# MATH 320 Notes

Sushrut Tadwalkar

# Contents

# 1 Cardinal Numbers

## 1.1 Sets and Logic (Rudin pp 24-30)

### 1.1.1 Russel's Paradox

We begin by classifying sets:

**Definition: Set classification**

A set $\mathcal{S}$ is called

- ***Normal*** if $\mathcal{S} \notin \mathcal{S}$.
- ***Abnormal*** if $\mathcal{S} \in \mathcal{S}$.

Every set is one or the other.

Consider $\mathscr{N}$, the collection of all normal sets. To classify $\mathscr{N}$:

- Consider the case when $\mathscr{N}$ is normal. By choice of $\mathscr{N}$, it has to be abnormal, which is a contradiction.
- In the case where $\mathscr{N}$ is abnormal, i.e., $\mathscr{N} \in \mathscr{N}$, we see that $\mathscr{N}$ is also normal.

Clearly, this is a contradiction. The problem here is that the words used to set up $\mathscr{N}$ sound like maths, but fall out of the scope for safe logical reasoning. To get around this we will employ the ZFC (Zermelo–Fraenkel Choice) axiomatic system, which we will analyse more comprehensively as the course proceeds.

# 2 Mappings

## 2.1 Cartesian products

**Definition: Cartesian product**

Given sets $\mathcal{X}$ and $\mathcal{Y}$, a ***Cartesian product*** builds a new set of ordered pairs

$$\mathcal{X} \times \mathcal{Y} = \{(x, y) : x \in \mathcal{X}, y \in \mathcal{Y}\}.$$

We can extend this:

**Definition: $n-$Cartesian product**

Given $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$, we define

$$\begin{aligned}\mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n =& \{(x_1, x_2, \ldots, x_n) : x_k \in \mathcal{X}_k \text{ for all } k \in \{1, 2, \ldots, n\}\} \\ =& \prod_{k=1}^{n} \mathcal{X}_k.\end{aligned}$$

## 2.2 Multifunctions

**Definition: Graph**

Given any sets $\mathcal{X}$, $\mathcal{Y}$ any subset $\mathcal{G}$ of $\mathcal{X}$, $\mathcal{Y}$ defines a set-values mapping (or multifunction): Given $x \in \mathcal{X}$ we define

$$\mathcal{G}(x) = \{y \in \mathcal{Y} : (x, y) \in \mathcal{G}\}.$$

$\mathcal{G}$ is called the ***graph*** of this mapping.

$\mathcal{G} : \mathcal{X} \rightrightarrows \mathcal{Y}$ is built as a subset $\mathcal{G}$ of $\mathcal{X} \times \mathcal{Y}$. Denote

$$\mathcal{G}(x) = \{y \in \mathcal{Y} : (x, y) \in \mathcal{G}\}.$$

We go on to invent a $\mathcal{G}^{-1} = \{(y, x) : (x, y) \in \mathcal{G}\} \subseteq \mathcal{Y} \times \mathcal{X}$ gives a related $\mathcal{G}^{-1} : \mathcal{Y} \rightrightarrows \mathcal{X}$ such that

**Definition: Pre-image**

$$\begin{aligned}\mathcal{G}^{-1}(y) =& \{x : (y, x) \in \mathcal{G}^{-1}\} \\ =& \{x : (y, x) \in \mathcal{G}\} \\ =& \{x : y \in \mathcal{G}\}.\end{aligned}$$

This is called the ***pre-image of y***.

Now, define $\operatorname{dom}(\mathcal{G}) = \{x \in \mathcal{X} : \mathcal{G}(x) \neq \phi\}$. Extend this notation to inputs that are *sets*: if $\mathcal{A} \subseteq \mathcal{X}$

$$\mathcal{G}(\mathcal{A}) = \bigcup_{a \in \mathcal{A}} \mathcal{G}(a),$$

if $\mathcal{B} \subseteq \mathcal{Y}$,

$$\begin{aligned}\mathcal{G}^{-1}(\mathcal{B}) =& \bigcup_{b \in \mathcal{B}} \mathcal{G}^{-1}(b) \\ =& \{x \in \mathcal{X} : \mathcal{G}(x) \cap \mathcal{B} \neq \emptyset\}.\end{aligned}$$

A set-valued map $\mathcal{G} : \mathcal{X} \rightrightarrows \mathcal{Y}$ is a *mapping* or *function* when $\mathcal{G}(x)$ is a *singleton* (one-point set) for each $x \in \mathcal{X}$ (this entails $\mathrm{dom}(\mathcal{G}) = \mathcal{X}$.)

**Notation 1.** Indicate this by writing $\mathcal{G} : \mathcal{X} \to \mathcal{Y}$, and simplifying $\mathcal{G}(x) = \{y\}$ down to $\mathcal{G}(x) = y$.

**Note.** A set $\mathcal{G}$ might give a function, but $\mathcal{G}^{-1}$ might not.

**Example 1.** Consider the equation $y = x^2$. Here, $\mathcal{X} = \mathbb{R}$, $\mathcal{Y} = \mathbb{R}$, and $\mathcal{X} \times \mathcal{Y} = \mathbb{R}^2$. The equation defines a set $\mathcal{G} = \{(x, y) | \; y = x^2\}$ in the plane that is the graph of the function $\mathcal{G} : \mathbb{R} \to \mathbb{R}$ defined by $\mathcal{G}(t) = t^2$, $t \in \mathbb{R}$. Note that

$$\begin{aligned} \mathcal{G}^{-1}(4) =& \{x : \mathcal{G}(x) = 4\} \\ =& \{-2, +2\}. \end{aligned}$$

Similarly,

$$\begin{aligned} \mathcal{G}^{-1}(0) =& \{x : \mathcal{G}(x) = 0\} \\ =& \{-0, 0\} = 0. \end{aligned}$$

Also, $\mathrm{dom}(\mathcal{G}^{-1}) = \{y : y \geq 0\} = [0, \infty)$.

Thus, we go on to define a function:

**Definition: Function**

Let $f : \mathcal{X} \to \mathcal{Y}$,

1. This $f$ is one-to-one (injective), i.e., different inputs give different outputs, i.e.,
   - for all $x_1, \; x_2 \in \mathcal{X}$, $x_1 \neq x_2$ if $f(x_1) \neq f(x_2)$.
   - We can also state the contrapositive: for all $x_1, \; x_2 \in \mathcal{X}$, $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$.
2. This $f$ is *onto* (surjective) when $f(\mathcal{X}) = \mathcal{Y}$, i.e.,
   - for all $y \in \mathcal{Y}$, there exists $x \in \mathcal{X}$ such that $f(x) = y$.
3. A function that is *both* one-to-one and onto is called *bijective*.

# 3 Countability

Recall that $\mathbb{N} = \{0, 1, 2, 3, \dots\}$, are defined to be the natural numbers.

**Definition: Finite sets**

A set $\mathcal{A}$ is ***finite*** if either $\mathcal{A} = \phi$ or there exists $n$ and a bijection $\varphi : \{1, 2, \dots, n\} \to \mathcal{A}$.

**Notation 2.** $|\emptyset| = 0$, $|\mathcal{A}| = n$.

This is more or less as expected, but *countable* sets are more interesting.

**Definition: Countable**

A set $\mathcal{S}$ is ***countable*** if there is a bijection $\varphi :\ \mathbb{N} \to \mathcal{S}$. Here, $|\mathcal{S}| = \aleph_0$.

**Example 2.** Consider

1. $\mathbb{N}$ itself; $(\varphi(x) = x)$.
2. Hilbert's hotel: $S = \mathbb{N} \cup \{0\}$; use $\varphi(n) = n - 1$.
3. $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.
   Use

$$\phi(n) = \begin{cases} -\dfrac{n}{2} & \text{if } n \text{ is even.} \\[2ex] \dfrac{n-1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

**Quotes of the day**

"All is fair, if you pre-declare." - Prof. Lowen, 09/11/23

## 3.1 Interlude: Well ordering property of $\mathbb{N}$

Let $\mathcal{S} \in \mathbb{N}$ such that $\mathcal{S} \neq \varphi$, i.e., there exists $\hat{s} \in \mathcal{S} : \hat{s} \leq s$ for all $s \in \mathcal{S}$. We will often use $\min(\mathcal{S})$ instead of $\hat{s}$. This is the basis for the principle of mathematical induction.

Going back to countable sets, we have the following property:

**Property 1.** Every subset of $\mathbb{N}$ is finite or countable.

*Proof.* Let $\mathcal{A} \subseteq \mathbb{N}$. If $\mathcal{A}$ is finite, we are done; assume $\mathcal{A}$ is infinite.

Define $\mathcal{A}_1 = \mathcal{A}$; since $\mathcal{A}$ is infinite, $\mathcal{A} \neq \phi$, so let $a_1 = \min(\mathcal{A}_1)$. Note that $a_1 \geq 1$; define $\varphi(1) = a_1$.

Now $\mathcal{A}$ is infinite, so $\mathcal{A}_2 := \mathcal{A} \backslash \{a_1\}$ is not empty. Let $a_2 = \min(\mathcal{A}_2)$; define $\varphi(2) = a_2$. Note $\varphi(2) = a_2 > a_1$, so $\varphi(s) \geq 2$. Continue with induction.

If step $n$ has been done, giving $\varphi(n) = a_n$ with $\varphi(n) \geq n$, proceed as follows:

Set $\mathcal{A}$ is infinite, so $\mathcal{A}_{n+1} = \mathcal{A} \backslash \{a_1, a_2, \dots, a_n\}$ is not empty. Let $a_{n+1} = \min(\mathcal{A}_{n+1})$, $\varphi(n+1) = a_{n+1}$. Note that $\varphi(n+1) = a_{n+1} \geq n + 1$.

Induction defines $\varphi : \mathbb{N} \to \mathcal{A}$. We can observe that $\varphi$ in injective since for $m \neq n$ (assume WLOG that $m < n$) $\varphi(m) < \phi(n)$ by construction. Similarly, we can observe that $\varphi$ is surjective; notice that for any $a \in \mathcal{A}$, we have $\varphi(a) \geq a$, so $a = \varphi(k)$ must have occurred for some stage $k$, with $k \geq a$. Hence, we have shown that $\varphi$ is a bijection, which by extension verifies the definition of "$\mathcal{A}$ is countable". $\square$

**Property 2.** Every subset of *any countable set* is finite or countable.

*Proof.* Let $\mathcal{X}$ be a countable set, with a subset $\mathcal{S}$. There exists a bijective function $\varphi : \mathbb{N} \to \mathcal{X}$. Define $f : \mathcal{S} \to \mathbb{N}$, such that

$$f(s) = \varphi^{-1}(s),\ s \in \mathcal{S}.$$

Thus, $f$ is a bijective map between $\mathcal{S}$ and $f(\mathcal{S}) \subseteq \mathbb{N}$. This implies that $f(\mathcal{S})$ is either finite or countable, and hence the same holds for $\mathcal{S}$. $\square$

**Theorem**

Given a set $\mathcal{A}$, either (a) or (b) below implies $\mathcal{A}$ is finite-or-countable:

(a) There exists a countable set $\mathcal{X}$ and an injective function $f : \mathcal{A} \to \mathcal{X}$.

(b) There exists a countable set $\mathcal{X}$ and a surjective function $g : \mathcal{X} \to \mathcal{A}$.

More about countable sets:

(a) If $\mathcal{A}$ and $\mathcal{B}$ are countable, then $\mathcal{A} \times \mathcal{B}$ is too.

(b) Let $\mathscr{S}_2$ be the collection of subsets of $\mathbb{N}$ with one or two elements; $\mathscr{S}_2$ is countable.

We see that this is true if we construct a map $f : \underset{\mathbb{N}\times\mathbb{N}}{(m, n)} \to \underset{\mathscr{S}_2}{\{m, n\}}$, and apply part (b) of the theorem.

**Quotes of the day**

"That's a little informal; I got tired of writing."- Philip Loewen, 09/13/2023

"Do they still teach LISP?" - Philip Loewen, 09/13/2023

**Claim 1.** The union of countably many sets, each one countable is also countable.

*Proof.* Let $\mathcal{A}^{(1)},\ \mathcal{A}^{(2)},\ \mathcal{A}^{(3)}, \ldots$ be a (countable) family of sets, each countable. This means each $k$ comes with some $\varphi_k : \mathbb{N} \to \mathcal{A}^{(k)}$, a bijection. So, define,

$$f : \mathbb{N} \times \mathbb{N} \to \bigcup_{k=1}^{\infty} \mathcal{A}^{(k)},$$

by $f(m, n) = \varphi_m(n)$. This is *surjective*, and input set $\mathbb{N} \times \mathbb{N}$ is countable; we are done. $\square$

**Claim 2.** The set $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ is countable.

*Proof.* By definition,

$$\mathbb{N} \times \mathbb{N} \times \mathbb{N} = \{(x_1,\ x_2,\ x_3) : x_k \in \mathbb{N}\}$$

is the image of

$$(\mathbb{N} \times \mathbb{N}) \times \mathbb{N} = \{((x_1,\ x_2),\ x_3) : x_1,\ x_2,\ x_3 \in \mathbb{N}\}$$

under $\varphi((x_1,\ x_2),\ x_3) = (x_1,\ x_2,\ x_3)$. Clearly, $\varphi$ is bijective and $(\mathbb{N} \times \mathbb{N}) \times \mathbb{N}$ is the cartesian product of two countable sets.

Extend by induction:

For any $n \in \mathbb{N}$, the n-fold product

$$\prod_{k=1}^{n} \mathbb{N} = \underbrace{\mathbb{N} \times \mathbb{N} \times \mathbb{N} \times \cdots \times \mathbb{N}}_{n-\text{copies}}$$

is countable. Thus,

$$\bigcup_{n\in\mathbb{N}} \left( \prod_{k=1}^{n} \mathbb{N} \right)$$

is countable. These are all the finite-length tuples with entries from $\mathbb{N}$. $\square$

**Definition: Uncountable**

A set $\mathcal{Y}$ is ***uncountable*** iff $\mathcal{Y}$ is infinite and not countable, i.e., $\mathcal{Y}$ is infinite and every function $\varphi : \mathbb{N} \to \mathcal{Y}$ fails to be bijective.

**Note.** We have to be careful with this definition. This is how Rudin defines uncountability,where exclusively infinite sets to be countable. For intuition, uncountable sets are all sets that are not countable.

**Note.** To prove that a set is $\mathcal{Y}$ is uncountable, we show that every $\varphi : \mathbb{N} \to \mathcal{Y}$ that is injective cannot be surjective.

**Example 3.** Let $\mathcal{I}$ be the set of sequences having the form

$$x = 0.d_1 d_2 d_3 \ldots,$$

where each $d_k \in \{0,\ 1,\ 2, \ldots,\ 9\}$ and for all $n \in \mathbb{N}$ there exists $m \geq n :\ d_m \neq 9$ (digit strings that do not "end" with an infinite list of 9's.) This $\mathcal{I}$ is uncountable. To prove this, pick some $\varphi : \mathbb{N} \to \mathcal{I}$ that is injective. We shall show that $\varphi$ is not surjective (this was first done by Cantor.) We note the list of $\varphi$ values:

$$\begin{aligned} \varphi(1) &= 0.\boxed{d_1^{(1)}} d_2^{(1)} d_3^{(1)} d_4^{(1)} \ldots \\ \varphi(2) &= 0.d_1^{(2)} \boxed{d_2^{(2)}} d_3^{(2)} d_4^{(2)} \ldots \\ \varphi(3) &= 0.d_1^{(3)} d_2^{(3)} \boxed{d_3^{(3)}} d_4^{(3)} \ldots \\ &\ \ \vdots \qquad\qquad\qquad \ddots \end{aligned}$$

Now, for each $k \in \mathbb{N}$, invent

$$x_k = \begin{cases} 5 & \text{if } d_k^{(k)} = 7 \\ 7 & \text{else.} \end{cases}$$

(Keep $x_i = d_i^{(k)}$ where $i \neq k$.)

Consider $x := 0.x_1 x_2 x_3 \ldots$. Now, $x \notin \varphi(\mathbb{N})$ because $x$ differs in position $k$ from the string $\varphi(k)$ so $\varphi$ is not surjective. $\square$

**Note.** Technically, we do not need to require that $\varphi$ is injective.

**Note.** Use $\mathbb{R}$ as known for now; we will build axiomatic connections later.

In this case, each $x = 0.x_1 x_2 x_3 \ldots$ in $\mathcal{I}$ defines the real number

$$\sum_{k=1}^{\infty} \frac{x_k}{10^k},$$

and the proof above shows $\mathcal{I} = [0, 1)$ is uncountable.

**Quotes of the day**

"Some of you have to do soul-searching and quit." - Philip Loewen, 09/15/2023

## 3.2 Cardinal numbers

**Definition**

We write $\mathcal{X} \sim \mathcal{Y}$ if there exists a bijection $\varphi : \mathcal{X} \to \mathcal{Y}$ (say "$\mathcal{X} \sim \mathcal{Y}$ under $\varphi$".)

**Proposition**

" $\sim$ " is called an ***equivalence relation***, i.e., for sets $\mathcal{X},\ \mathcal{Y},\ \mathcal{Z}$, we have

(a) $\mathcal{X} \sim \mathcal{X}$ (Reflexivity.)

(b) $\mathcal{X} \sim \mathcal{Y} \Rightarrow \mathcal{Y} \sim \mathcal{X}$ (Symmetricity.)

(c) $\mathcal{X} \sim \mathcal{Y},\ \mathcal{Y} \sim \mathcal{Z} \Rightarrow \mathcal{X} \sim \mathcal{Z}$ (Transitivity.)

**Notation 3.** $\mathcal{X} \sim \mathcal{Y} \iff \mathcal{X}$ and $\mathcal{Y}$ are equinumerous.

*Proof.* (R) Use $\varphi(x)$.

(S) If $\mathcal{X} \sim \mathcal{Y}$ under $\varphi$, $\mathcal{Y} \sim \mathcal{X}$ under $\varphi^{-1}$.

(T) If $\mathcal{X} \sim \mathcal{Y}$ under $\varphi$ and $\mathcal{Y} \sim \mathcal{Z}$ under $\psi$, then show $\mathcal{X} \sim \mathcal{Z}$ under $\psi \circ \varphi$ (would need to prove the composition of a bijection is also a bijection, but already did.) $\square$

Invent the symbol $|\mathcal{X}|$ for "the cardinal number of set $\mathcal{X}$" and encode $\mathcal{X} \sim \mathcal{Y}$ with $|\mathcal{X}| = |\mathcal{Y}|$.

**Note.** The equivalence relation here is a specific one; usually the concept of an equivalence relation is very general, but this is how Rudin defines it as well. I guess this is done because we don't require another notion of an equivalence relation in this course.

**Example 4.** Some examples are:

- $|\{1,\ 2,\ 3\}| = |\{a,\ b,\ c\}| = 3$.
- $|\mathbb{N}| = \aleph_0$.
- $|\mathbb{R}| = \mathfrak{c}$ ("the continuum".)

We have $\aleph_0 < \mathfrak{c}$. But we have not defined what inequality of cardinals mean.

**Example 5.** $|\mathbb{N}| = \aleph_0,\ |\mathbb{R}| = \mathfrak{c}$, such that $\mathfrak{c} > \aleph_0$.

**Definition: Cardinal inequalities**

Say $|\mathcal{X}| \leq |\mathcal{Y}|$ if there exists an injection $\varphi : \mathcal{X} \to \mathcal{Y}$ ($|\mathcal{Y}| \geq |\mathcal{X}|$ is the same.) Saying $|\mathcal{X}| < |\mathcal{Y}|$ (strict), i.e., there is no bijection from $\mathcal{X}$ to $\mathcal{Y}$, i.e., every $\varphi : \mathcal{X} \to \mathcal{Y}$ that's injective cannot also be surjective.

$|\mathcal{X}| < |\mathcal{Y}| \iff |\mathcal{X}| \leq |\mathcal{Y}|$ and there are no bijections of $\mathcal{X}$ into $\mathcal{Y}$, i.e., $\mathcal{X} \not\sim \mathcal{Y}$.

## 3.3 Beyond uncountability

Given a set $\mathcal{X}$, the power set $\mathscr{P}(\mathcal{X}) = \{\mathcal{S} : \mathcal{S} \subseteq \mathcal{X}\}$.

**Proposition**

For any set $\mathcal{X}$, $|\mathcal{X}| < |\mathscr{P}(\mathcal{X})|$.

So,

$$\aleph_0 < \mathfrak{c} < |\mathscr{P}(\mathbb{R})| < |\mathscr{P}(\mathscr{P}(\mathbb{R}))| < \dots$$

We can see that there is an infinite sequence of strictly increasing cardinal numbers. There are infinitely many "sizes of infinity" (we know that there are countably many cardinal numbers.) The question of whether there is a cardinal number between $\aleph_0$ and $\mathfrak{c}$ is undecidable; both whether there is or there is not is consistent with ZFC. The statement that there is no such cardinal number is called the *continuum hypothesis*. Also, note that $\mathscr{P}(\mathbb{N}) \sim \mathbb{R}$.

*Proof.* Suppose $\mathcal{X} \neq \emptyset$, and let $f : \mathcal{X} \to \mathscr{P}(\mathcal{X})$ be injective; we will show that $f$ is not surjective.

Define
$$\mathcal{S} = \{x \in \mathcal{X} : x \notin f(x)\}.$$

**Claim 3.** $\mathcal{S} \notin f(\mathcal{X})$.

For the sake of contradiction, suppose $\mathcal{S} \in f(\mathcal{X})$. Thus, there exists some $y \in X$ such that $f(y) = \mathcal{S}$; consider cases

1. $y \in \mathcal{S} \Rightarrow y \in f(y) \Rightarrow y \notin f(y)$.
2. $y \notin \mathcal{S} \Rightarrow y \notin f(y) \Rightarrow y \in \mathcal{S}$.

Thus, assuming $\mathcal{S} \in f(\mathcal{X})$ leads to a contradiction. $\square$

**Theorem: Schröder-Berstein theorem**

For sets $\mathcal{X}$ and $\mathcal{Y}$, if $|\mathcal{X}| \leq |\mathcal{Y}|$ and $|\mathcal{X}| \geq |\mathcal{Y}|$, then $|\mathcal{X}| = |\mathcal{Y}|$.

**Note.** Knowing $|\mathcal{X}| \leq |\mathcal{Y}|$ and $|\mathcal{X}| \geq |\mathcal{Y}|$ does not automatically force $|\mathcal{X}| = |\mathcal{Y}|$. It is in fact, but needs work, and thus the theorem.

**Quotes of the day**

"I like $\varepsilon$ over $\epsilon$ because $\epsilon$ kind of looks like set membership or Euler's constant…but you know we're in Vancouver; this is a place of diversity and inclusion." - Dr. Philip Loewen, 09/18/2023

"Have to stay on topic; the temptation to make cheeky remarks is overwhelming." - Dr. Philip Loewen, 09/18/2023

# 4 Numbers and vectors

## 4.1 Numbers

**Note.** For the moment, we will keep using $\mathbb{R}$ and $=,\ <,\ +,\ -,\ \times,\ \div,\ |\cdot|$ as usual. We will construct these things later.

**Property 3.** The Archimedean property:

For all $r \in \mathbb{R}$, there exists an $n \in \mathbb{N}$ such that $n > r$.

**Corollary 1.** We have the following corollaries for the Archimedean property:

(a) For each real $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$.

(b) Whenever $x,\ y \in \mathbb{R}$ obey $y - x > 1$, one has $(x, y) \cap \mathbb{Z} \neq \emptyset$.

(c) Whenever $a < b \in \mathbb{R}$, both $(a, b) \cap \mathbb{Q} \neq \emptyset$ and $(a, b) \backslash \mathbb{Q} \neq \emptyset$.

(a) *Proof.* Given $\varepsilon > 0$, define $r = \frac{1}{\varepsilon}$ and apply the Archimedean property: there exists $n \in \mathbb{N}$ such that

$$n > \frac{1}{\varepsilon} > 0 \Rightarrow \frac{1}{n} < \varepsilon.$$

□

(b) *Proof.* For $x, y$ as given, let

$$\mathcal{S} = \{n \in \mathbb{Z} : n \geq y\}.$$

Define $\hat{n} = \min(\mathcal{S})$. We claim that $z := \hat{n} - 1$ is in $(x, y)$. This is because

(i) Definition of $\hat{n}$ makes $\hat{n} - 1 \notin \mathcal{S}$, i.e., $z < y$.

(ii) $\hat{n} \in \mathcal{S} \Rightarrow \hat{n} \geq y \Rightarrow z = \hat{n} - 1 \geq y - 1 > x$. Thus, $z > x$.

□

(c) *Proof.* Given $a < b$, use (a) to get $n \in \mathbb{N}$ such that $\frac{1}{n} < b - a$. So, $1 < nb - na$ and (b) applies: $(na, nb) \cap \mathbb{Z} \neq \emptyset$.

Pick $m \in \mathbb{Z}$ such that $na < m < nb$, i.e., $a < \frac{m}{n} < b$. Now, $\frac{m}{n} \in (a, b) \cap \mathbb{Q}$. We are done.

In the case of $(a, b) \cap \mathbb{Q}$, it is at most countable, whereas $(a, b) = [(a, b) \cap \mathbb{Q}] \cup [(a, b) \backslash \mathbb{Q}]$ is uncountable. Thus, $(a, b) \backslash \mathbb{Q}$ is uncountable as well, in particular non-empty. □

**Property 4.** For any $a,\ b \in \mathbb{R}$, exactly one of $a < b$, $a = b$, and $a > b$ is true.

## 4.2 Vectors

**Notation 4.** We denote a vector "$x$" by $\underline{x}$.

As we have defined before,

$$\mathbb{R}^k = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \times \mathbb{R}}_{k-\text{times}} = \{(x_1, x_2, \ldots, x_k) : \text{each } x_j \in \mathbb{R} \text{ for } j = 1, \ldots, k\}.$$

Extend $+$ as always.

Define scalar multiplication as usual: Let

$$\underline{x} \cdot \underline{y} = \sum_{j=1}^{k} x_j y_j$$

$$|\underline{x}| = \sqrt{\sum_{j=1}^{k} x_j^2} = \sqrt{\underline{x} \cdot \underline{x}}.$$

**Property 5.** These are properties concerning dot products that are worth noting. We try to replace length with the dot product as often as possible since dealing with square roots is almost always trickier.

(i) $|\underline{x}|^2 = \underline{x} \cdot \underline{x}$.

(ii) $|\underline{x} \cdot \underline{y}| \leq |\underline{x}||\underline{y}|$ (Schwartz inequality.)

(iii) Triangle inequalities:

  (a) $|\underline{x} + \underline{y}| \leq |\underline{x}| + |\underline{y}|$.

  (b) $|\underline{y} - \underline{x}| = ||\underline{y}| - |\underline{x}||$.

**Note.** (i) The triangle inequalities are useful in proofs involving $|\underline{x}|$.

(ii) Given $\underline{x},\ \underline{y} \in \mathbb{R}^k$, consider

$$0 \leq |\underline{x} - t\underline{y}|^2 = (\underline{x} - t\underline{y}) \cdot (\underline{x} - t\underline{y}),$$

i.e.,

$$0 \leq \underline{x} \cdot \underline{x} - 2t\underline{x} \cdot \underline{y} + t^2 \underline{y} \cdot \underline{y} = |\underline{x}|^2 - 2t(\underline{x} + \underline{y}) + t^2|\underline{y}|^2.$$

Thus, this quadratic function of $t$ has at most one root. Quadratic formula states that discriminant must not be positive, i.e.,

$$(-2(\underline{x} \cdot \underline{y}))^2 - 4(|\underline{x}|^2|\underline{y}|^2) \leq 0$$
$$\iff 4(\underline{x} \cdot \underline{y})^2 \leq 4(|\underline{x}||\underline{y}|)^2$$
$$\iff |\underline{x} \cdot \underline{y}| \leq |\underline{x}||\underline{y}|.$$

(iii) Note that, using $t = -1$ above, we have

$$\begin{aligned} |\underline{x} + \underline{y}|^2 =& (\underline{x} + \underline{y}) \cdot (\underline{x} + \underline{y}) \\ =& |\underline{x}|^2 + 2|\underline{x}||\underline{y}| + |\underline{y}|^2 \\ \leq& |\underline{x}|^2 + 2|\underline{x}||\underline{y}| + |\underline{y}|^2 \\ =& (|\underline{x}| + |\underline{y}|)^2. \end{aligned}$$

We take square roots and substitute $\underline{y} = \underline{z} - \underline{x}$ to get

$$\begin{aligned} |\underline{x} + (\underline{z} - \underline{x})| \leq& |\underline{x}| + |\underline{z} + \underline{x}| \\ |\underline{z}| - |\underline{x}| \leq& |\underline{z} - \underline{x}|. \end{aligned}$$

This works for any $\underline{x}$, $\underline{y}$, $\underline{z}$. Furthermore,

$$\begin{aligned}|\underline{x}| - |\underline{z}| &\leq |\underline{x} - \underline{z}| \\ |\underline{x} - \underline{z}| &\geq \max\{|\underline{z}| - |\underline{x}|,\ |\underline{x}| - |\underline{z}|\} = ||\underline{x}| - |\underline{z}||.\end{aligned}$$

**Quotes of the day**

"Do you know about Dressew downtown on Hastings street? It's amazing. This where people who know how to make things with fabric go." - Dr. Philip Loewen, 09/20/2023

"There are politics jokes ready here – 'no matter how far to the right we start out, a point further right will escape the tolerance band.'" - Dr. Philip Loewen, 09/20/2023

# 5 Sequences and Limits

**Definition: Sequence**

A ***sequence*** in a given set $\mathcal{X}$ is simply a *function* $x : \mathbb{N} \to \mathcal{X}$. We will often write $x_n$ instead fo $x(n)$ and list the values.

**Example 6.** $x = (x_1, x_2, x_3, \dots)$.

The order of a sequence matters; the *sequence* $x_n = (x_1, x_2, x_3, \dots)$ is different from the *set* $\{x_1, x_2, x_3, \dots\}$.

**Example 7.** Consider $x_n = (-1)^{n+1}$ for $n \in \mathbb{N}$. This sequence has the set $\{(-1)^{n+1} : n \in \mathbb{N}\} = \{-1, +1\}$.

**Definition: Convergence**

Given a sequence $(x_n)_{n\in\mathbb{N}}$ and a point $\hat{x}$, all in $\mathcal{X} \in \mathbb{R}$, saying the sequence $(x_n)$ ***converges*** to $\hat{x}$ means for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, $|x_n - \hat{x}| < \varepsilon$.

In a simplified form, a real-valued sequence $(x_n)_{n\in\mathbb{N}}$ ***converges*** when

$$\text{there exists } \hat{x} \in \mathbb{R} \text{ such that } \lim_{n\to\infty} x_n = \hat{x}.$$

**Notation 5.** When this happens, we write

$$\hat{x} = \lim_{n\to\infty} x_n \quad \text{or} \quad \underset{\text{as } n\to\infty}{x_n \to \infty} \quad \text{or} \quad x_n \xrightarrow[n\to\infty]{} \hat{x}.$$

**Definition: Divergence**

A sequence is said to ***diverge*** when it *does not* converge.

Concretely, $(x_n)_{n\in\mathbb{N}}$ diverges iff for all $\hat{x} \in \mathbb{R}$ there exists $\varepsilon > 0$, and for all $N \in \mathbb{N}$ there exists $n > N$ such that $|x_n - \hat{x}| \geq \varepsilon$.

**Example 8.** Some simple examples that showcase the above definitions are:

(a) $x_n = \dfrac{1}{n}$ converges to $\hat{x} = 0$.

*Proof.* Given $\varepsilon > 0$, note $\dfrac{1}{\varepsilon} \in \mathbb{R}$ and Archimedes says there exists $n \in \mathbb{N}$ such that $N > \dfrac{1}{\varepsilon}$. For any $n > N$ we will have

$$|x_n - \hat{x}| = \left|\frac{1}{n} - 0\right| = \frac{1}{n} < \frac{1}{N} < \varepsilon.$$

□

(b) If $x_n = 1$ converges to $\hat{x} = 1$.

*Proof.* Given $\varepsilon > 0$ pick $N = 320$. Clearly, every $n > N$ makes $|x_n - \hat{x}| = 0 < \varepsilon$. □

Now,

**Example 9.** Consider some slightly harder examples:

(a) Suppose

$$x_n = \frac{\sin n}{1 + n + n^2 + n^3 + n^4 + n^5}.$$

For every $n \in \mathbb{N}$,

$$|x_n - \hat{x}| = \frac{|\sin n|}{1 + n + n^2 + \cdots + n^5} < \frac{1}{0 + n + 0 + \cdots + 0}.$$

Furthermore, $\frac{1}{n} < \varepsilon$ whenever $n > \frac{1}{\varepsilon}$. We pick some $N > \frac{1}{\varepsilon}$ and every $n > N$ will have $\frac{1}{n} < \varepsilon$ and make $|x_n - \hat{x}| < \varepsilon$.

**Note.** For efficiency, keeping $n^5$ rather than $n$ would give a much smaller $N$. However, we don't particularly care about efficiency, we prioritize *existence*.

(b) The sequence $x_n = \dfrac{n^2 - 320^{3/2}}{2n^2 - 801}$ converges to $\hat{x} = \dfrac{1}{2}$.

*Proof.* Given $\varepsilon > 0$, choose integer $N \geq \max\left\{30, \left(\frac{750}{\varepsilon}\right)^2\right\}$. We see that every for $n >$, $n > 30 \Rightarrow n^2 > 900$, giving us $2n^2 - 801 = n^2 + (n^2 - 801) > n^2$.

By Archimedean property, $\sqrt{n} > \dfrac{750}{\varepsilon}$, if $n > N$,

$$\begin{aligned}
|x_n - \hat{x}| &= \left|\frac{n^2 - 320n^{3/2}}{2n^2 - 801} - \frac{1}{2}\right| \\
&= \left|\frac{2(n^2 - 320^{3/}) - (2n^2 - 801)}{2(2n^2 - 801)}\right| \\
&\leq \frac{640n^{3/2} + 801}{2(2n^2 - 801)} \text{ (by triangle inequality)}, \\
&\leq \frac{640n^{3/2} + 801n^{3/2}}{2n^2} \\
&< \frac{1500n^{3/2}}{2n^2} = \frac{750}{\sqrt{n}} \\
&< \frac{750}{750/\varepsilon} = \varepsilon, \text{ as required.}
\end{aligned}$$

□

(c) $\lim\limits_{n\to\infty} n^{\frac{1}{n}} = 1$.

*Proof.* Define $x_n = n^{\frac{1}{n}} - 1$; each $x_n > 0$. Recall that by the binomial theorem, $(1+a)^n = 1 + na + \dfrac{n(n-1)}{2}a^2 + \cdots + na^{n-1} + a^n$; thus,

$$(1+a)^n \geq \frac{n(n-1)}{2}a^2 \text{ for all } a \geq 0.$$

Thus, when $n \geq 2$, we have $x_n > 0$, i.e., $n = (1 + x_n)^n \geq \frac{n(n-1)}{2} x_n^2 \Rightarrow 0 < x_n^2 < \frac{2}{n(n-1)} n$, i.e., $x_n \leq \sqrt{\frac{2}{n-1}}$. Thus, we solve for $\sqrt{\frac{2}{n-1}} < \varepsilon$ for $\frac{2}{n-1} < \varepsilon^2 \iff \frac{2}{\varepsilon^2} < n - 1$. So, choosing $N \geq \max\left\{2, 1 + \frac{2}{\varepsilon^2}\right\}$ is what we require. $\square$

**Quotes of the day**

Some girl: "Are you going to expect this on the exam?"

Dr. Loewen: "Yes."

Girl: "Oh no." - 22/09/2023

Recall that when we say that a limit *diverges*, we are saying that for all $\hat{x} \in \mathbb{R}$, there exists $\varepsilon > 0$ and for all $n \in \mathbb{N}$, there exists $n > N$ such that $|x_n - \hat{x}| \geq \varepsilon$.

**Example 10.** Show that $x_n = (-1)^n$ diverges.

*Proof.* To begin, we fix some $\hat{x} \in \mathbb{R}$ and pick $\varepsilon = 1$. Fix $N \in \mathbb{N}$. Now, consider an even $n_e$ and an odd $n_o$ such that $n_e > N$, $n_o > N$. Thus, $x_{n_e} = (-1)^{n_e} = 1$ and $x_{n_o} = (-1)^{n_o} = -1$. Thus,

$$\begin{aligned} 2 = |x_{n_e} - x_{n_o}| \leq& |(x_{n_e} - \hat{x}) + (\hat{x} - x_{n_o})| \\ \leq& |x_{n_e} - \hat{x}| + |x_{n_o} - \hat{x}|. \end{aligned}$$

One of the terms on the RHS is $\geq 1$. One of $n = n_e$ or $n = n_o$ completes the proof. $\square$

## 6 The Squeeze Theorem

Let $(a_n),\ (x_n),\ (b_n)$ be real-valued sequences, and $L \in \mathbb{R}$. Assume

(a) $a_n \to L$ as $n \to \infty$.

(b) $b_n \to L$ as $n \to \infty$.

(c) $a_n \leq x_n \leq b_n$ for all $n > N$.

Then, $x_n \to L$ as $n \to \infty$.

*Proof.* Given $\varepsilon > 0$, use (a) to get $N_a \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for all $n > N_a$. This implies $a_n > L - \varepsilon$. Use (b) to get $N_b \in \mathbb{N}$ such that $|b_n - L| < \varepsilon$ for all $n > N_b$. This implies $b_b < L + \varepsilon$. Use (c) to get $N_c \in \mathbb{N}$ such that $a_n \leq x_n \leq b_n$ for all $n > N_c$. Now, if $N = \max\{N_a, N_b, N_c\}$, every $n > N$ does 3 things: $L - \varepsilon < a_n \leq x_n \leq b_n < L + \varepsilon$, i.e., $|x_n - L| < \varepsilon$. □

**Lemma**

If $x_n \to \hat{x}$ (in $\mathbb{R}$), then there exists $N \in \mathbb{N}$ satisfying

(a) $|x_n| \leq |\hat{x}| + 1$, for all $n > N$.

(b) If, in addition, $\hat{x} \neq 0$, also $|x_n| \geq \dfrac{|\hat{x}|}{2}$, for all $n > N$.

(a) *Proof.* If we use the triangle inequality, we have $|x_n| \leq |x_n - \hat{x}| + |\hat{x}|$ for each $n$. Picking $\varepsilon = 1$, we apply the definition to find an $N \in \mathbb{N}$, such that $|x_n - \hat{x}| < 1$ for all $n \in \mathbb{N}$. This gives us $|x_n| < 1 + |\hat{x}|$ for all $n \in \mathbb{N}$. □

(b) *Proof.* Since $x_n = \hat{x} - (\hat{x} - x_n)$, using the triangle inequality, we have $|x_n| \geq |\hat{x}| - |\hat{x} - x_n|$. We let $\varepsilon = \dfrac{1}{2}|\hat{x}| > 0$ in the definition to get $\tilde{N}$ such that $|x_n - \hat{x}| < \dfrac{1}{2}|\hat{x}|$ when $n > \tilde{N}$. This does what we require, since

$$|x_n| \geq |\hat{x}| - \frac{1}{2}|\hat{x}| = \frac{|\hat{x}|}{2} \text{ for all } n \in \mathbb{N}.$$

Here we could have said $|x_n| \geq |\hat{x}| - \dfrac{1}{2}|\hat{x}| = \dfrac{|\hat{x}|}{2}$, however, we only require a non-strict inequality for the lemma. Notice that we can take $\max\{N, \tilde{N}\}$ to get both (a), (b) together. □

**Proposition 1.** If $x_n \to \hat{x}$, $y_n \to \hat{y}$, and $K \in \mathbb{R}$, then

(a) $x_n + Ky_n \to \hat{x} + K\hat{y}$.

(b) $x_n y_n \to \hat{x}\hat{y}$.

(c) $\dfrac{x_n}{y_n} \to \dfrac{\hat{x}}{\hat{y}}$, provided $\hat{y} \neq 0$.

(a) *Proof.* For each $n$,

$$|(x_n + Ky_n) - (\hat{x} + K\hat{y})| \leq |x_n - \hat{x}| + |K||y_n - \hat{y}|.$$

Given $\varepsilon > 0$, we define $\varepsilon' = \dfrac{\varepsilon}{2} > 0$ and $\varepsilon'' = \dfrac{\varepsilon}{2(|K|+1)} > 0$ and cite definitions of $x_n \to \hat{x}$, $y_n \to \hat{y}$ to get $N', N'' \in \mathbb{N}$ such that

$$|x_n - \hat{x}| < \varepsilon' = \frac{\varepsilon}{2}, \text{ for all } n > N'$$
$$|y_n - \hat{y}| < \varepsilon'' = \frac{\varepsilon}{2(|K|+1)}, \text{ for all } n > N''.$$

We define $N = \max\{N', N''\}$. Every $n > N$ obeys

$$|(x_n + Ky_n) - (\hat{x} + K\hat{y})| < \frac{\varepsilon}{2} + |K|\left(\frac{\varepsilon}{2(|K|+1)}\right) < 2\left(\frac{\varepsilon}{2}\right) = \varepsilon.$$

□

(b) *Proof.* For each $n$,

$$\begin{aligned}|x_ny_n - \hat{x}\hat{y}| =& |x_ny_n - x_n\hat{y} + x_n\hat{y} - \hat{x}\hat{y}| \\ \le& |x_n||y_n - \hat{y}| + |\hat{y}||x_n - \hat{x}| \text{ (which is from the triangle inequality)} \\ <& (|\hat{x}|+1)|y_n - \hat{y}| + |\hat{y}||x_n - \hat{x}| \text{ for all } n > N\text{, for some } N \text{ from previous lemma (a).}\end{aligned}$$

From part (a) of this proposition, the sequence in the RHS above converges to 0. Thus, because of the Squeeze theorem, we require that the LHS also converges to 0. □

**Note.** If sequences $a_n$, $b_n$ have $|a_n| \le b_n$ for all $n$, and $b_n \to 0$, then $-b_n < a_n < b_n$ forces $a_n \to 0$.

(c) *Proof.* At first, take $x_n = 1$, we have

$$\begin{aligned}\left|\frac{1}{y_n} - \frac{1}{\hat{y}}\right| =& \left|\frac{y_n - \hat{y}}{y_n\hat{y}}\right| \\ =& \frac{1}{|y_n|}\frac{1}{|\hat{y}|}|y_n - \hat{y}| \\ <& \frac{2}{|\hat{y}|^2}|y_n - \hat{y}|,\end{aligned}$$

for all $n > N$, where $N$ is given by part (b) of the previous lemma. We see that RHS$\to 0$ by part (a) of this proposition, so because of the Squeeze theorem, we require that LHS$\to 0$. Now, in general, $\frac{x_n}{y_n} = x_n\left(\frac{1}{y_n}\right)$ is covered by this part and part (b) combined together. □

**Note.** If given sequence $y_n$ has some $0-$elements, it will have some undefined terms. But when $\hat{y} \neq 0$, all $\frac{x_n}{y_n}$ "for $n$ sufficiently large" will be defined. We relax our interpretation to allow this.

**Example 11.** If $r > 0$, $r^{1/n} \to 1$ as $n \to \infty$.

*Proof sketch.* Squeeze theorem + reciprocal theorem using $n^{1/n} \to 1$. □

**Quotes of the day**

"How about the super on steroids version of the add and subtract trick?" - Dr. Philip Loewen, 09/27/2023

# 7 Completeness

"The property that makes $\mathbb{R}$ better than $\mathbb{Q}$."

## 7.1 Cauchy sequences

**Definition: Cauchy sequences**

Statement 1: A sequence $(x_n)$ is called ***Cauchy*** when for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$, we have $|x_m - x_n| < \varepsilon$.

Statement 2: An equivalent way of saying this is that for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and $p \in \mathbb{N}$, we have $|x_{n+p} - x_n| < \varepsilon$.

**Proposition 2.** Every convergent sequence is Cauchy.

*Proof.* We begin by picking a convergent sequence: let $(x_n)$ converge to $\hat{x}$. Estimate

$$\begin{aligned} |x_n - x_m| =& |(x_n - \hat{x}) + (\hat{x} - x_m)| \\ \leq& |x_n - \hat{x}| + |x_m - \hat{x}|. \end{aligned}$$

To show that this sequence Cauchy (Statement 1), let $\varepsilon > 0$ be given and use definition of $x_n \to \hat{x}$ with $\varepsilon' = \dfrac{\varepsilon}{2}$ to get $N \in \mathbb{N}$ such that $|x_k - \hat{x}| < \varepsilon'$ whenever $k > N$. This $N$ works in statement 1, since from what we have shown above, $m, n \geq \mathbb{N} \Rightarrow |x_m - x_n| < \varepsilon' + \varepsilon' = \varepsilon$. □

**Corollary 2.** Any sequence that is *not* Cauchy *must* diverge.

*Proof.* Contrapositive of the statement above; in general this is a great approach when proving divergence. □

**Theorem: Metric completeness**

Every Cauchy sequence converges (to a real limit) in $\mathbb{R}$.

The proof for this is something we will revisit after we have a bit more machinery, which we will now develop.

## 7.2 Bounded sets

**Theorem: Order completeness**

Given any non-empty $\mathcal{S} \subseteq \mathbb{R}$, let $\mathcal{A} := \{a \in \mathbb{R} : \text{for all } x \in \mathcal{S},\ a \leq x\}$, $\mathcal{B} := \{b \in \mathbb{R} : \text{for all } x \in \mathcal{S},\ x \leq b\}$, then:

(a) Either $\mathcal{A} = \emptyset$ or $\mathcal{A} = (-\infty, \alpha]$ for some $\alpha \in \mathbb{R}$.

(b) Either $\mathcal{B} = \emptyset$ or $\mathcal{B} = [\beta, \infty)$ for some $\beta \in \mathbb{R}$.

We say that $\mathcal{S}$ is *bounded above* when $\mathcal{B} \neq \emptyset$, and call each $b \in \mathcal{B}$ an *upper bound for* $\mathcal{S}$.

Similarly, $\mathcal{S}$ is *bounded below* when $\mathcal{A} \neq \emptyset$ each $a \in \mathcal{A}$ is a *lower bound for* $\mathcal{S}$. Just the word "bounded" means "bounded above" *and* "bounded below."

We now define one of the most important concepts of this course:

**Definition: Supremum**

When $\mathcal{B} \neq \emptyset$, we call $\beta$ the ***supremum*** of $\mathcal{S}$, i.e., $\beta = \sup(\mathcal{S})$.

Useful characterization:

(i) For all $x \in \mathcal{S}$, $x \leq \beta$ is the same as saying "$\beta$ is an *upper bound* for $\mathcal{S}$."

(ii) For all $\gamma < \beta$, there exists $x \in \mathcal{S}$ such that $\gamma < x$, which is the same as saying "nothing *less than* $\beta$ is an upper bound." This is why another name for the supremum is *the least upper bound*.

Similarly, we define:

**Definition: Infimum**

When $\mathcal{A} \neq \emptyset$, $\alpha = \inf(\mathcal{S})$ is the ***infimum*** or *the greatest lower bound* of $S$.

## 7.3 Monotonic sequences

**Theorem: Monotonic sequence property**

Given any sequence $(x_n)$ with $x_1 \leq x_2 \leq x_3 \leq \ldots$, either $x_n \to \infty$ *or* $x_n$ converges to a real limit.

**Note.** When we say that $x_n \to \infty$, we are saying more than just "the sequence diverges", we are commenting on specifically *how* it diverges.

**Note.** We will prove all these theorems at some point in this course, however, right now we will take them for granted.

**Note** (Linkages)**.** These 3 viewpoints on completeness contain equivalent information; each one implies the others.

Going back to metric completeness, we show one of these linkages:

**Theorem**

Metric completeness (which says Cauchy sequences must converge) implies order completeness (If $\mathcal{S} \neq \emptyset$ is bounded above, $\sup(\mathcal{S})$ exists.)

*Proof.* Let $\mathcal{S} \subseteq \mathbb{R}$ be non-empty; define $\mathcal{B} = \{b \in \mathbb{R} : \text{for all } s \in \mathcal{S},\ s \leq b\}$. Assume $\mathcal{B} \neq \emptyset$ and define sequence $b_n = \min\left\{\mathcal{B} \cap \left\{\frac{k}{2^n} : k \in \mathbb{Z}\right\}\right\}$. This is a Cauchy sequence (we just assert this but this is certainly something we would have to prove in a proof usually.) Thus, $\beta = \lim_{n\to\infty} b_n$ will have the properties defining $\sup(\mathcal{S})$. For each fixed $n$, we have

(i) $b_n - \frac{1}{2^n} \notin \mathcal{B} \Rightarrow$ there exists $s_n \in \mathcal{S}$ such that $s_n > b_n - \frac{1}{2^n}$.

(ii) $b_{n+1} \leq b_n$ (minimum over a larger set of points.)

(iii) Using $b_{n+1} \in \mathcal{B}$, we have $b_{n+1} \geq s_n$ for $s_n$ above. Using (ii),

$$b_{n+1} \geq s_n > b_n - \frac{1}{2^n} \iff 0 \leq b_n - b_{n+1} < \frac{1}{2^n}.$$

Now, we estimate

$$\begin{aligned}|b_{n+p} - b_n| =& b_n - b_{n+p} \\ =& (b_n - b_{n+1}) + (b_{n+1} - b_{n+2}) + \cdots + (b_{n+p-1} - b_{n+p}) \\ <& \frac{1}{2^n} + \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^{n+p-1}} + \frac{2}{2^n}.\end{aligned}$$

This is the key to showing $(b_n)$ is Cauchy. $\square$

**Quotes of the day: Dr. Zahl special**

"I won't intentionally say something wrong, but I might; I hit a car on my way to work today so I'm a bit rattled. Also, check your brakes before you, you know, brake." - Dr. Joshua Zahl, 09/29/2023.

# 8 Sub-sequences

Recall that sequences (of real numbers) are a function $f : \mathbb{N} \to \mathbb{R}$; a more practical way to of defining the notation for sequences is $x_n$ for $x \in \mathbb{N}$. This is because we are effectively "enumerating" the sequence, clarifying that it maps from the natural numbers.

**Definition: Sub-sequences**

If $x : \mathbb{N} \to \mathbb{R}$ is a sequence, then a ***sub-sequence*** of the $x$ is a sequence of the form $x \circ g : \mathbb{N} \to \mathbb{R}$, where $g : \mathbb{N} \to \mathbb{N}$ is *strictly increasing*.

**Notation 6.** This notation is clunky to work with, and the notation used in practice is $x_{n_1}, x_{n_2}, x_{n_3}, \dots, x_{n_k}$ such that $n_1 < n_2 < n_3 < \dots$

# 9 Completeness (Dr. Zahl)

**Property (a): Metric completeness**

Every *Cauchy sequence* of real numbers *converges*.

**Note.** Once we define a metric it becomes more apparent why this is useful; currently, both convergent and Cauchy are the same, but when we depart from the real numbers this fact is less obvious and needs some more work.

**Property (b): Order completeness (Least upper bound property)**

Consider $\mathcal{S} \subseteq \mathbb{R}$ ($\mathcal{S} \neq \emptyset$); we define $\mathcal{B} = \{b \in \mathbb{R} : \text{for all } s \in \mathcal{S},\ s \leq b\}$. Either $\mathcal{B} = \emptyset$ or $\mathcal{B} = [\beta, \infty)$ for some $\beta \in \mathbb{R}$.

**Note**:

(a) Every bounded sets have infinitely many upper bounds, however the least upper bound (supremum) tells us something about the structure of the set.

(b) The supremum is not necessarily in the set which it bounds.

(c) If we were to write the order completeness property for the rational numbers, this would not work. This is apparent if we define as $\mathcal{S}$ for the rationals, we will see that we get a contradiction because there will always be a rational number smaller than $\beta$ which will be an upper bound for $\mathcal{S}$, causing the supremum to never exist.

**Property (c): Monotone convergence property**

If $(x_n)$ is *monotone increasing*, either $x_n \to \infty$, or $(x_n)$ converges.

**Theorem**

Properties (a), (b), (c) are equivalent

We have already shown that (a)⇒(b). We proceed by showing that (b)⇒(c):

*Proof.* Let $(x_n)$ be a monotone increasing sequence. Let $\mathcal{S} = \{x_n : n \in \mathbb{N}\}$. If $\mathcal{S}$ is *not* bounded above, then $x_n \to \infty$. Otherwise, the set of upper bounds $\mathcal{B}$ is of the form $\mathcal{B} = [\beta, \infty)$. Let us show that $x_n \to \beta$.

We know $x_n \leq \beta$ for every $n$. Let $\varepsilon > 0$; there exists some $x_N \in \mathcal{S}$ such that $x_N > \beta - \varepsilon$. Notice that this has to be true since if it wasn't, we would have $x_n \leq \beta - \varepsilon$, i.e., $\beta - \varepsilon$ is an upper bound for $x_n$. However, since $\beta$ is the supremum, and $\beta - \varepsilon < \beta$, this is clearly not true. Thus, for all $n \geq N$, $x_n \geq x_N > \beta - \varepsilon$, i.e., $|x_n - \beta| < \varepsilon$. □

Now, we show that (c)⇒(a):

*Proof.* The sketch for this proof is as follows:

We begin by showing the following lemmas:

**Lemma-1**

Every sequence of real numbers has a *weakly* monotone sub-sequence.

This is fairly intuitive; consider $x_n = (-1)^n n$. This has uncountably many sub-sequences.

**Lemma-2**

Every *Cauchy sequence* is *bounded*.

The proof for this is basically pushing around definitions from before.

**Lemma-3**

If $(x_n)$ is Cauchy and $(x_{n_k})$ is a sub-sequence and $x_{n_k} \to \hat{x}$, then $x_n \to \hat{x}$.

To prove this we do require the fact that the sequence is Cauchy; every Cauchy sequence is bounded and from Lemma-1 we have a monotone increasing sub-sequence. We know this converges by the monotone convergence property, and thus, this forces the sequence itself to converge. So we combine all 3 lemmas to prove the claim. $\square$

**Quotes of the day**

"There's a Halloween joke here; 'limb soup.' could be a good math band name." - Dr. Philip Loewen, 10/04/2023

We now continue with actually doing the proof for (c)$\Rightarrow$(a):

*Proof.* Notice that

$$|x_n - L| \leq |x - x_{n_k}| + |x_{n_k} - L|. \tag{1}$$

By definition of a Cauchy sequence, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n > N$, we get $|x_m - x_n| < \dfrac{\varepsilon}{2}$. Since we have subsequence convergence, we can safely say that for sufficiently large $K \in \mathbb{N}$, all $k > K$ gives us $|x_{n_k} - L| < \dfrac{\varepsilon}{2}$. Pick some $\tilde{k} > K$ and $n > N$ such that $n_{\tilde{k}} > N$; using eq. (1), we get

$$\begin{aligned} |x_n - L| &< |x_n - x_{n_{\tilde{k}}}| + |x_{n_{\tilde{k}}} - L| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

$\square$

**Hunting License**

It is valid to use completeness in any of the three equivalent forms mentioned above to solve homework or test problems as long as it is cited.

**Corollary: Bolzano-Weierstrass Theorem**

Every bounded real sequence has a convergent subsequence.

**Note.** All of the three completeness properties and the corollary mentioned above only work for the real numbers; these fail for the rational numbers. This is quite apparent since the rational numbers do not have the least upper bound property.

## 9.1 More on the Supremum and Infimum

Given a set $\mathcal{S} \subseteq \mathbb{R}$, we let $\mathcal{B} = \{b \in \mathbb{R} : \text{for all } x \in \mathcal{S},\ \text{we have } x \leq b\}$. What are the possibilities for this set?

- If $\mathcal{S}$ has no upper bound (e.g. $\mathcal{S} = \mathbb{Z}$), then $\mathcal{B} = \emptyset$.
- If $\mathcal{S}$ has an upper bound, then $\mathcal{B} = [\beta, +\infty)$; here $\beta = \sup(\mathcal{S})$.
- If $\mathcal{S} = \emptyset$, then $\mathcal{B} = (-\infty, \infty)$.

Our aim now is to re-define the supremum to co-define all three cases:

- If $\mathcal{S}$ has no upper bound, then $\sup(\mathcal{S}) = +\infty$.
- If $\mathcal{S} = \emptyset$, then $\sup(\mathcal{S}) = -\infty$.

this is symmetric with the infimum:

- If $\mathcal{S}$ has no upper bound, then $\inf(\mathcal{S}) = -\infty$.
- If $\mathcal{S} = \emptyset$, then $\inf(\mathcal{S}) = +\infty$.

We have implicitly assumed something here: we have allowed the supremum and infimum to operate on sets that include the extended values $-\infty$ and $+\infty$. This is allowed, however it is something that should be acknowledged for the sake of rigour.

# 10 The Limes superior and Limes inferior (upper and lower limits)

Given a real sequence $(x_n)$, define

$$\limsup_{n\to\infty} x_n := \inf_{n\in\mathbb{N}}\left(\sup_{k\geq n} x_k\right)$$
$$\liminf_{n\to\infty} x_n := \sup_{n\in\mathbb{N}}\left(\inf_{k\geq n} x_k\right)$$

**Example 12.**

(a) Consider $x_n = \dfrac{1}{n}$; in this case

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n} &= \inf_{n\in\mathbb{N}}\left(\sup\left\{\frac{1}{k} : k\geq n\right\}\right)\\ &= \inf_{n\in\mathbb{N}}\left\{\frac{1}{n}\right\} = 0.\end{aligned}$$

(b) Consider $x_n = (-1)^n + \dfrac{1}{n}$; in this case

$$\begin{aligned}\limsup_{n\to\infty} x_n &= \inf_{n\in\mathbb{N}} \sup\left\{(-1)^k + \frac{1}{k} : k\geq n\right\}\\ &= \inf_{n\in\mathbb{N}}\begin{cases}1+\frac{1}{2j}, & \text{if } k=2j \text{ (even)}\\ 1+\frac{1}{2j-1}, & \text{if } k=2j-1 \text{ (odd)}\end{cases}\\ &= \inf_{n\in\mathbb{N}}\left\{1+\frac{1}{\lceil\frac{n}{2}\rceil}\right\} = 1.\end{aligned}$$

Similarly,

$$\begin{aligned}\liminf_{n\to\infty} x_n &= \sup_{n\in\mathbb{N}}\left(\inf_{k\geq n}(-1)^k + \frac{1}{k}\right)\\ &= \sup_{n\in\mathbb{N}}\{-1\} = -1.\end{aligned}$$

## 10.1 Laws for the limes superior and limes inferior

Let $x_n$ be a real-valued sequence:

(a)
$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n.$$

(b) For $L\in\mathbb{R}$, have $\lim_{n\to\infty} x_n = L$ iff $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = L$.

We will prove these facts in a later class.

**Proposition 3.** Let $(x_n)$ be a real-valued sequence; we define

$$\mu = \liminf_{n\to\infty} x_n$$
$$\nu = \limsup_{n\to\infty} x_n.$$

(a) If $\lim_{n\to\infty} x_{n_k} = l$ for some sub-sequence $(x_{n_k})$, then $\mu\leq l\leq\nu$.

(b) There exists sub-sequences $(x_{n_j})$ and $(x_{n_k})$ of $(x_n)$ along which $\mu = \lim_{j\to\infty} x_{n_j} = l$ and $\nu = \lim_{k\to\infty} x_{n_k} = l$.

*Proof.* Let $T_n = \{x_n, x_{n+1}, x_{n+2}, \dots\}$; define $i_n := \inf(T_n) = \inf_{k\geq n} x_k$ and $s_n := \sup(T_n) = \sup_{k\geq n} x_k$.

(a) For each $k \in \mathbb{N}$, $x_{n_k} \geq \inf\{x_j : j \geq n_k\} = i_{n_k}$. As $k \to \infty$, we have $x_{n_k} \to l$ and $i_{n_k} \to \mu \Rightarrow l \geq \mu$. Similarly, $x_{n_k} \leq s_{n_k}$ for all $k$, so $l \leq M$.

□

**Quotes of the day**

Some guy: "So, is that Minkowski addition?"

Dr. Loewen: "Pfff, I don't know! I know some famous names, and I know Minkowski, but I don't know!" - 10/11/2023

## Homework hints

Our aim here is to make $\liminf_{n\to\infty} x_n$ and $\limsup_{n\to\infty} x_n$ our main tools. Just writing $\lim_{n\to\infty} x_n$ requires prior work to show it exists. But $\limsup x_n$ and $\liminf x_n$ have values in $\mathbb{R} \cup \{\pm\infty\}$, so $\lim_{n\to\infty} x_n$ has meaning (in $\mathbb{R} \cup \{\pm\infty\}$) exactly when

$$\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n.$$

**Lemma**

If for real sequences $(x_n)$ and $(y_n)$ there exists $N \in \mathbb{N}$ such that $x_n \leq y_n$ for all $n \in \mathbb{N}$, then

$$\liminf_{n\to\infty} x_n \leq \liminf_{n\to\infty} y_n \tag{2}$$

$$\limsup_{n\to\infty} x_n \leq \limsup_{n\to\infty} y_n. \tag{3}$$

The proof for this is given in the canvas notes, and it is not too involved (just pushing around definitions of the sup and inf.)

**Note.** Given that $x_n < y_n$, it is not enough to say that eq. (2) and eq. (3) are strict inequalities; that requires more work and is not necessarily always true. One such example is $x_n = -\dfrac{1}{n}$ and $y_n = 0$.

Now, we want to show that $\lim_{n\to\infty} x_n = L$. For a given sequence $(x_n)$ and some $L \in \mathbb{R}$, it suffices to show that

$$L \leq \liminf_{n\to\infty} x_n \quad \text{and} \quad \limsup_{n\to\infty} x_n \leq L. \tag{4}$$

It always follows directly that

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n,$$

so that is not something we need to mention explicitly.

Going back to eq. (4), it is equivalent to show that for all $\varepsilon > 0$,

$$L - \varepsilon \leq \liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n \leq L + \varepsilon.$$

The idea here is to construct sequences $a_n \to L$ and $b_n \to L$, and show that for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$a_n - \varepsilon \leq x_n \leq b_n + \varepsilon, \text{ for all } n \geq N. \tag{5}$$

Once we get eq. (5) we can fix $\varepsilon > 0$ and take $\limsup / \liminf$ on $n$ to get eq. (4); this works for arbitrary $\varepsilon > 0$, i.e., it works for all $\varepsilon > 0$, so we are done.

# 11 Construction of $\mathbb{R}$

**Notation 7.** We start by defining required notation:

- Let $\mathrm{CS}(\mathbb{Q})$ be the set of Cauchy sequences with entries in $\mathbb{Q}$.
- $x, y, z$ will be typical sequence names, e.g., $x = (x_1, x_2, x_3, \dots)$.
- Let
$$R[x] = \{x' \in \mathrm{CS}(\mathbb{Q}) : \lim_{n\to\infty} |x'_n - x_n| = 0\}.$$
- Let
$$\mathcal{R} = \{R[x] : x \in \mathrm{CS}(\mathbb{Q})\}, \text{ which is our model for } \mathbb{R}.$$
- Let $\Phi : \mathbb{Q} \to \mathcal{R}$, such that $\Phi(q) = R[(q_1, q_2, q_3, \dots)]$.

## 11.1 Equality

For $x, x' \in \mathrm{CS}(\mathbb{Q})$ define a relation " $\sim$ " by

$$x' \sim x \iff \lim_{n\to\infty} |x'_n - x_n| = 0.$$

This is an "equivalence relation" (these relations are *reflexive, symmetric* and *transitive*), and the sets $R[x]$ are its equivalence classes.

## 11.2 Addition

For $x, y \in \mathrm{CS}(\mathbb{Q})$, we defined
$$x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots),$$
the result is in $\mathrm{CS}(\mathbb{Q})$ (we proved this in homework 4, problem 6(a).) Now we wish to extend the definition to $\mathbb{R}$:

$$R[x] + R[y] = R[x + y], \quad \text{where } x, y \in \mathrm{CS}(\mathbb{Q}).$$

**Proposition 4.** This " $+$ " is well defined, i.e., if $x, x', y, y' \in \mathrm{CS}(\mathbb{Q})$ obey $R[x] = R[x']$ and $R[y] = R[y']$, then $R[x + y] = R[x' + y']$.

*Proof.* With $x, x', y, y'$ as above, pick any $z' \in R[x' + y']$ to show $z' \in R[x + y]$; use the properties:

$$\begin{aligned} z' \in R[x' + y'] &\iff z'_n - (x'_n + y'_n) \to 0 \\ x' \in R[x'] = R[x] &\iff x'_n - x_n \to 0 \\ y' \in R[y'] = R[y] &\iff y'_n y_n \to 0. \end{aligned}$$

Thus, we get

$$z'_n - (x_n + y_n) = \underbrace{z'_n - (x'_n + y'_n)}_{\to 0} + \underbrace{(x'_n - x_n)}_{\to 0} + \underbrace{(y'_n - y_n)}_{\to 0},$$

which is the sum rule for limits, giving us $z'_n - (x_n + y_n) \to 0$, i.e., $z' \in R[x + y]$; since $z'$ is arbitrary, so $R[x' + y'] \subseteq R[x + y]$. The other inclusion can be shown by swapping $(x, y, z) \leftrightarrow (z', y', z')$ above to get

$$R[x' + y'] \supseteq R[x + y] \Rightarrow R[x' + y'] = R[x + y].$$

$\square$

**Proposition**

We call $(\mathcal{R}, +)$ is an Abelian group, i.e., for all $x, y, z \in \mathrm{CS}(\mathbb{Q})$, we have

(A) $R[x] + R[y]$ is a well-defined element of $\mathcal{R}$.

(B) $R[x] + R[y] = R[y] + R[x]$.

(C) $(R[x] + R[y]) + R[z] = R[x] + (R[y] + R[z])$.

(D) $R[x] + \Phi(0) = R[x]$.

(E) $\mathcal{R}$ contains another element "$-R[x]$", satisfying $R[x] + (-R[x]) = \Phi(0)$.

*Proof.* The proof for (A) to (D) follows pretty much directly from definitions, and some of them are already proved before.

For part (E): Given $x = (x_1, x_2, x_3, \ldots)$, define $-x = (-x_1, -x_2, -x_3, \ldots) \in \mathrm{CS}(\mathbb{Q})$, and note

$$R[x] + R[-x] = R[x + (-x)] = R[(0, 0, 0, \ldots)] = \Phi(0).$$

$\square$

## 11.3 Multiplication

We now extend "$\cdot$" to $\mathcal{R}$ by lifting "$\cdot$" defined in $\mathrm{CS}(\mathbb{Q})$. Define

$$\begin{aligned} R[x] \cdot R[y] &:= R[x \cdot y] \\ &:= R[(x_1 y_1, x_2 y_2, x_3 y_3, \ldots)]. \end{aligned}$$

We have shown before in Homework 4 problem 6(b) that $x \cdot y \in \mathrm{CS}(\mathbb{Q})$.

**Proposition**

This "$\cdot$" is well-defined, i.e., if $x, x', y, y' \in \mathrm{CS}(\mathbb{Q})$ with $R[x] = R[x']$ and $R[y] = R[y']$ then $R[x] \cdot R[y] = R[x'] \cdot R[y']$.

*Proof.* Given $x, x', y, y'$, as in the setup, $x' \in R[x]$ and $y' \in R[y]$, i.e., $x'_n - x_n \to 0$ and $y'_n - y_n \to 0$. Now, we rearrange

$$\begin{aligned} x'_n y'_n - x_n y_n =& [(x'_n - x_n) + x_n] y'_n - x_n y_n \\ =& (x'_n - x_n) y'_n + x_n (y'_n - y_n). \end{aligned}$$

**Note.** Another trick like this can be found in the proof for Theorem 3.3 part (c) in Rudin.

Now, since every Cauchy sequence is bounded, there exists $M_0, M_1$ such that

$$|x'_n y'_n - x_n y_n| \leq M_0 |x'_n - x_n| + M_1 |y'_n - y_n|;$$

by squeeze theorem, LHS $\to 0$, i.e., $x' \cdot y' \in R[x \cdot y]$. However, $x' \cdot y' \in R[x' \cdot y']$. A non-empty intersection implying inequality is something that we prove on Homework 5 problem 3. $\square$

**Proposition**

We call $(\mathcal{R}^*, \cdot)$ is an Abelian group (where $\mathcal{R}^* = \mathcal{R}\backslash\{\Phi(0)\}$), i.e., for all $x, y, z \in \mathrm{CS}(\mathbb{Q})$, we have

(A) $R[x] \cdot R[y]$ is a well-defined element of $\mathcal{R}$.

(B) $R[x] \cdot R[y] = R[y] \cdot R[x]$.

(C) $(R[x] \cdot R[y]) \cdot R[z] = R[x] \cdot (R[y] \cdot R[z])$.

(D) $R[x] \cdot \Phi(1) = R[x]$.

(E) $\mathcal{R}$ contains another element "$\dfrac{1}{R[x]}$", satisfying $R[x] \cdot \left(\dfrac{1}{R[x]}\right) = \Phi(1)$.

*Proof.* Similar to the Abelian group under addition, the proof for parts (A) to (D) follow from the definition or have already been shown before. The proof for part (E) is Homework 5 problem 6. $\square$

## 11.4 Distribution

**Proposition**

GIven any $a, b, c \in \mathrm{CS}(\mathbb{Q})$, we have

$$R[a] \cdot (R[b] + R[c]) = (R[a] \cdot R[b]) + (R[a] \cdot R[c]).$$

*Proof.* Using the definitions, we get

$$\begin{aligned} \text{LHS} =& R[a] \cdot R[b+c] = R[a \cdot (b+c)] \\ \text{RHS} =& (R[a \cdot b]) + (R[a \cdot c]) = R[a \cdot b + a \cdot c]. \end{aligned}$$

Inside the Cauchy sequences, we have the $n^{\text{th}}$ terms

$$\left.\begin{aligned} [a \cdot (b+c)]_n &= a_n(b_n + c_n) \\ [a \cdot b + a \cdot c]_n &= a_n b_n + a_n c_n \end{aligned}\right\} \text{ same for each } n \in \mathbb{N}.$$

So indeed both LHS and RHS share this representative sequence, i.e., they must be equal. $\square$

## 11.5 Ordering

**Definition**

Given $x, y \in \mathrm{CS}(\mathbb{Q})$ define $R[x] < R[y]$ when there exists $r > 0$ (where $r \in \mathbb{Q}$), there exists $n \in \mathbb{N}$ such that $x_n + r < y_n$, for all $n \geq N$.

**Proposition**

This definition for " $<$ " is unambiguous, i.e., independent of representatives selected from $R[x], R[y]$. Thus,

(a) Every $x \in \mathrm{CS}(\mathbb{Q})$ obeys exactly one of $R[x] < \Phi(0)$, or $R[x] = \Phi(0)$, or $R[x] > \Phi(0)$.

(b) $R[x] < R[y]$ and $R[y] < R[z]$ implies $R[x] < R[z]$.

*Proof.* The proof for this is Homework 5 problem 5. $\square$

**Quotes of the day**

"Inside the Sauder school of business…don't get me started; but outside Sauder professors only wear ties on special occasions. Today is a special occasion." - Dr. Loewen wearing a tie, 10/13/2023

**Proposition: Order components- OC**

Let $a, b \in \mathrm{CS}(\mathbb{Q})$;

(a) If $R[a] > R[b]$ then there exists $N \in \mathbb{N}$ such that $R[a] > \Phi(b_N)$.

(b) If $\Phi(b_k) \geq R[a]$ for all $k \in \mathbb{N}$, then $R[a] \leq R[b]$.

*Proof.* Note that (b) $\iff$ (a) by contrapositive, so we just prove (a). Given $R[a] > R[b]$, there exists $N_0 \in \mathbb{N}$ and $\mathbb{Q} \ni r > 0$ such tat $a_n > b_n + r$ for all $n \geq N_0$. Also, $a, b \in \mathrm{CS}(\mathbb{Q})$ given $N_a, N_b \in \mathbb{N}$ such that $|a_m - a_n| < \dfrac{r}{10}$ for $m, n \geq N_a$ and $|b_m - b_n| < \dfrac{r}{10}$ for $m, n \geq N_b$. Let $N = \max\{N_0, N_a, N_b\}$; for any $m \geq N$,

$$\begin{aligned} b_N =& b_m + (b_N - b_m) \\ \leq& b_m + \frac{r}{10} \\ \leq& (a_m - r) + \frac{r}{10} \\ =& a_m - \frac{9r}{10} < a_m - \frac{r}{2}. \end{aligned}$$

Thus, $b_N < a_m - \dfrac{1}{2}$ for all $m \geq N$; we are done. □

## 11.6 Completeness of $\mathcal{R}$

**Notation 8.** We let $\alpha = R[a]$, $\beta = R[b]$ and $\gamma = R[c]$.

**Reflection.** The mapping $\Phi : \mathbb{Q} \to \mathcal{R}$ embeds a "working copy of $\mathbb{Q}$" in $\mathcal{R}$. We confirmed $\Phi(p+q) = \Phi(p) + \Phi(q)$, $\Phi(p \cdot q) = \Phi(p) \cdot \Phi(q)$, and $p < q \iff \Phi(p) < \Phi(q)$; everything we want from the rationals is mirrored in $\mathrm{CS}(\mathbb{Q})$. Notice that there are some equivalence classes of Cauchy sequences that are not in here: $\mathcal{R}$ includes many elements not of the form $\Phi(q)$ where $q \in \mathbb{Q}$.

**Example 13.** A member of $\mathcal{R}$ : $\pi = R[(3, 3.1, 3.141, 3.1415, \dots)]$ is outside $\Phi(\mathbb{Q})$.

**Theorem: Completeness of $\mathcal{R}$**

Let $\mathcal{A}$ be a non-empty subset of $\mathcal{R}$ with an upper bound, i.e., there exists $\mu \in \mathcal{R}$ such that for all $\alpha \in \mathcal{A}$ we have $\alpha \leq \mu$; there exists $\beta \in \mathcal{R}$ such that

(a) For all $\alpha \in \mathcal{A}$, $\alpha \leq \beta$ ($\beta$ is an upper bound.)

(b) For all $\gamma \in \mathcal{R}$ such that $\gamma < \beta$, there exists $\alpha \in \mathcal{A}$ such that $\gamma < \alpha$ ($\beta$ is te supremum.)

*Proof.* For each $n \in \mathbb{N}$, define $b_n = \min(S_n)$ where $S_n = \left\{\dfrac{k}{2^n} : k \in \mathbb{Z},\ \text{for all } \alpha \in \mathcal{A},\ \Phi\left(\dfrac{k}{2^n} \geq \alpha\right)\right\}$. By the hypothesis, $\mu$ is an upper bound for $\mathcal{A}$. From Homework 5, we know that there exists some $K \in \mathbb{Z}$ such that $\mu < \Phi(K)$. Every $\dfrac{k}{2^n} \geq K$ is in $S_n$. Additionally $S_n$ has a lower bound (showing this is left as an exercise.) For each $S_n \subseteq S_{n+1}$; so we have $b_n \geq b_{n+1} \geq b_n - \dfrac{1}{2^{n+1}}$, or $0 \leq b_n - b_{n+1} \leq \dfrac{1}{2^{n+1}}$. Therefore if $p \in \mathbb{N}$, we have

$0 \le b_n - b_{n+p} \le (b_n - b_{n+1}) + \cdots + (b_{n+p-1} - b_{n+p})$, which means

$$\begin{aligned} 0 \le& b_n - b_{n+p} \\ \le& \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots + \frac{1}{2^{n-p}} \\ \le& \frac{1}{2^n}\left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^p}\right) \\ <& \frac{1}{2^n}. \end{aligned}$$

Hence, $b = (b_n)$ is a Cauchy sequence. Since $\beta = R[b]$, from the definition of $b_n \in S_n$, $\Phi(b_n) \ge \alpha$ for all $\alpha \in \mathcal{A}$. from OC (b), $\beta \ge \alpha$ for all $\alpha \in \mathcal{A}$; this is conclusion (a).

For (b), let $\gamma < \beta$ and say $\gamma = R[c]$. This comes with $\mathbb{Q} \ni r > 0$, $N \in \mathbb{N}$ where for all $n \ge N$, $c_n + r < b_n$. Increase $N$ as needed to get $\frac{1}{2^N} < \frac{r}{2}$. Then, since $b_n \ge b_{n+1}$,

$$c_n + \frac{r}{2} = (c_n + r) - \frac{r}{2} < b_n - \frac{r}{2} < b_n - \frac{1}{2^N} \le b_N - \frac{1}{2^N}.$$

Thus, $\Phi(c) < \Phi(b_N)$.

**Note.** He forgot about $\alpha$ and has mentioned that we should check his notes on canvas.

Therefore, every monotone bounded sequence of these converge. Every Cauchy sequence of these converges. The Archimedean property follows (have to read Canvas notes for this part.) □

**Quotes of the day**

"After the math conference mathematicians go to the bar. The first one says can I get a beer, the second one says can I get half a beer, the third one says can I get a quarter of a beer, and so on. The bartender slams two beers on the counter, and says 'figure it out yourself!'." - Dr. Loewen, 10/16/2023

**Theorem: Archimedean property**

The set $\mathbb{N}$ has no upper bound in $\mathbb{R}$.

*Proof.* For the sake of contradiction, suppose that $\mathbb{N}$ has an upper bound; consider $\alpha = \sup(\mathbb{N})$. In this case, $\alpha - \frac{1}{2}$ cannot be an upper bound (by the definition of the supremum), i.e., there exists some $n \in \mathbb{N}$ such that $\alpha - \frac{1}{2} < n$. However, this gives us $n+1 > n + \frac{1}{2} > \alpha \Rightarrow \alpha$ is not an upper bound, and therefore not the supremum; contradiction. Hence, $\mathbb{N}$ is not bounded. □

**Note.** This seems like a very obvious fact, but we need some work because there can be cases where this breaks down. One of them is mentioned in the example that follows.

**Example 14.** Let $\mathscr{F}$ be the set of rational functions $f : \mathbb{R} \to \mathbb{Q}[x]$ (here $\mathbb{Q}[x]$ denotes the set of all polynomials with rational coefficients) such that,

$$f(x) = \frac{p_0 + p_1 x + \cdots + p_m x^m}{q_0 + q_1 x + \cdots + x q_n x^n}$$

To define an *order*, say "$f > 0$" when some representation as above has $\frac{p_m}{q_n} > 0$; equivalently, $f(x) > 0$ for sufficiently large $x$. The constants in the numerator and denominator show that $\mathbb{Q}$ is a sub-field of $\mathscr{F}$ with a well defined "$<$". However, we get a contradiction, since the function $f(x) = x$ is an upper bound for $\mathbb{N}$. The proof for this is fairly elementary where for each $n$, we have $f(x) - n > 0$ for $x > n$. We are not saying that the set $\mathbb{N}$ has a supremum; this is clearly false. However, this does not mean that we cannot find an upper bound. Since the range of the function $f(x) = x$ is the rational numbers, it includes the natural numbers, so for every natural number we have a greater natural number (this fact needs no proof.) However, this does not imply there's a supremum because while this upper bound is technically "increasing" (it is not exactly; we just find larger values in the range), the exists no numerical value that can serve as the supremum.

## 12 Series

**Definition: Series**

Given a sequence $(a_n)$ in $\mathbb{R}$, the corresponding ***series*** is the new sequence $s_1, s_2, \ldots$ defined by

$$s_n = \sum_{k=1}^{n} a_k.$$

The "sum" is

$$S' = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k,$$

denoted

$$S = \sum_{k=1}^{\infty} a_k, \quad \text{or } S = \sum_{k\in\mathbb{N}} a_k.$$

A series *converges* when $S \in \mathbb{R}$ and *diverges* otherwise. Some divergent series can be described with extended values $\pm\infty$.

**Note.** The key point is "some" can be described this way because saying that they diverge *and* are one of $\pm\infty$ is more than saying that they diverge; we are describing *how* they diverge.

### 12.1 Geometric series

For any real $r$,

$$\begin{aligned}(1-r)(1+r+r^2+\cdots+r^n) =& 1+r+r^2+\cdots+r^n-r-r^2-\cdots-r^n-r^{n+1}\\ =& 1-r^{n+1}\\ \Rightarrow 1+r+r^2+\cdots+r^n =& \frac{1-r^{n+1}}{1-r} \quad (\text{when } r\neq 1).\end{aligned}$$

Thus,

$$\sum_{k\in\mathbb{N}} r^k = \begin{cases} \lim\limits_{n\to\infty} \dfrac{1-r^{n+1}}{1-r}, & \text{if } r\neq 1\\ \lim\limits_{n\to\infty} (n+1), & \text{if } r=1\end{cases},$$

i.e.,

$$\sum_{k\in\mathbb{N}} r^k = \begin{cases} \dfrac{1}{1-r}, & \text{if } |r|<1\\ +\infty, & \text{if } r>1\\ \text{DNE}, & \text{if } r<-1\end{cases}.$$

Alternatively, we define $f(x) = \sum_{k\in\mathbb{N}} x^k$. Then, $\text{Domain}(f) = (-1,1)$, and $f(x) = \dfrac{1}{1-x}$.

**Note.** It is generally tricky to look at a series and find the value it exactly converges to, but we can often talk about the domain it converges in.

**Example 15.** Consider

$$\sum_{n\in\mathbb{N}} \frac{2}{4n^2-1} = 1.$$

Using partial fractions, we have

$$\frac{2}{4n^2-1} = \frac{1}{2n-1} - \frac{1}{2n+1},$$

so clearly, this is "telescoping", i.e., all the terms except a select few cancel out:

$$\begin{aligned}
\sum_{n=1}^{n} \frac{2}{4n^2-1} =& \sum_{n\in\mathbb{N}} \left[\frac{1}{2n-1} - \frac{1}{2n+1}\right] \\
=& \frac{1}{2} - \cancel{\frac{1}{3}} + \cancel{\frac{1}{3}} - \cancel{\frac{1}{4}} + \cancel{\frac{1}{4}} + \cdots - \cancel{\frac{1}{2n-1}} + \cancel{\frac{1}{2n-1}} - \frac{1}{2n+1} \\
=& \frac{1}{2} - \frac{1}{2n+1},
\end{aligned}$$

as as we take the limit as $n \to \infty$, we get

$$\sum_{n\in\mathbb{N}} \frac{2}{4n^2-1} = \frac{1}{2}.$$

## 13 Testing for convergence

**Theorem: Monotone convergence**

If $a_n \geq 0$ for all $n$, then $\sum_{n\in\mathbb{N}} a_n$ converges iff $S_N = \sum_{n=1}^{N} a_n$ is bounded.

*Proof.* Note that $S_{N+1} - S_N = a_{N+1} \geq 0$ shows that $S_N$ is a non-decreasing sequence; rest follows from monotone convergence property. In this case "$\sum_{n\in\mathbb{N}} a_n$ converges." $\square$

**Theorem: Cauchy's Criterion**

The series $S = \sum_{n\in\mathbb{N}} a_n$ converges iff for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $m \geq N$, and for all $p \in \mathbb{N} \cup \{0\}$, we have

$$|a_m + a_{m+1} + \cdots + a_{m+p}| < \varepsilon.$$

*Proof.* Notice that $a_m + \cdots + a_{m+p} = S_{m+p} - S_{m-1}$; this states condition is just a reformulation of Cauchy's criterion for seq of partial sums, where we have already shown that the sequence of partial sums is Cauchy. $\square$

**Theorem: Test for divergence**

If $\lim_{n\to\infty} a_n \neq 0$, or the limit does not exist, then $\sum_{n\in\mathbb{N}} a_n$ diverges.

**Note.** This ***does not*** say that the sequence converges, and is merely a relatively quick test to check whether a sequence *diverges* or not; equivalently, we are saying that the converse of this statement is not generally true.

*Proof.* We prove this by contrapositive. If $\sum_{n\in\mathbb{N}} a_n$ converges, pick any $\varepsilon > 0$ and use Cauchy's criterion to get a $N \in \mathbb{N}$ such that $|a_m + \cdots + a_{m+p}| < \varepsilon$ for all $m \geq N$, and for all $p \in \mathbb{N} \cup \{0\}$. Use $p = 0$: for all $m \geq N$, $|a_m| < \varepsilon \Rightarrow a_n \to 0$. $\square$

**Theorem: Comparison test**

(a) If $0 \leq |a_n| \leq b_n$ for all $n \in \mathbb{N}$ and $\sum_{n\in\mathbb{N}} b_n < +\infty$, then $\sum_{n\in\mathbb{N}} a_n$ converges.

(b) If $\sum_{n\in\mathbb{N}} |a_n| = +\infty$, then $\sum_{n\in\mathbb{N}} |b_n| = +\infty$ as well.

*Proof.* For part (a), we use Cauchy's criterion and the triangle inequality to get

$$\begin{aligned}|a_m + \cdots + a_{m+p}| \leq& |a_m| + |a_{m+1}| + \cdots + |a_{m+p}| \\ \leq& b_m + b_{m+1} + \cdots + b_{m+p}.\end{aligned}$$

We now use Cauchy's criterion for $(b_n)$ to provide requirements for $\sum_{n\in\mathbb{N}} a_n$ to converge.

The proof for (b) is left as an exercise. $\square$

**Corollary 3.** Absolute convergence implies convergence; if $\sum_{n\in\mathbb{N}} |a_n| < +\infty$, then $\sum_{n\in\mathbb{N}} a_n < +\infty$.

Proof for this is the same as for the theorem where we set $b_n = |a_n|$.

**Note.** Convergence for sequence $S_N = \sum_{n=1}^{N} a_n$ holds iff we have convergence fo each $S_N^m = \sum_{n=m}^{N} a_i$; writing

$$\sum_{n\in\mathbb{N}} a_n = \sum_{n}^{\infty} a_n$$
$$= \sum a_n$$

is abuse of notation.

**Example 16.** Harmonic series $\sum_n \frac{1}{n}$ diverges.

*Proof.* Show Cauchy's criterion fails: there exists $\varepsilon > 0$ such that for all $N \in \mathbb{N}$, there exists $m \geq N$ and $p \in \mathbb{N} \cup \{0\}$ such that

$$|a_m + \cdots + a_{m+p}| \geq \varepsilon.$$

Pick $\varepsilon = \frac{1}{2}$, for any $N \in \mathbb{N}$, choose $n = N$, $p = N$ which gives us

$$\frac{1}{N} + \frac{1}{N+1} + \cdots + \frac{1}{N+N} \geq \frac{N+1}{2N} > \frac{1}{2} = \varepsilon.$$

$\square$

**Theorem: The root test**

Consider $S = \sum_{n\in\mathbb{N}} a_n$, we let $\alpha := \limsup_{n\to\infty} |a_n|^{1/n}$

(a) If $\alpha < 1$, then $S$ converges absolutely.

(b) If $\alpha > 1$, then $S$ diverges.

(a) *Proof.* Given $\alpha < 1$, pick $r \in (\alpha, 1)$, for all $n \leq N$ $|a_n|^{1/n} < r$, so

$$\sum_{n=N}^{\infty} |a_n| < \sum_{n=N}^{\infty} r^n < 1 \quad \text{(geometric series)}.$$

$\square$

(b) *Proof.* Pick $R \in (1, \alpha) \Rightarrow |a_n|^{1/n} > R$ for infinitely many $n$, $|a_n| > R^n > 1$, for all those $n$, $S$ diverges by crude test. $\square$

**Theorem: Ratio test**

Consider $\sum_{n\in\mathbb{N}} a_n$ with $a_n \neq 0$, for all $n$

(a) If $\overline{\alpha} = \limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} < 1$, then $S$ converges absolutely.

(b) If $\underline{\alpha} = \liminf_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} > 1$, then $S$ diverges.

(a) *Proof.* Choose $r \in (\overline{\alpha}, 1)$. Since $r < \overline{\alpha}$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, we have $\dfrac{|a_{n+1}|}{|a_n|} < r \iff |a_{n+1}| < r|a_n|$. So, $|a_{n+k}| < r|a_{n+k-1}| < \cdots < r^k|a_n|$ for all $k \in \mathbb{N}$. Thus,

$$\sum_{k\in\mathbb{N}} |a_{N+k}| < |a_N| \sum_{k\in\mathbb{N}} r^k < +\infty;$$

this implies absolute convergence. □

(b) *Proof.* The proof for this is left as an exercise. □

## 13.1 Comparing these tests

Given $S = \sum_{n\in\mathbb{N}} a_n$, define $\alpha := \limsup_{n\to\infty} |a_n|^{1/n}$, $\overline{\alpha}$ and $\underline{\alpha}$ as above

(i) $\overline{\alpha} < 1 \Rightarrow \alpha < 1 \Rightarrow \sum_{n\in\mathbb{N}} |a_n| < +\infty$.

(ii) $\underline{\alpha} > 1 \Rightarrow \alpha > 1 \Rightarrow S$ diverges.

(iii) If $\alpha = 1 \Rightarrow (\underline{\alpha} \leq 1 \leq \overline{\alpha})$, anything can happen.

Let us prove the first implication of (i):

*Proof.* When $\overline{\alpha} = +\infty$, we are done.

When $\overline{\alpha} < +\infty$, it suffices to show that for all $\varepsilon > 0$, $\alpha \leq \overline{\alpha} + \varepsilon$. Fix some $\varepsilon' > 0$ and define $\beta = \overline{\alpha} + \varepsilon'$. Now, $\beta > \overline{\alpha}$; there exists $N \in \mathbb{N}$ such that $n \geq N$,

$$\frac{|a_{n+1}|}{|a_n|} < \beta \Rightarrow |a_{n+1}| \leq \beta |a_n|.$$

For $p \in \mathbb{N}$, we have

$$|a_{N+p}| < \beta|a_{N+p-1}| < \cdots < \beta^p |a_N|,$$

so for all $m > N \in \mathbb{N}$, we have $|a_m|^{1/m} < (\beta^{-N}|a_N|)^{1/m}(\beta^m)^{1/m}$. Therefore, we have

$$\limsup_{m\to\infty} |a_m|^{1/m} \leq \beta.$$

□

**Theorem: Cauchy condensation**

If $a_n \geq a_{n+1} \geq 1$, for all $n \in \mathbb{N}$, the following are equivalent:

(a) $S = \sum_{n=N} a_n < +\infty$.

(b) $T = \sum_{k\in\mathbb{N}} 2^k a_{2^k} < +\infty$.

We start by showing that (b)⇒(a):

*Proof.* Notice that

$$\begin{aligned} a_1 + a_2 + \cdots + a_n =& (a_1) + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \ldots \\ \leq& a_1 + 2a_2 + 4a_4 + \ldots \\ \leq& T < +\infty; \end{aligned}$$

this holds for all $n$. □

**Quotes of the day**

"I love WeBWork (sarcastic), I wrote some of those questions, and some of them really lit up Piazza." - Dr. Loewen, 10/25/2023

Now we show that (a)⇒(b):

*Proof.* If $S$ converges, consider a partial sum for $T$:

$$\begin{aligned}
t_n =& \sum_{k=0}^{n} 2^k a_{2^k} \\
=& a_1 + 2a_2 + 4a_4 + \cdots + 2^n a_{2^n} \\
=& 2\left[\frac{1}{2}a_1 + a_2 + 2a_4 + \cdots + 2^{n-1}a_{2^n}\right] \\
\le& 2\left[a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \cdots + (a_{2^{n-1}+1} + \cdots + a_{2^n})\right] \\
\le& 2s_{2^n} \le 2S.
\end{aligned}$$

Partial sums for $T$ are bounded, which implies that $T$ converges. □

## 13.2 $p$-series

For fixed $p$, let

$$\zeta(p) := \sum_{n=1}^{\infty} \frac{1}{n^p}.$$

We ask ourselves for which $p$ does this function converge? We can discard all $p \le 0$ by the Crude test for divergence. For $p > 0$, the summand decreases, so Cauchy condensation gives us

$$\sum_{k=0}^{\infty} \frac{2^k}{(2^k)^p} = \sum_{k=0}^{\infty} \frac{1}{(2^{p-1})^k}.$$

This is a geometric series with common ratio $r = \dfrac{1}{2^{p-1}}$; we know this converges iff $r < 1$ or $p > 1$.

**Note.** The ratio and root test will not help with $p$-series; the series converges too slow to detect geometrically, and ratio and root test are not sharp enough to detect this. Because of this, we in fact get

$$\alpha = \limsup_{n\to\infty} |a_n|^{1/n} = \limsup_{n\to\infty} \frac{1}{(n^p)^{1/n}} = \limsup_{n\to\infty} \frac{1}{(n^{1/n})^p} = 1.$$

## 13.3 Kummer's test

**Theorem**

Consider $S = \sum_{n=1}^{\infty} a_n$ such that $a_n > 0$. Let $(D_n)$ be a sequence such that $D_n > 0$ for all $n$. We define

$$\begin{aligned}
\overline{L} :=& \limsup_{n\to\infty} \frac{D_k a_k - D_{k+1}a_{k+1}}{a_{k+1}} \\
\underline{L} :=& \liminf_{n\to\infty} \frac{D_k a_k - D_{k+1}a_{k+1}}{a_{k+1}}.
\end{aligned}$$

(a) If $\underline{L} > 0$, then $S$ converges.

(b) If $\overline{L} < 0$ and $\sum_{n=1}^{\infty} \frac{1}{D_n} = +\infty$, then $S$ diverges.

(a) *Proof.* If $\underline{L} > 0$, then pick $r \in (0, \underline{L})$. By definition of $\liminf$, there exists $N \in \mathbb{N}$ such that for all $k \geq N$, $r < \dfrac{D_k a_k - D_{k+1}a_{k+1}}{a_{k+1}}$, i.e., $ra_{k+1} < D_k a_k - D_{k+1}a_{k+1}$. Thus,

$$\begin{aligned} ra_{N+1} &< D_N a_N - D_{N+1}a_{N+1} \\ ra_{N+2} &< D_{N+1}a_{N+1} - D_{N+2}a_{N+2} \\ &\vdots \\ ra_{N+p} &< D_{N+p-1}a_{N+p-1} - D_{N+p}a_{N+p}. \end{aligned}$$

Thus, $r[a_{N+1} + \cdots + a_{N+p}] < D_N a_N - D_{N+p}a_{N+p} < D_N a_N$. Since $N$ is fixed and $p \in \mathbb{N}$ is arbitrary, we get that the partial sums are bounded, and thus the series converges. □

(b) *Proof.* The proof is on canvas, and the idea is similar to the proof for (a). □

**Example 17.** Take $D_n = 1$ in Kummer's test:

$$\begin{aligned} \underline{L} &= \liminf_{n\to\infty} \frac{a_k - a_{k+1}}{a_{k+1}} = \liminf_{n\to\infty} \left(\frac{a_k}{a_{k+1}} - 1\right) \\ \overline{L} &= \limsup_{n\to\infty} \left(\frac{a_k}{a_{k+1}} - 1\right) = \left(\frac{1}{\liminf_{n\to\infty}\left(\frac{a_{k+1}}{a_k}\right)} - 1\right), \end{aligned}$$

as we have shown before on the homework. So, $\underline{L} = \dfrac{1}{\overline{\alpha}} - 1$, $\overline{L} = \dfrac{1}{\underline{\alpha}} - 1$, and we see that we get convergence if $\underline{\alpha} < 1$, and divergence if $\underline{\alpha} > 1$; we have recovered the ratio test. So Kummer's test covers the ratio test.

**Example 18.** Take $D_n = n$ in Kummer's test; note that

$$\sum_{n\in\mathbb{N}} \frac{1}{D_n} = \sum_{n\in\mathbb{N}} \frac{1}{n} = +\infty.$$

Therefore,

$$\begin{aligned} \frac{D_k a_k - D_{k+1}a_{k+1}}{a_{k+1}} &= \frac{ka_k - (k+1)a_{k+1}}{a_{k+1}} \\ &= \frac{k(a_k - a_{k+1}) - a_{k+1}}{a_{k+1}} \\ &= k\left(\frac{a_k}{a_{k+1}} - 1\right) - 1. \end{aligned}$$

Testing the limits of this leads to the refined test named after Raabe; proving this is Homework 7 problem 7.

**Quotes of the day**

"We love our add and subtract trick, we plan to use it for Homework 7 problem 4. We love our telescoping series. What is we did both?" - Dr. Loewen, 10/27/2023

## 13.4 Alternating series

Most of the series we've been looking at have had all positive terms, now we have ones that include negative terms.

**Theorem: Alternating series test**

If $S = \sum_{n=0}^{\infty} (-1)^n a_n$ and

(a) $a_0 \geq a_1 \geq a_2 \geq \dots$

(b) $\lim_{k \to \infty} a_k = 0$

then $S$ converges.

*Proof.* Let $s_n = \sum_{k=0}^{\infty} (-1)^k a_k$ for $n \geq 0$. We can envision this as the partial sums going back and forth (alternating) and shrinking at the same time:

$$a_1 \leq s_3 \leq s_5 \leq \cdots \leq s_6 \leq s_4 \leq s_2 \leq s_0.$$

Note that the odd partial sums form a monotonically increasing sequence, and the even partial sums form a monotonically decreasing sequence, and clearly both sequences are bounded. Thus, they both converge. However,

$$0 \leq s_{2n} - s_{2n+1} = a_{2n+1},$$

which has limit 0, so Squeeze theorem (with $a_k \to 0$) shows both have the same limit. $\square$

**Note.** Recall from Math 101: any partial sum gives a lower or upper bound on the final value that $S$ converges to (depending on if it is even or odd); this is a strategy for calculation (not very useful in MATH 320.)

## 13.5 Summation by parts

Consider $\sum_{k=0}^{n} A_k b_k$. Define $A'_k := A_k - A_{k-1}$, $B_n := b_0 + b_1 + \cdots + b_n$, and $b_k = B'_k = B_k - B_{k-1}$. Therefore, we have

$$\begin{aligned}
\sum_{k=0}^{n} A_k b_k =& \sum_{k=0}^{n} A_k B'_k \\
=& A_0 b_0 + A_1 b_1 + A_2 b_2 + \cdots + A_n b_n \\
=& A_0 B_0 + A_1(B_1 - B_0) + A_2(B_2 - B_1) + \cdots + A_n(B_n - B_{n-1}) \\
=& (A_0 - A_1)B_0 + (A_1 - A_2)B_1 + \cdots + (A_{n-1} - A_n)B_{n-1} + A_n B_n \\
=& (-A'_1)B_0 + (-A'_2)B_1 + \cdots + (-A'_n)B_{n-1} + A_n B_n \\
=& A_n B_n - \sum_{k=1}^{n} A'_k B_{k-1}.
\end{aligned}$$

An analogue to this in integration is integration by parts:

$$\int_a^b u\,dv = uv|_a^b - \int_a^b v\,du;$$

hence the name summation by parts.

**Theorem: Dirichlet's test**

Consider $S = \sum_{n=0}^{\infty} a_n b_n$. If

(a) $a_n \geq a_{n+1}$ for all $n$, and $a_n \to 0$ as $n \to \infty$,

(b) $B_n = b_0 + b_1 + \cdots + b_n$ form a bounded sequence.

Then, $S$ converges as well.

**Note.** If $b_n = (-1)^n$, this will give us th alternating series test.

*Proof.* Use $A_k = a_k$ in the summation by parts formula. Look at the partial sums:

$$S_n = \sum_{k=0}^{n} a_k b_k = a_n b_n - \sum_{k=1}^{n} \underbrace{(a_k - a_{k-1})}_{a'_k} B_{k-1}.$$

Both the right hand side sums converge as $n \to \infty$. Prove this using assumption (b) first. Let $C = \sup_k |B_k|$. Then $|a_n B_n| \leq C|a_n| \to 0$ by (a). For the second piece, use monotonicity:

$$\sum_{k=1}^{n} |(a_k - a_{k-1}) B_{k-1}| \leq C \sum_{k=1}^{n} (a_{k-1} - a_k) = C(a_0 - a_n) \leq C a_0,$$

where the equality is because this is a telescoping series. Thus, the series $\sum_{k=1}^{\infty} (a_k - a_{k-1}) B_{k-1}$ converges absolutely; hence it must converge. $\square$

**Note.** Dirichlet's test applies to any monotone sequence; it does not have to be monotonically increasing. This makes sense since we can just multiply signs to flip inequalities as required.

**Note.** Professor said that using convergence tests is mostly a homework activity; proving them, however, might show up on the final.

## 13.6 Absolute convergence vs Conditional convergence

Recall if $\sum_{n=1}^{\infty} |a_n| < +\infty$ (absolute convergence), then $\sum_{n=1}^{\infty} a_n$ converges (and we say it is absolutely convergent.) The converse is not true. Alternating series test shows $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ converges, yet $\sum_{n=1}^{\infty} \left| \frac{(-1)^n}{n} \right|$ famously diverges. Any series where $\sum a_n$ converges but $\sum |a_n| = +\infty$ are called conditionally convergent.

### 13.6.1 Rearrangement

Reordering terms is valid for absolutely convergent series, but strange for conditionally convergent ones; for example, let $S = \sum_{n=0}^{\infty} \frac{(-1)^n}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \ldots$. We build $\tilde{S}$ using the same pieces, but we shuffle the order;

$$
\begin{aligned}
\tilde{S} =& 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \ldots \\
=& \left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} + \frac{1}{10}\right) - \\
=& \frac{1}{2}\left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots\right] \\
=& \frac{S}{2}.
\end{aligned}
$$

This is quite an interesting result; one might even say that the sum is “not abelian” (this means nothing, it is just a group theory trauma-dump joke.)

**Quotes of the day**

"The shape of water – woman falls in love with monster; fine. However, the shape of water is a beautiful phrase, because water has no shape!" - Dr. Loewen, 10/30/2023

"Metric. . . is that Greek for measurement? Might be." - Dr. Loewen, 10/30/2023

"Why are 320 lectures so slow? Well I've been teaching calculus for decade, I've had to slow down for the civilians." Dr. Loewen, 10/30/2023

# 14 Point-set topology

## 14.1 Open sets

**Step 1: Euclidean $\mathbb{R}^k$**

Recall $\mathbb{R}^k = \{(x_1, \ldots, x_k) : x_j \in \mathbb{R}\}$ with $|\underline{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_k^2}$.

**Definition: Open set**

A subset $\mathcal{U} \subseteq \mathbb{R}^k$ is ***open*** iff for all $x \in \mathcal{U}$, there exists $\varepsilon > 0$ such that for all $x'$ with $|x' - x| < \varepsilon$, $x' \in \mathcal{U}$.

The following diagram provides intuition:

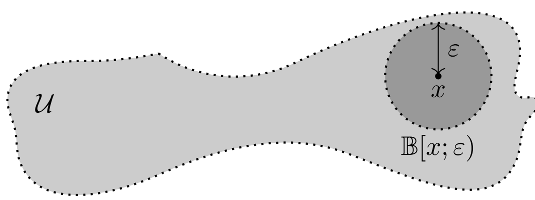



Figure 1: Visualization of an open set.

**Note.** Call $\mathbb{B}[x;\varepsilon)$ with $\varepsilon > 0$ an "open ball with centre $x$ and radius $\varepsilon$." It does indeed deserve to be called open: pick some $y \in \mathbb{B}[x;\varepsilon)$; thus $r = |y - x| < \varepsilon$. Here, $\varepsilon - r > 0$, and $\mathbb{B}[y;\varepsilon - r) \subseteq \mathbb{B}[x;\varepsilon)$. To verify, pick some $z \in \mathbb{B}[y;\varepsilon - r)$. Then

$$\begin{aligned} |z - x| \leq & |z - y| + |y - x| \\ < & (\varepsilon - r) + r = \varepsilon. \end{aligned}$$

The following diagram provides intuition:

**Notation 9.** We define:

(a) An open ball as

$$\mathbb{B}[x;\varepsilon) := \{x' \in \mathbb{R}^k : |x' - x| < \varepsilon\}.$$

(b) A closed ball as

$$\mathbb{B}[x;\varepsilon] := \{x' \in \mathbb{R}^k : |x' - x| \leq \varepsilon\}.$$

(c) A punctured ball as

$$\mathbb{B}(x;\varepsilon) := \{x' \in \mathbb{R}^k : 0 < |x' - x| < \varepsilon\}.$$

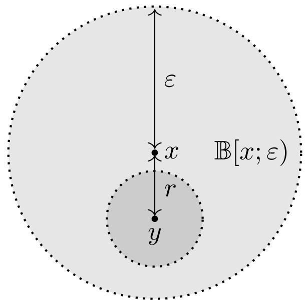



Figure 2: Visualization of an open ball.

**Definition: Topology**

We call $\mathscr{T}$ a ***topology***, where

$$\mathscr{T} := \{\mathcal{U} \subseteq \mathbb{R}^k : \mathcal{U} \text{ is open}\}.$$

Here, $\mathscr{T} \subseteq \mathscr{P}(\mathbb{R}^k)$.

There are a few key properties of $\mathscr{T}$:

(HTS 1) $\emptyset \in \mathscr{T}$ and $\mathbb{R}^k \in \mathscr{T}$.

(HTS 2) For *any* collection $\mathscr{G}$ of open sets,

$$\bigcup \mathscr{G} = \bigcup_{\mathcal{G} \in \mathscr{G}} \mathcal{G}$$

is also open.

(HTS 3) For any *finite* collection of open sets $\mathcal{G}_1, \ldots, \mathcal{G}_N$, the set

$$\bigcap_{k=1}^{N} \mathcal{G}_k = \mathcal{G}_1 \cap \mathcal{G}_2 \cap \cdots \cap \mathcal{G}_N$$

is also open.

*Proof.* Pick $x \in \bigcap_{k=1}^{N} \mathcal{G}_k$. For each $k$, $x \in \mathcal{G}_k \Rightarrow$ there exists $\varepsilon_k > 0$ such that $\mathbb{B}[x;\varepsilon_k) \subseteq \mathcal{G}_k$. Now, set $\varepsilon = \min\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N\}$ to see that $\mathbb{B}[x;\varepsilon) \subseteq \mathcal{G}_k$ for *all* $k$. Therefore, $\mathbb{B}[x;\varepsilon) \subseteq \bigcap_{k=1}^{N} \mathcal{G}_k$ . ☐

(HTS 4) For any $x, y \in \mathbb{R}^k$ – where $x \neq y$ – there exists $\mathcal{U}, \mathcal{V} \in \mathscr{T}$ such that $x \in \mathcal{U}$, $y \in \mathcal{V}$, $\mathcal{U} \cap \mathcal{V} = \emptyset$.

*Proof.* Let $r = |x - y| > 0$ and take $\mathcal{U} = \mathbb{B}[x; r/3)$, $\mathcal{V} = \mathbb{B}[y; r/3)$. Picking some $z \in \mathcal{U} \cap \mathcal{V}$ is an immediate contradiction:

$$r = |x - y| \leq |x - z| + |z - y| < \frac{r}{3} + \frac{r}{3}.$$

The following diagram provides intuition:

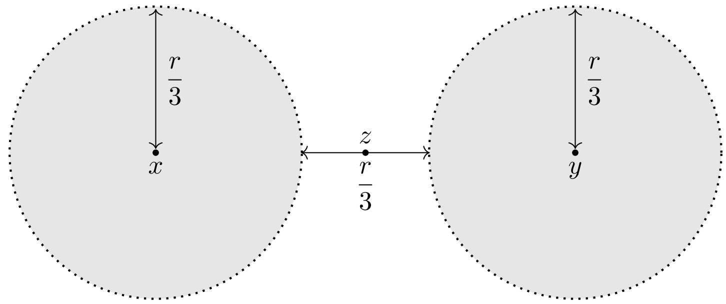



Figure 3: Visualization of HTS 4.

□

**Note.** HTS stands for *Hausdorff Topological Space*. The condition for a topological space for being Hausdorff is just (HTS 4); rest are true for general topological spaces.

**Step 2: Metric spaces**

**Definition: Metric**

Given any set $\mathcal{X} \neq \emptyset$, a function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is called a ***metric*** exactly when

(a) $d(x, y) \geq 0$ for all $x, y \in \mathcal{X}$ with
$$d(x, y) = 0 \iff x = y.$$

(b) $d(x, y) = d(y, x)$ for all $x, y \in \mathcal{X}$ (Symmetry).

(c) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in \mathcal{X}$ (Triangle inequality).

**Example 19.** The following are examples are metrics:

(a) Euclidean $\mathbb{R}^k$ has $d(x, y) = |y - x|$; this is called the "Euclidean metric".

(b) Recall
$$\ell^2 = \left\{ x = (x_1, x_2, \dots) : \sum_{k=1}^{\infty} x_k^2 < \infty \right\}$$
with $\|y - x\| = \left[ \sum_{k=1}^{\infty} (y_k - x_k)^2 \right]^{1/2} = d(x, y)$ (we proved this in homework 7 problem 3.)

(c) For the set $\mathcal{X}$ of bounded functions, $f : [0, 1] \to \mathbb{R}$, $d(x, y) = \sup\{|x(t) - y(t)|\}$ works (we proved this in homework 5 problem 8.)

(d) For any $\mathcal{X} \neq \emptyset$, take
$$d(x, y) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{if } x \neq y \end{cases}.$$
This is called the "discrete metric".

**Reflection.** Why do we care about all this anyways? The main reason – at the moment at least – is to extend the idea of *convergence* to cover many structures at once.

**Quotes of the day**

*After someone sneezed* "I used to teach engineers; when someone sneezed, the whole class would say 'bless you'." - Dr. Loewen, 11/1/2023

"As is the norm in mathematics, I know the name [Hausdorff] and nothing else." - Dr. Loewen, 11/1/2023

**Definition: Metric Space**

A ***metric space*** is a pair $(\mathcal{X}, d)$ combining a non-empty set $\mathcal{X}$ and a function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ obeying the conditions mentioned in the definition of a metric.

Using $d$ for "distance" extends ideas/notations beyond Euclidean case. For example:

$$\mathbb{B}[x; r) = \{x' \in \mathcal{X} : d(x', x) < r\}$$

defines a "ball" with centre $x$ and radius $r > 0$. We declare a set $\mathcal{U} \subseteq \mathcal{X}$ to be *open* exactly when for all $x \in \mathcal{U}$, there exists $\varepsilon > 0$ such that $\mathbb{B}[x; \varepsilon) \subseteq \mathcal{U}$. Let $\mathscr{T}$ denote the set of all open sets in $\mathcal{X}$ that the (metric) topology. As before,

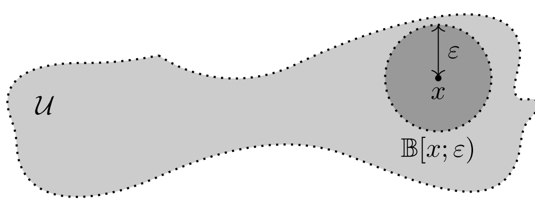



Figure 4: Visualization of an open set.

$\mathbb{B}[x; \varepsilon)$ is an open set for any $x \in \mathcal{X}$, $\varepsilon > 0$, and the set $\mathscr{T}$ has properties (HTS 1) $-$ (HTS 4).

**Reflection.** As before, why do we care about this? Primarily we can think about this as a way to capture convergence of sequence ideas.

**Definition: Convergence**

Given a sequence $x_1, x_2, \ldots$ in $\mathcal{X}$ and a point $\hat{x} \in \mathcal{X}$ to say

$$\lim_{n\to\infty} x_n = \hat{x} \text{ or } x_n \to \hat{x} \text{ as } n \to \infty,$$

we say that for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n > N$, $d(x_n, \hat{x}) < \varepsilon$.

**Proposition**

In a metric space $(\mathcal{X}, d)$, a set $\mathcal{U}$ is open iff every point $x \in \mathcal{U}$ obeys the following: whenever any $x_n$ has $x = \lim_{n\to\infty} x_n$ then $x_n \in \mathcal{U}$ for all $n$ sufficiently large.

*Proof.* ($\Rightarrow$) Given $\mathcal{U} \in \mathscr{T}$ and a point $x \in \mathcal{U}$, definition of open set gives some $\varepsilon > 0$ such that $\mathbb{B}[x; \varepsilon) \subseteq \mathcal{U}$. So, if $(x_n)$ is any sequence with $x_n \to x$, definition of convergence gives $N \in \mathbb{N}$ such that for all $n > N$, $d(x_n, x) < \varepsilon$, where $x_n \in \mathbb{B}[x; \varepsilon) \subseteq \mathcal{U}$.

($\Leftarrow$) We prove this by contrapositive; assume set $\mathcal{U}$ is *not* open. Then some $x \in \mathcal{U}$ breaks the defining property: there exists $\varepsilon > 0$, $\mathbb{B}[x; \varepsilon) \subseteq \mathcal{U}$. So, for each $n \in \mathbb{N}$, $\varepsilon = \dfrac{1}{n}$ here makes $\mathbb{B}\left[x; \dfrac{1}{n}\right) \not\subseteq \mathcal{U}$. That is, some point

$x_n \in \mathbb{B}\left[x; \dfrac{1}{n}\right)$ obeys $x_n \notin \mathcal{U}$. Now $d(x_n, x) < \dfrac{1}{n}$ (an "analogue" of the Squeeze theorem in $\mathbb{R}$), so sequence $x_n$ obeys $x_n \to x$ and $x_n \notin \mathcal{U}$. $\square$

**Note.** This is one of the pros of doing the course not in the same order as Rudin; we get many concrete examples for dealing with convergence, so metric spaces make a lot more sense.

## 14.2 Ball shapes in $\mathbb{R}^2$

Note that $\mathbb{R}^2 = \{x = (x_1, x_2) : x_1, x_2 \in \mathbb{R}\}$ can be given several metrics, some of which are:

- $d_2(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$; we recover the most popular ball $\mathbb{B}[0; 1)$:

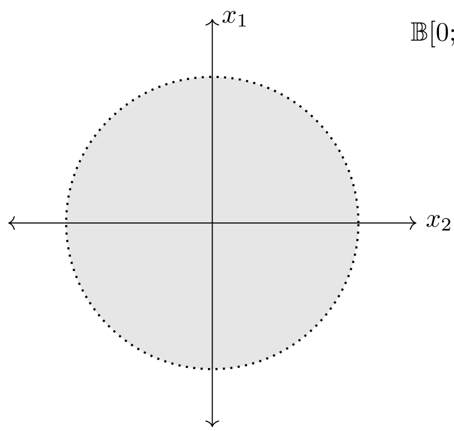



Figure 5: The $\mathbb{B}[0; 1)$ ball; it is actually a ball in this case.

- $d_1(x, y) = |x_1 - y_1| + |x_2 - y_2|$; our ball is now shaped like a diamond:

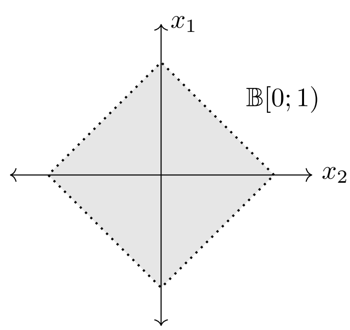



Figure 6: The $\mathbb{B}[0; 1)$ ball; it is a diamond in this case.

- $d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$; our ball is now a square:

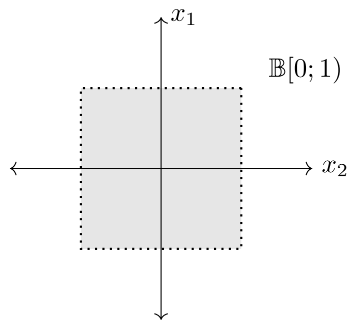



Figure 7: The $\mathbb{B}[0;1)$ ball; it is a square in this case.

- Can interpolate: $d_p(x,y) = (|x_1 - y_1|^p + |x_2 - y_2|^p)^{1/p}$, which is a valid metric for any $p \geq 1$, including $p = +\infty$.

In principle, each different metric gives a different family of open sets, and different story with convergence. However, it is quite bizarre that they all define a compatible idea of convergence: the set of open set are the same regardless of the metric.

## 14.3 Hausdorff topological space

**Definition**

A HTS is an ordered pair $(\mathcal{X}, \mathscr{T})$ where $\mathcal{X}$ is a nonempty set, and $\mathscr{T} \subseteq \mathscr{P}(\mathcal{X})$ with these 4 properties:

(HTS 1) $\emptyset \in \mathscr{T},\ \mathcal{X} \in \mathscr{T}$.

(HTS 2) For any collection $\mathcal{J} \subseteq \mathscr{T}$ one has $\bigcup \mathcal{J} \in \mathscr{T}$ (any union of open sets is open.)

(HTS 3) If $\mathcal{U}_1, \ldots, \mathcal{U}_N \in \mathscr{T}$ (and $N \in \mathbb{N}$) then $\displaystyle\bigcap_{k=1}^{N} \mathcal{U}_k \in \mathscr{T}$ (any finite intersection of open sets in open.)

(HTS 4) Whenever $x \neq y$ in $\mathcal{X}$, there exists $\mathcal{U}, \mathcal{V} \in \mathscr{T}$ obeying $x \in \mathcal{U},\ y \in \mathcal{V}$, and $\mathcal{U} \cap \mathcal{V} = \emptyset$ (enough open sets to separate points.)

**Example 20.**

- Any metric topology.
- Discrete topology: $\mathscr{T} = \mathscr{P}(\mathcal{X})$ (“every set if open”.)

**Quotes of the day**

"*chuckles* I'm sorry, but all of this seems obvious." - Dr. Loewen, 11/3/2023

## 14.4 Neighbourhoods and interiors

**Definition: Neighbourhood**

Given a HTS $(\mathcal{X}, \mathscr{T})$ and an $x \in \mathcal{X}$, a set $\mathcal{S} \subseteq \mathcal{X}$ is a ***neighbourhood*** of $x$ exactly when there exists $\mathcal{U} \subseteq \mathscr{T}$ such that $x \in \mathcal{U} \subseteq \mathcal{S}$. We write $\mathscr{N}(x)$ to be the set of all such $\mathcal{S}$.

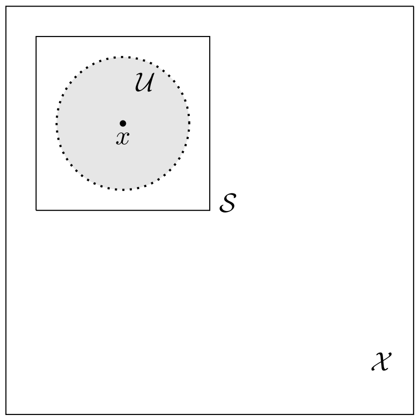



Figure 8: Visualization of the definition.

**Lemma**

In any HTS $(\mathcal{X}, \mathscr{T})$ with $\mathcal{A} \subseteq \mathcal{X}$, the following are equivalent:

(a) $\mathcal{A}$ is open.

(b) For all $x \in \mathcal{A}$, we have $\mathcal{A} \in \mathscr{N}(x)$.

*Proof sketch.* Just pushing around definitions, nothing too complex. □

**Definition: Interior points**

In a HTS $(\mathcal{X}, \mathscr{T})$ with $\mathcal{A} \subseteq \mathcal{X}$, a point $x$ is an ***interior point of*** $\mathcal{A}$ if there exists $\mathcal{U} \in \mathscr{N}(x)$ such that $\mathcal{U} \subseteq \mathcal{A}$.

The collection of interior points is

$$\mathcal{A}^\circ = \text{"the interior of } \mathcal{A}.\text{"}$$

**Note.** If $\mathcal{A} \subseteq \mathcal{B}$, then $\mathcal{A}^\circ \subseteq \mathcal{B}^\circ$.

**Proposition**

In a HTS $(\mathcal{X}, \mathscr{T})$ with $\mathcal{A} \subseteq \mathcal{X}$,

(a) $\mathcal{A}^\circ$ is *open*, with $\mathcal{A}^\circ \subseteq \mathcal{A}$.

(b) If $\mathcal{G}$ is open, and $\mathcal{G} \subseteq \mathcal{A}$, then $\mathcal{G} \subseteq \mathcal{A}^\circ$; $\mathcal{A}^\circ$ is the **largest open subset** of $\mathcal{A}$.

(c) $\mathcal{A}$ is open iff $\mathcal{A} = \mathcal{A}^\circ$.

*Proof.* The proofs for these are very short and given in the notes, but the professor suggested we try to do them ourselves. □

**Note.** The shape of our neighbourhood depends on the topology we lay down on the space; if we look at the real line, the topology on it *does not* allow a point to be its *own* neighbourhood.

**Example 21.** If $a < b$ in $\mathbb{R}$, we have

$$[a, b]^\circ = [a, b)^\circ = (a, b]^\circ = (a, b)^\circ = (a, b).$$

Observe that $\mathbb{Q}^\circ = \emptyset$.

**Note.** The "largest open subset" is as shown above, and the "smallest open subset" does not exist. We illustrate this in $\mathbb{R}$ as follows: Consider $\mathcal{A} = [0, 1]$; If $\mathcal{U}$ is open, and $\mathcal{U} \supset \mathcal{A}$, then $0 \in \mathcal{U}$ and $1 \in \mathcal{U}$ imply that there exists a sufficiently large $n \in \mathbb{N}$ such that $\left(-\frac{1}{n}, \frac{1}{n}\right) = \mathbb{B}\left[0; \frac{1}{n}\right) \subseteq \mathcal{U}$ and $\left(1 - \frac{1}{n}, 1 + \frac{1}{n}\right) = \mathbb{B}\left[1; \frac{1}{n}\right) \subseteq \mathcal{U}$. So, $\mathcal{U} \supseteq \left(-\frac{1}{n}, 1 + \frac{1}{n}\right)$. Increasing $n$ gives us a smaller alternative that is still open and covers $[0, 1]$; no smallest such subset exists.

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right) = [0, 1].$$

## 14.5 Closed sets; Closure

**Definition: Closed set**

In a HTS $(\mathcal{X}, \mathscr{T})$, a set $\mathcal{F} \subseteq \mathcal{X}$ is ***closed*** iff $\mathcal{F}^c = \mathcal{X} \backslash \mathcal{F}$ is open.

Imagine defining $\mathscr{F} := \{\mathcal{U}^c : \mathcal{U} \in \mathcal{J}\}$ is the set of all closed subsets in $\mathcal{X}$. The HTS axioms could be set up starting with closed sets and $\mathscr{F}$ instead of open sets and $\mathscr{T}$, and this would be logically equivalent:

(HTS 1) $\emptyset, \mathcal{X} \in \mathscr{F}$; note that these sets are both closed and open (by definition). These can also be called "clopen" sets.

(HTS 2) Arbitrary intersection of closed sets is closed: if $\mathscr{K} \subseteq \mathscr{F}$ then $\bigcap \mathscr{K} \in \mathscr{F}$, where

$$\begin{aligned} \bigcap \mathscr{K} &= \bigcap_{\mathcal{K} \in \mathscr{K}} \mathcal{K} \\ &= \left( \bigcup_{k \in \mathscr{K}} \mathcal{K}^c \right)^c. \end{aligned}$$

(HTS 3) If $\mathcal{F}_1, \dots, \mathcal{F}_N$ are closed are closed (where $N \in \mathbb{N}$), then $\bigcup_{j=1}^{N} \mathcal{F}_j$ is closed as well, where

$$\left( \bigcup_{j=1}^{N} \mathcal{F}_j \right)^c = \bigcap_{j=1}^{N} (\mathcal{F}_j^c) \in \mathscr{T}.$$

(HTS 4) If $x_1 \neq x_2$ in $\mathcal{X}$, there exist $\mathcal{F}_1, \mathcal{F}_2 \in \mathscr{F}$ such that $x_1 \notin \mathcal{F}_1$, $x_2 \notin \mathcal{F}_2$, $\mathcal{F}_1 \cup \mathcal{F}_2 = \mathcal{X}$.

**Note.** As HTS 1 already suggests, if a set is not open, it ***does not*** mean it is closed; it could be clopen, which is purely a definition. This is also why I think the names are not good because "sets are not doors" - Dr. Jim Bryan, 2023.

**Definition: Closure**

In a HTS $(\mathcal{X}, \mathscr{T})$ with $\mathcal{A} \subseteq \mathcal{X}$, the ***closure*** of $\mathcal{A}$ is the set

$$\overline{\mathcal{A}} := \left((\mathcal{A}^c)^\circ\right)^c,$$

which is the complement of an open set $\left((\mathcal{A}^c)^\circ\right)$, so it is closed:

$$\begin{aligned} &(\mathcal{A}^c)^\circ \subseteq \mathcal{A}^c \\ \Rightarrow &\left((\mathcal{A}^c)^\circ\right)^c \supseteq (\mathcal{A}^c)^c = \mathcal{A} \\ \Rightarrow &\overline{\mathcal{A}} \supseteq \mathcal{A}. \end{aligned}$$

Indeed $\overline{\mathcal{A}}$ is the smallest closed superset of $\mathcal{A}$; if $\mathcal{F}$ is closed, and $\mathcal{F} \supseteq \mathcal{A}$, then $\mathcal{F} \supseteq \overline{\mathcal{A}}$ as well.

**Note.** If $\mathcal{A} \subseteq \mathcal{B}$, then $\overline{\mathcal{A}} \subseteq \overline{\mathcal{B}}$. Example: $\overline{(ab)} = [a, b]$.

## 14.6 Boundary points

**Definition: Boundary**

In a HTS $(\mathcal{X}, \mathscr{T})$ with $\mathcal{A} \subseteq \mathcal{X}$, a point $x \in \mathcal{X}$ belongs to $\partial\mathcal{A}$, the ***boundary*** of $\mathcal{A}$, iff for all $\mathcal{G} \in \mathscr{N}(x)$, we have $\mathcal{G} \cap \mathcal{A} \neq \emptyset$ and $\mathcal{G} \cap \mathcal{A}^c \neq \emptyset$.

**Example 22.** Consider:

(a) $\partial(a, b) = \{a, b\}$.

(b) $\partial\mathbb{Q} = \mathbb{R}$ (very interesting).

## 14.7 Limit points

**Definition: Limit points**

Given a HTS $(\mathcal{X}, \mathscr{T})$ with set $\mathcal{A} \subseteq \mathcal{X}$, a point $z \in \mathcal{X}$ is a ***limit point*** for $\mathcal{A}$ iff

$$\text{for all } \mathcal{U} \in \mathscr{N}(z), (\mathcal{U} \backslash \{z\}) \cap \mathcal{A} \neq \emptyset.$$

The set of all such $z$ is denoted by $\mathcal{A}'$.

**Notation 10.** Some synonyms for "limit point" are: cluster point, accumulation point, etc.

**Lemma**

In a metric space $(\mathcal{X}, d)$ with $\mathcal{A} \subseteq \mathcal{X}$, the following are equivalent:

(a) $x \in \mathcal{A}'$.

(b) $x = \lim_{n\to\infty} x_n$ for some sequence $(x_n)$ of *distinct points* all in $\mathcal{A}$.

*Proof.* $(a \Rightarrow b)$: We build a sequence like in (b): pick $x_1 \in \mathbb{B}(x;1) \cap \mathcal{A}$. Pick $x_2 \in \mathbb{B}\left(x; \min\left\{\frac{1}{2}, d(x_1,x)\right\}\right) \cap \mathcal{A}$ and then $x_3 \in \mathbb{B}\left(x; \min\left\{\frac{1}{3}, d(x_2,x)\right\}\right) \cap \mathcal{A}$, and continue like this, so we get a sequence $(x_n)$ such that all distinct $x_n \in \mathcal{A}$ for all $n$, and $d(x,x_n) < \frac{1}{n} \Rightarrow x_n \to x$.

**Note.** Imagine $\mathcal{A} = (0,1]$ and we want to show $0 \in \mathcal{A}'$; choosing $x_n = \frac{1}{2} + \frac{1}{2n}$ gives a decreasing $x_n$, but $x_n \to \frac{1}{2}$, not $0$.

$(b \Rightarrow a)$ We assume (b), and let $\mathcal{U} \in \mathscr{N}(x)$. By definition of $\mathscr{N}(x)$, there exists $\varepsilon > 0$ such that $\mathbb{B}[x;\varepsilon) \subseteq \mathcal{U}$. Use the fact that "$x_n \to x$" to get $N \in \mathbb{N}$ such that for all $n > N$ we have $\underbrace{d(x_n,x) < \varepsilon}_{x_n \in \mathbb{B}[x;\varepsilon) \subseteq \mathcal{U}}$. So we get many of these (all different, since $x_n \not\to x$ for all $(x_n)$) $x_n \in (\mathcal{U} \backslash \{x\}) \cap \mathcal{A} \neq \emptyset$, as required. ☐

The following are some facts, the proofs for which are in the canvas notes:

(i) If $\mathcal{A} \subseteq \mathcal{B}$, then $\mathcal{A}' \subseteq \mathcal{B}'$.

(ii) $z \notin \mathcal{A}' \iff$ there exists $\mathcal{U} \in \mathscr{N}(x)$ such that $(\mathcal{U} \backslash \{z\}) \cap \mathcal{A} = \emptyset$.

(iii) $\mathcal{G} \subseteq \mathcal{X}$ is open $\iff (\mathcal{G}^c)' \subseteq \mathcal{G}^c$.

(iv) $\mathcal{F} \subseteq \mathcal{X}$ is closed $\iff \mathcal{F}' \subseteq \mathcal{F}$.

(v) For any $\mathcal{A} \subseteq \mathcal{X}$, set $\mathcal{A}'$ is closed.

(vi) For any $\mathcal{A} \subseteq \mathcal{X}$, $\overline{\mathcal{A}} = \mathcal{A} \cup \mathcal{A}'$.

**Definition: Isolated point**

For $\mathcal{A} \subseteq \mathcal{X}$ in a HTS, the points of $\mathcal{A} \backslash \mathcal{A}'$ are called ***isolated***.

**Example 23.** In $\mathbb{R}$, $(0,1)' = [0,1]$, $(\mathbb{Q} \cap (0,1))' = [0,1]$, and $\mathcal{A} = [\mathbb{Q} \cap (-\infty,0)] \cup \mathbb{Z}$ such that $\mathcal{A}' = (-\infty,0]$ and the isolated points are $\mathcal{A} \backslash \mathcal{A}' = \mathbb{N}$.

## 14.8 Subspaces

For any metric space $(\mathcal{X},d)$, the same $d$ works as a metric in any subset $\mathcal{Y} \subseteq \mathcal{X}$. So $(\mathcal{Y},d)$ is a metric space too. Topology in $\mathcal{Y}$ will have sets "open in $\mathcal{Y}$" that are subsets of $\mathcal{X}$ but may *fail* to be "open in $\mathcal{X}$".

# 15 Compactness, Convergence, and Completeness

## 15.1 Compactness

**Definition: Compact**

Given a HTS $(\mathcal{X}, \mathscr{T})$, let $\mathcal{K} \subseteq \mathcal{X}$. We say that $\mathcal{K}$ is ***compact*** means for *every* collection $\mathscr{G}$ of open sets with $\mathcal{K} \subseteq \bigcup \mathscr{G}$ there exists $N \in \mathbb{N}$ and $\mathcal{G}_1, \ldots, \mathcal{G}_N \in \mathscr{G}$ satisfying

$$\mathcal{K} \subseteq \mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_N.$$

**Note.** Every open cover for $\mathcal{K}$ has a finite subcover.

**Corollary 4.** Any finite set is compact.

*Proof.* Let $S = \{x_1, \ldots, x_N\}$ be a finite set. Given any $\mathscr{G} \subseteq \mathscr{T}$ with $\mathcal{S} \subseteq \bigcup \mathscr{G}$; for each $k =, \ldots, N$, pick some $\mathcal{G}_k \in \mathscr{G}$. Then, $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_N$ obeys $\mathcal{S} \subseteq \mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_N$. □

### Failure modes

A set $\mathcal{K}$ is compact iff every open cover has a finite subcover. Thus, a set $\mathcal{S}$ fails to be compact iff some *open* cover has *no* finite subcover.

**Lemma**

In $(\mathbb{R}, |\cdot|)$, the set $\mathbb{Z}$ is not compact.

*Proof.* Let $\mathscr{G} = \{(n-1, n+1) : n \in \mathbb{Z}\}$. Clearly, each element $\mathscr{G}$ is an open set, and $\mathbb{Z} \subseteq \bigcup \mathscr{G}$. However, any finite subset $\mathcal{G}_1, \ldots, \mathcal{G}_N$ of $\mathscr{G}$ will cover only finite subsets of $\mathbb{Z}$, so $\mathbb{Z} \subseteq \mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_N$.

Another open cover could be $\mathcal{G} = \{(-n, n) : n \in \mathbb{N}\}$. □

**Definition: Bounded**

In a metric space $(\mathcal{X}, d)$, we say that a set $\mathcal{A} \subseteq \mathcal{X}$ is ***bounded*** exactly when there exists $x \in \mathcal{X}$ and $R > 0$ such that $A \subseteq \mathbb{B}[x; R)$.

**Proposition**

In any metric space $(\mathcal{X}, d)$, every compact set is bounded.

*Proof.* Let $\mathcal{K} \subseteq \mathcal{X}$ be compact. Pick any $x \in \mathcal{X}$ and let $\mathscr{G} = \{\mathbb{B}[x; n) : n \in \mathbb{N}\}$. Then, $\bigcup \mathscr{G} \supseteq \mathcal{X} \supseteq \mathcal{K}$, so $\mathscr{G}$ is an open cover for $\mathcal{K}$. Hence, it must have a finite subcover $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_N$ with each $\mathcal{G}_k = \mathbb{B}[x; n_k)$. Let $R = \max\{n_1, n_2, \ldots, n_N\}$ to get $\mathbb{B}[x; R) = \mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_N \supseteq \mathcal{K}$. □

**Lemma**

In $\mathbb{R}$, let $\mathcal{S} = \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$.

(a) Set $\mathcal{S}$ is *not* compact.

(b) Set $\overline{\mathcal{S}} = \mathcal{S} \cup \{0\}$ is compact.

(a) *Proof.* For each $n$, note that $\frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)} > \frac{1}{(n+1)^2}$.

Let $\mathcal{G}_n = \mathbb{B}\left[\frac{1}{n}; \frac{1}{(n+1)^2}\right)$ to get an open interval with $\mathcal{G}_n \cap \mathcal{S} = \left\{\frac{1}{n}\right\}$. Use $\mathscr{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ as an open cover for $\mathcal{S}$. No finite subcover can include all points of $\mathcal{S}$ since each $\mathcal{G}_k$ only holds one point of $\mathcal{S}$. $\square$

(b) *Proof.* Let $\mathscr{G}$ be any open cover for $\overline{S}$. Thus, there must be some open $\mathcal{G}_0 \in \mathscr{G}$ with $0 \in \mathcal{G}_0$. Being open, $\mathcal{G}_0$ must contain $\mathbb{B}[0;\varepsilon)$ for some $\varepsilon > 0$. Pick any integer $N > \frac{1}{\varepsilon}$. Then, $\frac{1}{n} < \varepsilon$ for all $n > N$, so all these points lie in $\mathcal{G}_0$. For indices $1, 2, \dots, N$, pick $\mathcal{G}_k \in \mathscr{G}$ such that $\frac{1}{k} \in \mathcal{G}_k$. Hence, we conclude that

$$\begin{aligned}\overline{S} = &\left\{1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{N}\right\} \cup \left\{\frac{1}{N+1}, \frac{1}{N+2}, \dots\right\} \\ \subseteq &(\mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_N) \cup \mathcal{G}_0,\end{aligned}$$

which is a finite subcover. $\square$

**Proposition**

In any HTS $(\mathcal{X}, \mathscr{T})$, every compact set is closed.

*Proof.* Let $\mathcal{K} \subseteq \mathcal{X}$ be compact. We will show that $\mathcal{K}^c$ is open. Pick any $z \in \mathcal{K}^c$. Now, for each $x \in \mathcal{K}$, HTS 4 implies that there exists $\mathcal{U}_x, \mathcal{V}_x \in \mathscr{T}$ with $x \in \mathcal{U}_x$, $z \in \mathcal{V}_x$ such that $\mathcal{U}_x \cap \mathcal{V}_x = \emptyset$.

Now, let $\mathscr{G} = \{\mathcal{U}_x : x \in \mathcal{K}\}$ is clearly an open cover for $\mathcal{K}$, so by compactness, it must have a finite subcover:

$$\mathcal{K} \subseteq \mathcal{U}_{x_1} \cup \mathcal{U}_{x_2} \cup \cdots \cup \mathcal{U}_{x_N}$$

for some points $x_1, \dots, x_N \in \mathcal{K}$. Thus,

$$\begin{aligned}\mathcal{K}^c \supseteq &\mathcal{U}_{x_1}^c \cap \mathcal{U}_{x_2}^c \cap \cdots \cap \mathcal{U}_{x_N}^c \\ \supseteq &\mathcal{V}_{x_1} \cap \mathcal{V}_{x_2} \cap \cdots \cap \mathcal{V}_{x_N} \supseteq \{z\}.\end{aligned}$$

Therefore, since $\mathcal{V}_{x_1} \cap \mathcal{V}_{x_2} \cap \cdots \cap \mathcal{V}_{x_N}$ is open (HTS 3), we conclude that $z \in (\mathcal{K}^c)^\circ$. $\square$

### 15.1.1 Ultimate end-goal:

In a metric space $(\mathcal{X}, d)$

$[\mathcal{K}$ is compact$] \iff [\mathcal{K}$ is closed$]$ and $[\mathcal{K}$ is bounded$]$ and $[??]$ (where this depends on what $(\mathcal{X}, d)$ we study.)

**Note.** In $\ell^2$, $\mathcal{S} = \{\hat{e}_p = \underbrace{(0, 0, \dots, 1, 0, \dots, 0)}_{1 \text{ at } p^{\text{th}} \text{ position}} : p \in \mathbb{N}\}$ is closed and bounded and ***not compact***; we need more conditions for compactness.

**Proposition**

In a HTS $(\mathcal{X}, \mathscr{T})$, if $\mathcal{K}$ is compact, every closed subset of $\mathcal{K}$ is compact.

*Proof.* Proof in canvas notes. $\square$

**Corollary 5.** In any HTS, if $\mathcal{K}$ is compact, and $\mathcal{F}$ is closed, then $\mathcal{K} \cap \mathcal{F}$ is compact.

*Proof.* $\mathcal{K} \cap \mathcal{F}$ is closed, and $\mathcal{K} \cap \mathcal{F} \subseteq \mathcal{K}$. $\square$

**Corollary 6.** Let $\mathcal{K}$ be compact in some HTS, any infinite set $\mathcal{A} \subseteq \mathcal{K}$ must have $\mathcal{A}' \neq \emptyset$.

*Proof.* By prove this by contrapositive. Suppose $\mathcal{S} \subseteq \mathcal{K}$ has $\mathcal{S}' = \emptyset$. For each $x \in \mathcal{S}$, $x \notin \mathcal{S}'$, which implies some open $\mathcal{G}_x$ obeys $\mathcal{G}_x \cap \mathcal{S} = \{x\}$. Thus, $\mathscr{G} = \{\mathcal{G}_x : x \in \mathcal{S}\}$ is an open cover for $\mathcal{S}$. Observe that $\overline{\mathcal{S}} = \mathcal{S} \cup \mathcal{S}' = \mathcal{S}$ is closed, so it is compact; hence $\mathscr{G}$ has a finite subcover $\mathcal{G}_{x_1}, \mathcal{G}_{x_2}, \dots, \mathcal{G}_{x_N}$, i.e., $\mathcal{S} \subseteq \mathcal{G}_{x_1} \cup \mathcal{G}_{x_2} \cup \cdots \cup \mathcal{G}_{x_N}$. Therefore, by construction, $\mathcal{S} = \{x_1, x_2, \dots, x_N\}$ is finite. $\square$

### 15.1.2 Complementary view of compactness

**Definition: Finite intersection property**

A family of sets $\mathscr{F}$ has the ***finite intersection property*** (F.I.P.) if every finite choice of $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_N \in \mathscr{F}$ gives $\displaystyle\bigcap_{k=1}^{N} \mathcal{F}_k \neq \emptyset$.

**Theorem**

In any HTS $(\mathcal{X}, \mathscr{T})$ with subset $\mathcal{K} \subseteq \mathcal{X}$, assume $\mathcal{K}$ is closed. Then, the following are equivalent:

(a) $\mathcal{K}$ is compact.

(b) Every family $\mathscr{F}$ of *closed* subsets of $\mathcal{K}$ with F.I.P. has $\bigcap \mathscr{F} \neq \emptyset$.

*Proof sketch.* Left as practice (advised to do proof by contrapositive). □

## 15.2 Convergence

**Theorem**

In a metric space $(\mathcal{X}, d)$ with $\mathcal{K} \subseteq \mathcal{X}$, the following are equivalent:

(a) $\mathcal{K}$ is compact.

(b) Every sequence $(x_n)$ in $\mathcal{K}$ has a subsequence that converges to a point in $\mathcal{K}$.

*Proof.* $(a \Rightarrow b)$ Let $(x_n)$ be a sequence in $\mathcal{K}$. Consider $\mathcal{A} = \{x_n : n \in \mathbb{N}\}$ be the range of that sequence. If $\mathcal{A}$ is finite, a constant subsequence exists (some point of $\mathcal{A}$ is "hit" by infinitely many $x_n$). Otherwise, $\mathcal{A}' \neq \emptyset$; any $x \in \mathcal{A}'$ will have $\mathcal{A} \cap \mathbb{B}\left(x; \dfrac{1}{n}\right) \neq \emptyset$. Standard methods will give subsequence of $(x_n)$ converging to $x$. Furthermore, $x \in \mathcal{A}' \subseteq \mathcal{K}$ because $\mathcal{K}$ is closed.

$(b \Rightarrow a)$ Let $\mathcal{K}$ have property in (b). Given arbitrary open cover $\mathscr{G}$ for $\mathcal{K}$, for each $x \in \mathcal{K}$, some $\mathcal{G} \subseteq \mathscr{G}$ obeys $x \in \mathcal{G}$. Consider

$$R(x) = \begin{cases} \sup\{\varepsilon > 0 : \mathbb{B}[x; \varepsilon) \subseteq \mathcal{G}\}, \text{ for some } \mathcal{G} \in \mathscr{G}, & \text{if the RHS is not } +\infty. \\ 1, & \text{otherwise.} \end{cases}$$

Then, let $r(x) = \dfrac{1}{2} R(x)$; for all $x \in \mathcal{K}$, $\mathbb{B}[x; r(x)) \subseteq \mathcal{G}$ holds for some $\mathcal{G} \in \mathscr{G}$.

$$\begin{aligned} &\text{Pick any } x_1 \in \mathcal{K};\ \text{write } r_1 = r(x_1). \\ &\text{Pick any } x_2 \in \mathcal{K} \backslash \mathbb{B}[x_1; r_1) \text{ write } r_2 = r(x_2). \\ &\text{Pick any } x_3 \in \mathcal{K} \backslash \mathbb{B}[x_1; r_1) \cup \mathbb{B}[x_2; r_2));\ \text{write } r_3 = r(x_3) \\ &\vdots \end{aligned}$$

Expect a sequence of $x_1, x_2, \dots$ with corresponding $r_1, r_2, r_3, \dots$ such that if $q > p$, $x_q \notin \mathbb{B}[x_p; r_p)$, i.e., $d(x_q, x_p) \geq r_p$.

**Claim 4.** This construction *cannot* run forever.

*Proof.* For the sake of contradiction, suppose this does work and produce a sequence $(x_n)$ in $\mathcal{K}$. Use (b) to get a subsequence $(x_{n_k})$ and $\hat{x} \in \mathcal{K}$ such that $x_{n_k} \to \hat{x}$ as $k \to \infty$. Note that $(p = n_k,\ q = n_{k+1}$ above)

$$\begin{aligned} r_{n_k} \leq & d(x_{n_{k+1}}, x_{n_k}) \\ \leq & d(x_{n_{k+1}}, \hat{x}) + d(\hat{x}, x_{n_k}) \to 0 + 0 \text{ as } k \to \infty, \end{aligned}$$

so $r_{n_k} \to 0$. Now, $\hat{x} \in \mathcal{K}$, so $r(\hat{x})$ is defined and $\mathbb{B}[\hat{x}; r(\hat{x})) \subseteq \hat{\mathcal{G}}$ for some $\hat{\mathcal{G}} \subseteq \mathscr{G}$. Use $x_{n_k} \to \hat{x}$ to say that for all sufficiently large $k$, $d(x_{n_k}, \hat{x}) < \dfrac{1}{2} r(\hat{x})$. So,

$$\mathbb{B}\left[x_{n_k}; \frac{r(\hat{x})}{2}\right) \subseteq \mathbb{B}\left[\hat{x}; r(\hat{x})\right) \subseteq \hat{\mathcal{G}},$$

and hence $R(x_{n_k}) \geq \dfrac{1}{2} r(\hat{x})$, so

$$r_{n_k} = \frac{1}{2} R(x_{n_k}) \geq \frac{1}{4} r(\hat{x}),$$

which is a contradiction. So construction must fail at some stage. Therefore,

$$\mathcal{K} \setminus \left(\mathbb{B}[x_1; r_1) \cup \mathbb{B}[x_2; r_2) \cup \cdots \cup \mathbb{B}[x_M; r_M)\right) = \emptyset,$$

or

$$\mathcal{K} \subseteq \mathbb{B}[x_1; r_1) \cup \cdots \cup \mathbb{B}[x_M; r_M).$$

Each of these balls fits inside some corresponding $\mathcal{G}_k$ from $\mathscr{G}$; finite subcover has been found. $\square$

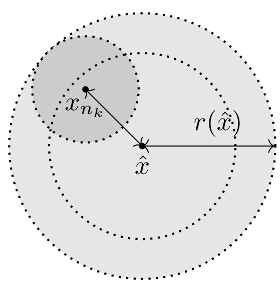



Figure 9: Visualization of the construction in the proof.

$\square$

**Theorem: Heine-Borel Theorem**

In $\mathbb{R}^k$ (with the usual topology), for a subset $\mathcal{K} \subseteq \mathbb{R}^k$, the following are equivalent:

(a) $\mathcal{K}$ is compact.

(b) $\mathcal{K}$ is closed and bounded.

*Proof.* (a⇒b) This is a general fact about metric spaces.

(b⇒a) We know that compactness $\iff$ sequential compactness in a metric space, so we start with a list of sequences $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)})$ in $\mathcal{K}$. The sequence in the first component is $(x_1^{(n)})$ is a bounded sequence in $\mathbb{R}$, so by Bolzano-Weierstrass theorem, there exists a convergent sub-sequence; let this sub-sequence converge to $\hat{x}_1$. Similarly, the $i^{\text{th}}$ component is also bounded, so there exists a sub-sequence converging to $\hat{x}_i$. After repeating this process of extracting and nesting sub-sequences $k$ times, we get that $x^{(n_j)} \to (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_k)$ as $j \to \infty$. The list of limits of sequences is in $\mathcal{K}$ because $\mathcal{K}$ is closed, and hence contains all of its limit points. Thus, $\mathcal{K}$ is sequentially compact $\iff$ $\mathcal{K}$ is compact. $\square$

## Things to be cautious about

1. In any HTS $(\mathcal{X}, \mathscr{T})$, set $\mathcal{X}$ is closed. What about the HTS where $\mathcal{X} = (0,1)$ in $\mathbb{R}$ and $\mathscr{T}$ comes from the usual metric on $\mathbb{R}$. This is a metric space, so $\mathcal{X}$ is closed *in itself, but **not** in* $\mathbb{R}$; it is bounded and a subset of $\mathbb{R}$, yet ***not compact***. We require closed in $\mathbb{R}^k$ to make this work.

2. Convergence of sequences can be defined in a general HTS, but in that context compactness is not equivalent to sequential compactness.

## 15.3 Completeness

**Definition: Cauchy sequence and Complete space**

Let $(\mathcal{X}, d)$ be a metric space.

(a) A sequence $(x_n)$ in $\mathcal{X}$ is ***Cauchy*** iff for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n > N$, $d(x_m, x_n) < \varepsilon$.

(b) The space $(\mathcal{X}, d)$ is called ***complete*** exactly when every Cauchy sequence in $\mathcal{X}$ converges in $\mathcal{X}$.

**Example 24.** $(\mathbb{R}, |\cdot|)$ is complete, but $(\mathbb{Q}, |\cdot|)$ is not complete.

Some facts about Cauchy sequences:

(i) Every convergent sequence is Cauchy.

(ii) Every Cauchy sequence is bounded.

(iii) If a Cauchy sequence has a subsequence that converges, then the full original sequence converges to the same limit.

**Note.** We have already done all this before in $\mathbb{R}$; the proofs for these are very similar in general metric spaces.

**Theorem: Rudin 3.11 (b)**

Every compact metric space is complete.

*Proof sketch.* Compact $\iff$ sequentially compact; we start with a Cauchy sequence, get a convergent subsequence (since the space is compact), we use property (iii) to conclude the proof. $\square$

**Example 25.** Consider the sequence space

$$\ell^\infty = \left\{ x = (x_1, x_2, x_3, \dots) : \sup_j |x_j| < +\infty \right\}$$

Let $d(x,y) = \sup_j |x_j - y_j|$, which we have shown is a metric in Homework 5 problem 8. This $(\ell^\infty, d)$ is complete.

*Proof.* Let $(x^{(n)})_{n \in \mathbb{N}}$ be a Cauchy sequence in $\ell^\infty$. We must show it converges. Convergence in $\ell^\infty$ is:

$$\text{For all } \varepsilon > 0, \text{ there exists } N \in \mathbb{N} \text{ such that for all } m, n > N,\ \sup_j \left| x_j^{(m)} - x_j^{(n)} \right| < \varepsilon.$$

For each particular $j \in \mathbb{N}$, this implies the Cauchy property for the real-values sequence in slot $j$, i.e., $(x_j^{(n)})_{n \in \mathbb{N}}$. Hence, by completeness of $\mathbb{R}$, $x_j = \lim_{n \to \infty} x_j^{(n)}$ exists in $\mathbb{R}$. Repeat for each $j$ to get $\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \dots)$

**Claim 5.** $\hat{x} \in \ell^\infty$.

*Proof.* We use Cauchy property (ii): the original sequence $(x^{(n)})$ must be bounded, so some $M > 0$ obeys $x^{(n)} \in \mathbb{B}[0; M]$, where $0 = (0, 0, 0, \dots) \in \ell^\infty$, i.e., $\sup_j \left| x_j^{(n)} \right| \leq M$ for each $n$. We fix $j$ and let $n \to \infty$ to get $|\hat{x}_j| \leq M$. Hence, we see that indeed $\hat{x} \in \mathbb{B}[0; M] \subseteq \ell^\infty$. □

**Claim 6.** $d(x^{(n)}, \hat{x}) \to 0$.

*Proof.* We require:

$$\text{For all } \varepsilon > 0, \text{ there exists } N \in \mathbb{N} \text{ such that for all } n > N, \ \sup_j \left| x_j^{(n)} - \hat{x}_j \right| < \varepsilon.$$

Fix some $\varepsilon > 0$, and use $\varepsilon' = \dfrac{\varepsilon}{2}$ in the original Cauchy property. Estimate

$$\begin{aligned} \left| x_j^{(n)} - \hat{x}_j \right| &\leq \left| x_j^{(n)} - x_j^{(m)} \right| + \left| x_j^{(m)} - \hat{x}_j \right| \\ &< \varepsilon' + \left| x_j^{(m)} - \hat{x}_j \right| \end{aligned}$$

for all $m, n > N'$, where $N'$ comes from $\varepsilon'$ via the Cauchy property shown above. Now, let $m \to \infty$ on both sides:

$$\left| x_j^{(n)} - \hat{x}_j \right| \leq \frac{\varepsilon}{2} < \varepsilon;$$

this holds for all $n > N'$. Next, taking $\sup_j$ on both sides, we get $d(x^{(n)}, \hat{x}) < \varepsilon$, as required. □

□

**Definition: Diameter**

We define the ***diameter*** of the set as:

$$\text{diam}(\mathcal{S}) = \sup\{d(x, y) : x, y \in \mathcal{S}\}$$

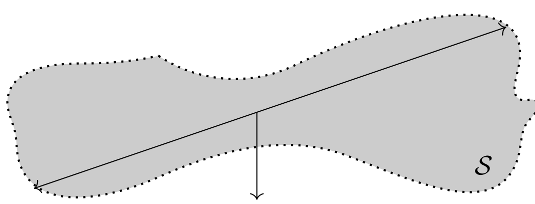



Figure 10: Visualization of the diameter.

**Theorem: Cantor's intersection theorem**

Let $(\mathcal{X}, d)$ be a metric space; the following are equivalent:

(a) $(\mathcal{X}, d)$ is complete – every Cauchy sequence converges in $\mathcal{X}$.

(b) For every sequence of nested, closed, and non-empty sets $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \mathcal{F}_3 \supseteq \dots$ in $\mathcal{X}$ having $\text{diam}(\mathcal{F}_n) \to 0$ as $n \to \infty$; the set $\mathcal{F} = \bigcap_{n=1}^{\infty} \mathcal{F}_n$ contains exactly one point.

*Proof.* (a⇒b) Givens sets $(\mathcal{F}_n)$ as in setup (b). Pick any $x_n \in \mathcal{F}_n$ for each $n$, this defining a sequence $(x_n)$; this sequence is Cauchy: let $\varepsilon > 0$ be given. We use the fact that $\operatorname{diam}(\mathcal{F}_n) \to 0$ to get $N \in \mathbb{N}$ such that $\operatorname{diam}(\mathcal{F}_n) < \varepsilon$ for all $n > N$. In our sequence, if $m, n > N$, then both $x_m, x_n \in \mathcal{F}_{\min\{m,n\}} \subseteq \mathcal{F}_{N+1}$; hence $d(x_m, x_n) < \varepsilon$. By completeness, some $\hat{x} \in \mathcal{X}$ satisfies $\hat{x} = \lim_{n\to\infty} x_n$. Note that:

(i) $\hat{x} \in \mathcal{F}$, because for each $n$, *closed set* $\mathcal{F}_n$ contains each point $x_{n+p}$ for $p \in \mathbb{N}$, so $\hat{x} = \lim_{p\to\infty} x_{n+p}$ lies in $\mathcal{F}_n$.

(ii) If $y \neq \hat{x}$, then $y \notin \mathcal{F}$, because if $y \neq \hat{x}$, then $d(y, \hat{x}) > 0$; however, $\mathcal{F} \subseteq \mathcal{F}_n \subseteq \mathbb{B}[\hat{x}; \operatorname{diam}(\mathcal{F}_n))$. Hence, $y \in \mathcal{F}$ is excluded by the fact that $\operatorname{diam}(\mathcal{F}) \to 0$.

This tells us that $\mathcal{F} = \hat{x}$.

(b⇒a) Given a Cauchy sequence $(x_n)$ in $\mathcal{X}$, let

$$\mathcal{F}_n = \overline{\{x_n, x_{n+1}, x_{n+2}, \dots\}}.$$

Now, each $\mathcal{F}_n$ is closed, non-empty, and $\mathcal{F}_n \supseteq \mathcal{F}_{n+1}$. Furthermore, $\operatorname{diam}(\mathcal{F}_n) \to 0$ as $n \to \infty$, since $(x_n)$ is Cauchy: let $\varepsilon > 0$ be given; we get $N \in \mathbb{N}$ such that for all $m, n > N$, we have $d(x_m, x_n) < \varepsilon$. This makes $\operatorname{diam}(\mathcal{F}_n) \leq \varepsilon$ for all $n \in \mathbb{N}$, so using (b), let $\{\hat{x}\} = \bigcap_{n\in\mathbb{N}} \mathcal{F}_n$. We show that $x_n \to \hat{x}$: given any $\varepsilon > 0$, we pick a sufficiently large $N$ such that $\operatorname{diam}(\mathcal{F}_n) < \varepsilon$ for all $n > N$; then $x_n, \hat{x} \in \mathcal{F}_n$ which tells us that $d(x_n, \hat{x}) < \varepsilon$. □

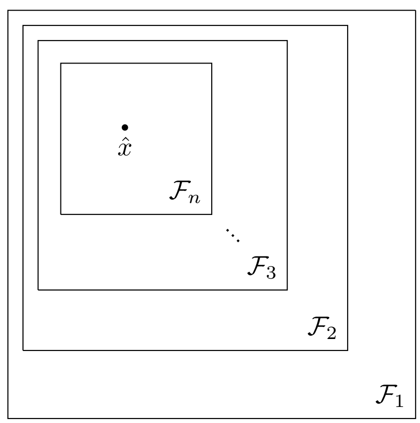



Figure 11: Visualization of the theorem.

## 15.4 Completing a metric space

Let $(\mathcal{X}, d)$ be a metric space. We can construct a ***complete*** metric space $(\hat{\mathcal{X}}, D)$ such hat $\mathcal{X}$ is dense in $(\hat{\mathcal{X}}, D)$, and $D(x, y) = d(x, y)$ for all $x, y \in \mathcal{X}$.

**Note** (Analogy)**.** We built $\mathbb{R}$ from $\mathbb{Q}$ by exactly these methods ($\mathbb{R} = \hat{\mathbb{Q}}$).

**Note** ("Weasel words")**.** $(\hat{\mathcal{X}}, D)$ actually contains a "working copy" of $(\mathcal{X}, d)$... not the exact points.

### Outline of the process

Let $\operatorname{CS}(\mathcal{X})$ be the set of all Cauchy sequences $a = (a_1, a_2, \dots)$ with elements in $\mathcal{X}$; we will call them "vectors". Elements of $\hat{\mathcal{X}}$ will be sets like this:

$$P[a] = \left\{ b \in \operatorname{CS}(\mathcal{X}) : \lim_{n\to\infty} d(a_n, b_n) = 0 \right\}.$$

Observe:

(i) Every $a \in \mathrm{CS}(\mathcal{X})$ lies in $P[a]$.

(ii) For given $a, b \in \mathrm{CS}(\mathcal{X})$, the sets $P[a], P[b]$ are either disjoint or equal.

(iii) Different "representatives" $a, b \in \mathrm{CS}(\mathcal{X})$ can give the same $P[a] = P[b]$ in $\hat{X}$.

**Example 26.** If $\mathcal{X} = \mathbb{Q}$, observe that $a = (0, 0, 0, \dots)$ and $b = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right)$ have $P[a] = P[b]$.

The metric in $\hat{\mathcal{X}}$ is defined as: if $P[a], P[b] \in \hat{\mathcal{X}}$, let

$$D\left(P[a], P[b]\right) = \lim_{n\to\infty} d(a_n, b_n).$$

Things to check:

(i) This limit actually exists:

$$\text{Show } \delta_n = d(a_n, b_n) \text{ is Cauchy in } \mathbb{R}.$$

(ii) Different representatives $a' \in P[a]$, $b' \in P[b]$ give the same $\lim_{n\to\infty} d(a'_n, b'_n)$.

(iii) $D(\cdot, \cdot)$ is truly a metric on $\hat{\mathcal{X}}$.

(iv) $\mathcal{X}$ – or a suitable copy – is dense in $\hat{\mathcal{X}}$; we use constant sequences for this.

Now $\hat{\mathcal{X}} = \{P[x] : x \in \mathrm{CS}(\mathcal{X})\}$ with metric $D$ is a *complete* space implies that every Cauchy sequence converges. To show this, we start with a Cauchy sequence in $(\hat{\mathcal{X}}, D)$, say $A_1, A_2, \dots$, i.e., $A_p = P[x^{(p)}]$ for some $x^p \in \mathrm{CS}(\mathcal{X})$, i.e., $x^{(p)} = (x_1^{(p)}, x_2^{(p)}, x_3^{(p)}, \dots)$. To show that $(A_p)$ converges in $(\hat{\mathcal{X}}, D)$, we first identify a candidate $\hat{A} \in \hat{X}$ for the limit. We need $\hat{A} \in P[\hat{x}]$ for some $\hat{x} \in \mathrm{CS}(\mathcal{X})$, so we build that. We use the $\mathrm{CS}(\mathcal{X})$ property for each sequence $x^{(1)}, x^{(2)}, x^{(3)}, \dots$.

$$\begin{gathered}
\text{Pick a sufficiently large } n_1 \text{ such that } d(x_j^{(1)}, x_k^{(1)}) < 1, \text{ for all } j, k > n_1 \\
\text{Pick a sufficiently large } n_2 > n_1 \text{ such that } d(x_j^{(2)}, x_k^{(2)}) < \frac{1}{2}, \text{ for all } j, k > n_2 \\
\text{Pick a sufficiently large } n_3 > n_2 \text{ such that } d(x_j^{(3)}, x_k^{(3)}) < \frac{1}{3}, \text{ for all } j, k > n_3 \\
\vdots
\end{gathered}$$

Each stage gives an element of $\hat{x}$:

$$\begin{aligned}
\hat{x}_1 &= x_{n_1}^{(1)} \\
\hat{x}_2 &= x_{n_2}^{(2)} \\
&\vdots \\
\hat{x}_p &= x_{n_p}^{(p)} \\
&\vdots
\end{aligned}$$

**Claim 7.** Sequence $\hat{x}_p$ is Cauchy.

*Proof sketch.* The bulk of the proof is left as an exercise. Start with

$$\begin{aligned}
d(\hat{x}_p, \hat{x}_q) = d(x_{n_p}^{(p)}, x_{n_q}^{(q)}) &\le d(x_{n_p}^{(p)}, x_j^{(p)}) + d(x_j^{(p)}, x_j^{(q)}) + d(x_j^{(q)}, \hat{x}_{n_q}^{(q)}) \\
&\le \frac{1}{p} + d(x_j^{(p)}, x_j^{(q)}) + \frac{1}{q}
\end{aligned}$$

provided $j \ge \max\{n_p, n_q\}$. Consider limit as $j \to \infty$, which tells us

$$d(\hat{x}_p, \hat{x}_q) \le \frac{1}{p} + D(P[x^{(p)}], P[x^{(q)}]) + \frac{1}{q} \dots$$

$\square$

**Claim 8.** $\hat{A} = P[\hat{x}]$ obeys $\lim_{p\to\infty} A_p = \hat{A}$, i.e., $D(A_p, \hat{A}) \to 0$ as $p \to \infty$.

*Proof.* Left as an exercise. $\square$

**Note.** After the construction succeeds, think of an original $\mathcal{X}$ as a subset of $\hat{\mathcal{X}}$ (true embedding is $\{\Phi[x] : x \in \mathcal{X}\}$, but they are functionally indistinguishable). Then $\overline{\mathcal{X}} = \hat{X}$. However, before $\hat{\mathcal{X}}$ is built, original $\mathcal{X}$ is closed as a subset of $(\mathcal{X}, d)$, so in that *original* setup, $\overline{\mathcal{X}} = \mathcal{X}$.

## 15.5 Cantor set

Consider the subset of $C_0 = [0, 1]$ that we obtain by throwing away the middle third $\Omega_0 = \left(\frac{1}{3}, \frac{2}{3}\right)$, i.e., $C_1 = C_0 \backslash \Omega_0$. This $C_1$ has 2 closed intervals: let $\Omega_1$ be the 2 open middle third pieces: $\Omega_1 = \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)$. We have $C_2 = C_1 \backslash \Omega_1$, and we keep doing this.

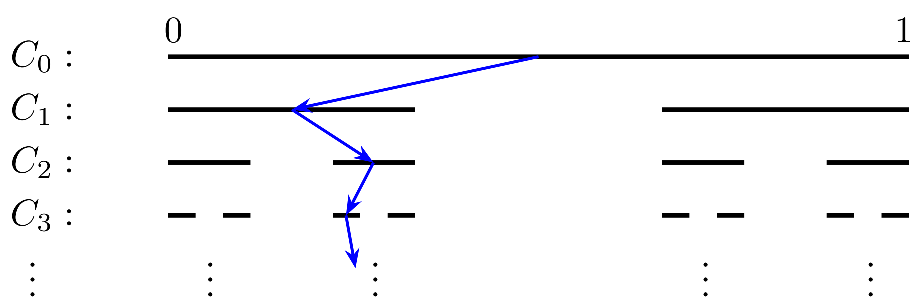



Figure 12: Visualization of the middle-thirds Cantor set; the blue path is explained in part (ii) of the notable properties of the cantor set.

The cantor set is defined as

$$\mathcal{C} := \bigcap_{k=0}^{\infty} C_k,$$

which is considered a very rich example in analysis.

Some notable properties of the cantor set:

(i) $C \neq \emptyset$ since clearly $0, 1 \in \mathcal{C}$. By self-similarity, endpoints of closed intervals in set $C_k$ *all* lie in $\mathcal{C}$.

(ii) $|\mathcal{C}| = |\mathbb{R}| \Rightarrow \mathcal{C}$ is uncountable. This is because any $0-1$ sequence defines a left-right path down the tree, as shown in the diagram above, that selects a nested sequence of closed intervals with a $1-$point intersection. Different sequences select different points of $\mathcal{C}$ (number of $0-1$ sequences equals $|\mathbb{R}|$).

(iii) $\mathcal{C}' = \mathcal{C}$.

(iv) $\mathcal{C}^\circ = \emptyset$.

(v) The total length of open sets removed from $[0, 1]$ to define $\mathcal{C}$ equal $1$. If length has any meaning for set $\mathcal{C}$, the only possible value is $0$.

# 16 Continuity

## 16.1 The big picture

Let $(\mathcal{X}, \mathscr{T}_{\mathcal{X}})$, $(\mathcal{Y}, \mathscr{T}_{\mathcal{Y}})$, $(\mathcal{Z}, \mathscr{T}_{\mathcal{Z}})$ be HTS's.

**Definition: Continuous**

Let $f : \mathcal{X} \to \mathcal{Y}$ be a (single valued) mapping. To call $f$ ***continuous*** (on $(\mathcal{X}, \mathscr{T}_{\mathcal{X}})$) means for all $\mathcal{G} \in \mathscr{T}_{\mathcal{Y}}$, $f^{-1}(\mathcal{G}) \in \mathscr{T}_{\mathcal{X}}$, i.e., every open set $\mathcal{G}$ in $\mathcal{Y}$ has an open pre-image $f^{-1}(\mathcal{G}) = \{x \in \mathcal{X} : f(x) \in \mathcal{G}\}$.

**Example 27.** For any metric space $(\mathcal{X}, d)$ with $p \in \mathcal{X}$, the function $f : \mathcal{X} \to \mathbb{R}$ such that $f(x) = d(x, p)$ is continuous.

*Proof sketch.* Given open $\mathcal{G} \subseteq \mathbb{R}$, consider $f^{-1}(\mathcal{G})$, a set in $\mathcal{X}$. E.g., if $\mathcal{G} = (a, b)$, and $a \geq 0$,

$$\begin{aligned} f^{-1}((a,b)) =& \{x \in \mathcal{X} : a < d(x,p) < b\} \\ =& \mathbb{B}[p; b) \backslash \mathbb{B}[p; a] \\ =& \mathbb{B}[p; b) \cap (\mathbb{B}[p; a])^c, \end{aligned}$$

which is an intersection of two open sets, hence open. Rest of the proof is left as an exercise. $\square$

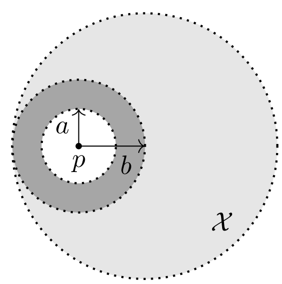



Figure 13: Visualization of the pre-image in the proof above.

**Note.** Continuity does not allow functions with "holes" to exist: something that you have probably seen before in MATH 100 or equivalent courses.

**Proposition: Continuity and Density**

Given $f_1, f_2 : \mathcal{X} \to \mathcal{Y}$ both continuous, and some set $\mathcal{Q} \subseteq \mathcal{X}$,

$$f_1(q) = f_2(q) \text{ for all } q \in \mathcal{Q} \Rightarrow f_1(q) = f_2(q) \text{ for all } q \in \overline{\mathcal{Q}}.$$

*Proof.* Pick any $x \in \overline{\mathcal{Q}}$ and let $y_1 = f_1(x)$, $y_2 = f_2(x)$. We shall show that $y_1 = y_2$. Pick any open neighbourhoods $\mathcal{U}_1 \in \mathscr{N}(y_1), \mathcal{U}_2 \in \mathscr{N}(y_2)$ such that

$$\Omega_1 = f_1^{-1}(\mathcal{U}_1),\ \Omega_2 = f_2^{-1}(\mathcal{U}_2)$$

are open by continuity. Let $\Omega = \Omega_1 \cap \Omega_2$; clearly $x \in \Omega$, so since $x \in \overline{\mathcal{Q}}$, $\Omega \cap \mathcal{Q} \neq \emptyset$. Thus, for any $q \in \Omega \cap \mathcal{Q}$, $f_1(q) = f_2(q)$, and this point lies in $\mathcal{U}_1 \cap \mathcal{U}_2$. However, this shows that for all $\mathcal{U}_1 \in \mathscr{N}(y_1), \mathcal{U}_2 \in \mathscr{N}(y_2), \mathcal{U}_1 \cap \mathcal{U}_2 \neq \emptyset$. Finally, all that remains is to compare HTS: the negation of this reveals that $y_1 = y_2$ $\square$

## 16.2 Continuity and Compactness

**Theorem**

Suppose $\mathcal{X}$ is a *compact* HTS, and $f : \mathcal{X} \to \mathcal{Y}$ is continuous (on $\mathcal{X}$), then $f(\mathcal{X}) = \{f(x) : x \in \mathcal{X}\}$ is compact in $\mathcal{Y}$.

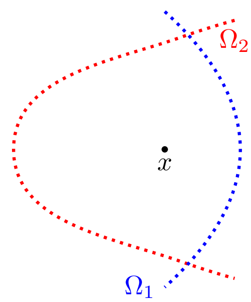



Figure 14: Visualization of the proof.

*Proof.* Let $\mathscr{G}$ be an arbitrary open cover for $f(\mathcal{X})$. We construct $\mathscr{G}_0 = \{f^{-1}(\mathcal{G}) : \mathcal{G} \in \mathscr{G}\}$. Each element in $\mathscr{G}_0$ is *open* by continuity of $f$, and clearly each $x \in \mathcal{X}$ is in at least one $\mathcal{G}$, so $\mathcal{G}_0$ is an open cover of $\mathcal{X}$. Compactness of $\mathcal{X}$ guarantees that for some $N \in \mathbb{N}$, $f^{-1}(\mathcal{G}_1), f^{-1}(\mathcal{G}_2), \dots, f^{-1}(\mathcal{G}_N)$ form a finite subcover for $\mathcal{X}$. In turn, this would mean that $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_N$ is an open subcover of $f(\mathcal{X})$ selected from $\mathscr{G}$. $\square$

**Note** (Consequences)**.** For continuous $f : \mathcal{X} \to \mathcal{Y}$ and $\mathcal{X} \neq \emptyset$ compact,

(1) If $\mathcal{Y}$ is a metric space, $f(\mathcal{X})$ is closed and compact.

(2) If $\mathcal{Y} = \mathbb{R}$, then $f(\mathcal{X})$ *includes* the numbers $\inf f(\mathcal{X})$, $\sup f(\mathcal{X})$, i.e., $\mathcal{X}$ contains points $\underline{x}, \overline{x} \in \mathcal{X}$ obeying $f(\underline{x}) \leq f(x) \leq f(\overline{x})$, for all $x \in \mathcal{X}$, where $\underline{x} := \min\{f(x)\}_{x\in\mathcal{X}}$ and $\overline{x} := \max\{f(x)\}_{x\in\mathcal{X}}$.

**Theorem: Inverse mappings**

Suppose $\mathcal{X}$ is compact and $f : \mathcal{X} \to \mathcal{Y}$ is bijective and continuous. Then, $f^{-1} : \mathcal{Y} \to \mathcal{X}$ is continuous.

*Proof.* For simplicity, let $h = f^{-1}$. To check continuity, we show one of two equivalent statements:

(i) $h^{-1}(\mathcal{U})$ is *open* for any *open* $\mathcal{U} \subseteq \mathcal{X}$.

(ii) $h^{-1}(\mathcal{C})$ is *closed* for any *closed* $\mathcal{C} \subseteq \mathcal{X}$.

For this proof, we will show (ii): if $\mathcal{C} \subseteq \mathcal{X}$ is closed,

$$\begin{aligned} h^{-1}(\mathcal{C}) &= \{y \in \mathcal{Y} : h(y) \in \mathcal{C}\} \\ &= \{y \in \mathcal{Y} : f^{-1}(y) \in \mathcal{C}\} \\ &= \{y \in \mathcal{Y} : y \in f(\mathcal{C})\} = f(\mathcal{C}); \end{aligned}$$

however, from our hypothesis, $\mathcal{C}$ is compact, so $f(\mathcal{C})$ is compact, hence closed. $\square$

## 16.3 Continuity at a point

**Definition: Continuous at a point**

Let $(\mathcal{X}, \mathscr{T}_\mathcal{X})$, $(\mathcal{Y}, \mathscr{T}_\mathcal{Y})$, $(\mathcal{Z}, \mathscr{T}_\mathcal{Z})$ be given HTS's, with $f : \mathcal{X} \to \mathcal{Y}$ and $x \in \mathcal{X}$. To say "$f$ ***is continuous at*** $x$" is to say that

$$\text{for all } \mathcal{W} \in \mathscr{N}_\mathcal{Y}(f(x)), \text{ one has } f^{-1}(\mathcal{W}) \in \mathscr{N}_\mathcal{X}(x).$$

Equivalently, for all $\mathcal{W} \in \mathscr{T}_\mathcal{Y}$ with $f(x) \in \mathcal{W}$, there exists $\mathcal{U} \in \mathscr{T}_\mathcal{X}$ with $x \in \mathcal{U}$ and $f(\mathcal{U}) \subseteq \mathcal{W}$.

**Lemma**

For $f : \mathcal{X} \to \mathcal{Y}$ as above, the following are equivalent:

(a) $f$ is continuous (on $\mathcal{X}$), i.e., $f^{-1}(\Omega)$ is open in $\mathcal{X}$, for each $\Omega$ open in $\mathcal{Y}$.

(b) $f$ is continuous at $x$, for each $x \in \mathcal{X}$.

*Proof.* (a⇒b) This is immediate.

(b⇒a) Given any open $\mathcal{W} \subseteq \mathcal{Y}$, define $\mathcal{U} = f^{-1}(\mathcal{W})$. To show $\mathcal{U}$ is open, pick any $x \in \mathcal{U}$ and show $x \in \mathcal{U}^\circ$. Consider $y = f(x) \in \mathcal{W}$; by definition of "continuous at $x$", $\mathcal{U} \in \mathscr{N}(x)$, i.e., $\mathcal{U}$ contains an open set $\mathcal{V}$ with $x \in \mathcal{V} \subseteq \mathcal{U}$. Thus, $x \in \mathcal{U}^\circ$, as required. □

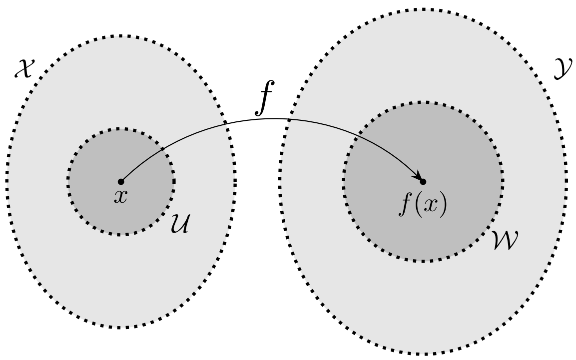



Figure 15: Visualization of the proof.

**Notation 11.** Going forward, if it is not clarified what $\mathcal{X}$, $\mathcal{Y}$ or $\mathcal{Z}$ are, they will always be HTS's.

**Proposition**

If $f : \mathcal{X} \to \mathcal{Y}$ and $g : \mathcal{Y} \to \mathcal{Z}$, and $f$ is continuous at $x_0$, $g$ is continuous at $y_0 = f(x_0)$, then $h = g \circ f$ is continuous at $x_0$.

*Proof.* Pick any open $\mathcal{W} \subseteq \mathcal{Z}$ with $h(x_0) \in \mathcal{W}$. Then

$$\begin{aligned} h^{-1}(\mathcal{W}) =& \{x \in \mathcal{X} : g \circ f(x) = h(x) \in \mathcal{W}\} \\ =& \{x \in \mathcal{X} : f(x) \in g^{-1}(\mathcal{W})\} \quad (g^{-1}(W) \text{ is an open neighbourhood of } y_0 \text{ by continuity of } g.) \\ =& f^{-1}(g^{-1}(\mathcal{W})) \quad (\text{open neighbourhood of } x_0 \text{ by continuity of } f.) \end{aligned}$$

□

**Proposition**

If $f, g : \mathcal{X} \to \mathbb{R}$ both continuous at $x_0 \in \mathcal{X}$, then as are the new functions

- $(f + cg)(x) = f(x) + cg(x)$ for all $x \in \mathcal{X},$ and any $c \in \mathbb{R}$.
- $(fg)(x) = f(x)g(x)$ for all $x \in \mathcal{X}$.
- $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$ for all $x \in \mathcal{X}$, provided $g(x) \neq 0$.

*Proof.* Left as an exercise. □

## 16.4 The metric case

**Proposition**

Let $(\mathcal{X}, d_{\mathcal{X}})$, $(\mathcal{Y}, d_{\mathcal{Y}})$ be metric spaces, $x \in \mathcal{X}$ and $f : \mathcal{X} \to \mathcal{Y}$. The following are equivalent:

(a) $f$ is continuous at $x$.

(b) For all $\varepsilon > 0$, there exists $\delta > 0$ such hat for all $x'$ with $d(x, x') < \delta$, $d_{\mathcal{Y}}(f(x), f(x')) < \varepsilon$.

(c) For any sequence $(x_n)$ in $\mathcal{X}$ with $x_n \to x \in \mathcal{X}$, one has $f(x_n) \to f(x) \in \mathcal{Y}$.

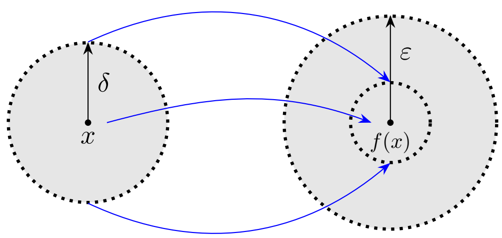



Figure 16: Visual representation showing that anything within $\delta$ of $x$ gets mapped to an open neighbourhood around $f(x)$.

*Proof.* (a$\Rightarrow$b) Assume (a); pick an arbitrary $\varepsilon > 0$. Let $\mathcal{V} = \mathbb{B}_{\mathcal{Y}}[f(x); \varepsilon)$; this is open, so $f^{-1}(\mathcal{V}) \in \mathscr{N}_{\mathcal{X}}(x)$, i.e., for some radius $\delta > 0$, we have $\mathbb{B}_{\mathcal{X}}[x; \delta) \subseteq f^{-1}(\mathcal{V})$. Then, $f(x') \in \mathcal{V}$ for all $x' \in \mathbb{B}_{\mathcal{X}}[x; \delta)$. Finally, express using $d_{\mathcal{Y}}$ to recover (b).

(b$\Rightarrow$c) is left as an exercise.

(c$\Rightarrow$a) We show this by contrapositive, i.e., $(\neg a) \Rightarrow (\neg c)$. Assume $(\neg a)$, i.e., $f$ is *not continuous* at $x$. Then, for some $\mathcal{V} \in \mathscr{N}_{\mathcal{Y}}(f(x))$, we have $x \notin \left(f^{-1}(\mathcal{V})\right)^c$. We then shrink $\mathcal{V}$ as necessary to say $\mathcal{V} = \mathbb{B}[f(x); \varepsilon)$ for some $\varepsilon > 0$. Now, if $f^{-1}(\mathcal{V})$ is *not* a neighbourhood of $x$, each ball $\mathbb{B}_{\mathcal{X}}\left[x; \frac{1}{n}\right)$ must contain a point of $\left[f^{-1}(\mathcal{V})\right]^c$; pick one such point and call it $x_n$: $f(x_n) \notin \mathcal{V}$, i.e., $d_{\mathcal{Y}}\left(f(x_n), f(x)\right) \geq \varepsilon$, and yet $d_{\mathcal{X}}(x_n, x) < \frac{1}{n}$. This sequence $(x_n)$ has $x_n \to x \in \mathcal{X}$, but $f(x_n) \not\to f(x) \in \mathcal{Y}$. $\square$

**Note** (Proof strategies)**.** To *prove* continuity, use (b). To *disprove* continuity, use (c) ($\neg(c)$ just requires one sequence.)

**Quotes of the day**

(Said with hopium) "Everything is fine; we still have three lectures, we can do the derivative." - Dr. Loewen, 11/29/2023

**Example 28.** Some examples of the two definitions of continuity are:

(a) On $\mathcal{X} = (0, +\infty)$, $f(x) = \dfrac{1}{x}$.

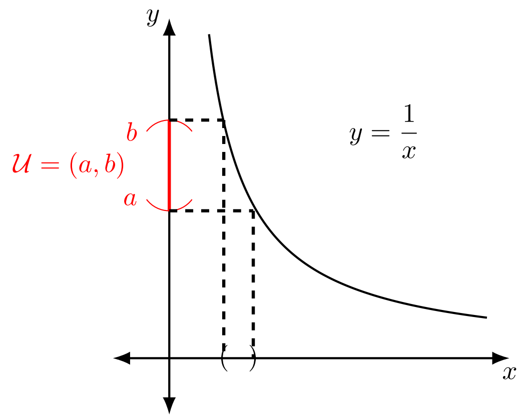



Figure 17: Plot of $y = \dfrac{1}{x}$.

It is clear that every open interval $\mathcal{U} = (a, b)$ in $\mathbb{R}$ has $f^{-1}(\mathcal{U})$, an open interval in $\mathcal{X}$. Thus, $f$ is continuous on $\mathcal{X} \Rightarrow f$ is continuous at every point in $\mathcal{X}$.

(b) (Dirichlet's function) On $\mathcal{X} = \mathbb{R}$,

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

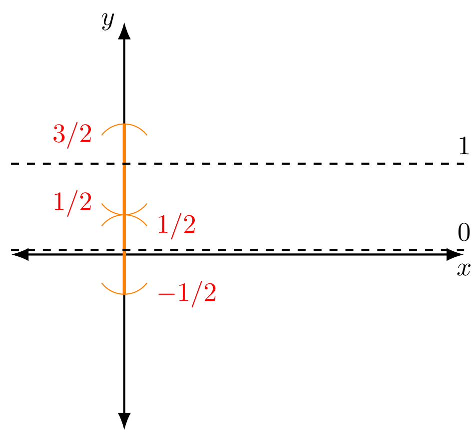



Figure 18: Plot of Dirichlet's function.

This is *not* "continuous on $\mathcal{X}$", because intervals $\left(\dfrac{1}{2}, \dfrac{3}{2}\right)$, having $f^{-1}\left(\left(\dfrac{1}{2}, \dfrac{3}{2}\right)\right) = \mathbb{Q}$, and $\left(-\dfrac{1}{2}, \dfrac{1}{2}\right)$, having $f^{-1}\left(\left(-\dfrac{1}{2}, \dfrac{1}{2}\right)\right) = \mathbb{R}\backslash\mathbb{Q}$, both have pre-images that are *not open*.

We can a bit better here and say that $f$ fails to be continuous at every point in $\mathcal{X}$. We will illustrate this with discontinuity at 0: $f(0) = 1$ because $0 \in \mathbb{Q}$, and sequence

$$x_n = \begin{cases} \frac{1}{n} & \text{if } n \equiv 1 \mod 2 \\ \frac{\sqrt{2}}{n} & \text{if } n \equiv 0 \mod 2 \end{cases}$$

has $x_n \to 0$, but

$$f(x_n) = \begin{cases} 1 & \text{if } n \equiv 1 \mod 2 \\ 0 & \text{if } n \equiv 0 \mod 2 \end{cases}$$

fails to obey "$\lim_{n\to\infty} f(x_n) = f(0)$", as required.

(c) Use $f(x)$ above to build $g(x) = xf(x)$.

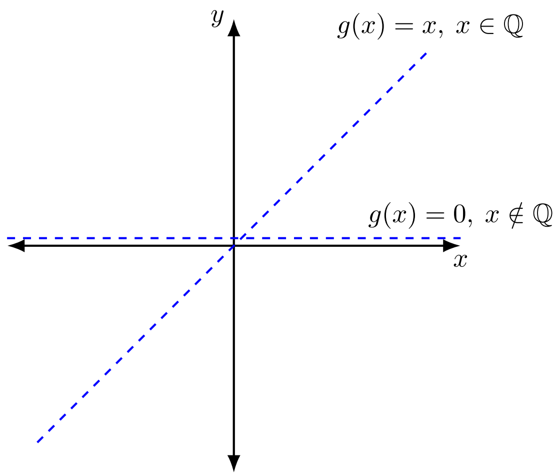



Figure 19: Plot of $g(x) = xf(x) = x$(Dirichlet's function).

This $g$ is continuous at $0$, but discontinuous at every other point. How do we show this? Well, this is equivalent to saying discontinuous at $x \neq 0$, so find sequences $q_n \in \mathbb{Q}$ and $z_n \notin \mathbb{Q}$ with $q_n \to x$ and $z_n \to x$; for one of these sequences, the values of $f(q_n)$ and $f(z_n)$ will have a limit different from $f(x)$.

When $x = 0$, let $\varepsilon > 0$ be given. Pick $\delta = \varepsilon$; every $x \in \mathbb{B}[0;\delta)$ has

$$|g(x) - g(0)| = |g(x)| \leq |x| < \delta = \varepsilon.$$

## 16.5 Uniform continuity

This only makes sense in metric spaces.

**Definition: Uniform continuity**

A function $f : \mathcal{X} \to \mathcal{Y}$ is *uniformly continuous on* $\mathcal{X}$ ($\mathcal{X}, \mathcal{Y}$ metric spaces) exactly when:

$$\textit{For all } \varepsilon > 0, \text{ there exist } \delta > 0 \text{ such that, for all } s \in \mathcal{X},\ t \in \mathbb{B}_{\mathcal{X}}[s;\delta),\ f(t) \in \mathbb{B}_{\mathcal{Y}}[f(s);\varepsilon).$$

**Note.** The main point here is that the same $\delta$ covers all $s \in \mathcal{X}$.

Contrasting this with "$f$ is continuous on $X$", which means

$$\text{For all } s \in \mathcal{X},\ \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that, for all } t \in \mathbb{B}_{\mathcal{X}}[s;\delta),\ f(t) \in \mathbb{B}_{\mathcal{Y}}[f(s);\varepsilon),$$

we see that in non-uniform continuity, we allow $\delta = \delta(\varepsilon, s)$ to depend on base point $s$; this changes the meaning, and makes uniform continuity visibly better.

**Example 29.** In $\mathcal{X} = (0,\infty)$, $f(x) = \dfrac{1}{x}$ is continuous on $\mathcal{X}$, but *not* uniformly continuous on $\mathcal{X}$.

*Proof.* We start by noting that

$$|f(s) - f(t)| = \left|\frac{1}{s} - \frac{1}{t}\right| = \frac{|s-t|}{st}.$$

For fixed $s > 0$, every $t \in \left(\frac{s}{2}, \frac{3s}{2}\right)$ will obey $\frac{1}{t} < \frac{2}{s}$. Hence,

$$|f(s) - f(t)| = \frac{|s-t|}{st} < \left(\frac{2}{s^2}\right)|s-t|;$$

to get RHS $< \varepsilon$ when $|s-t| < \delta$, we should choose $\delta = \left(\frac{s^2}{2}\right)\varepsilon$. However, to defend the assumption $t > \frac{\varepsilon}{2}$, we have to make sure $\delta < \frac{s}{2}$, so we choose $\delta = \min\left\{\frac{s}{2}, \left(\frac{s^2}{2}\right)\varepsilon\right\}$. This concludes the proof, since

$$t \in (s-\delta, s+\delta) \Rightarrow |f(s) - f(t)| < \varepsilon.$$

□

**Note.** Notice that $\delta$ depends on $s$; smaller $s$ demands a smaller $\delta$. This dependence is not "uniform", however, this is not yet a proof that this function is not uniformly continuous, because we have not yet falsified the definition yet, but merely shown that our methods are too weak to confirm that the function is uniformly continuous.

*Proof.* To prove that uniform continuity fails, we look back at the definition and negate that:

$$\text{There exists } \varepsilon > 0, \text{ such that for all } \delta > 0, \text{ there exists } s \in \mathcal{X},\ t \in \mathbb{B}[s;\delta), \text{ with } |f(t) - f(s)| \geq \varepsilon.$$

This holds for $\varepsilon = 1$. Given any $\delta > 0$, pick any $s \in (0, \min\{\delta, 1\})$, and any $t \in \left(0, \frac{s}{2}\right)$: $0 < s - t < s < \delta$ and

$$f(t) - f(s) = \frac{1}{t} - \frac{1}{s} > \frac{2}{s} - \frac{1}{s} = \frac{1}{s} > 1 = \varepsilon.$$

□

Where exactly to get uniform continuity then?

**Idea.** Pointwise continuity at points $x_1, x_2, \ldots, x_N \in \mathcal{X}$ give $N$ choices for $\delta_1, \delta_2, \ldots, \delta_N$ for any $\varepsilon > 0$. Using $\delta = \min\{\delta_1, \ldots, \delta_N\}$ gives something like uniform continuity.

**Theorem**

If $f : \mathcal{X} \to \mathcal{Y}$ is continuous, and $\mathcal{X}$ is compact, then $f$ is uniformly continuous on $\mathcal{X}$.

*Proof.* Given any $\varepsilon > 0$, use point-wise continuity at each $x \in \mathcal{X}$ to get $\delta = \delta(x) > 0$ such that $x' \in \mathbb{B}[x;\delta(x))$, i.e., $d(f(x), f(x')) < \frac{\varepsilon}{5}$. Now, $\mathscr{G} := \left\{\mathbb{B}\left[x; \frac{1}{7}\delta(x)\right) : x \in \mathcal{X}\right\}$ is an open cover for $\mathcal{X}$; compactness gives us a finite subcover with labels $x_1, x_2, \ldots x_N$. Let $\delta_k(x) := \delta(x_k)$ and thus $\delta := \frac{1}{7}\min\{\delta_1, \ldots, \delta_N\}$. Now, we pick any $x \in \mathcal{X}, x' \in \mathbb{B}[x;\delta)$. From the finite subcover, our $x \in \mathbb{B}\left[x_k; \frac{1}{7}\delta_k\right)$. Also, $x'$ has

$$\begin{aligned} d(x', x_k) \leq& d(x', x) + d(x, x_k) \\ <& \delta + \frac{1}{7}\delta_k \leq \frac{2}{7}\delta_k < \delta_k, \end{aligned}$$

so

$$\begin{aligned} d(f(x'), f(x)) \leq& d(f(x'), f(x_k)) + d(f(x_k), f(x)) \\ <& \frac{\varepsilon}{5} + \frac{\varepsilon}{5} < \varepsilon. \end{aligned}$$

□

**Example 30.** An increasing function $f : \mathbb{R} \to \mathbb{R}$ that is continuous at $x$ if and only if $x \notin \mathbb{Q}$ (we have encountered this function before in Homework 11 problem 6.) We enumerate the rationals $\mathbb{Q} = \{q_1, q_2, \dots\}$):

$$f(x) = \sum_{i \in I(x)} \frac{1}{2^i},$$

where $I(x) := \{i \in \mathbb{N} : q_i < x\}$.

**Note.** If $a < b$, then

$$f(b) - f(a) = \sum_{I(a) \setminus I(b)} \frac{1}{2^i} > 0,$$

where $I(b) \setminus I(a) := \{i \in \mathbb{N} : a \leq q_i < b\} \neq \emptyset$.

If $x \in \mathbb{Q}$, then we have $x = q_N$ for some $N$. For any sequence $(x_n)$ of rationals with $x_n \to q_N$ (decreasing), $f(x_n) > f(q_N) + \dfrac{1}{2^N}$, so "$f(x_n) \to f(x)$" is impossible. However, if $x \notin \mathbb{Q}$, continuity at $x$ holds. Indeed, given any $\varepsilon > 0$, pick $N$ to make

$$\sum_{i=N+1}^{\infty} \frac{1}{2^i} < \varepsilon.$$

Then, let $\delta := \min\{|x - q_1|, |x - q_2|, \dots, |x - q_N|\}$. For any $x'$ with $|x' - x| < \delta$, all of $q_1, \dots, q_N$ lie *outside* $(x - \delta, x + \delta)$, so

$$|f(x') - f(x)| \leq \sum_{i=N+1}^{\infty} \frac{1}{2^i} < \varepsilon.$$

## 16.6 Connectedness and Intermediate Value Theorem

**Proposition 5.** Let $(\mathcal{X}, \mathscr{T})$ be a HTS, and suppose $f : \mathcal{X} \to \mathbb{R}$ is continuous. For any $q \in \mathbb{R}$, let

$$\Omega(q) := \{x \in \mathcal{X} : f(x) < q\}.$$

Then, $\partial \Omega(q) \subseteq \{x \in \mathcal{X} : f(x) = q\}$.

*Proof.* The proof is left as an exercise while using the canvas notes as a reference. $\square$

**Note.** Strict inclusion is possible.

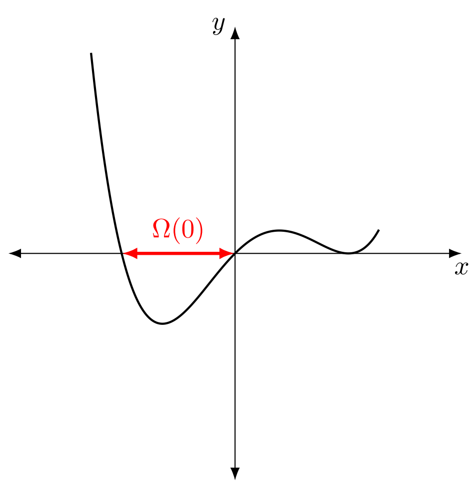



Figure 20: Plot of the function $y = x(x + 1)(x - 1)^2$, where $\Omega(0) = (-1, 0)$ is highlighted.

**Note** (Comments on the plot)**.** In the plot above, $\Omega(0) = (-1, 0)$, $\partial\Omega(0) = \{-1, 0\}$, but $f^{-1}(\{0\}) = \{-1, 0, 1\}$.

**Corollary 7.** In the setup above, if $\Omega(q) \neq \emptyset$, and yet $f(x) \neq q$ for all $x \in \mathcal{X}$, then $\Omega(q)$ is both open and closed in $\mathcal{X}$.

*Proof.* $\Omega(q)$ is *open* by continuity;

$$\overline{\Omega(q)} = \Omega(q) \cup \partial\Omega(q) = \Omega(q).$$

□

# 17 Limits of functions

Closely related to the idea of continuity is the concept of limits:

**Definition: Limit of a function**

Suppose $\mathcal{X}, \mathcal{Y}$ are metric spaces with $f : \mathcal{X} \to \mathcal{Y}$ and $x_0 \in \mathcal{X}$; we say

$$\lim_{x \to x_0} f(x_0) = y_0$$

exactly when

(i) $x_0 \in \mathcal{X}'$.

(ii) $\varphi(x) = \begin{cases} f(x) & \text{if } x \neq x_0 \\ y_0 & \text{if } x = x_0 \end{cases}$ is continuous at $x_0$.

Equivalently,

(i) $x_0 \in \mathcal{X}'$.

(ii) For all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in \mathbb{B}(x_0; \delta)$, $d_Y(f(x), y_0) < \varepsilon$.

**Note.** For (ii), a sequential alternative exists, namely: $f(x_n) \to y_0$ for any sequence $x_n \to x_0$, $x_n \neq x_0$ for all $n$.

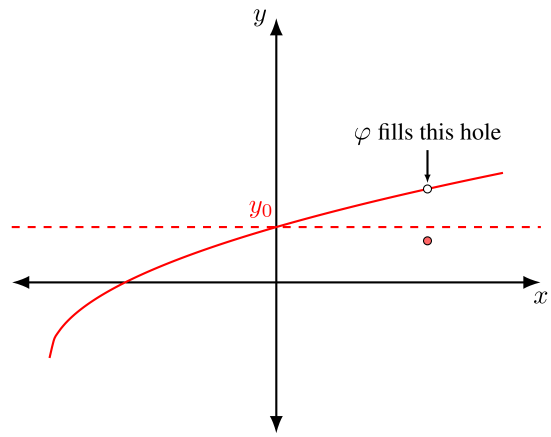



Figure 21: Plot demonstrating the definition above.

# 18 Differentiation

## 18.1 Basics

**Definition: Differentiability**

Given an interval $[a, b] \subseteq \mathbb{R}$, a point $c \in [a, b]$, and $f : [a, b] \to \mathbb{R}$, let

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

When this limit converges, this is called the "the derivative of $f$ at $c$". Say "$f$ is ***differentiable*** at $c$" when this happens.

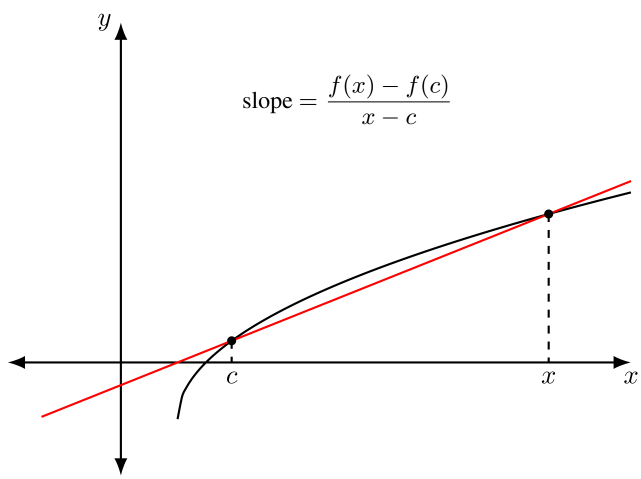



Figure 22: Plot demonstrating the definition above.

### 18.1.1 Instantaneous slope

When $f'(c)$ exists, it tells us the slope of the "best linear approximation" for $f(x)$ near $x = c$. Comparing with $\ell(x) = f(c) + m(x - c)$:

**Theorem: Butterfly lemma**

In the setup above,

(i) If $m < f'(c)$, then there exists $\delta > 0$ such that

$$\begin{aligned} \ell(x) > f(x) \quad &\text{for all } x \in (c - \delta, c) \\ \ell(x) < f(x) \quad &\text{for all } x \in (c, c + \delta). \end{aligned}$$

(ii) If $m > f'(c)$, then there exists $\delta > 0$ such that

$$\begin{aligned} \ell(x) < f(x) \quad &\text{for all } x \in (c - \delta, c) \\ \ell(x) > f(x) \quad &\text{for all } x \in (c, c + \delta). \end{aligned}$$

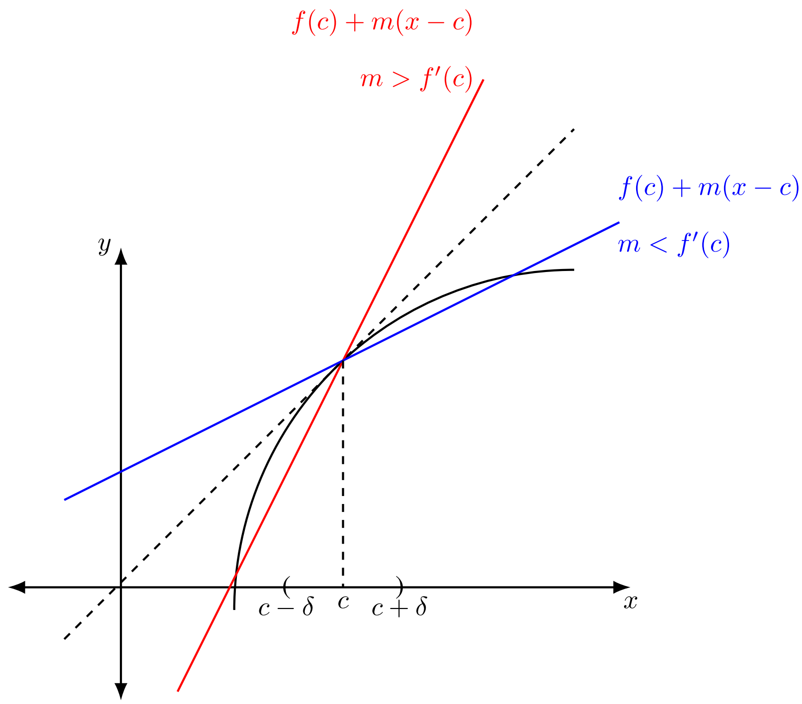



Figure 23: Visualization of the theorem above.

(i) *Proof.* Suppose $m < f'(c)$. Define $\varepsilon = \frac{1}{2}(f'(c) - m)$. We use this in the limit definition to get $\delta > 0$ such that

$$\left|\frac{f(x)-f(c)}{x-c} - f'(c)\right| < \epsilon$$

for all $x \in \mathbb{B}[c;\delta)$. Therefore,

$$\begin{aligned}
&-\varepsilon|x-c| < f(x) - f(c) - f'(c)(x-c) < \varepsilon|x-c| \\
\Rightarrow &-\varepsilon|x-c| < f(x) - [\ell(x) + (f'(c)-m)(x-c)] < \varepsilon|x-c| \\
\Rightarrow &(f'(c)-m)(x-c) - \varepsilon|x-c| < f(x) - \ell(x) < \varepsilon|x-c| + (f'(c)-m)(x-c).
\end{aligned}$$

Now, if $x > c$, $|x-c| = x-c$, so the left inequality tells us $0 < \varepsilon(x-c) < f(x) - \ell(x)$ for $x \in (c, c+\delta)$.

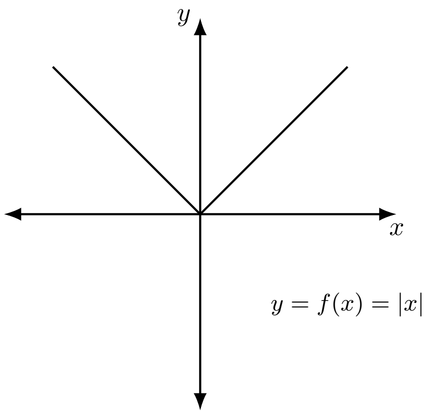
If $x < c$, $|x-c| = -(x-c)$, and the right inequality tells us $f(x) - \ell(x) < \varepsilon(x-c) < 0$ for $x \in (c-\delta, c)$. □

(ii) *Proof.* The proof is very similar to that of (i), and is left as an exercise. □

**Corollary 8.** If $f$ is differentiable at $c$, then $f$ must be continuous at $c$.

*Proof sketch.* Apply squeeze theorem to the theorem above. □

**Note.** The converse of the corollary is false: $f(x) = |x|$ is continuous at $x = 0$, but *not* differentiable at $x = 0$.



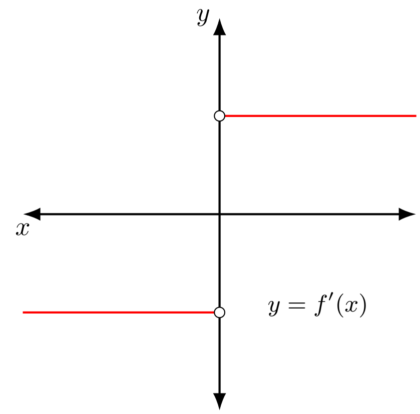




Figure 24: Plots of the counterexample.

Some derivatives are continuous – but in ways more interesting than $f(x) = |x|$

**Example 31.** Consider

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}.$$

Here,

$$\begin{aligned} f'(x) =& 2x \sin\left(\frac{1}{x}\right) + x^2 \cos\left(\frac{1}{x}\right)\left(-\frac{1}{x^2}\right) \\ =& -\cos\left(\frac{1}{x}\right) + 2x \sin\left(\frac{1}{x}\right), \text{ for } x \neq 0, \end{aligned}$$

and

$$\begin{aligned} f'(0) =& \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} \\ =& \lim_{x \to 0} x \sin\left(\frac{1}{x}\right) \\ =& 0, \text{ by Squeeze theorem.} \end{aligned}$$

Hence, $f'$ exists for all $x$, and is discontinuous at 0.

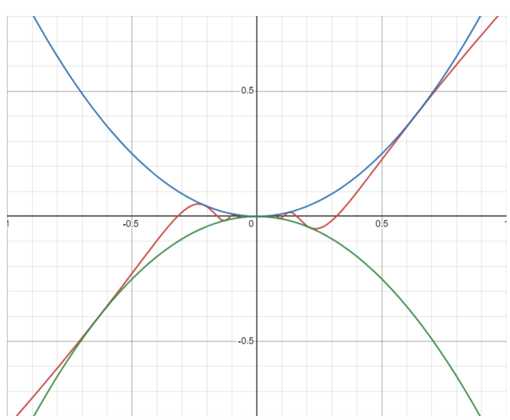



Figure 25: Plot of $y = x^2 \sin\left(\frac{1}{x}\right)$ from desmos; the red curve is $y$.

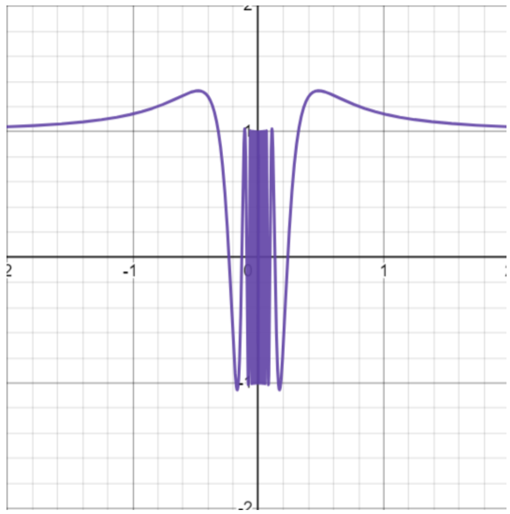



Figure 26: Plot of $y = f'(x)$ from Desmos.

**Quotes of the day**

"Let's back up and see how nice it would be to be so smart that you're right all the time." - Dr. Philip Loewen, 12/06/2023

"A tragedy, the butterfly! We killed the butterfly! Ah, it'll come back in the spring." - Dr. Philip Loewen, 12/06/2023

"Job done." - Dr. Philip Loewen, 12/06/2023

"The final exam, it's worth a lot right? It's worth 50%, so you probably want to study." - Dr. Philip Loewen, 12/06/2023

Some common misconceptions:

1. The tangent line touches the graph at one point only.

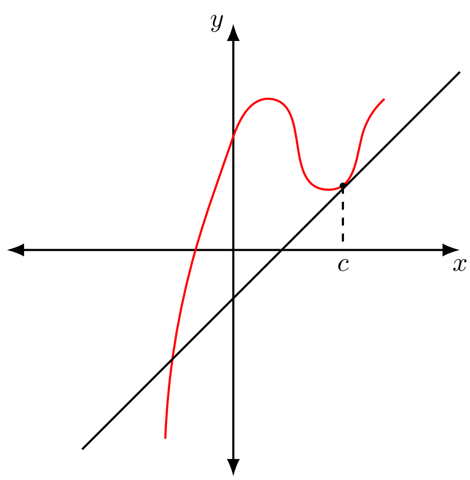



Figure 27: Example where the tangent line at a point passes through the curve at another point.

2. The tangent line touches the curve but doesn't cross it.

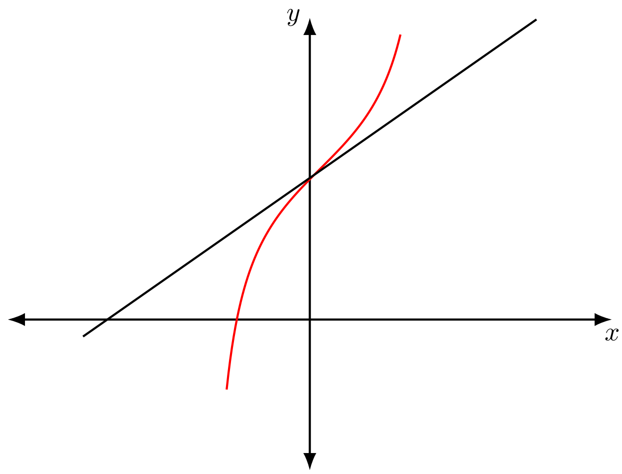



Figure 28: Example where the tangent line touches the curve and crosses it.

3. If $f'(c) > 0$, then $f$ is increasing on some open interval with midpoint $c$. Consider the example

$$f(x) = \begin{cases} \frac{1}{2}x + x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{at } x = 0. \end{cases}$$

Note that

$$f'(x) = \frac{1}{2} - \cos\left(\frac{1}{x}\right) + 2x \sin\left(\frac{1}{x}\right),$$

has positive values in the interval $(-\delta, \delta)$.

## 18.2 Optimization

**Theorem**

Suppose $f : (a, b) \to \mathbb{R}$ has a minimum at some point $c \in (a, b)$, i.e., $f(c) \leq f(x)$ for all $x \in (a, b)$, then $f'(c) = 0$.

*Proof.* We will show this proof by contrapositive, and use butterfly lemma.

Suppose $f'(c) \neq 0$. Without loss of generality, assume $f'(c) > 0$. Pick $m = 0$, and use butterfly lemma to get $\delta > 0$ such that

$$f(x) < f(c), \text{ for all } x \in (c - \delta, c)$$

Thus, $f(c)$ is *not* a minimizer on $(a, b)$. $\square$

**Theorem: (Darboux)**

Derivatives have the intermediate value property:

If $f$ is differentiable at all points of the closed interval $[a, b]$ and $\mu$ lies between $f'(a)$ and $f'(b)$, then there exists $c \in (a, b)$ where $f'(c) = \mu$.

*Proof.* The result is obvious if $f'(a) = \mu$ or $f'(b) = \mu$.

Suppose without loss of generality, $f'(a) < \mu < f'(b)$. Consider the new function

$$g(x) = f(x) - \mu x :$$

this is a continuous function on $[a, b]$ with

$$g'(a) = f'(a) - \mu$$
$$g'(b) = f'(b) - \mu.$$

Hence, by butterfly lemma, that the absolute minimum of function $g$ on the set $[a, b]$ cannot occur at either end. So, it must be at some point $c \in (a, b)$, and thus, $0 = g'(c) = f'(c) - \mu$. $\square$

**Theorem: Mean Value Theorem**

Given $f : [a, b] \to \mathbb{R}$ is a continuous function, suppose $f'(x)$ exists for all $x \in (a, b)$. Then, there exists $c \in (a, b)$ such that $\dfrac{f(b) - f(a)}{b - a} = f'(c)$.

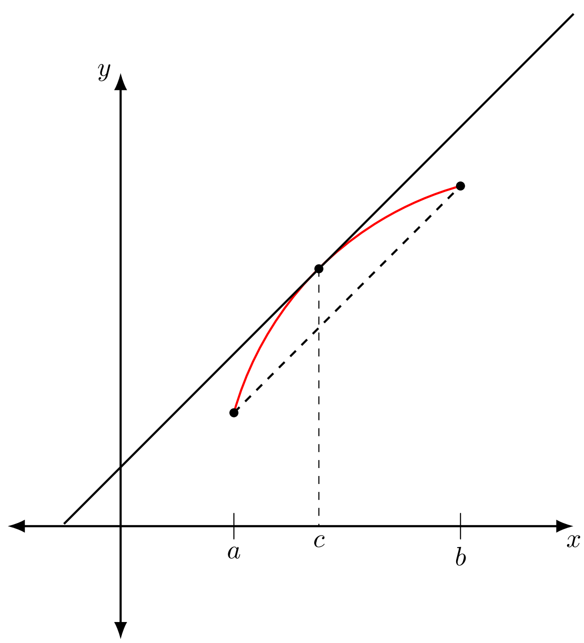



Figure 29: Visualization of the mean value theorem.

*Proof.* Define $m := \dfrac{f(b) - f(a)}{b - a}$, and let $g(x) = f(x) - mx$. Evaluate

$$g(a) = f(a) - ma$$
$$g(b) = f(b) - mb.$$

Note,

$$g(b) - g(a) = [f(b) - f(a)] - m(b - a)$$
$$= 0.$$

Since $g$ is continuous with a compact domain, it has absolute maximum and absolute minimum $[a, b]$.

If $g$ is constant, those are equal, but

$$0 = g'(c) = f'(c) - m, \text{ for all } c \in (a, b).$$

If $g$ is not constant, then some $c \in (a, b)$ will provide a local extremum and give

$$0 = g'(c) = f'(c) - m.$$

$\square$

**Corollary 9.** If $f'(x) > 0$ for all $x \in (a, b)$, then $f$ is increasing on $(a, b)$.

*Proof.* If $\tilde{a} < \tilde{b}$ lie in $(a, b)$; MVT says

$$\frac{f(\tilde{b}) - f(\tilde{a})}{\tilde{b} - \tilde{a}} = f'(c) > 0, \text{ for some } c \in (\tilde{a}, \tilde{b}).$$

Hence, $f(\tilde{b}) > f(\tilde{a})$. $\square$

**End of MATH 320 :-)**