let $\varepsilon > 0$, $\delta > 0$ be arbitrary for now. Let $x, u \in [a, \infty)$ such that $|x - u| < \delta$. Then:

$\left|\frac{1}{x} - \frac{1}{u}\right| = \left|\frac{u-x}{ux}\right| < \frac{\delta}{xu}$

Since $x, u \geq a$, $\frac{1}{x}, \frac{1}{u} \leq \frac{1}{a} \implies \frac{1}{xu} \leq \frac{1}{a}^2$

$\implies \frac{\delta}{xu} \leq \frac{\delta}{a^2} < \varepsilon \iff \delta < \varepsilon a^2$

Since δ does not depend on x, only on a, we can say $\forall x, u \in [a, \infty)$, $|x-u| < \delta \implies |f(x)-f(u)| < \varepsilon$ and f is uniformly continuous on this interval.

To show f is not uniformly continuous on $(0, \infty)$, let $(x_n) = \frac{1}{n}$ and $(u_n) = \frac{1}{2n}$. Then $(x_n - u_n) = \frac{1}{2n}$ which as limit 0, and $|f(x_n) - f(u_n)| = |n - 2n| = n \geq 1$. $\varepsilon := 1$ then satisfies the sequential criterion, and f fails to be uniformly continuous. □

Uniform Continuity, Convergence and Topology

If (x_n) is a convergent sequence in D, and $f : D \to \mathbb{R}$ is a continuous function, does the image sequence $f(x_n)$ always converge? No for ordinary continuity. If $\lim(x_n) \in D$, the sequential definition of continuity allows this to work, since $(x_n) \to x \implies f(x_n) \to f(x) \in f(D)$. But if $\lim(x_n) \notin D$, say $f : (0, \infty) \to \mathbb{R}$, $(x_n) = \frac{1}{n}$, (x_n) is in D but its limit 0 is not. Then $f(x_n) = n$ which diverges.

Theorem 5.35
Let $f : D \to \mathbb{R}$ be uniformly continuous and (x_n) a convergent sequence in D. Then $f(x_n)$ is convergent. Equivalently, if (x_n) is cauchy, $f(x_n)$ is also cauchy.

Proof. Let $\varepsilon > 0$. Then $\exists\delta > 0 : \forall x, u \in D, |x - u| < \delta \implies |f(x) - f(u)| < \varepsilon$. Let (x_n) be convergent and in D, and thus cauchy. Then $\exists N \in \mathbb{N}$ such that $\forall n, m \geq N$, $|x_n - x_m| < \delta$. Since $\forall n \in \mathbb{N}, x_n, x_m \in D$, $|f(x_n) - f(x_m)| < \varepsilon$, so $f(x_n)$ is cauchy and thus convergent. □

Example 5.36
(Assume $\sin x$ is continuous) Let $f : (0, 1] \to \mathbb{R}$, $f(x) = \sin(\frac{1}{x})$. Show that f is not uniformly continuous.

Proof. Since f is the composition of $\sin(x)$ and $\frac{1}{x}$, we know $\sin(\frac{1}{x})$ is continuous. However, consider $(x_n) = \frac{1}{(2n-1)\frac{\pi}{2}}$. $\lim(x_n) = 0$ by algebraic limit laws, but $f(x_n) = \sin((2n-1)\frac{\pi}{2})$, which is 1 if

n is odd and -1 if n is even. Thus $f(x_n)$ diverges, which ensures f cannot be uniform continuous by last theorem. □

We saw that continuous functions only preserve compactness, and not boundedness on its own; but uniform continuous functions have the ability to preserve boundedness on domains that are not necessarily closed

Theorem 5.37
Let $f : D \to \mathbb{R}$ be uniformly continuous, and D be bounded. Then $f(D)$ is bounded.

Proof. Assume $f(D)$ is unbounded. Then $\forall n \in \mathbb{N}, \exists x_n \in D : f(x_n) \geq n$. Consider (x_n); since D is bounded, by Bolzano-Weierstrass, $\exists (x_{n_k})$ convergent. By last theorem, $f(x_{n_k})$ is also convergent and thus bounded, but we just said $\forall k \in \mathbb{N}, |f(x_{n_k})| \geq n_k \geq k$ which means $f(x_{n_k})$ is unbounded. Contradiction, and thus $f(D)$ must be bounded. □

Theorem 5.38
Let $D \subseteq \mathbb{R}$ be compact, $f : D \to \mathbb{R}$ be continuous. Then f is uniformly continuous on D.

Proof. By sequential compactness, D has a convergent subsequence x_{n_k} whose limit x is in D. Assume f is not uniformly continuous. Then the two sequence criterion says $\exists (x_n), (u_n)$ in D such that $\lim(x_n - u_n) = 0$ and $|f(x_n) - f(u_n)| \geq \varepsilon$ $(*)$.

Consider a subsequence of (u_n), (u_{n_k}). Note that $u_{n_k} = x_{n_k} - (x_{n_k} - u_{n_k})$

$$\implies \lim(u_{n_k}) = \lim(x_{n_k}) - \underbrace{\lim(x_{n_k} - u_{n_k})}_{=0} = x$$

$\implies$ (u_{n_k}) has the same limit as (x_{n_k}), so by the sequential definition of continuity, $\lim(f(x_{n_k})) = \lim(f(u_{n_k})) = f(x)$

$$\implies \lim\left(f(x_{n_k}) - f(u_{n_k})\right) = 0$$

$$\implies \exists K \in \mathbb{N} : \forall k \geq K, |f(x_{n_k}) - f(u_{n_k})| < \varepsilon$$

But it follows from $(*)$ that $\forall k \in \mathbb{N} : |f(x_{n_k}) - f(u_{n_k})| \geq \varepsilon$, which is a contradiction, so f must be uniformly continuous. □

This theorem essentially means that on compact domains, the properties of continuous and uniformly continuous functions are identical.

Lipschitz Continuity

Definition 5.39 (Lipschitz Continuity)
$f : D \to \mathbb{R}$ is Lipschitz/Lipschitz Continuous if $\exists k > 0 : \forall x, u \in D, |f(x) - f(u)| \leq k|x - u|$. k is referred to as a Lipschitz constant of f.

Since k is positive, this actually holds for all constants greater than k as well, so it really only makes senes to talk about the *smallest* Lipschitz constant.

Theorem 5.40
Every Lipschitz continuous function is uniformly continuous.

Proof. Let $f : D \to \mathbb{R}$ be Lipschitz, $\varepsilon > 0$, and $\delta = \frac{\varepsilon}{k}$ so that $|x - u| < \delta$. Then $|f(x) - f(u)| \leq k|x - u| < k\delta = \varepsilon$, so f is uniformly continuous. □

The converse does not hold; uniform continuous functions are not always Lipschitz.

Example 5.41
Show that $\sqrt{x}$ is not Lipschitz on $[0, 1]$.

Proof. Assume f is Lipschitz, with constant k. Let $x := \frac{1}{4k^2}$ and $u = 0$.

$\implies k|x - u| = k\frac{1}{4k^2} = \frac{1}{4k}$, and:

$|f(x) - f(u)| = \left|\frac{1}{2k}\right| > \frac{1}{4k} = |x - u|$

which is not possible if f is Lipschitz. □

6 Metric Spaces

Definition 6.1 (Metric)
Given a set X, we say that a function $d : X \times X \to [0, \infty)$, $(x, y) \mapsto d(x, y)$ is a metric or distance function if:

1. $\forall (x, y) \in X \times X, d(x, y) > 0$ if $x \neq y$, and $d(x, y = 0)$ if and only if $x = y$
2. $d(x, y) = d(y, x)$
3. $\forall x, y, z \in X, d(x, z) \leq d(x, y) + d(y, z)$

Any set X equipped with a metric d is called a metric spaces (X, d).

Example 6.2
Some examples of metric spaces we will encounter:

1. For every $n \in \mathbb{N}$, $p \in [1, \infty]$, define $d_p : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$, $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$, where:

$$(x, y) \longmapsto \begin{cases} \left(\sum_{i=1}^n |x_i - y_i|^p\right)^{1/p} & \text{if } p < \infty \\ \max\limits_{1 \le i \le n} |x_i - y_i| & \text{if } p = \infty \end{cases}$$

2. Complex modulus: $\mathbb{C} \times \mathbb{C} \to (0, \infty)$, $(z_1, z_2) \mapsto |z_1 - z_2| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}$ if $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$
3. The discrete metric: for any set X, $x, y \in X$,

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

To see that (3), is indeed a metric, axiom 1 and 2 are satisfied by definition. To check the triangle inequality, consider the four cases:

$$\forall x, y, z \in X, d(x, z) = \begin{cases} 0 = 0 + 0 = d(x, z) + d(z, y) & \text{if } x = y = z \\ 1 = 0 + 1 = d(x, y) + d(y, z) & \text{if } x = y \neq z \\ 1 = 1 + 0 = d(x, y) + d(y, z) & \text{if } x \neq y = z \\ 1 \le 1 + 1 = d(x, y) + d(y, z) & \text{if } x \neq y \neq z \end{cases}$$

In all cases, $d(x, z)$ is no larger than $d(x, y) + d(y, z)$, so the discrete metric satisfies the triangle inequality.

Further, the metric shown in (1) is a special class of metrics called a norm.

Definition 6.3 (Normed Vector Space)
Given a vector space X over $\mathbb{R}$ or $\mathbb{C}$, we say that a function $||.|| : x \mapsto ||x||$ is a norm if:

1. $\forall x \in X$, $||x|| > 0$ if $x \neq 0$ and $||x|| = 0$ if and only if $x = 0$
2. $\forall x \in X$, $\lambda \in \mathbb{R}$ or $\mathbb{C}$, $||\lambda x|| = |\lambda|||x||$
3. $\forall x, y \in X$, $||x + y|| \le ||x|| + ||y||$

We say $(X, ||.||)$ is a normed vector space.

Some examples may include $\mathbb{R}^n$ equipped with the p-norm we defined earlier $||.||_p$, or on $C([a, b])$ the set of continuous functions, which forms a vector space, define $||.||_\infty : f \mapsto \max_{a \le x \le b} |f(x)| = ||f||_\infty$. $||f||_\infty$ is known as the uniform norm on $C([a, b])$:

- $||f||_\infty \ge 0$ since $|f(x)| \ge 0 \; \forall x \in [a, b]$
- $||f||_\infty = 0 \iff \max |f(x)| 0 \iff |f(x)| \le 0 \iff f(x) = 0 \; \forall x \in [a, b]$
- $||\lambda f||_\infty = |\lambda f(x)| = |\lambda| \max f(x) = |\lambda|||f||_\infty$

- $\forall f, g \in C([a,b]), |f(x)+g(x)| \leq |f(x)|+|g(x)| \leq \max|f(x)|+\max|g(x)| = \|f\|_\infty + \|g\|_\infty$

This proves our statement.

Theorem 6.4 (Hölder's Inequality)
$\forall p \in [1,\infty], x, y \in \mathbb{R}^n$,

$$\sum_{i=1}^n |x_i y_i| \leq \|x\|_p \|y\|_{p'}$$

where $p' \in [1,\infty]$ is the Hölder conjugate, satisfying $\frac{1}{p} + \frac{1}{p'} = 1$, using the convention that $\frac{1}{\infty} = 0$.

Note that p' can be rewritten into cases as follows:

$$p' = \begin{cases} \frac{p}{p-1} & \text{if } 1 < p < \infty \\ 1 & \text{if } p = \infty \\ \infty & \text{if } p = 1 \end{cases}$$

Proof. (Case 1, $p = \infty$) $\forall i \in \{1,\ldots,n\}$, $|x_i y_i| = |x_i||y_i| \leq (\max_{1\leq i\leq n}|x_i|)\,|y_i|$. Therefore we have:

$$\sum_{i=1}^n |x_i y_i| \leq \sum_{i=1}^n \|x\|_\infty |y_i| = \|x\|_\infty \sum_{i=1}^n |y_i| = \|x\|_\infty \|y\|_1$$

(Case 2, $p = 1$) Same process as above.

(Case 3, $1 < p < \infty$) Assume that $x \neq 0$ and $y \neq 0$ (otherwise the argument is trivial). Let $x' = \frac{1}{\|x\|_p} x$, $y' = \frac{1}{\|y\|_{p'}} y$. Observe that by the homogeneity of norms, $\|x'\|_p = 1$ and $\|y'\|_{p'} = 1$. Secondly,

$$\sum_{i=1}^n |x'_i y'_i| = \sum_{i=1}^n \frac{|x_i y_i|}{\|x\|_p \|y\|_{p'}}$$

which means that $\sum_{i=1}^n |x_i y_i| \leq \|x\|_p \|y\|_{p'} \iff \sum_{i=1}^n |x'_i y'_i| \leq 1$. Young's inequality states that $\forall a, b \geq 0$, $ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}$, which follows from the convexity of the exponential function. Since

$$ab = e^{\ln a} e^{\ln b} = e^{1/p \ln a^p} e^{1/p' \ln b^{p'}} = e^{1/p \ln a^p + 1/p' \ln b^{p'}}$$

let $t = \frac{1}{p}$, $1 - t = p'$, $A = \ln a^p$, $B = \ln b^{p'}$. We then have $ab = e^{tA+(1-t)B}$. Now, since $\left(\frac{d}{dx}\right)^2 e^x = e^x > 0\ \forall x$, we apply Jensen's inequality, which means $e^{tA+(1-t)B} \leq te^A + (1-t)e^B$. Assuming this fact for now, we have proven $ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}$ as desired. Finally, taking $a = |x'_i|$ and $b = |y'_i|$, we get:

$$\begin{aligned} |x'_i y'_i| &\leq \frac{1}{p}|x'_i|^p + \frac{1}{p'}|y'_i|^{p'} \\ \implies \sum_{i=1}^n |x'_i y'_i| &\leq \sum \frac{1}{p}|x'_i|^p + \sum \frac{1}{p'}|y'_i|^{p'} \\ &= \frac{1}{p}\|x'\|_p^p + \frac{1}{p'}\|y'\|_{p'}^{p'} \\ &= \frac{1}{p} + \frac{1}{p'} = 1 \implies \sum_{i=1}^n |x'_i y'_i| \leq 1 \end{aligned}$$

which concludes the proof that $\sum_{i=1}^{n} |x_i y_i'| \leq \|x\|_p \|y\|_{p'}$. □

Theorem 6.5 (Minkowski's Inequality)
$\forall p \in [1, \infty], x, y \in \mathbb{R}^n$, $\|x+y\|_p \leq \|x\|_p + \|y\|_p$.

Proof. (Case 1, $p = \infty$) $\forall 1 \leq i \leq n, |x_i + y_i| \leq |x_i| + |y_i| \leq \max |x_i| + \max |y_i|$. Therefore $\max |x_i + y_i| \leq \max |x_i| + \max |y_i|$, which means that $\|x+y\|_\infty \leq \|x\|_\infty + \|x\|_\infty$.

(Case 2, $p = 1$) Follows from triangle inequality of absolute value.

(Case 3, $1 < p < \infty$) Note that

$$\|x+y\|_p^p = \sum_{i=1}^{n} |x_i + y_i|^p = \sum_{i=1}^{n} |x_i + y_i|^{p-1} |x_i + y_i| \leq \sum_{i=1}^{n} |x_i + y_i|^{p-1} |x_i| + \sum_{i=1}^{n} |x_i + y_i|^{p-1} |y_i|$$

Then, by Hölder's Inequality, it follows that:

$$\|x+y\|_p^p \leq \left(\sum_{i=1}^{n} |x_i + y_i|^{(p-1)p'}\right)^{1/p'} \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |x_i + y_i|^{(p-1)p'}\right)^{1/p'} \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}$$

Since $(p-1)p' = p$, we have:

$$\|x+y\|_p^p \leq \|x+y\|_p^{p/p'} \|x\|_p + \|x+y\|_p^{p/p'} \|y\|_p = \|x+y\|_p^{p/p'} (\|x\|_p + \|y\|_p)$$

Thus

$$\|x+y\|_p^{p-p/p'} \leq \|x\|_p + \|y\|_p$$

Again, we note that $p - \frac{p}{p'} = p - (p-1) = 1$, so we convieniently have:

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p$$

□

Inner product spaces

Theorem 6.6 (Cauchy-Schwartz Inequality)

Proposition 6.7 (Induced Norm of Inner Product)

6.1 Balls, Open Sets, and Closed Sets

Definition 6.8 (Open Ball)
Let (X, d) be a metric space, $x \in X, r > 0$. We call the open ball of center x and radius r, denoted $B(x, r)$, or $B_d(x, r) = \{y \in X : d(x, y) < r\}$.

This definition changes drastically based on what metric we are using. For example, in $(\mathbb{R}^n, \|.\|_p)$, and $(\mathbb{R}^n, \|.\|_q)$ for $p < q$, $B_{\|.\|_p} \subseteq B_{\|.\|_q}$ since $\|x-y\|_p \leq \|x-y\|_q$. When $n = 1$, for any p, $B_{\|.\|_p}$ is the open interval $(x-r, x+r)$. Lastly, given (X, d) where d is the discrete metric we defined earlier,

$$B_d = \begin{cases} \{x\} & \text{if } r \leq 1 \\ X & \text{if } r > 1. \end{cases}$$

Definition 6.9 (Open and Closed Sets)
Let (X, d) be a metric space. $A \subseteq X$ is open if $\forall a \in A, \exists r > 0$ such that $B(x, r) \subseteq A$. We say that A is closed in (X, d) if $X \setminus A$ is open in (X, d).

Proposition 6.10
Let (X, d) be a metric space, $x_0 \in X, r_0 > 0$.

1. $B(x_0, r_0)$ is open in (X, d)
2. $\{y \in X : d(x_0, y) \leq r_0\}$ is closed in (X, d)

Proof. 1. Let $x \in B(x_0, r_0)$ and define $r = r_0 - d(x, x_0)$. Then $\forall y \in B(x, r)$, $d(x_0, y) \leq d(x_0, x) + d(x, y) < d(x_0, x) + r_0 - d(x, x_0) = r_0$, i.e. $y \in B(x_0, r_0)$, which shows $B(x, r) \subseteq B(x_0, r_0)$.

2. Define $A = \{y \in X : d(x_0, y) \leq r_0\}$. Let $x \in X \setminus A$ and $r = d(x, x_0) - r_0$. Then, if $y \in B(x, r)$, $d(x, x_0) \leq d(x, y) + d(y, x_0) < r + d(y, x_0) = d(x, x_0) - r_0$. Rearranging, we have that $r_0 < d(y, x_0)$, so $y \in X \setminus A$. This proves that $X \setminus A$ is open, and A is closed. □

Proposition 6.11
Let (X, d) be a metric space.

1. X and $\varnothing$ are both open and closed in (X, d)
2. If $(A_i)_{i \in I}$ is a collection of open sets in (X, d), $\bigcup_{i \in I} A_i$ is open in (X, d)
3. If $(A_i)_{1 \leq i \leq n}$ is a finite collection of open sets in (X, d), $\bigcup_{i=1}^n A_i$ is open in (X, d)
4. If $(A_i)_{i \in I}$ is a collection of closed sets in (X, d), $\bigcap_{i \in I} A_i$ is open in (X, d)
5. If $(A_i)_{1 \leq i \leq n}$ is a finite collection of closed sets in (X, d), $\bigcap_{i=1}^n A_i$ is closed in (X, d)
6. If A_1 is open and A_2 is closed in (X, d), then $A_1 \setminus A_2$ is open in (X, d)
7. If A_1 is closed and A_2 is open in (X, d), then $A_1 \setminus A_2$ is closed in (X, d)

Proof. 1. $\forall x \in X, B(x, 1) \subseteq X$ so X is open in X. $\varnothing$ is open since $\nexists x \in \varnothing$. At the same time X is closed since $X \setminus X = \varnothing$ which is open, $\varnothing$ is open as $X \setminus \varnothing = X$ which is open.

2. Let $(A_i)_{i \in I}$ be a collection of open sets in (X, d), and $x \in \bigcup_{i \in I} A_i$. Then $\exists i_x \in I$ such that $x \in A_{i_x}$, i.e. $\exists r_x > 0$ so that $B(x, r_x) \subseteq A_{i_x} \subseteq \bigcup_{i \in I} A_i$, which shows that $\bigcup_{i \in I} A_i$ is open.

3. Let $(A_i)_{1 \leq i \leq n}$ be a finite collection of open sets in (X, d), $x \in \bigcap_{i=1}^n A_i$. Then $\forall i \in I, x \in A_i$, and $\exists r_1, r_2, \ldots, r_n$ so that $B(x, r_i) \subseteq A_i$. Let $r = \min\{r_1, r_2, \ldots, r_n\}$. Then $\forall 1 \leq i \leq n, B(x, r) \subseteq A_i$, meaning $B(x, r) \subseteq \bigcap_{i=1}^n A$. Thus, $\bigcap_{i=1}^n A_i$ is open.

4. Let $(A_i)_{i\in I}$ be a collection of closed sets in (X, d). By De Morgan's law $\bigcap_{i\in I} A_i = X \setminus \bigcup_{i\in I}(X \setminus A_i)$. Since each $X \setminus A_i$ is open, and by (3) arbitary unions of open sets are open. Hence $\bigcap_{i\in I} A_i$ is closed.

5. Let $(A_i)_{1\leq i\leq n}$ be a finite collection of closed sets in (X, d). Then $\bigcup_{i=1}^n A_i = X \setminus \bigcap_{i=1}^n (X \setminus A_i)$ which is open by (3).

6. Let A_1 be open, A_2 closed. Then $A_1 \setminus A_2 = A_1 \cap (X \setminus A_2)$ which is open by (3) since A_1 and $X \setminus A_2$ is open.

7. $A_1 \setminus A_2 = A_1 \cap (X \setminus A_2)$, and since A_1 and A_2 are closed, by (4) $A_1 \setminus A_2$ is closed. □

In the case of the discrete metric, defined on any set X, all subsets $A \subseteq X$ are open since $\forall x \in A, B(x, 1) = \{x\} \subseteq A$. In this fashion, all sets are closed as well.

Definition 6.12 (Closure, Interior, Boundary)
Let (X, d) be a metric space, $A \subseteq X$.

1. We call the closure of A in (X, d) the set $\overline{A} = \{x \in X : \forall r > 0, B(x, r) \cap A \neq \varnothing\}$
2. We call the interior of A in (X, d) the set $\mathring{A} = \{x \in X : \exists r > 0 \text{ s.t. } B(x, r) \subseteq A\}$
3. We call the boundary of A in (X, d) the set $\partial A = \overline{A} \setminus \mathring{A}$.

We will examine what the closure, interior, and boundary of the open balls we defined earlier are in a normed vector space, as well as a space equipped with the discrete metric. In any normed vector space $(X, \|.\|)$, $x \in X, r_0 > 0$, $\overline{B(x_0, r_0)} = \{x \in X : \|x - x_0\| \leq r_0\}$. Indeed, $\forall x \in X$:

- if $\|x - x_0\| > r_0$, then $B(x, \|x - x_0\| - r_0) \cap B(x_0, r_0) = \varnothing$
- if $\|x - x_0\| = r_0$, then $\forall r > 0, B(x, r) \cap B(x_0, r_0) \neq \varnothing$, since

This is not true in most cases: with the discrete metric (X, d), taking $r_0 = 1$, $B(x_0, r_0) = \{x_0\}$, and it holds that $\forall x \in X, r \in (0, 1), B(x, r) \cap B(x_0, r_0) = \{x_0\} \cap \{x\}$. This is nonempty only if $x_0 = x$, meaning $\overline{B(x_0, r_0)} = \{x_0\}$ in (X, d).

Next, we claim that in $(X, \|.\|)$, $\mathring{\overline{B(x_0, r_0)}} = B(x_0, r_0)$. Let $A = \{x \in X : \|x - x_0\| \leq r_0\}$. If $\|x - x_0\| > r_0$, then $B(x, r) \nsubseteq A$ and thus $B(x, r) \nsubseteq \mathring{A}$, so $x \notin \mathring{A}$. If $\|x - x_0\| = r_0$, $\forall r > 0$, $B(x, r) \nsubseteq A$, since $(1 + \varepsilon)(x - x_0) + x_0$. Nonetheless, if $\|x - x_0\| < r_0$, $B(x, r_0 - \|x - x_0\|) \subseteq A$, so $x \in \mathring{A}$. With the discrete metric and $r_0 = 1$, $B(x, r_0) = \{x \in X : d(x, x_0) \leq 1\} = X$, and $\mathring{X} = X$ since $\forall x \in X, B(x, 1) \subseteq X$. On the other hand, $B(x_0, r_0) = \{x_0\}$.

Finally, in $(X, \|.\|)$, $\partial B(x_0, r_0) = \overline{B(x_0, r_0)} \setminus B(x_0, r_0) = \{x \in X : \|x - x_0\| = r_0\}$. In (X, d), $\partial B(x_0, r_0) = \{x_0\} \setminus \{x_0\} = \varnothing$.

Proposition 6.13
Let (X, d) be a metric space and $A \subseteq X$.

1. $\overline{X \setminus A} = X \setminus \mathring{A}$ and $\overbrace{X \setminus A}^{\circ} = X \setminus \overline{A}$
2. $\partial A = \overline{A} \cap \overline{X \setminus A} = \partial(X \setminus A)$

Proof. 1. $x \in \overline{X \setminus A} \iff \forall r > 0, \underbrace{B(x,r) \cap (X \setminus A)}_{B(x,r)\setminus A} \neq \varnothing$

$\equiv \neg(\exists r > 0 : B(x,r) \setminus A = \varnothing)$

$\equiv \neg(\exists r > 0 : B(x,r) \subseteq A)$

$\equiv \neg(x \in \mathring{A})$, i.e. $x \in X \setminus \mathring{A}$. If we apply the same process to $X \setminus A$ instead of A, we obtain $\overline{X \setminus (X \setminus A)} = X \setminus \overbrace{X \setminus A}^{\circ}$, which gives $X \setminus \overline{A} = X \mathring{\setminus} A$.

2. $\partial A = \overline{A} \setminus \mathring{A} = \overline{A} \cap (X \setminus \mathring{A}) = \overline{A} \cap \overline{X \setminus A}$ by part (1). Since this holds for A and $X \setminus A$, this also proves that $\partial A = \partial(X \setminus A)$. □

Proposition 6.14
Let (X,d) be a metric space and $A \subseteq X$. Then $\mathring{A}$ is the largest open subset of A in the sense that:

1. $\mathring{A} \subseteq A$
2. $\mathring{A}$ is open in (X,d)
3. $B \subseteq \mathring{A}$ for every $B \subseteq A$ that is open

Moreover, 4. $\mathring{A} = A$ if and only if A is open.

Proof. 1. Let $x \in \mathring{A}$. Then $\exists r > 0$ such that $B(x,r) \subseteq A$. Since $x \in B(x,r)$, $x \in A$, so $\mathring{A} \subseteq A$

2. Let $x \in \mathring{A}$, i.e. $\exists r > 0$ such that $B(x,r) \subseteq A$. Let $y \in B(x,r)$; since $B(x,r)$ is open, $\exists s > 0$ such that $B(y,s) \subseteq B(x,r) \subseteq A$. This proves that $y \in \mathring{A}$, hence $B(x,r) \subseteq \mathring{A}$ and $\mathring{A}$ is open.

3. Let $B \subseteq A$ be open in in (X,d). Then $\exists r > 0$ such that $B(x,r) \subseteq B \subseteq A$, so $x \in \mathring{A}$. Thus $B \subseteq A$.

4. Let $A = \mathring{A}$. Then A is open by (2). Conversely, if A is open, (1) tells us $\mathring{A} \subseteq A$, and by (3), $A \subseteq \mathring{A}$. Thus $\mathring{A} = A$. □

Proposition 6.15
Let (X,d) be a metric space and $A \subseteq X$. Then $\overline{A}$ is the smallest closed set in (X,d) that contains A, in the sense that:

1. $A \subseteq \overline{A}$
2. $\overline{A}$ is closed in (X,d)
3. $\overline{A} \subseteq B$ for every $A \subseteq B$ that is closed.

Moreover , 4. $A = \overline{A}$ if and only if A is closed.

Proof. We apply the previous results to $X \setminus A$:

1. $\overbrace{X \setminus A}^{\circ} \subseteq X \setminus A$, so $X \setminus \overline{A} \subseteq X \setminus A$ by our earlier proposition, which means $A \subseteq \overline{A}$

2. $\overbrace{X \setminus A}^{\circ}$ is open, so $\overline{A}$ is closed.

3. Let B be closed in (X, d) such that $A \subseteq B$. Then $X \setminus B$ is open and $X \setminus B \subseteq X \setminus A$, implying that $X \setminus B \subseteq \overbrace{X \setminus A}^{\circ} = X \setminus A$. Therefore $\overline{A} \subseteq B$.

4. $A\ overline A \iff X \setminus A = X \setminus \overline{A} = \overbrace{X \setminus A}^{\circ} \iff X \setminus A$ is open $\iff A$ is closed □

Corollary 6.16
Let (X, d) be a metric space and $A \subseteq X$. Then ∂A is closed in (X, d).

Proof. $\partial A = \overline{A} \setminus \mathring{A}$ is closed since $\overline{A}$ is closed and $\mathring{A}$ is open. □

6.2 *Sequences and Convergence*

Definition 6.17
Let (X, d) be a metric space, $x \in X$, and $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements in X. We say that (x_k) converges to x or that x is the limit of (x_k) in (X, d), denoted $\lim_{k\to\infty}(x_k) = x$ if $\forall \varepsilon > 0, \exists k_\varepsilon \in \mathbb{N}$ such that $\forall k \geq k_\varepsilon, d(x, x_k) < \varepsilon$ (i.e. $x_k \in B(x, \varepsilon)$).

Proposition 6.18
Let (X, d) be a metric space, $x \in X$, and $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements in X.

1. $\lim_{k\to\infty}(x_k) = x$ in $(X, d) \iff \lim_{k\to\infty} d(x, x_k) = 0$ as a sequence in $(\mathbb{R}, |.|)$
2. If $\lim_{k\to\infty}(x_k)$ exists, it is unique
3. $\lim_{k\to\infty}(x_k) = x \iff \lim_{j\to\infty}(x_{k_j}) = x$ for every subsequence $(x_{k_j})_{j\in\mathbb{N}}$.

Proof. 1. Recall that $\lim_{k\to\infty} d(x, x_k) = 0 \iff \forall \varepsilon > 0, \exists k_\varepsilon \in \mathbb{N} : |d(x, x_k) - 0| < \varepsilon \iff d(x, x_k) < \varepsilon$

2. Asume that $\lim(x_k) = x$ and $\lim_{k\to\infty}(x_k) = x'$. Then, by the triangle inequality, $d(x, x') \leq d(x, x_k) + d(x_k, x')$, and since both $d(x, x_k) + d(x_k, x') \to 0$ as $k \to \infty$, we have $d(x, x') \to 0$ i.e. $x = x'$.

3. $\lim_{k\to\infty}(x_k) = x \iff \lim_{k\to\infty} d(x, x_k) \to 0$. Since $(d(x, x_k))_{k\in\mathbb{N}}$ is a sequence in $\mathbb{R}$, we know from analysis 1 that $\lim_{j\to\infty} d(x, x_{k_j}) = 0$ for every subsequence $(x_{k_j})_{j\in\mathbb{N}}$. Thus by (1), $\lim_{j\to\infty}(x_{k_j}) = x$. □

A few examples of convergence under different metrics may enlighten this discussion: in $(\mathbb{R}^n, \|.\|_p)$, we have component wise convergence, i.e. $\lim_{k\to\infty}(x_k)_i = x_i, i = 1, \ldots, n$, where $(x_k)_i$ and x_i are the i^{th} coordinate of (x_k) and (x). Indeed, observe that $|(x_k - x)_i| \leq \|x_k - x\|_p$. Hence if $\|x_k - x\|_p \to 0$, $(x_k - x)_i \to 0$. Conversely if $|(x_k - x)_i| \to 0 \forall i = 1, \ldots, n$, then $\max |(x_k - x)_i| \to 0$ and $\|x_k - x\|_p \to 0$ for $p < \infty$ (we will prove on the assignment that $\|.\|_p \leq \|.\|_q$ for $p > q$).

In the case of the discrete metric on (X, d), $\lim_{k\to\infty}(x_k) = x$ in $(X, d) \iff \exists k_1 \in \mathbb{N}$ such that $\forall k \geq k_1, x_k = x$. If this is achieved, then $d(x, x_k) = 0 \leq \varepsilon\ \forall \varepsilon > 0$.

Finally, in $C([a,b], \|.\|_\infty)$, $\lim_{k\to\infty}(f_k) = f \iff \exists k_\varepsilon \in \mathbb{N}$ so that $\forall k \geq k_\varepsilon$, $\|f_k - f\|_\infty < \varepsilon$

$\iff \forall\varepsilon > 0, k_\varepsilon \in \mathbb{N}\forall k \geq k_\varepsilon, \max_{a\leq x\leq b}|f_k(x) - f(x)| < \varepsilon$, i.e. $\forall x \in [a,b] f_k(x) - f(x)| < \varepsilon$. We say that $(f_k)_k$ converges uniformly to f on $[a,b]$, or that f is the uniform limit of $(f_k)_k$ on $[a,b]$. A concrete example may be if $a = 0$, $0 < b \leq 1$, and $f(x) = x^k$. We know that $0 \leq x^k \leq b^k$ $\forall x \in [a,b]$, hence $\|f_k\|_\infty \leq b^k$. If $b < 1$, $b^k \to 0$ so $\|f_k\|_\infty \to 0$ as $k \to \infty$. If $b = 1$, then $\|f_k\|_\infty \leq |f_k(1|$, so $\|f_k\|_\infty$ does not converge to 0.

Proposition 6.19 (Sequential Characterization of Closure)
Let (X,d) be a metric space, $A \subseteq X$. Then $x \in \overline{A}$ if and only if $\exists(x_k)_{k\in\mathbb{N}}$ of element in A such that $\lim_{k\to\infty}(x_k) = x$ in (X,d).

Proof. ($\implies$) Assume that $x \in \overline{A}$. Then $\forall k \in \mathbb{N}$, $A \cap B(x, \frac{1}{k}) \neq \varnothing$, i.e. $x_k \in A \cap B(x, \frac{1}{k})$, which gives $x_k \in A$ and $d(x_k, x) < \frac{1}{k} \to 0$. It follows that $\lim_{k\to\infty} d(x_k, x) = 0$ and thus $\lim_{k\to\infty}(x_k) = x$

($\impliedby$) Assume $\exists(x_k)_{k\in\mathbb{N}}$ of element in A such that $\lim_{k\to\infty}(x_k) = x$. Then $\forall r > 0$, $\exists k_r \in \mathbb{N}$ such that $\forall k \geq k_r$, $d(x, x_k) < r$, i.e. $x_k \in B(x,r)$. Since $x_k \in A$ as well, $x_k \in B(x,r) \cap A$, which means $A \cap B(x,r) \neq \varnothing$ $\forall r > 0$. This proves $x \in \overline{A}$. □

Theorem 6.20 (Sequential Criterion for Closedness)
Let (X,d) be a metric space, and $A \subseteq X$. Then the following are equivalent:

1. A is closed in (X,d)
2. For every sequence $(x_k)_{k\in\mathbb{N}}$ of elements in A, $x \in X$, if $\lim_{k\to\infty}(x_k) = x$ in (X,d) then $x \in A$.

Proof. (1) $\implies$ (2): Assume that A is closed. Then $\overline{A} = A$. Let $(x_k)_k$ be a sequence in A and $x \in X$ such that $\lim_{k\to\infty}(x_k) = x$. By the previous result, $x \in \overline{A}$, so $x \in A$

(2) $\implies$ (1): Assume (2) holds. Since $A \subseteq \overline{A}$, it suffices to show $\overline{A} \subseteq A$. Let $x \in \overline{A}$. By the previous result $\exists(x_k)_k$ in A such that $\lim_{k\to\infty}(x_k) = x$ in (X,d). By (2), $x \in A$, meaning $\overline{A} \subseteq A$, so $\overline{A} = A$ and A is closed. □

Definition 6.21 (Cauchy Sequence)
Let (X,d) be a metric space and $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements in X. We say that $(x_k)_k$ is a Cauchy sequence in (X,d) if $\forall\varepsilon > 0$, $\exists k_\varepsilon \in \mathbb{N}$ such that $\forall k, k' \geq k_\varepsilon$, $d(x_k, d_{k'}) < \varepsilon$.

Proposition 6.22
Every convergence sequence in a metric space (X,d) is Cauchy in (X,d).

Proof. Let $(x_k)_{k\in\mathbb{N}}$ be a squence in X such that $\lim_{k\to\infty}(x_k) = x$ for some $x \in X$. Then $\forall\varepsilon > 0$, $\exists k_\varepsilon \in \mathbb{N}$ such that $\forall k \geq k_\varepsilon$, $d(x, x_k) < \varepsilon$. Then $\forall k, k' \geq k_{\varepsilon/2}$, $d(x_k, x_{k'}) \leq d(x_k, x) + d(x, x_{k'}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Hence $(x_k)_k$ is Cauchy. □

Definition 6.23 (Completeness)
We say that a metric space (X, d) is complete if every Cauchy sequence in (X, d) is convergent. A complete normed vector space is known as a Banach Space, and a complete inner product space is a Hilbert Space.

For example, $(\mathbb{R}, |.|)$ is a Banach space, but $(\mathbb{Q}, |.|)$ is not complete (Analysis 1). For every $p \in [1, \infty]$, $(\mathbb{R}^n, \|.\|_p)$ is also Banach. Indeed, let $(x_k)_k$ be a Cauchy sequence in the space. Observe that $\forall k, k' \in \mathbb{N}$, and $i = 1, \ldots, n$, $|(x_k - x_{k'})_i| \leq \|x_k - x_{k'}\|$, and $(x_k)_i$ is a cauchy sequence in $\mathbb{R}$, which is convergent, with limit say x_i. Then, it follows that $\lim_{k\to\infty}(x_k) = (x_1, \ldots, x_n)$ in $(\mathbb{R}^n, \|.\|_p)$.

In the case of the discrete metric d, if $(x_k)_k$ is a Cauchy sequence in (X, d), then $\exists k_1 \in \mathbb{N}$ such that $\forall k, k' \geq k_1$, $d(x_k, x_{k'}) < 1$, i.e. $d(x_k, x_{k'}) = 0$ which implies that $x_k = x_{k'} = x_{k_1}$. Since this is true for all terms after k_1, the limit of (x_k) is x_{k_1} and therefore any space with the discrete metric is complete.

Proposition 6.24
Let (X, d) be a complete metric space and $Y \subseteq X$. Then the metric subspace (Y, d) is complete if and only if Y is closed in (X, d).

Proof. ($\Longrightarrow$) Assume that (Y, d) is complete. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in Y and $x \in X$ such that $\lim_{k\to\infty}(x_k) = x$. Then $(x_k)_k$ is Cauchy in (X, d), and since $(x_k)_k$ is in Y, it is also Cauchy in (Y, d). By completeness of (Y, d), $(x_k)_k$ converges to some $\tilde{x} \in Y$. Since limits are unique, $\tilde{x} = x \in Y$, so by the sequence criterion for closedness, Y is closed.

($\Longleftarrow$) Assume that Y is closed in (X, d). Let $(x_k)_k$ be a Cauchy sequence in Y. Since $Y \subseteq X$, $(x_k)_k$ is Cauchy in (X, d). Since (X, d) is complete, $\lim_{k\to\infty}(x_k) = x \in Y$ by the sequence criterion. Thus (Y, d) is complete. □

Definition 6.25 (Sequential Compactness)
Let (X, d) be a metric space and $A \subseteq X$. We say that A is compact or sequentially compact in (X, d) if for every sequence $(x_k)_{k\in\mathbb{N}}$ in A, $\exists x \in A$ and a subsequence $(x_{k_j})_{j\in\mathbb{N}}$ such that $\lim_{j\to\infty}(x_{k_j}) = x$ in (X, d). If this is true for $A = X$, then we say that (X, d) is compact.

Proposition 6.26
Let (X, d) be a metric space and $A \subseteq X$.

1. If A is compact in (X, d) then A is closed in (X, d)
2. If A is compact in (X, d) then A is bounded in (X, d), in the sense that $\exists C > 0$ such that $\forall x, y \in A, d(x, y) < C$
3. If (X, d) is compact, then it is complete.

Proof. 1. Assume that A is compact in (X, d). Let $(x_k)_k$ be a sequence in A and $x \in X$ such that $\lim_{k\to\infty}(x_k) = x$. Since A is compact, $\exists(x_{k_j})$ and $\tilde{x} \in A$ such that $\lim_{j\to\infty}(x_{j_k}) = \tilde{x}$. By uniqueness of the limit, $\tilde{x} = x \in A$ and A is closed.

2. First we observe that for a fixed $x_0 \in X$, a set A is bounded in (X, d) if and only if $\exists C > 0$ such that $\forall x \in A, d(x_0, x) < C$. Indeed, if A is bounded, then fix $x_1 \in A$ and write $d(x, x_0) \leq d(x, x_1) + d(x_0, x_1) < C + d(x_0, x_1) = C_0$. Hence $d(x, x_0) < C_0$. Now, assume A is compact and by contradiction, that A is unbounded. Then, it follows from above that $\nexists C > 0$ so that $\forall x \in A, d(x, x_0) < C$, i.e. $\forall k \in \mathbb{N}, \exists x \in A$ where $d(x, x_0) \geq k$. Let $(x_k)_k$ be a sequence in A and $(x_{k_j})_j \to x$ be its convergent subsequence. Then we have that $k_j \leq d(x_{k_j}, x_0) \leq d(x_{k_j}, x) + d(x, x_0)$. However, since $\lim_{j\to\infty} d(x_{k_j}, x) = 0$ and $\lim_{j\to\infty} k_j = \infty$, we obtain a contradiction. Therefore A is bounded.

3. Assume that (X, d) is compact, and let $(x_k)_k$ now be a Cauchy sequence in (X, d). By compactness of (X, d), $existsx \in X$ and a subsequence $(x_{k_j})_j$ such that $\lim_{j\to\infty}(x_{k_j}) = x$, i.e. $\forall \varepsilon > 0, \exists j_\varepsilon \in \mathbb{N} : d(x_{k_j}, x) < \varepsilon$. Further, since (x_k) is Cauchy, $\forall \varepsilon > 0, \exists k_\varepsilon : \forall k, k' \geq k_\varepsilon, d(x_k, x_{k'}) < \varepsilon$. Thus, $\forall \varepsilon > 0$, letting $k' = k_j$ for $j > j_\varepsilon$, and $k_j \geq k_{\varepsilon/2}$, we obtain that $\forall k \geq k_{\varepsilon/2}, d(x_k, x) \leq d(x_{k_j}, x_k) + d(x_{k_j}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. This proves that $\lim_{k\to\infty}(x_k) = x$ and that (X, d) is complete. □

Proposition 6.27
Let (X, d) be a compact metric space and $A \subseteq X$. Then A is compact if and only if A is closed in (X, d).

Proof. ($\Longrightarrow$) Follows from previous result.

($\Longleftarrow$) Assume that A is closed in (X, d). Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in A. By of compactness of (X, d), $existsx \in X$ and a subsequence $(x_{k_j})_j$ such that $\lim_{j\to\infty}(x_{k_j}) = x$. Since A is closed, $x_{k_j} \in A \ \forall j \in \mathbb{N}$, it follows that $x \in A$. Therefore A is compact in (X, d). □

Theorem 6.28 (Heine-Borel)
Let $n \in \mathbb{N}, p \in [1, \infty]$, and $A \subseteq \mathbb{R}^n$. Then A is compact in $(\mathbb{R}^n, \|.\|_p)$ if and only if A is closed and bounded in $\mathbb{R}^n, \|.\|_p)$.

Proof. ($\Longrightarrow$) Always true, already proven earlier.

($\Longleftarrow$) Assume A is closed and bounded in $(\mathbb{R}^n, \|.\|_p)$. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in A. Recall that $(x_k)_i \leq \|.\|_p \ \ forallk \in \mathbb{N}, i = 1, \ldots, n$, where as usual $(x_k)_i$ is the i^{th} coordinate of $(x_k)_k$. Since x_k is in A and A is bounded, we obtain that $((x_k)_i)_{k\in\mathbb{N}} \ \forall i$ are bounded sequences in $(\mathbb{R}, |.|)$. By Bolzano-Weierstrass Theorem, it follows that there exists a subsequence of $(x_k)_k$ such that the first coordinate converges. Likewise, there exists a subsequence of that subsequence so that the second coordinate converges, and so on until we obtain an $n-$subsequence of $(x_{k_j})_j$ such that all coordinates converge, i.e. $\lim_{j\to\infty}(x_{k_j}) = x$ for some $x \in \mathbb{R}^n$. Since A is closed and $x_{k_j} \in A \ \forall j \in \mathbb{N}$, we have that $x \in A$. This proves that A is compact in $(\mathbb{R}^n, \|.\|_p)$. □

Note
Heine-Borel Theorem is not true in infinite dimensional vector spaces, which we will prove later.

Proposition 6.29
Let (X, d) be a metric space and $A \subseteq X$. If A is compact in (X, d), then for every open cover of A, namely a collection $(U_i)_{i \in I}$ of open sets in (X, d) such that $A \subseteq \bigcup_{i \in I} U_i$, there exists a finite subcover of A, namely a subcollection $(U_{i_j})_{1 \le j \le N}$ where $N \in \mathbb{N}$ and $i_1, \ldots, i_N \in I$ such that $A \subseteq \bigcup_{i=1}^{N} U_{i_j}$

Proof. Assume A is compact and let $(U_i)_{i \in I}$ be an open cover.

Claim 6.30. $\exists r > 0$ *such that* $\forall x \in A, \exists i_x \in I$ *such that* $B(x, r) \subseteq U_i$

We will finish the proof assuming this claim, and prove it at the end. Choose $x_1 \in A$ (provided $A \neq \emptyset$ but the proof is trivial otherwise) and by induction for each $j \in \mathbb{N}$ □

Proof. □

6.3 *Continuity*

Definition 6.31 (Limit and Continuity)
Let (X, d_X) and (Y, d_Y) be metric spaces, $A \subseteq X$, $x_0 \in \overline{A}$, $y_0 \in Y$. We say that $f(x)$ converges to y_0 as $x \to x_0$, and denote $\lim_{x \to x_0} f(x) = y_0$ if $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0$ such that $\forall x \in X, 0 < d_X(x, x_0) < \delta_\varepsilon \implies d_Y(f(x), y_0) < \varepsilon$. In other words, $f\left(A \cap (B_X^*(x_0, \delta_\varepsilon))\right) \subseteq B_Y(y_0, \varepsilon)$ [a]

Moreover, we say that f is continuous at x_0 if $\lim_{x \to x_0} f(x) = f(x_0)$.

[a] Here, $B^*(x, r)$ denotes the "punctured" ball $B(x, r) \setminus \{x\}$.

Some examples may include:

- the function $d(., y_0) : X \to \mathbb{R}, x \mapsto d(x, y_0)$ for a fixed y_0. We can show by the triangle inequality that $d(x, y_0) \le d(x, x_0) + d(x_0, y_0)$; and $|d(x, y_0) - d(x_0, y_0)| < d(x, x_0) < \delta_\varepsilon = \varepsilon$ works. In particular, this shows that in a normed vector space $(X, \|.\|)$, the norm $\|.\| : X \to \mathbb{R}$ is continuous since $\|x\| = d(x, 0)$.
- In an inner product space $(X, \langle ., . \rangle)$, for $y_0 \in X$ fixed, the function $X \to \mathbb{R}, x \mapsto \langle x, y_0 \rangle$ is continious. Indeed, $\forall x, x_0 \in X, |\langle x, y_0 \rangle - \langle x_0, y_0 \rangle| = |\langle x - x_0, y_0 \rangle| \le \|x - x_0\| \|y_0\|$ by Cauchy-Schwartz. Therefore letting $\delta_\varepsilon = \varepsilon / \|y\|$ as long as $y \neq 0$ works.
- For each $n_1, n_2 \in \mathbb{N}, p_1, p_2 \in [1, \infty]$ and $A \subseteq \mathbb{R}^{n_1}$, the function $f : (A, \|.\|_{p_1}) \to (\mathbb{R}^{n_2}, \|.\|_{p_2})$ that takes $x \mapsto (f_1(x), \ldots, f_n(x))$ is continuous at x_0 if and only if f_i is continuous $\forall i \in \{1, \ldots, n\}$. Indeed, if f is continuous, then $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0$ such that $\|x - x_0\|_{p_1} < \delta_\varepsilon \implies \|f(x) - f(x_0)\|_{p_2} < \varepsilon$. Since for all coordinates, $|(x - x_0)_i| \le \|x - x_0\|_{p_1}, |f_i(x) - f_i(x_0)| \le \|\|$

Proposition 6.32
Let (X, d_Y) and (Y, d_X) be metric spaces, $A \subseteq X$, $x_0 \in \overline{A \setminus \{x_0\}}$[a], and $f : A \to Y$. Then if $\lim_{x \to x_0} f(x)$ exists, it is unique.

[a] this condition ensures that x_0 is not an isolated point, in the sense that $\exists r > 0$ such that $B(x_0, r) \cap A = \{x_0\}$

Proof. Assume that by contradiction, $\exists y_0, \tilde{y_0}$ where $y_0 \neq \tilde{y_0}$ and $\lim_{x \to x_0} f(x) = y_0$, $\lim_{x \to x_0} f(x) = \tilde{y_0}$, i.e.

- $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0$ such that $\forall x \in A, 0 < d_X(x, x_0) < \delta_\varepsilon \implies d_Y(f(x), y_0) < \varepsilon$
- $\forall \varepsilon > 0, \exists \tilde{\delta_\varepsilon} > 0$ such that $\forall x \in A, 0 < d_X(x, x_0) < \tilde{\delta_\varepsilon} \implies d_Y(f(x), \tilde{y_0}) < \varepsilon$

Let $\delta'_\varepsilon = \min\{\delta_\varepsilon, \tilde{\delta_\varepsilon}\}$. Since $x_0 \in \overline{A \setminus \{x_0\}}$, $\exists x_\varepsilon \in B(x_0, \delta'_\varepsilon) \cap A \setminus \{x_0\}$ □

Proposition 6.33
Let (X, d_x), (Y, d_y), (Z, d_z) be metric spaces, $x_0 \in X, f : X \to Y, g : Y \to Z$. If f is continuous at x_0, and g is continuous at $f(x_0)$, then $g \circ f$ is continuous at x_0.

Proposition 6.34
Let (X, d_X) and Y, d_Y be metric spaces, $A \subseteq X$, and $f : X \to Y$ be continuous. If A is compact in (X, d_X), then $f(A)$ is compact in (Y, d_Y).

Proof. Let $(y_k)_{k\in\mathbb{N}}$ be a sequence in $f(A)$. Then $exists x_k \in X$ such that $f(x_k) = y_k$. Since A is compact, $(x_k)_k$ has a convergent subsequence $(x_{k_j})_j$ such that $\lim_{j\to\infty}(x_{k_j}) = x$ for some $x \in A$. By continuity of f, it follows that $\lim_{j\to\infty} y_{k_j} = f(x)$. Moreover, $f(x) \in f(A)$. Therefore $f(A)$ is compact. □

6.4 Uniform and Lipschitz Continuity

Definition 6.35
Let (X, d) be a metric space, $A \subseteq X$ and $f : X \to X$. We say that f is a contraction on A if $exists C \in (0, 1)$ such that $\forall x, y \in A, d(f(x), f(y)) < Cd(x, y)$.

Note that a contraction is obviously Lipschitz continuous, and when $C < 1$, $f(x)$ is "closer" to $f(y)$ than x is to y.

Theorem 6.36 (Banach Fixed Point Theorem)
Let (X, d) be a nonempty complete metric space, and $f : X \to X$ a contraction. Then f has a unique fixed point in X, namely there eixsts $x_0 \in X$ such that $f(x_0) = x_0$.

A common example is any function whose derivative is small enough: let I be a closed interval (so that $(I, |.|)$ is complete in $\mathbb{R}$), and $f : I \to I$ be differentiable on I such that $\sup_{x\in I} |f'(x)| < 1$. Then f has a unique fixed point in I since it is a contraction.

Proof. We will start by proving existence of the fixed point. Since X is nonempty, we can choose $x_1 \in X$. Then define $(x_k)_{k\in\mathbb{N}}$ by $x_{k+1} = f(x_k)$. Since f is a contraction, $exists C \in (0, 1)$ such that $\forall x, y \in X, d(f(x), f(y)) < Cd(x, y)$. Hence $\forall n \in \mathbb{N}, d(x_{k+2}, x_{k+1}) < Cd(x_{k+1}, x_k)$, which gives

$$d(x_{k+1}, x_k) < Cd(x_k, x_{k-1}) < C^2 d(x_{k-1}, x_{k-2}) < \cdots < C^{k-1} d(x_2, x_1)$$

It follows that for all $k' > k$:

$$d(x_k, x_{k'}) \leq d(x_k, x_{k+1}) + d(x_{k+1}, x_{k+1}) + \dots$$
$$\leq (C^{k-1} + C^k + \dots + C^{k'-2}) d(x_1, x_2) = C^{k-1} \sum_{j=0}^{k'-k-1} C^j d(x_1, x_2)$$
$$C^{k-1} \sum_{j=0}^{k'-k-1} C^j = C^{k-1} \frac{1 - C^{k'-k}}{1-C} \leq \frac{C^{k-1}}{1-C} \to 0 \text{ as } k \to \infty$$

Therefore $\forall \varepsilon > 0, \exists k_\varepsilon \in \mathbb{N}$ such that $\forall k \geq k_\varepsilon$, $\frac{C^{k-1}}{1-C} d(x_1, x_2) < \varepsilon$, i.e. $\forall k, k' \geq k_\varepsilon, d(x_k, x_{k'}) < \varepsilon$. Hence $(x_k)_k$ is Cauchy in (X, d), and thus converges to some $x_0 \in X$ by completeness of (X, d). Since f is continuous, $\lim_{k\to\infty} f(x_k) = f(x_0)$, and as $f(x_k) = x_{k+1}$, we have that $x_0 = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} f(x_k) = f(x_0)$.

To show uniqueness, assume by contradiction that $\exists x_0' \neq x_0$ such that $f(x_0) = x_0, f(x_0') = x_0'$. Since f is a contraction, $d(x_0, x_0') = d(f(x_0), f(x_0')) < Cd(x_0, x_0') < d(x_0, x_0')$. Therefore there is only one fixed point. □

7 Infinite Series

Definition 7.1

Let $(X, \|.\|)$ be a normed vector space, $k_0 \in \mathbb{Z}$, and $(a_k)_{k\in\mathbb{N}}$ be a sequence of elements in X. For each $n \geq k_0$, we call $\sum_{k=k_0}^{n} a_k$ the partial sum of order n of the series $\sum_{k=k_0}^{\infty} a_k$. If $\lim_{n\to\infty} s_n = s$ in $(X, \|.\|)$ for some $s \in X$, then we denote $s = \sum_{k=k_0}^{\infty} a_k$ and say that the series converges. If $(s_n)_n$ does not converge to any limit, we say the series diverges.

Examples of series:

- Let $r \in \mathbb{R}$, $a_k = r^k$ for all $k \in \mathbb{N}$. Then
$$\sum_{k=0}^{\infty} r^k = \begin{cases} \frac{1-r^{n+1}}{1-r} & r \neq 1 \\ n+1 & r = 1 \end{cases}$$
If $|r| < 1$, $\sum_{k=0}^{\infty} r^k$ converges with sum $\frac{1}{1-r}$.
- Let $M(n, \mathbb{R})$ be the vector space of $n \times n$ matrices with real entries, equipped with the norm
$$\|A\|_{2,2} = \sup_{\|x\|_2 = 1} \|Ax\|_2$$
Remark that $\forall x \in \mathbb{R}^n$, $\|Ax\|_2 \leq \|A\|_{2,2}\|x\|_2$, which we proved elsewhere. Then, observe that $(I_n - A)\sum_{k=0}^{n} A^k = \sum_{k=0}^{n} A^k - \sum_{k=0}^{n+1} A^k = I_n - A^{n+1}$. Suppose that $\|A\|_{2,2} < 1$. Then $I_n - A$ is invertible, since $(I_n - A)x = 0 \iff Ax = x$, and so $\|x\|_2 = \|Ax\|_2 \leq \|A\|_{2,2}\|x\|$ which contradicts $\|A\|_{2,2} < 1$ unless $\|x\|_2 = 0$. It follows that $\sum_{k=0}^{n} A^k = (I_n - A)^{-1}(I_n - A^{n+1})$. We will show that $\sum_{k=0}^{\infty} A^k = (I_n - A)^{-1}$: $\forall x \in X$,
$$\|(\sum_{k=0}^{n} A^k - (I_n - A))x\|_2 = \|(I_n - A)^{-1}(I_n A^{n+1} - I_n)x\|_2$$
$$= \|(I_n - A)^{-1} A^{n+1} x\|_2 \leq \|(I_n - A)^{-1}\|_{2,2} \|A^{n+1} x\|_2$$
$$\leq \dots \leq \|(I_n - A)^{-1}\|_{2,2} \|A\|_{2,2}^{n+1} \|x\|_2$$

which gives $\| \sum_{k=0}^{\infty} A^k - (I_n - A)^{-1}\|_{2,2} \to 0$ as $n \to \infty$ since $\|A\|_{2,2} < 1$.

Proposition 7.2 1. $\forall k' \geq k_0$, $\sum_{k=k'}^{\infty} a_k$ converges if and only if $\sum_{k=k_0}^{\infty} a_k$ converges

2. If $\sum_{k=k_0}^{\infty} a_k$ converges, $\lim_{k\to\infty} a_k = 0$

Proof. 1. $\forall n \geq k'$, $\sum_{k=k_0}^{n} = a_{k_0} + \cdots + a_{k'} + \sum_{k=k'}^{n} a_k$, where the first terms and independent of n. Hence $\sum_{k=k_0}^{\infty} a_k$ converges if and only if $\sum_{k=k'}^{\infty} a_k$ converges.

2. $\forall n \geq k_0$, $a_n = s_n - s_{n-1}$. Hence if $\lim_{n\to\infty} s_n = s$, then $\lim_{n\to\infty} a_n = s - s = 0$. □

Theorem 7.3 (Cauchy Criterion for Convergence)
Let $(X, \|.\|)$ be a Banach space. Then $\sum_{k=k_0}^{\infty} a_k$ converges if and only if $\forall \varepsilon > 0$, $\exists n_\varepsilon \geq k_0$ such that $\forall n' > n \geq n_\varepsilon$, $\| \sum_{k=n+1}^{n'} a_k\| < \varepsilon$.

Proof. If $(X, \|.\|)$ is Banach, then $\sum_{k=k_0}^{\infty} a_k$ converges if and only if $(s_n)_n$ is cauchy in $(X, \|.\|)$, i.e. $\forall \varepsilon > 0$, $\exists n_\varepsilon \geq k_0$ such that $\forall n' > n \geq n_\varepsilon$, $\|s_{n'} - s_n\| = \| \sum_{k=n+1}^{n'} a_k\| < \varepsilon$. □

Definition 7.4
We say that the series $\sum_{k=k_0}^{\infty} a_k$ converges absolutely if $\sum_{k=k_0}^{\infty} \|a_k\|$ converges in $(\mathbb{R}, |.|)$.

Note that since $\tilde{s}_n = \sum_{k=k_0}^{n} \|a_k\|$ is increasing, so $(\tilde{s}_n)_n$ converges if and only if it is bounded from above. Moreover, if $(\tilde{s}_n)_n$ is bounded from above, $\sum_{k=k_0}^{n} \|a_k\| = \sup_{n\geq k_0} \tilde{s}_n$.

Proposition 7.5
If $(X, \|.\|)$ is Banach, then for every series $\sum_{k=k_0}^{\infty} a_k$, if $\sum_{k=k_0}^{\infty} a_k$ converges absolutely, then $\sum_{k=k_0}^{\infty} a_k$ converges in $(X, \|.\|)$. Morever, $\forall n \geq k_0$, $\| \sum_{k=n}^{\infty} a_k\| \leq \sum_{k=n}^{\infty} \|a_k\|$.

Proof. Since $\sum_{k=k_0}^{\infty} a_k$ converges absolutely, by applying the Cauchy Criterion to the series in $(\mathbb{R}, |.|)$, we obtain $\forall \varepsilon > 0$, $\exists n_\varepsilon \geq k_0$ such that $\forall n \geq n_\varepsilon$,

$$\sum_{k=n+1}^{\infty} \|a_k\| = \left| \sum_{k=k_0}^{\infty} \|a_k\| - \sum_{k=k_0}^{n} \|a_k\| \right| < \varepsilon$$

It follows that $\forall n' > n \geq n_\varepsilon$,

$$\left\| \sum_{k=n}^{n'} a_k \right\| \leq \sum_{k=n}^{n'} \|a_k\| \leq \sum_{k=n}^{\infty} \|a_k\|$$

Hence by the Cauchy Criterion, since $(X, \|.\|)$ is Banach, we obtain that $\sum_{k=k+0}^{\infty} a_k$ converges. Moreover $\forall n' > n \geq k_0$, $\| \sum_{k=n}^{n'} a_k\| \leq \sum_{k=n}^{n'} \|a_k\| \leq \sum_{k=n}^{\infty} \|a_k\|$. Since $\sum_{k=n}^{\infty} a_k$ converges, by continuity of $\|.\|$, it follows that

$$\left\| \sum_{k=n}^{\infty} a_k \right\| \leq \sum_{k=n}^{\infty} \|a_k\|$$

□

REARRANGEMENT PROOF

7.1 *Convergence Tests for Series with Non-negative Terms*

Theorem 7.6 (Comparison Test)
Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be series in $(\mathbb{R}, |.|)$ such that $0 \leq a_k \leq b_k$ forall $k \in \mathbb{N}$. Then:

1. if $\sum_{k=1}^{\infty} b_k$ converges then $\sum_{k=1}^{\infty} a_k$ converges
2. if $\sum_{k=1}^{\infty} a_k$ diverges then $\sum_{k=1}^{\infty} b_k$ diverges

Proof. 1. If $\sum_{k=1}^{\infty} b_k$ converges, it is bounded from above. Since $a_k \leq b_k$, $\sum_{k=1}^{\infty} a_k$ is bounded from above as well, so it converges

2. Follows from (1) □

Theorem 7.7 (Root Test)
Let $\sum_{k=1}^{\infty} a_k$ be a series in $(\mathbb{R}, |.|)$ such that $a_k \geq 0$ forall $k \in \mathbb{N}$.

1. if $\limsup_{k\to\infty} \sqrt[k]{a_k} < 1$, then $\sum_{k=1}^{\infty} a_k$ converges
2. if $\limsup_{k\to\infty} \sqrt[k]{a_k} > 1$, then $\sum_{k=1}^{\infty} a_k$ converges

Proof. Let $r = \limsup_{k\to\infty} \sqrt[k]{a_k} = \lim_{k\to\infty} \sup_{k\geq n} \sqrt[k]{a_k}$

1. Assume $r < 1$. Then $\forall \varepsilon > 0, \exists n_\varepsilon \in \mathbb{N}$ such that $\forall n \geq n_\varepsilon$, $\sup_{k\geq n} \sqrt[k]{a_k} < r + \varepsilon$, which gives $\forall k \geq n_\varepsilon$, $\sqrt[k]{a_k} < r + \varepsilon$. Letting $\varepsilon \in (0, 1-r)$, we obtain that $\sum_{k=1}^{\infty}(r+\varepsilon)^k$ converges, and thus by comparison $\sum_{k=1}^{\infty} a_k$ converges.

2. Assume $r > 1$. Then $\exists n_1 \in \mathbb{N}$ such that $\forall n \geq n_1$, $\sup_{k\geq n} \sqrt[k]{a_k} > 1$. Hence $\forall n \geq n_1, \exists k \in \mathbb{N}$ such that $\sqrt[k]{a_k} > 1$. Collect these into a subsequence $(a_{n_k})_{k\in\mathbb{N}}$. It follows that $\sum_{n=1}^{n_k} a_n \geq \sum_{j=1}^{j} a_{n_j} \geq \sum_{j=1}^{k} 1 = k \to \infty$ as $k \to \infty$. Hence $\sum_{k=1}^{\infty} a_k$ diverges. □

Theorem 7.8 (Ratio Test)
Let $\sum_{k=0}^{\infty} a_k$ be a series in $(R, |.|)$ such that $\exists k_0 \in \mathbb{N}$ so that $\forall k \geq k_0$, $a_k \geq 0$. Then:

1. If $\limsup_{k\to\infty} \frac{a_{k+1}}{a_k} < 1$, then $\sum_{k=0}^{\infty} a_k$ converges
2. If $\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} > 1$ then $\sum_{k=0}^{\infty} a_k$ diverges.

Proof. 1. Let $r = \limsup_{k\to\infty} \frac{a_{k+1}}{a_k}$. Then $\forall \varepsilon > 0, \exists k_\varepsilon \in \mathbb{N}$ such that $\forall n \geq k_\varepsilon$, $\sup_{k\geq n} \frac{a_{k+1}}{a_k} < r + \varepsilon$. Since $r < 1$, let ε be such that $r + \varepsilon < 1$. Then we obtain $a_{k+1} \leq (r+\varepsilon)a_k \leq (r+\varepsilon)^2 a_{k-1} \leq \cdots \leq (r+\varepsilon)^{k-k_\varepsilon} a_{k_\varepsilon}$. By comparison to $\sum_{k=0}^{\infty}(r+\varepsilon)^k$, $\sum_{k=0}^{\infty} a_k$ converges.

2. Since $\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} > 1$, then $\exists k_1 \in \mathbb{N}$ such that $\forall n \geq k$, $\inf_{k\geq n} \frac{a_{k+1}}{a_k} > 1$, which implies $\forall k \geq k_1$, $\frac{a_{k+1}}{a_k} > 1$, i.e. $a_{k+1} > a_k$. Therefore $\lim_{k\to\infty} a_k \neq 0$, and $\sum_{k=0}^{\infty} a_k$ diverges. □

7.2 *Series that Converge but not Absolutely*

Theorem 7.9 (Alternating Series Test)
Let $\sum_{k=0}^{\infty}(-1)^k a_k$ be a series in $(\mathbb{R}, |.|)$ such that $\lim_{k\to\infty} a_k = 0$ and $\forall k \in \mathbb{N}, 0 \leq a_{k+1} \leq a_k$. Then $\sum_{k=0}^{\infty}(-1)^k a_k$ converges and

$$0 \leq \sum_{k=0}^{\infty}(-1)^k a_k \leq a_0$$

Proof. Observe that $\forall n \in \mathbb{N}$,

$$\begin{aligned} s_{2n} = \sum_{k=0}^{2n}(-1)^k a_k &= a_0 - a_1 + a_2 - \cdots - a_{2n-1} + a_{2n} \\ &= a_0 - (a_1 - a_2) - (a_3 - a_4) - \cdots - (a_{2n+1} - a_{2n}) \\ &= a_0 - \sum_{k=0}^{n} \underbrace{(a_{2k-1} - a_{2k})}_{\geq 0} \end{aligned}$$

Likewise,

$$s_{2n+1} = (a_0 - a_1) + (a_2 - a_3) + \cdots + (a_{2n} - a_{2n+1}) = \sum_{k=0}^{n} \underbrace{(a_{2k} - a_{2k_1})}_{\geq 0}$$

Moreover, $s_{2n+1} = s_{2n} - a_{2n+1} \leq s_{2n}$. It follows that

$$\begin{cases} (s_{2n})_n \text{ is decreasing} \\ (s_{2n+1})_n \text{ is increasing} \\ 0 \leq s_{2n+1} \leq s_{2n} \leq a_0 \end{cases}$$

Further, $(s_{2n})_n$ is bounded below by 0 and $(s_{2n+1})_n$ is bounded above by a_0. Hence $\exists s, s' \in [0, a_0]$ such that $\lim_{n\to\infty} s_{2n} = s$, $\lim_{n\to\infty} s_{2n+1} = s'$. Since $s_{2n+1} = s_{2n} - a_{2n+1}$, and $\lim_{n\to\infty} a_n = 0$ by assumption, we have that $s = s'$, which implies that $\sum_{k=0}^{\infty}(-1)^k a_k = s = s'$. □

Proposition 7.10
Let $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ be series in $(\mathbb{R}, |.|)$ such that $\sum_{k=0}^{\infty} a_k$ converges absolutely and $\sum_{k=0}^{\infty} b_k$ converges. Define for each $n \in \mathbb{N}$, $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Then $\sum_{k=0}^{\infty} c_n$ converges and

$$\sum_{k=0}^{\infty} c_n = \left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{k=0}^{\infty} b_k\right)$$

Proof. □

7.3 *Power Series*

Definition 7.11
Let $(a_k)_{k\in\mathbb{N}}$ be a sequence in $(\mathbb{R}, |.|)$, and $x_0 \in \mathbb{R}$. Then the function

$$x \longmapsto \sum_{k=0}^{\infty} a_k (x - x_0)^k$$

(at points where it converges) is called a power series.

Proposition 7.12
Let $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ be a power series in $(\mathbb{R}, |.|)$ or $(\mathbb{C}, |.|)$. Define

$$R = \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|a_k|}}$$

with the convention $\frac{1}{\infty} = 0$, $\frac{1}{0} = \infty$. Then:

1. For every $x \in \mathbb{R}$, if $|x - x_0| < R$, then $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ converges absolutely
2. If $|x - x_0| > R$, then $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ diverges

Definition 7.13
Let I be an interval in $\mathbb{R}$, $x_0 \in I$, and $f : I \to \mathbb{R}$ be infinitely differentiable at x_0. Then the series

$$\sum_{k=0}^{\infty} (x - x_0)^k \frac{1}{k!} f^{(k)}(x_0)$$

is called the Taylor Series of f at x_0.

Theorem 7.14 (Taylor's Theorem with Mean Value Remainder)
Let $x, x_0 \in \mathbb{R}$,

$$I = \begin{cases} [x, x_0] \text{ if } x < x_0 \\ [x_0, x] \text{ if } x_0 < x \end{cases}, \mathring{I} = \begin{cases} (x, x_0) \text{ if } x < x_0 \\ (x_0, x) \text{ if } x_0 < x \end{cases}$$

and $f : I \to \mathbb{R}$ be n-times differentiable on I, as well as $n+1$-times differentiable on $\mathring{I}$. Then $\exists x' \in \mathring{I}$ such that

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k + \frac{1}{(n+1)!} f^{(n+1)}(x')(x - x_0)^{n+1}$$

Remark that the usual mean value theorem corresponds to the case $n = 0$, where $f(x) = f(x_0) + f'(x')(x - x_0)$.

Proof. Define

$$g(t) = f(x) - \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(t)(x - t)^k \ \forall t \in I$$

We want to show that $\exists x' \in \mathring{I}$ such that $g(x_0) = \frac{1}{(n+1)!} f^{(n+1)}(x - x_0)^{n+1}$. g is continuous on I since

f is differentiable on I, and differentiable on $\mathring{I}$ since f is differentiable on $\mathring{I}$. Further,

$$\begin{aligned}
g'(t) &= -\sum_{k=0}^{n} \frac{1}{k!}\left(f^{(k+1)}(t)(x-t)^k - kf^{(k)}(t)(x-t)^{k-1}\right) \\
&= -\sum_{k=0}^{n} \frac{1}{k!} f^{(k+1)}(t)(x-t)^k + \underbrace{\sum_{k=0}^{n} \frac{1}{(k-1)!} f^{(k)}(t)(x-t)^{k-1}}_{=\sum_{k=0}^{n-1} \frac{1}{k!} f^{(k+1)}(t)(x-t)^k} \\
&= -\frac{1}{n!} f^{(n+1)}(t)(x-t)^n
\end{aligned}$$

Now define

$$h(t) = g(t)(x-x_0)^{n+1} - g(x_0)(x-t)^{n+1}$$

so that

$$h(x_0) = g(x_0)(x-x_0)^{n+1} - g(x_0)(x-x_0)^{n+1} = 0$$

Then we also have

$$h(x) = g(x)(x-x_0)^{n+1} - g(x_0)(x-x)^{n+1} = 0$$

since $g(x) = f(x) - f(x) = 0$. Therefore by Rolle's theorem, $\exists x' \in \mathring{I}$ such that $h'(t) = 0$, i.e.

$$\begin{aligned}
-\frac{1}{n!} f^{(n+1)}(t)(x-x')^n (x-x_0)^{n+1} + (n+1)g(x_0)(x-x')^n &= 0 \\
(n+1)(x-t)^n \left(-\frac{1}{(n+1)!} f^{(n+1)}(x')(x-x_0)^{n+1} + g(x_0)\right) &= 0 \\
g(x_0) &= \frac{1}{(n+1)!} f^{(n+1)}(x')(x-x_0)^{n+1}
\end{aligned}$$

which is what we needed to show. □

8 Riemann-Stieltjes Integral

We will start with some definitions and eventually build up to the definition of a Riemann-Stieltjes Integral.

Definition 8.1
Let $a, b \in \mathbb{R}$ suh that $a < b$. We say that a finite set P $\{x_i\}_{i=0}^n$ is a partition of $[a, b]$ if $a \le x_0 \le x_1 \le \cdots \le x_n \le b$.

Let $\alpha : [a, b] \to \mathbb{R}$ be increasing on $[a, b]$, and $f : [a, b] \to \mathbb{R}$ be bounded on $[a, b]$. For every partition $P = (x_0, \ldots, x_n)$, we define:

$$L(P, f, \alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i, U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$$

where

$$m_i = \inf_{x_{i-1} \le x \le x_i} f(x), M_i = \sup_{x_{i-1} \le x \le x_i} f(x), \Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$

We then define

$$\underline{\int_a^b} f d\alpha = \sup_P L(P, f, \alpha)$$

and

$$\overline{\int_a^b} f d\alpha = \inf_P U(P, f, \alpha)$$

$\underline{\int_a^b} f d\alpha$ and $\overline{\int_a^b} f d\alpha$ are the lower and upper Riemann Sums of f over $[a, b]$ with respect to α. In the case where they are equal, we denoted it $\int_a^b f d\alpha$. We say that f is Riemann-Stieltjes Integrable over $[a, b]$ with respect to α if $\int_a^b f d\alpha$ exists. The set of all bounded functions $f : [a, b] \to \mathbb{R}$ which are Riemann-Stieltjes integrable is denoted $\mathscr{R}([a, b], \alpha)$.

Remark that since α is monotone, it is bounded: $\forall x \in [a, b], \alpha(a) \leq \alpha(x) \leq \alpha(b)$. Further, we have that

$$\left(\inf_{a \leq x \leq b} f(x)\right) \underbrace{\sum_{i=1}^n \Delta\alpha_i}_{\alpha(b)-\alpha(a)} \leq L(P, f\alpha) \leq U(P, f, \alpha) \leq \left(\sup_{a \leq x \leq b} f(x)\right) \sum_{i=1}^n \Delta\alpha_i$$

since $\inf_{a \leq x \leq b} f(x) \leq m_i \leq M_i \leq \sup_{a \leq x \leq b} f(x)$ and $\Delta\alpha_i \leq 0$ (α monotone). Hence

$$\underline{\int_a^b} f d\alpha, \overline{\int_a^b} f d\alpha \in \left[\left(\inf_{a \leq x \leq b} f(x)\right)(\alpha(b) - \alpha(a)), \left(\sup_{a \leq x \leq b} f(x)\right)(\alpha(b) - \alpha(a))\right]$$

and if $f \in \mathscr{R}([a, b], \alpha)$,

$$\left(\inf_{a \leq x \leq b} f(x)\right)(\alpha(b) - \alpha(a)) \leq \int_a^b f d\alpha \leq \left(\sup_{a \leq x \leq b} f(x)\right)(\alpha(b) - \alpha(a))$$

Proposition 8.2
f is Riemann-Stieltjes Integrable over $[a, b]$ with respect to α if and only if $\forall \varepsilon > 0$, there exists a partition P_ε of $[a, b]$ such that $U(P_\varepsilon, f, \alpha) - L(P_\varepsilon, f, \alpha) < \varepsilon$

Proof. Assume that $f \in \mathscr{R}([a, b], \alpha)$. Then

$$\underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha = \int_a^b f d\alpha$$

and$\forall \varepsilon > 0, \exists P_\varepsilon, P'_\varepsilon$ partitions of $[a, b]$ such that

$$\int_a^b f d\alpha + \varepsilon \leq L(P_\varepsilon, f\alpha) \leq \int_a^b f d\alpha \leq U(P'_\varepsilon, f, \alpha) \leq \int_a^b f d\alpha + \varepsilon$$

Letting $P^*_\varepsilon = P_\varepsilon \cup P'_\varepsilon$, and noting that $U(P^*_\varepsilon, f, \alpha) \leq U(P'_\varepsilon, f, \alpha)$ □

Theorem 8.3
Let $a, b \in \mathbb{R}$ such that $a < b$ and $f \in \mathcal{R}([a, b])$. Then the function $F : [a, b] \to \mathbb{R}$ defined by

$$F(x) = \int_a^x f$$

is Lipschitz continuous on $[a, b]$. Moreover, if f is continuous at $x_0 \in [a, b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Proof. For all $x, y \in [a, b]$ with $x < y$,

$$\begin{aligned}|F(y) - F(x)| = \left|\int_a^y f - \int_a^x f\right| = \left|\int_x^y f dy\right| \leq \int_x^y |f| \\ \leq \left(\sup_{a \leq x \leq b} |f(x)|\right) \underbrace{(y - x)}_{=|y-x|}\end{aligned}$$

Hence F is Lipschitz continuous. Assume now that f is continuous at $x_0 \in [a, b]$, i.e. $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0$ such that $\forall x \in [a, b]$, $|x - x_0| < \delta_\varepsilon \implies |f(x) - f(x_0)| < \varepsilon$. Hence $\forall t \in (-\delta_\varepsilon, \delta_\varepsilon) \setminus \{0\}$, if $x_0 + t \in [a, b]$, then

$$\left|\frac{1}{t}(F(x_0 + t) - F(x_0)) - f(x_0)\right| = \left|\frac{1}{t}\int_{x_0}^{x_0+t} f - f(x_0)\right|$$

Since $f(x_0)$ is constant, we can write it as an integral from x_0 to $x_0 + t$ of itself. Hence we obtain

$$\left|\frac{1}{t}\int_{x_0}^{x_0+t} f(s) - f(x_0) ds\right| \leq \frac{1}{t}\int_{x_0}^{x_0+t} |f(s) - f(x_0)| ds \leq \frac{1}{t}\int_{x_0}^{x_0+t} \varepsilon = \varepsilon$$

It follows that $\limsup_{t\to 0} |\frac{1}{t}(F(x_0 + t) - F(x_0)) - f(x_0)| \leq \varepsilon$. Letting $\varepsilon \to 0$ we obtain $F'(x_0) = f(x_0)$. $\square$

Theorem 8.4 (Fundamental Theorem of Calculus)
Let $a, b \in \mathbb{R}$ such that $a < b$ and $F : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, differentiable on (a, b) and such that $F' = f \in \mathcal{R}([a, b])$. Then:

$$F(b) - F(a) = \int_a^b f$$

Proof. Since $f \in \mathcal{R}([a, b])$, $\forall \varepsilon > 0$, $\exists P_\varepsilon = \{x_0, \dots, x_{n_\varepsilon}\}$ partition of $[a, b]$ such that

$$U(P_\varepsilon, f) - L(P_\varepsilon, f) < \varepsilon$$

Hence, given this partition, we can write

$$F(b) - F(a) = \sum_{i=0}^{n_\varepsilon} (F(x_i) - F(x_{i-1})) = \sum_{i=1}^{n_\varepsilon} f(s_i) \underbrace{(x_i - x_{i-1})}_{=\Delta x_i}$$

for some $s_i \in [x_{i-1}, x_i]$ by the Mean Value Theorem given continuity of F on $[a, b]$, and differentiability of F on (a, b), and since $F' = f$. Then

$$\begin{aligned}\left|F(b) - F(a) - \int_a^b f\right| = \left|\sum_{i=1}^{n_\varepsilon} f(s_i)\Delta x_i - \sum_{i=1}^{n_\varepsilon}\int_{x_{i-1}}^{x_i} f\right| &= \left|\sum_{i=1}^{n_\varepsilon}\int_{x_{i-1}}^{x_i} f(s_i) - f(s)ds\right| \\ &\leq \sum_{i=1}^{n_\varepsilon}\int_{x_{i-1}}^{x_i} |f(s_i) - f(s)|ds \leq \sum_{i=1}^{n_\varepsilon}\int_{x_{i-1}}^{x_i} M_i - m_i \\ &= \sum_{i=1}^{n_\varepsilon}(M_i - m_i)\Delta x_i = U(P_\varepsilon, f) - L(P_\varepsilon, f) < \varepsilon\end{aligned}$$

with M_i, m_i defined as usual. Letting $\varepsilon \to 0$, we obtain

$$\int_a^b f = F(b) - F(a)$$

□

Theorem 8.5 (Taylor's Theorem with Integral Remainder)
Let $x, x_0 \in \mathbb{R}$,

$$I = \begin{cases}[x, x_0] \text{ if } x < x_0 \\ [x_0, x] \text{ if } x_0 < x\end{cases}, \mathring{I} = \begin{cases}(x, x_0) \text{ if } x < x_0 \\ (x_0, x) \text{ if } x_0 < x\end{cases}$$

and $f : I \to \mathbb{R}$ be n-times continuously differentiable on I, as well as $n+1$-times differentiable on $\mathring{I}$, and $t \mapsto f^{(n+1)}(t)(x-t)^n$ be Riemann-Integrable over I. Then

$$f(x) = \sum_{k=0}^{n}\frac{1}{k!}f^{(k)}(x_0)(x - x_0)^k + \frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(t)(x-t)^n dt$$

Theorem 8.6 (Integral Test for Convergence)
Let $n_0 \in \mathbb{N}$ and $f : [n_0, \infty) \to [0, \infty)$ be decreasing and such that $\lim_{x\to\infty} f(x) = 0$. For each $n \geq n_0$, define:

$$s_n = \sum_{k=n_0}^{n} f(k),\ t_n = \int_{n_0}^{n} f,\ d_n = s_n - t_n$$

Then:

1. $0 \leq f(n+1) \leq d_{n+1} \leq d_n \leq f(n_0)$
2. $\lim_{n\to\infty} d_n = d$ for some $d \in [0, f(n_0)]$
3. $\sum_{k=n_0}^{\infty} f(k)$ converges if and only if f is Riemann Integrable over $[n_0, \infty)$
4. $0 \leq d_n - d \leq f(n)\ \forall n \geq n_0$

Proof. 1. By definition of d_n and d_{n+1}, given that f is decreasing, we have

$$\begin{aligned} d_{n+1} = \sum_{k=n_0}^{n+1} f(k) - \int_{n_0}^{n} f &= f(n+1) + \sum_{k=n_0}^{n} \underbrace{\left(f(k) - \int_k^{k+1} f \right)}_{=\int_k^{k+1}(f(k)-f(s))ds \,\geq\, 0} \\ &\geq f(n+1) \end{aligned}$$

Further,

$$\begin{aligned} d_n = \sum_{k=n_0}^{n} f(k) - \int_{n_0}^{n} f &= f(n_0) + \sum_{k=n_0+1}^{n} \underbrace{\left(f(k) - \int_k^{k+1} f \right)}_{=\int_{k-1}^{k}(f(k)-f(s))ds \,\leq\, 0} \\ &\leq f(n_0) \end{aligned}$$

Finally

$$d_{n+1} - d_n = f(n+1) - \int_n^{n+1} f = \int_n^{n+1} (f(n+1) - f(s))\, ds \leq 0$$

2. Since $(d_n)_n$ is monotone and bounded by 0 and $f(n_0)$, we obtain that $\lim_{n\to\infty} d_n = d$ for some $d \in [0, f(n_0)]$.

3. Since $f \geq 0$, we have that $x \mapsto \int_{n_0}^{x} f$ is increasing. Therefore

$$\lim_{n\to\infty} \int_{n_0}^{n} f$$

exists, i.e. f is Riemann Integrable over $[n_0, \infty)$, and is finite if and only if t_n exists and is finite. Moreover, $\lim_{n\to\infty} d_n = \lim_{n\to\infty}(s_n - t_n) = d$ by (2), so we obtain that $(s_n)_n$ converges, i.e. $\sum_{k=n_0}^{\infty} f(k)$ converges if and only if $(t_n)_n$ converges. Thus $\sum_{k=n_0}^{\infty} f(k)$ converges if and only if f is Riemann Integrable over $[n_0, \infty)$.

4. $\forall n' > n \geq n_0$, we have

$$d_n - d_{n'} = \sum_{k=n}^{n'-1} (d_k - d_{k+1}) = \sum_{k=n}^{n'-1} \int_k^{k+1} (f(k) - f(s))\, ds$$

Since f is decreasing, $f(s) - f(k+1) \in [0, f(k) - f(k+1)]$ for $x \in [k, k+1]$. Hence

$$\begin{aligned} 0 \leq d_n - d_{n'} &\leq \sum_{k=n}^{n'} f(k) - f(k+1) \\ &= f(n) - f(n') \end{aligned}$$

Since $\lim_{n\to\infty} d_n = d$ and $\lim_{n\to\infty} f(n) = 0$, it follows that $0 \leq d_n - d \leq f(n)$. □