Honours Analysis 1 (Math 254) Tutorial Notes

Edward Chernysh

(Dated: February 7, 2020)

McGill University, Montréal, Québec, Canada.

Email: edward.chernysh@mail.mcgill.ca

Contents

1 First Tutorial

During this first tutorial, we try to give several worked out examples of proofs. More importantly, we will cover some of the most common *methods* of proof. As these techniques are rather indispensable, it is a good idea for us to give as many detailed examples as possible.

1.1 Proof by Induction

When it comes to mathematical induction, I think that the best way to understand the concept is to work out examples. Nonetheless, it is important to understand the logic behind this method. Say we have a statement $P(n)$, depending on $n \in \mathbb{N}$, that we wish to prove. More precisely, let's assume we want to show that $P(n)$ is true for all $n \geq 1$. For example, $P(n)$ could be the proposition

$$\text{“}5^n - 1 \text{ is divisible by } 4\text{”}.$$

In any case, our (hoped for) proof by induction should include the following:

(i) The *base case*. That is, we should check that $P(n)$ is true for the "base value" $n = 1$.

(ii) The *inductive step*. Here, we assume that $P(n)$ is true for some $n \geq 1$. Using this assumption, we then try to prove that $P(n + 1)$ is *also* true.

If we can establish both these parts, we will have proven that $P(n)$ holds true for all $n \geq 1$. Why? Since we know that $P(1)$ is true by our base case, the inductive step with $n = 1$ ensures that $P(1 + 1) = P(2)$ is also true. But then, the inductive step would again imply that $P(2 + 1) = P(3)$ is true, and so on.

Having given this explanation, we are ready to give some detailed examples in which the technique of mathematical induction proves to be very useful.

Proposition 1.1. *For all integers $n \geq 4$, one has $2^n < n!$.*

Proof. We proceed by induction on $n \geq 4$. Since we want to prove a statement for all $n \geq 4$, our base case should be verifying the case $n = 4$.

Base case. For $n = 4$ we manually verify the claim:

$$2^4 = 16 < 24 = 4!.$$

That is, the statement $2^n < n!$ is valid when $n = 4$. This completes our base case.

Inductive step. Now we make the induction hypothesis. For some $n \geq 4$ let us assume that

$$2^n < n!. \tag{1.1}$$

We want to show that the statement also holds for $n+1$. More precisely, we want to prove the following inequality:

$$2^{n+1} < (n+1)!$$

based off of the information in (1.1). By the inductive hypothesis (1.1), we have

$$(n+1)! = (n+1) \cdot n! \overset{(1.1)}{>} (n+1) \cdot 2^n.$$

Because $n \geq 4$, we also get $(n+1) \geq 5 \geq 2$. Using this together with the equation above yields

$$(n+1)! > (n+1) \cdot 2^n \geq 2 \cdot 2^n = 2^{n+1}.$$

This completes the inductive step. $\square$

Proposition 1.2. $5^n - 1$ *is divisible by* 4 *for all* $n \in \mathbb{N}$.

Proof. Since we wish to prove a claim for all $n \geq 1$, we should take $n = 1$ as the base case.

Base case. Clearly, $5^1 - 1 = 4$ which is certainly divisible by 4. Hence, the base case is true.

Inductive step. Assume that $5^n - 1$ is divisible by 4 for some $n \geq 1$ (this is the inductive hypothesis). We must show that $5^{n+1} - 1$ is also divisible by 4. But, this is easily seen by writing

$$\begin{aligned} 5^{n+1} - 1 = 5 \cdot 5^n - 1 &= \underbrace{5^n + 5^n + 5^n + 5^n}_{4 \text{ times}} + (5^n - 1) \\ &= 4 \cdot 5^n + (5^n - 1). \end{aligned}$$

Clearly, $4 \cdot 5^n$ is divisible by 4. By the induction hypothesis, so is $5^n - 1$. Hence, 4 must be a divisor of their sum, which is equal to $5^{n+1} - 1$ by the calculations above. $\square$

1.2 Sets, Functions, and Direct Proofs

Informally, a set should be understood as an unambiguous collection of objects. For example, $\mathbb{N} = \{1, 2, 3, \dots\}$ is the set of all natural numbers. Given two sets X and Y, a *function* $f : X \to Y$ is a rule which assigns to each $x \in X$ a *unique* element $f(x) \in Y$. Note that we do not require that f be described by an explicit formula here.

Definition 1.1 (Image & Pre-image)**.** Let X and Y be sets and $f : X \to Y$ a function. For any set $A \subseteq X$, we define

$$f(A) := \{f(x) : x \in A\}$$

which is a subset of Y. This set $f(A)$ is called the *image of A under f*. Similarly, given $B \subseteq Y$, the *pre-image of B under f* is the set

$$f^{-1}(B) := \{x \in X : f(x) \in B\} .$$

That is, $f^{-1}(B)$ consists of those $x \in X$ that f takes to B.

Remark 1.1. The notation f^{-1} should **not** be confused with the inverse function of f, which in general does not exist. Even when f does not have an inverse function, we can still make sense of $f^{-1}(B)$!

In practice one often has to deal with a large collection of sets. For such purposes, the concept of an index set is *very* useful. Informally, an index set I is a set that operates as the labeling scheme for a family of sets. By way of an example, consider for each $x \in \mathbb{R}$ the singleton set $\{x\}$.[1] Then, we may want to consider the set of all these singletons, i.e.

$$\Sigma := \Big\{\{x\} : x \in \mathbb{R}\Big\}.$$

Here, Σ is a set whose elements are those sets $\{x\}$ with $x \in \mathbb{R}$. However, when dealing with more complicated sets, this notation rapidly becomes cumbersome and awkward. Instead, we could use the more general concept of an index set. Defining $A_x := \{x\}$ for $x \in \mathbb{R}$, the set Σ can instead be written as $\{A_x\}_{x \in \mathbb{R}}$. Here, the real line $\mathbb{R}$ serves as our index set.

[1] Here, $\mathbb{R}$ denotes the real number line $(-\infty, \infty)$.

The concepts of unions and intersections carry over nicely when dealing with index sets. Let I be an index set and $\{X_\alpha\}_{\alpha\in I}$ be an indexed family of sets[2], we define:

$$\bigcup_{\alpha\in I} X_\alpha := \{x : \text{there exists } \alpha \in I \text{ such that } x \in X_\alpha\}\,,$$
$$\bigcap_{\alpha\in I} X_\alpha := \{x : x \in X_\alpha \text{ for all } \alpha \in I\}\,.$$

Finally, let us recall a piece of notation: given a set X and a subset $A \subseteq X$, we denote by A^{c} the set $X \setminus A$. That is,

$$A^{\mathsf{c}} := \{x \in X : x \notin A\}\,.$$

Especially, we have $X = A \cup A^{\mathsf{c}}$ with $A \cap A^{\mathsf{c}} = \varnothing$. It should also be noted that A^{c} depends on the "parent set" X. Equipped with these concepts, we are ready to introduce de Morgan's laws:

Theorem 1.3 (de Morgan's Laws)**.** *Let X be a set and let $\{X_\alpha\}_{\alpha\in I}$ be an indexed family of subsets of X, i.e. $X_\alpha \subseteq X$ for each $\alpha \in I$. Then,*

$$\left(\bigcup_{\alpha\in I} X_\alpha\right)^{\mathsf{c}} = \bigcap_{\alpha\in I} X_\alpha^{\mathsf{c}}, \tag{1.2}$$
$$\left(\bigcap_{\alpha\in I} X_\alpha\right)^{\mathsf{c}} = \bigcup_{\alpha\in I} X_\alpha^{\mathsf{c}}. \tag{1.3}$$

Proof. We begin by establishing (1.2). Here, we have the following two inclusions to demonstrate:

$$\left(\bigcup_{\alpha\in I} X_\alpha\right)^{\mathsf{c}} \subseteq \bigcap_{\alpha\in I} X_\alpha^{\mathsf{c}} \quad \text{and} \quad \bigcap_{\alpha\in I} X_\alpha^{\mathsf{c}} \subseteq \left(\bigcup_{\alpha\in I} X_\alpha\right)^{\mathsf{c}}.$$

(1) Let $x \in (\bigcup_{\alpha\in I} X_\alpha)^{\mathsf{c}}$ be given. By definition, this means that $x \notin \bigcup_{\alpha\in I} X_\alpha$. Consequently, $x \notin X_\alpha$ for each $\alpha \in I$.[3] Put otherwise, we have $x \in X_\alpha^{\mathsf{c}}$ for every $\alpha \in I$. Therefore, $x \in \bigcap_{\alpha\in I} X_\alpha^{\mathsf{c}}$. Hence, we have proven that $(\bigcup_{\alpha\in I} X_\alpha)^{\mathsf{c}} \subseteq \bigcap_{\alpha\in I} X_\alpha^{\mathsf{c}}$.

[2]Remember, this only means that we have associated to each $\alpha \in I$ a set X_α. That is, we are using the elements of I to label the X_α's. This is particularly general because, as we shall see later on, some sets are so large that they cannot be systematically enumerated!

[3]Indeed, if x were an element of some X_α, it would also have to be an element of $\bigcup_{\alpha\in I} X_\alpha$.

(2) For the reverse inclusion, let $x \in \bigcap_{\alpha \in I} X_\alpha^{\mathsf{c}}$. By definition, this implies that $x \in X_\alpha^{\mathsf{c}}$ for every $\alpha \in I$. Equivalently,

$$x \notin X_\alpha \text{ for each index } \alpha \in I.$$

Now, $x \in \bigcup_{\alpha \in I} X_\alpha$ implies $x \in X_\alpha$ for some α. Since this contradicts the above, we must have $x \notin \bigcup_{\alpha \in I} X_\alpha$. By definition, this is simply the statement $x \in (\bigcup_{\alpha \in I} X_\alpha)^{\mathsf{c}}$. It follows that $\bigcap_{\alpha \in I} X_\alpha^{\mathsf{c}} \subseteq (\bigcup_{\alpha \in I} X_\alpha)^{\mathsf{c}}$.

This completes the proof of (1.2). Next, we turn our attention to (1.3). As above, we have two inclusions to prove:

(1) Fix a point $x \in (\bigcap_{\alpha \in I} X_\alpha)^{\mathsf{c}}$ and note that, by definition, $x \notin \bigcap_{\alpha \in I} X_\alpha$. Hence, there must exist an index $\alpha \in I$ such that $x \notin X_\alpha$. Otherwise, we would have $x \in X_\alpha$ for all $\alpha \in I$ whence $x \in \bigcap_{\alpha \in I} X_\alpha$ – which is a clear contradiction. Now, since $x \notin X_\alpha$ for some $\alpha \in I$, we also have $x \in X_\alpha^{\mathsf{c}}$ for this same α. In particular, $x \in \bigcup_{\alpha \in I} X_\alpha^{\mathsf{c}}$. This proves the inclusion $(\bigcap_{\alpha \in I} X_\alpha)^{\mathsf{c}} \subseteq \bigcup_{\alpha \in I} X_\alpha^{\mathsf{c}}$.

(2) Conversely, let $x \in \bigcup_{\alpha \in I} X_\alpha^{\mathsf{c}}$ be given. By definition of the union, this means that $x \in X_\alpha^{\mathsf{c}}$ for some $\alpha \in I$. That is, $x \notin X_\alpha$ for some index α. Consequently, we cannot have $x \in \bigcap_{\alpha \in I} X_\alpha$. Hence, $x \in (\bigcap_{\alpha \in I} X_\alpha)^{\mathsf{c}}$. We have therefore shown that $\bigcup_{\alpha \in I} X_\alpha^{\mathsf{c}} \subseteq (\bigcup_{\alpha \in I} X_\alpha)^{\mathsf{c}}$.

With this, (1.3) has been established. □

Proposition 1.4. *Let X and Y be sets and $f : X \to Y$ a function. Let I be an index set and suppose that $\{V_\alpha\}_{\alpha \in I}$ is an indexed family of subsets of Y. Then,*

$$f^{-1}\left(\bigcup_{\alpha \in I} V_\alpha\right) = \bigcup_{\alpha \in I} f^{-1}(V_\alpha).$$

Proof. There are two inclusions to be shown:

$$f^{-1}\left(\bigcup_{\alpha \in I} V_\alpha\right) \subseteq \bigcup_{\alpha \in I} f^{-1}(V_\alpha) \quad \text{and} \quad \bigcup_{\alpha \in I} f^{-1}(V_\alpha) \subseteq f^{-1}\left(\bigcup_{\alpha \in I} V_\alpha\right).$$

We begin by establishing the former. Let $x \in f^{-1}(\bigcup_{\alpha \in I} V_\alpha)$ be given. By definition, this means that

$$f(x) \in \bigcup_{\alpha \in I} V_\alpha.$$

Again, by definition, there exists an index $\alpha \in I$ such that $f(x) \in V_\alpha$. From this, it follows that $x \in f^{-1}(V_\alpha)$. This implies that $x \in \bigcup_{\alpha\in I} f^{-1}(V_\alpha)$. The first inclusion has thus been proven.

Conversely, let $x \in \bigcup_{\alpha\in I} f^{-1}(V_\alpha)$. There then exists an index $\alpha \in I$ such that $x \in f^{-1}(V_\alpha)$. By definition of the set $f^{-1}(V_\alpha)$, this is simply the statement that $f(x) \in V_\alpha$. Hence,

$$f(x) \in \bigcup_{\alpha\in I} V_\alpha$$

so that $x \in f^{-1}\left(\bigcup_{\alpha\in I} V_\alpha\right)$. This verifies the second inclusion and the proof is complete. $\square$

Proposition 1.5. *Let X, Y and Z be sets. Suppose that*

$$f : X \to Y \quad \textit{and} \quad g : Y \to Z$$

are functions. Then, for every $A \subseteq Z$, we have

$$(g \circ f)^{-1}(A) = f^{-1}\left(g^{-1}(A)\right).$$

Here, $g \circ f$ denotes the composite function $x \mapsto g(f(x))$.

Proof. First, observe that the setting of the problem makes sense because $g \circ f$ is a function $X \to Z$. Let $x \in (g \circ f)^{-1}(A)$ be given. This is to say that

$$(g \circ f)(x) = g(f(x)) \in A.$$

Hence, $f(x) \in g^{-1}(A)$. But this implies that $x \in f^{-1}(g^{-1}(A))$. This verifies the inclusion

$$(g \circ f)^{-1}(A) \subseteq f^{-1}(g^{-1}(A)).$$

Conversely, fix $x \in f^{-1}(g^{-1}(A))$. It follows that

$$f(x) \in g^{-1}(A) \implies g(f(x)) \in A.$$

Put otherwise, $(g \circ f)(x) = g(f(x)) \in A$. Thus, $x \in (g \circ f)^{-1}(A)$. Since x was arbitrary, it follows that

$$f^{-1}(g^{-1}(A)) \subseteq (g \circ f)^{-1}(A).$$

$\square$

Here are some extra (solved) problems that we may not have the time to cover during the tutorials:

Proposition 1.6. *For any sets A and B, we have*

$$(A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

Remark 1.2. Given sets A and B, the **symmetric difference** of A and B is defined by the equation:

$$A \vartriangle B := (A \setminus B) \cup (B \setminus A).$$

$A \vartriangle B$ consists of all elements that belong to A or B, but not both.

Proof. We have two inclusions to prove:

$$(A \setminus B) \cup (B \setminus A) \subseteq (A \cup B) \setminus (A \cap B)$$
$$(A \cup B) \setminus (A \cap B) \subseteq (A \setminus B) \cup (B \setminus A).$$

(1) We begin with the first inclusion. Let $x \in (A \setminus B) \cup (B \setminus A)$ so that it belongs to $(A \setminus B)$ or $(B \setminus A)$. Without loss of generality, we can assume that $x \in A \setminus B$.[4] Then, $x \in A$ but $x \notin B$. This means that

$$x \in A \cup B \quad \text{and} \quad x \notin A \cap B.$$

Hence, $x \in (A \cup B) \setminus (A \cap B)$.

(2) Conversely, let $x \in (A \cup B) \setminus (A \cap B)$. This simply means that

$$x \in A \cup B \quad \text{and} \quad x \notin A \cap B.$$

Without loss of generality, we can assume that $x \in A$. Since $x \notin A \cap B$, we cannot have $x \in B$. That is, $x \notin B$. This implies that $x \in A \setminus B$. More generally,

$$x \in (A \setminus B) \subseteq (A \setminus B) \cup (B \setminus A).$$

This completes the proof. □

[4]Otherwise, $x \in B \setminus A$. In this case, we just relabel A as B and B as A!

1.3 More Examples of Induction

Depending on much time was left, we may not have touched upon all (if any) of the following problems. Nonetheless, it is probably useful to try these out.

Proposition 1.7. *There holds*

$$3 + 11 + \cdots + (8n - 5) = 4n^2 - n$$

for all $n \in \mathbb{N}$.

Proof. For $n \in \mathbb{N}$, the proposition/statement depending on n we wish to prove is the equality

$$3 + 11 + \cdots + (8n - 5) = 4n^2 - n.$$

Since we wish to prove the above for all $n \geq 1$, our base case should be $n = 1$.

Base case. We verify directly that the statement is true for $n = 1$:

$$3 = 8 \cdot 1 - 5 = 4 \cdot 1^2 - 1.$$

Inductive step. Our induction hypothesis will be that the statement is true for some $n \geq 1$. Our job is then to show that the claim holds for $n + 1$. The induction hypothesis means that for this n:

$$3 + 11 + \cdots + (8n - 5) = 4n^2 - n. \tag{1.4}$$

We now deduce that:

$$3 + 11 + \cdots + (8(n + 1) - 5) = 4(n + 1)^2 - (n + 1).$$

First, our inductive hypothesis in (1.4) gives

$$\begin{aligned} 3 + 11 + \cdots + (8(n + 1) - 5) &= \underbrace{3 + 11 + \cdots + (8n - 5)}_{=4n^2-n \text{ by } (1.4)} + (8(n + 1) - 5) \\ &= 4n^2 - n + 8(n + 1) - 5 \\ &= 4n^2 + 7n + 3. \end{aligned}$$

On the other hand,

$$\begin{aligned} 4(n + 1)^2 - (n + 1) &= 4\left(n^2 + 2n + 1\right) - n - 1 \\ &= 4n^2 + 8n + 4 - n - 1 \\ &= 4n^2 + 7n + 3. \end{aligned}$$

Thus,

$$3 + 11 + \cdots + (8(n+1) - 5) = 4n^2 + 7n + 3 = 4(n+1)^2 - (n+1).$$

This completes the inductive step. $\square$

Lemma 1.8. *$2n + 1 \leq 2^n$ for all integers $n \geq 3$.*

Proof. This looks like something we should try to prove with induction.

Base case. With $n = 3$, we check directly that

$$2(3) + 1 = 7 < 8 = 2^3.$$

This means that the base case is true.

Inductive step. Suppose that the statement is true for some $n \geq 3$; we must prove that it is also true for $n + 1$. Namely, our induction hypothesis is:

$$2n + 1 \leq 2^n. \tag{1.5}$$

We want to show that $2(n+1) + 1 \leq 2^{n+1}$. First, write

$$2(n+1) + 1 = 2n + 2 + 1 = (2n+1) + 2 \overset{(1.5)}{\leq} 2^n + 2. \tag{1.6}$$

In the last step, we used the induction hypothesis (1.5) to get $2n + 1 \leq 2^n$. Since $n \geq 3$, it is clear that $2 \leq 2^n$. Therefore, (1.6) implies that

$$2(n+1) + 1 \leq 2^n + 2 \leq 2^n + 2^n = 2 \cdot 2^n = 2^{n+1}.$$

This completes the inductive step and the claim is proven. $\square$

Using this proposition, we prove a far more interesting identity.

Proposition 1.9. *If $n \geq 5$ is an integer, then $n^2 < 2^n$.*

Proof. Since we are asked to prove an equality involving integers, it is a good idea to try and use mathematical induction. We start with the base case $n = 5$.

Base case. With $n = 5$ we check that

$$5^2 = 25 < 32 = 2^5.$$

Hence, the claim holds for $n = 5$.

<u>Inductive step</u>. Suppose the claim is true for $n \geq 5$, we must show that is is also true for $n + 1$. Our inductive hypothesis is therefore the following:

$$n^2 < 2^n. \tag{1.7}$$

Using this, we must show that

$$(n+1)^2 < 2^{n+1}.$$

Using the induction hypothesis (1.7), we calculate

$$(n+1)^2 = n^2 + 2n + 1 \overset{(1.7)}{<} 2^n + (2n+1). \tag{1.8}$$

Because $n \geq 5 \geq 3$, we can apply the previous lemma to (1.8) above. Doing so gives $(2n + 1) \leq 2^n$ whence

$$(n+1)^2 < 2^n + (2n+1) \leq 2^n + 2^n = 2 \cdot 2^n = 2^{n+1}.$$

This completes the inductive step. □

2 *Second Tutorial*

We would first like to recall some notions regarding functions. If X and Y are sets, a *bijection* from X to Y is a function $f : X \to Y$ satisfying each of the following properties:

(i) f is *injective*, i.e. $f(x) = f(x')$ implies $x = x'$;

(ii) f is *surjective*, i.e. for every $y \in Y$ there exists $x \in X$ such that $f(x) = y$.

Put otherwise, a function $f : X \to Y$ is a bijection *if and only if* for every $y \in Y$ there exists a *unique* point $x \in X$ such that $f(x) = y$.

Example 2.1. The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is not a bijection because it is neither injective nor surjective. Certainly, injectivity fails because $f(-1) = 1 = f(1)$. Surjectivity does not hold because there does not exist $x \in \mathbb{R}$ such that $f(x) = -1$.

Even so, we must be careful as a function that is not surjective can become surjective if we restrict its range. Consider the function

$$f : \mathbb{R} \to [0, \infty), \quad f(x) = x^2.$$

This is the same function as in the previous example, but with a "smaller" range. Although the function is still not injective ($f(-1) = f(1)$), it is surjective! Indeed, given $y \in [0, \infty)$, we have that $f(\sqrt{y}) = y$. Similarly, a function can become injective when we restrict its domain. Certainly, the function

$$f : [0, \infty) \to [0, \infty), \quad f(x) = x^2$$

is actually a bijection.

2.1 Properties of Functions

The concept of an inverse function arises naturally when discussing bijective functions. If X and Y are sets and $f : X \to Y$ is a bijection, there exists for each $y \in Y$ a *unique* point x with the property that $f(x) = y$. Therefore, we can construct a function $f^{-1} : Y \to X$ by simply defining $f^{-1}(y) := x$. This map f^{-1} is called the *inverse function* of f.

The following property of the inverse function is almost immediate from the definition we have just given above:

Proposition 2.1. *Let X and Y be sets and $f : X \to Y$ a bijection. Let f^{-1} be the inverse of f, as defined above. Then,*

(1) $(f^{-1} \circ f)(x) = x$ for all $x \in X$;

(2) and $(f \circ f^{-1})(y) = y$ for all $y \in Y$.

Proof. The claim in (1) follows at once from the definition of f^{-1}. We therefore need only verify (2). Given $y \in Y$, there exists (since f is bijective) a unique point $x \in X$ such that $f(x) = y$. But, by definition of f^{-1}, we must have $f^{-1}(y) = x$. Hence, $(f \circ f^{-1})(y) = f(f^{-1}(y)) = f(x) = y$. $\square$

Of course, the next result is also to be expected:

Proposition 2.2. *Let X and Y be sets and $f : X \to Y$ a bijection. Let $f^{-1} : Y \to X$ be the inverse function of f. Then, f^{-1} is a bijection $Y \to X$.*

Proof. We first check that f^{-1} is surjective. This amounts to showing that for every $x \in X$, there exists $y \in Y$ with the property that $f^{-1}(y) = x$. Given $x \in X$, we simply take $y = f(x) \in Y$. Then, the previous result tells us that

$$f^{-1}(y) = f^{-1}(f(x)) = x.$$

Therefore, f is surjective. To prove injectivity we assume that $f^{-1}(y) = f^{-1}(y')$ for some $y, y' \in Y$. We want to show that this forces $y = y'$. To this end, we "apply" the function f to both sides of the equality $f^{-1}(y) = f^{-1}(y')$. Doing so gives

$$f(f^{-1}(y)) = f(f^{-1}(y')) \implies y = y'$$

by virtue of (2) in Proposition 2.1. Hence, f^{-1} is injective. □

Proposition 2.3. *Let A, B and C be sets, and let $f : A \to B$ and $g : B \to C$ be functions. There hold the following:*

(1) if f and g are injective, then so is $g \circ f$;

(2) if $g \circ f$ is injective, then so is f;

(3) if f and g are surjective, then so is $g \circ f$;

(4) if $g \circ f$ is surjective, then so is g.

Proof.

(1) Let $a, a' \in A$ and suppose that $(g \circ f)(a) = (g \circ f)(a')$; our goal is to prove that $a = a'$. Now, this equation is equivalent to $g(f(a)) = g(f(a'))$. Because g is injective, this implies $f(a) = f(a')$. But, f is also injective! Consequently, it follows that $a = a'$.

(2) Suppose that $g \circ f$ is injective and let $a, a' \in A$ be such that $f(a) = f(a')$. We want to deduce that $a = a'$. Let us now apply the function g to both sides of the equality $f(a) = f(a')$. Doing so gives

$$(g \circ f)(a) = g(f(a)) = g(f(a')) = (g \circ f)(a').$$

Because $g \circ f$ is injective, it follows that $a = a'$. We conclude that f is injective.

(3) Let f and g both be surjective. For any point $c \in C$, we can choose $b \in B$ such that $g(b) = c$. Since f is surjective, there exists $a \in A$ with the property that $f(a) = b$. Then,

$$(g \circ f)(a) = g(f(a)) = g(b) = c.$$

(4) If $g \circ f$ is surjective, given any point $c \in C$ there exists $a \in A$ such that $(g \circ f)(a) = c$. This is simply the statement that $g(f(a)) = c$. Taking $b := f(a) \in B$, we see that $g(b) = c$. Hence, g is surjective.

□

Remark 2.1. Despite the above, there do exist functions f and g such that
(i) f is not surjective;
(ii) g is not injective;
(iii) and $g \circ f$ is a bijection.
Indeed, let $\mathbb{E}$ denote the set of all even natural numbers $\{2, 4, 6, \ldots, \}$. Consider the functions:

$$f : \mathbb{E} \to \mathbb{N}, \quad f(n) = n,$$
$$g : \mathbb{N} \to \mathbb{E}, \quad g(n) = \begin{cases} n & \text{if } n \text{ is even} \\ 2n & \text{if } n \text{ is odd.} \end{cases}$$

Clearly, f is injective but not surjective. Because $g(1) = g(2) = 2$, we see that g is not injective. But, $g \circ f$ is simply the identity map

$$g \circ f : \mathbb{E} \to \mathbb{E}, \quad (g \circ f)(n) = n$$

which is a bijection (check this yourself).

Proposition 2.3 has many corollaries, some of which are listed below.

Corollary 2.4. *Composition to the right or left by an injection or a surjection does not affect the injectivity or surjectivity of a bijection, respectively. Namely, the following properties hold:*

(1) *Let A, B and C be sets. Suppose that $f : A \to B$ is a function and $g : B \to C$ is a bijection. Then, $g \circ f$ is injective whenever f is. Moreover, $g \circ f$ is surjective whenever f is.*

(2) *Let A, B and C be sets. Suppose that $f : A \to B$ is a bijection and $g : B \to C$ a general function. Then, $g \circ f$ is injective whenever g is. Furthermore, $g \circ f$ is surjective if g is.*

Proof. Let us prove (1). First, note that g is both injective and surjective. If f is injective, then $g \circ f$ is injective by part (1) of the previous problem. If f is surjective, then so is $g \circ f$ (by part (3) of the last problem). Part (2) is verified in a similar way. We leave the details as an exercise to the reader. □

Similarly, we can deduce the following.

Corollary 2.5. *Let A, B and C be sets and $f : A \to B$, $g : B \to C$ functions. If any two of f, g and $g \circ f$ are bijective, then so is the third.*

Proof. If f and g are bijective, then parts (1) and (3) of Proposition 2.3 tell us that $g \circ f$ is also a bijection. Suppose that f and $g \circ f$ are bijections. Then, f has an inverse function f^{-1}. By Proposition 2.2, this inverse is also a bijection. Finally, part (1) of Corollary 2.4 implies that

$$g = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1}$$

is bijective. Similarly, f can be shown to be bijective whenever g and $g \circ f$ are. □

Theorem 2.6. *Let X and Y be sets and $f : X \to Y$ a function. Then, f is injective if and only if there exists a function (called a left-inverse) $g : Y \to X$ such that*

$$g \circ f : X \to X$$

satisfies $(g \circ f)(x) = x$ for all $x \in X$.

Proof. Suppose that f is injective. If $y \in f(X)$, then there exists a unique point $x \in X$ such that $y = f(x)$. Let us call this *unique* point $f^{-1}(y)$. Define a function

$$g : Y \to X, \quad g(y) := \begin{cases} f^{-1}(y), & \text{if } y \in f(X), \\ x_0, & \text{if } y \notin f(X). \end{cases}$$

Here, $x_0 \in X$ is **any** fixed point. Let $x \in X$ be given; we want to verify that

$$(g \circ f)(x) = g(f(x)) = x.$$

Clearly, $y = f(x)$ belongs to $f(X)$. By definition of the function g,

$$g(f(x)) = g(y) = f^{-1}(y) = x.$$

Conversely, suppose that such a function g exists. The composition $g \circ f$ is then equal to the identity map $h(x) = x$ on X. Thus, $g \circ f$ is an injection $X \to X$. By part (2) of Proposition 2.3, we see that f is injective. □

Using a similar argument, one can prove the following:

Exercise 2.1. Let X and Y be sets and let $f : X \to Y$ be a function. Then, f is surjective if and only if there exists a function $g : Y \to X$ (called a right-inverse) such that

$$(f \circ g)(y) = y$$

for all $y \in Y$.

2.2 *Cauchy-Schwarz and the Triangle Inequality*

The remaining problems for this tutorial are meant to serve as extra examples that use some of what you've proven in the first assignment. As an added bonus, we will be proving inequalities that you will probably find to be useful throughout your entire mathematical career. At the very least, the first inequality should help you in Advanced Calculus and Honours Analysis 2.

Let us recall Problem 4 in your first assignment, which states that

$$\left(\sum_{j=1}^{n} a_j b_j\right)^2 \leq \left(\sum_{j=1}^{n} a_j^2\right)\left(\sum_{j=1}^{n} b_j^2\right) \tag{2.1}$$

for all $n \in \mathbb{N}$ and all $a_j, b_j \in \mathbb{R}$. In fact, this implies that

$$\sum_{j=1}^{n} a_j b_j \leq \left|\sum_{j=1}^{n} a_j b_j\right| \leq \left(\sum_{j=1}^{n} a_j^2\right)^{1/2} \left(\sum_{j=1}^{n} b_j^2\right)^{1/2}. \tag{2.2}$$

Let $x = (x_1, \ldots, x_n)$ be a point in $\mathbb{R}^n$ for $n \geq 1$. The *norm* of this point x is defined via the equation

$$\|x\| := \left(x_1^2 + \cdots + x_n^2\right)^{1/2} \geq 0 \tag{2.3}$$

If $y = (y_1, \ldots, y_n)$ is another point in $\mathbb{R}^n$, we define the *dot product* of x and y according to the formula

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n. \tag{2.4}$$

This is an example of what is often called an *inner product*. Returning to the equation above, we see that

$$x \cdot x = x_1^2 + \cdots + x_n^2 \geq 0.$$

Note that $\|x\|^2 = x \cdot x$ whence $\|x\| = \sqrt{x \cdot x}$. This relationship means that the norm $\|x\|$ is *induced* by the dot product in $\mathbb{R}^n$ (this very important topic will be further explored in Honours Analysis 4). For now, we can at least give a proof of the triangle inequality:

Theorem 2.7 (Triangle Inequality)**.** *For any $n \geq 1$, one has $\|x + y\| \leq \|x\| + \|y\|$ for all vectors $x, y \in \mathbb{R}^n$.*

Proof. By our earlier observations,

$$\begin{aligned}
\|x + y\|^2 = (x + y) \cdot (x + y) &= \sum_{j=1}^{n} (x_j + y_j)^2 \\
&= \sum_{j=1}^{n} x_j^2 + 2 \sum_{j=1}^{n} x_j y_j + \sum_{j=1}^{n} y_j^2 \\
&= \|x\|^2 + 2 \sum_{j=1}^{n} x_j y_j + \|y\|^2 .
\end{aligned}$$

Invoking (2.2), we get

$$\begin{aligned}
\|x + y\|^2 &\leq \|x\|^2 + 2 \left(\sum_{j=1}^{n} x_j^2 \right)^{1/2} \left(\sum_{j=1}^{n} y_j^2 \right)^{1/2} + \|y\|^2 \\
&= \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2 .
\end{aligned}$$

Taking square roots gives us the desired inequality. □

2.3 *The AM-GM Inequality*

Theorem 2.8. *For each $n \geq 1$, and all $x_1, \dots, x_n > 0$, there holds*

$$\frac{x_1 + \cdots x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n}. \tag{2.5}$$

Remark 2.2. The quantity on the left hand side of (2.5) is called the *arithmetic mean* of $(x_1, \dots, x_n)$. The right hand term is instead called the *geometric mean* of $(x_1, \dots, x_n)$.

To prove this result, we will first require a lemma that is very interesting in its own right.

Lemma 2.9. *Let $n \in \mathbb{N}$ and $x_1, \ldots, x_n > 0$. If $x_1 \cdots x_n = 1$, then*

$$x_1 + \cdots + x_n \geq n.$$

Proof. We prove this inequality by induction on $n \geq 1$. Luckily for us, the base case with $n = 1$ is obvious.

Inductive Step. Suppose that the claim holds for an $n \geq 1$. We will prove that the statement holds for $n + 1$. Thus, our inductive hypothesis is the statement:

$$\text{“}(\forall x_1, \ldots, x_n > 0),\ (x_1 \cdots x_n = 1 \implies x_1 + \cdots + x_n \geq n.)\text{”}$$

Let now $x_1, \ldots, x_n, x_{n+1} > 0$ be real numbers such that $x_1 \cdots x_n x_{n+1} = 1$. We will show that

$$x_1 + \cdots + x_n + x_{n+1} \geq n + 1.$$

If every $x_j = 1$, then the statement is obvious. Thus, we can assume that there exist indices $1 \leq i, j \leq n + 1$ such that $x_i > 1$ and $x_j < 1$.[5] After a relabeling, we can assume that $x_1 < 1$ and $x_{n+1} > 1$. Because $(x_{n+1} - 1)(1 - x_1) > 0$, we must have

$$x_1 + x_{n+1} > 1 + x_1 x_{n+1}. \tag{2.6}$$

Defining $y := x_1 x_{n+1} > 0$, the inequality above can then be written as

$$x_1 + x_{n+1} > 1 + y. \tag{2.7}$$

Note that $y x_2 \cdots x_n = 1$. Since $\{y, x_2, \ldots, x_n\}$ consists of n positive numbers, our inductive hypothesis implies that

$$y + x_2 + \cdots + x_n \geq n. \tag{2.8}$$

Using this together with (2.7) yields

$$\begin{aligned}
x_1 + x_2 + \cdots + x_n + x_{n+1} &= (x_1 + x_{n+1}) + x_2 + \cdots + x_n \\
&> (1 + y) + x_2 + \cdots + x_n \qquad \text{(by (2.7))} \\
&= 1 + (y + x_2 + \cdots + x_n) \\
&\geq 1 + n \qquad \text{(by (2.8))} \\
&= n + 1.
\end{aligned}$$

[5] This is because $x_1 x_2 \cdots x_{n+1} = 1$. Indeed, if $x_i > 1$ for some index i and $x_j \geq 1$ for all indices $j \neq i$, then we would have $x_1 \cdots x_{n+1} > x_1 \cdots x_{i-1} x_{i+1} \cdots x_{n+1} \geq 1$ which is a contradiction. Similarly, if some $x_i < 1$, there must be another x_j with $x_j > 1$.

This completes the inductive step. $\square$

We are finally ready to give a nice proof of Theorem 2.8.

Proof of Theorem 2.8. Given $x_1, \ldots, x_n > 0$ let us put $A := (x_1 \cdots x_n)^{\frac{1}{n}} > 0$. For each $1 \leq j \leq n$ define

$$a_j := \frac{x_j}{A} > 0$$

Then, $a_1 \cdots a_n = 1$. By the previous lemma, we get that

$$a_1 + \cdots + a_n \geq n.$$

Or, rather, that

$$\frac{x_1 + \cdots + x_n}{A} = \frac{x_1 + \cdots + x_n}{(x_1 \cdots x_n)^{\frac{1}{n}}} \geq n.$$

Therefore,

$$\frac{x_1 + \cdots + x_n}{n} \geq (x_1 \cdots x_n)^{\frac{1}{n}}$$

as was asserted. $\square$

3 Third Tutorial

Let X and Y be non-empty sets. We say that X and Y have the same *cardinality* if there exists a bijection $f : X \to Y$. In this case, we write $|X| = |Y|$. If there exists an injective function $X \to Y$ but no bijection $X \to Y$ exists, we write $|X| < |Y|$. Remember that X is called *countably infinite* if $|X| = |\mathbb{N}|$. A set X is said to be **countable** if it is either finite or countably infinite.

Proposition 3.1. *The set* $\mathbb{N} \times \mathbb{N}$ *is countable.*

First Proof. Clearly, $\mathbb{N} \times \mathbb{N}$ is not finite. Thus, our only option is to show that $\mathbb{N} \times \mathbb{N}$ is countably infinite. Next, we observe that the elements of $\mathbb{N} \times \mathbb{N}$ can be listed according to the following table:

$$\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & \cdots & (1,n) & \cdots \\
(2,1) & (2,2) & (2,3) & \cdots & (2,n) & \cdots \\
(3,1) & (3,2) & (3,3) & \cdots & (3,n) & \cdots \\
(4,1) & (4,2) & (4,3) & \cdots & (4,n) & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \cdots \\
(m,1) & (m,2) & (m,3) & \cdots & (m,n) & \cdots \\
\vdots & \vdots & \vdots & & \vdots & \ddots
\end{array}$$

We then enumerate the elements of $\mathbb{N} \times \mathbb{N}$ (i.e. the elements of the list above) according to the so-called "diagonal rule". Namely, we list the elements of $\mathbb{N} \times \mathbb{N}$ in the following order:

$$(1,1),\ (2,1),\ (1,2),\ (3,1),\ (2,2),\ (1,3),\ldots$$

Pictorially, we are drawing arrows through the diagonals of the list and "enumerating" the elements of $\mathbb{N} \times \mathbb{N}$ by following the arrows. In doing so, we are actually constructing a function $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$. This function would satisfy

$$1 \mapsto (1,1), \quad 2 \mapsto (2,1), \quad 3 \mapsto (1,2)\ldots.$$

Now, this function is clearly surjective[6]. Thus, we have shown that $\mathbb{N} \times \mathbb{N}$ is countable. $\square$

Second Proof. We define a function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ according to the formula $f(m,n) = 2^m 3^n$. Suppose that $2^m 3^n = 2^r 3^s$ for natural numbers m, n, r, s. By the uniqueness of factorization into primes, we must have $m = r$ and $n = s$. Thus, f is injective whence $\mathbb{N} \times \mathbb{N}$ is countable (see Theorem 1.3.10 in Bartle). $\square$

Using the countability of $\mathbb{N} \times \mathbb{N}$, we are able to deduce the following easy fact.

Proposition 3.2. *If X and Y are countable sets, then so is $X \times Y$.*

Proof. Because X and Y are both countable, we can find surjective functions (see Theorem 1.3.10 in Bartle)

$$f : \mathbb{N} \to X \quad \text{and} \quad g : \mathbb{N} \to Y.$$

Now consider the function $h : \mathbb{N} \times \mathbb{N} \to X \times Y$ given by the formula

$$h(m,n) = (f(m), g(n)).$$

Let $(x,y) \in X \times Y$ be a point. Because f and g are surjective, there exist $m, n \in \mathbb{N}$ such that $f(m) = x$ and $g(n) = y$. Thus, $h(m,n) = (x,y)$. Hence, h is surjective. By the previous proposition, we know there exists a surjection $\psi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$. Then, the composite $h \circ \psi$ will be a surjective function $\mathbb{N} \to X \times Y$. $\square$

Using a simple inductive argument, we can actually conclude that finite products of countable sets are again countable. However, we should first discuss some of the unfortunate intricacies of notation.

[6] Since every element of $\mathbb{N} \times \mathbb{N}$ can be found on one of these diagonal lines, our function will "eventually reach" any given element of $\mathbb{N} \times \mathbb{N}$.

3.1 Cartesian Products

Let $X_1, \ldots, X_n$ be sets. We define $X_1 \times \cdots \times X_n$ to be the set of all n-tuples

$$\{(x_1, \ldots, x_n) : x_j \in X_j,\ 1 \le j \le n\}.$$

To simplify notation, we often denote the product $X_1 \times \cdots \times X_n$ by

$$\prod_{k=1}^{n} X_k.$$

Thus, when we have $n+1$ sets $X_1, \ldots, X_n, X_{n+1}$,

$$X_1 \times \cdots \times X_n \times X_{n+1} \ne (X_1 \times \cdots \times X_n) \times X_{n+1}. \tag{3.1}$$

Indeed, the set on the left has elements the form $(x_1, \ldots, x_{n+1})$ but the product on the right consists of tuples of the form $((x_1, \ldots, x_n), x_{n+1})$. However, there always exists a bijection between these two sets. Certainly, consider the function

$$f : (X_1 \times \cdots \times X_n) \times X_{n+1} \to X_1 \times \cdots \times X_{n+1}$$

given by $f((x_1, \ldots, x_n), x_{n+1}) = (x_1, \ldots, x_n, x_{n+1})$. Clearly, f is surjective. To see that it is injective, suppose that

$$(x_1, \ldots, x_{n+1}) = (x'_1, \ldots, x'_{n+1}) .$$

Then, $x_j = x'_j$ for all $1 \le j \le n+1$. In particular, $x_{n+1} = x'_{n+1}$ and

$$(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n).$$

Therefore, $((x_1, \ldots, x_n), x_{n+1}) = \left((x'_1, \ldots, x'_n), x'_{n+1}\right)$. This shows that f is injective, and hence a bijection.

Since both the sets in (3.1) are always in bijection, there is no reason to distinguish between the two when discussing cardinality.

Proposition 3.3. *If $X_1, \ldots, X_n$ are countable sets, then so is $X_1 \times \cdots \times X_n$.*

Proof. We argue by induction on n. The base case $n = 1$ is obvious. Thus, assume that the claim holds true for $n \ge 1$ and let $X_1, \ldots, X_n, X_{n+1}$ be countable sets. By the inductive hypothesis, the set

$$X_1 \times \cdots \times X_n$$

is countable. By Proposition 3.2, so must be the product

$$(X_1 \times \cdots \times X_n) \times X_{n+1} \text{ “=” } X_1 \times \cdots \times X_n \times X_{n+1}.$$

This completes the proof. Here we are writing “=” instead of = because, although the sets are not equal per se, there exists a bijection between the two. Thus, these sets can be treated as the “same set” when discussing cardinality. □

Proposition 3.4. *Let X and Y be non-empty sets with $Y \subseteq X$. There exists a surjective function $X \to Y$.*

Proof. This is not difficult to prove. Fix any point $y_0 \in Y$ and define a function

$$\phi : X \to Y, \quad \phi(x) := \begin{cases} x & \text{if } x \in Y, \\ y_0 & \text{if } x \notin Y. \end{cases}$$

Obviously, ϕ is the desired surjective map. □

Proposition 3.5. *Let X and Y be sets with $Y \subseteq X$. If X is countable, then so is Y.*

Proof. Because X is countable (i.e. finite or countably infinite) there exists a surjective function $f : \mathbb{N} \to X$. By Proposition 3.4, we may choose a surjective map $g : X \to Y$. Then, the composite $g \circ f$ is a surjective map $\mathbb{N} \to Y$. Thus, Y is countable. □

3.2 The Power Set

Let X be a set (possibly empty). The *power set of* X, denoted $\mathcal{P}(X)$, is defined to be the **set** of all subsets of X. Symbolically,

$$\mathcal{P}(X) := \{A : A \subseteq X\}.$$

Note that $\varnothing$ is always a subset of X, regardless of the set X. Even if X is empty, $\varnothing$ will be a subset of X. This means that $\varnothing \in \mathcal{P}(X)$ for all sets X. In particular, $\mathcal{P}(X)$ is never empty.

Example 3.1. We compute the power set for very small sets. First, we consider $\mathcal{P}(\varnothing)$. We have already seen that $\varnothing \in \mathcal{P}(\varnothing)$. Because $\varnothing$ is empty, it cannot have any non-empty subsets. Therefore, $\mathcal{P}(\varnothing) = \{\varnothing\}$.[7]

[7] Be warned that $\{\varnothing\}$ is a set containing the empty set. It is **not** the empty set, but rather a set whose single element is the empty set!

Example 3.2. We compute the power set of a singleton $\{x\}$. As always, $\varnothing$ is a subset of $\{x\}$ and hence $\varnothing \in \mathcal{P}(X)$. But, $\{x\}$ is also a subset of X. Thus,

$$\mathcal{P}(\{x\}) = \{\varnothing, \{x\}\}.$$

This argument shows that for a general set X there holds $\varnothing, X \in \mathcal{P}(X)$.

These examples tell us that $|\mathcal{P}(\varnothing)| = 1$ and $|\mathcal{P}(\{x\})| = 2$. The next problem generalizes this property.

Proposition 3.6. *If X is a finite set, then $|\mathcal{P}(X)| = 2^{|X|}$. Here, $|X|$ denotes the cardinality of the set X.*

Proof. As this is a statement depending on $n \in \mathbb{N}_0$, it is a good idea to prove the claim by induction.

Base case(s). In the previous examples we verified directly the cases $n = 0$ and $n = 1$. Therefore, our base case is already complete.

Inductive Step. We now make the induction hypothesis: assume that $|\mathcal{P}(X)| = 2^n$ for all sets X having n-elements. We want to show that $|\mathcal{P}(Y)| = 2^{n+1}$ for all sets Y having $(n+1)$-elements. Let Y be a set with $(n+1)$-elements. To count the elements of $\mathcal{P}(Y)$ is to count the subsets of Y. Fix a point $y \in Y$ and define $X := Y \setminus \{y\}$; this set X has n-elements. Now, there are **exactly** two types of subsets of Y:

(i) subsets of Y that contain y;

(ii) subsets of Y that do not contain y.

Clearly,

$$\begin{aligned} |\mathcal{P}(Y)| &= \#\text{subsets of } Y \\ &= \#\text{subsets of } Y \text{ that contain } y + \#\text{subsets of } Y \text{ not containing } y. \end{aligned}$$

Let $A \subseteq Y$ be a subset *not* containing y. Then, $A \subseteq X = Y \setminus \{y\}$. By our induction hypothesis, $|\mathcal{P}(X)| = 2^n$. Thus, there are exactly 2^n possibilities for A or, equivalently, 2^n-subsets of Y that don't contain y.

On the other hand, suppose that B is a subset of Y that does contain y. Then, $B \setminus \{y\}$ does not contain y. By the previous part, there are exactly 2^n-distinct possibilities for $B \setminus \{y\}$. But this means that there are precisely 2^n-distinct possibilities for B. Hence, there are 2^n-subsets of Y that contain y. This implies that

$$\begin{aligned} |\mathcal{P}(Y)| &= \#\text{subsets of } Y \text{ that contain } y + \#\text{subsets of } Y \text{ not containing } y \\ &= 2^n + 2^n = 2^{n+1}. \end{aligned}$$

This completes the inductive step. ☐

Corollary 3.7. $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n$ *for all* $n \geq 0$.

Proof. Let X be any set containing n-elements.[8] For a natural number $0 \leq k \leq n$, let N_k denote the number of *distinct* subsets of X having k-elements. That is, put

$$N_k := |\{A \subseteq X : |A| = k\}|.$$

Then, since $\binom{n}{k} = N_k$ for each $0 \leq k \leq n$, we see from the previous proposition that

$$\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} N_k = |\mathcal{P}(X)| = 2^n.$$

This completes the proof. ☐

3.3 Cantor's Theorem

We have already discussed the power set of a given set X. We will now show that there never exists a surjective function between X and its power set $\mathcal{P}(X)$.

Theorem 3.8 (Cantor)**.** *Let X be a set. There does not exist a surjection $X \to \mathcal{P}(X)$. In particular, no bijection $X \to \mathcal{P}(X)$ exists. Hence, $|X| < |\mathcal{P}(X)|$.*[9]

Proof. We argue by contradiction. Let $f : X \to \mathcal{P}(X)$ be a surjection. Consider the following subset of X:

$$A := \{x \in X : x \notin f(x)\}.$$

[8] For instance, take $X = \{1, 2, \ldots, n\}$.

[9] The map $f : X \to \mathcal{P}(X)$ given by $f(x) = \{x\}$ is clearly an injection from $X \to \mathcal{P}(X)$. Since no bijection $X \to \mathcal{P}(X)$ exists, we may indeed write $|X| < |\mathcal{P}(X)|$.

Because f is surjective, there exists $a \in X$ such that $f(a) = A$. There are now two cases to be considered.

(1) Suppose $a \in A$. By definition of A, this forces $a \notin f(a)$. But, $f(a) = A$. Therefore, $a \notin f(a)$ implies $a \notin A$. This is a contradiction.

(2) Assume $a \notin A$. That is, $a \notin f(a)$ whence $a \in A$. This is also a contradiction.

In either case we have a contradiction. Hence, no surjective function f can exist. $\square$

This subtle theorem has grandiose implications. Essentially, it states that one can build arbitrarily large sets by simply taking more and more power sets. Consider for instance the infinite set $\mathbb{N}$. You have probably seen that there does not exist a bijection $\mathbb{N} \to \mathbb{R}$. In this sense, we say that $\mathbb{R}$ is of a "larger infinity" than $\mathbb{N}$.[10] Cantor's theorem states that $\mathcal{P}(\mathbb{N})$ is of a "larger infinity" than $\mathbb{N}$. But then, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$ is of an "even larger infinity" than $\mathcal{P}(\mathbb{N})$. Continuing in this way, we see that there are "infinitely many distinct infinities".

Corollary 3.9 (Russel's Paradox)**.** *The set of all sets does not exist.*

Proof. We argue by contradiction. Suppose that the set of all sets exists, and call it X. Then, $\mathcal{P}(X)$ is a set of subsets of X. Hence, every element of $\mathcal{P}(X)$ is a set. Thus, $\mathcal{P}(X) \subseteq X$. By Proposition 3.4, there exists a surjective function $X \to \mathcal{P}(X)$. But this contradicts Cantor's theorem. $\square$

Proposition 3.10. *A binary sequence is a "list of points"*

$$a_1, a_2, a_3, \ldots, a_n, \ldots$$

such that each a_j belongs to the set $\{0, 1\}$. Usually, one denotes such a sequence by (a_n). Let $\mathcal{B}$ be the set of all binary sequences. Then, $\mathcal{B}$ is uncountable.

Proof. We argue by contradiction. Suppose that $\mathcal{B}$ is countable, then we can

[10]This can be formalized in the sense of cardinal numbers. However, this is a topic for discussion in a dedicated set theory or logic course.

enumerate the elements of $\mathcal{B}$:

$$\begin{gathered}
(a_n^{(1)}) = a_1^{(1)}, a_2^{(1)}, a_3^{(1)} \dots \\
(a_n^{(2)}) = a_1^{(2)}, a_2^{(2)}, a_3^{(2)} \dots \\
\vdots \\
(a_n^{(m)}) = a_1^{(m)}, a_2^{(m)}, a_3^{(m)} \dots \\
\vdots
\end{gathered}$$

Here, each $(a_n^{(m)})$ represents a sequence in $\mathcal{B}$. To derive a contradiction, we will construct an element in $\mathcal{B}$ (i.e. a binary sequence) that does not belong to the list above. For a given $n \in \mathbb{N}$, let b_n be given by the following formula:

$$b_n := \begin{cases} 1 & \text{if } a_n^{(n)} = 0, \\ 0 & \text{if } a_n^{(n)} = 1. \end{cases}$$

Grouping together all the b_n into a list gives us a sequence (b_n) in $\mathcal{B}$. We claim that (b_n) is not equal to any $(a_n^{(m)})$ appearing in the list above. If $(b_n) = (a_n^{(m)})$ for some $m \geq 1$, then $b_m = a_m^{(m)}$ for this m. But this would directly contradict the construction of (b_n). Hence, $(b_n) \notin \mathcal{B}$ which contradicts our assumption that $\mathcal{B}$ was countable. □

Sets of Functions

Proposition 3.11. *Let X be a finite set and Y a countable set. Let $\mathcal{F}$ be the set of all functions $f : X \to Y$. Then $\mathcal{F}$ is countable.*

Proof. Since X is finite, it can be expressed as $\{x_1, \dots, x_n\}$ for some $n \in \mathbb{N}$. We now claim that there exists a bijection

$$\mathcal{F} \to \underbrace{Y \times \cdots \times Y}_{n \text{ times}} = \prod_{k=1}^{n} Y$$

To achieve this, consider the map $\Phi : \mathcal{F} \to \prod_{k=1}^{n} Y$ given by

$$\Phi(f) := (f(x_1), \dots, f(x_n)) \in \prod_{k=1}^{n} Y.$$

Fix now a point $(y_1, \ldots, y_n) \in \prod_{k=1}^{n} Y$. Define a function $f : X \to Y$ according to the rule:

$$f(x_k) = y_k, \quad \forall 1 \le k \le n.$$

Then, $f \in \mathcal{F}$ satisfies $\Phi(f) = (y_1, \ldots, y_n)$. Hence, Φ is surjective. To see that Φ is also injective, suppose that $\Phi(f) = \Phi(g)$ for $f, g \in \mathcal{F}$. This means that

$$(f(x_1), \ldots, f(x_n)) = (g(x_1), \ldots, g(x_n)) .$$

Hence, $f(x_k) = g(x_k)$ for all $1 \le k \le n$. But this means that $f = g$ on all of X. This is precisely the statement that f and g are the same functions. We conclude that Φ is also injective. By Problem 3.3, we know that $\prod_{k=1}^{n} Y$ is countable. Since $\mathcal{F}$ is in bijection with this set, we see that $\mathcal{F}$ is countable. □

3.4 Other Inequalities

We now prove some basic inequalities that do not rely on induction.

Proposition 3.12. *Prove the following inequalities:*

(1) for all $x \in \mathbb{R}$ one has $x(1-x) \le \frac{1}{4}$;

(2) for all $0 \le x, y \le 1$ at least one of the following hold true:

$$xy \le \frac{1}{4} \quad \text{or} \quad (1-x)(1-y) \le \frac{1}{4}.$$

Proof.

(1) Clearly,

$$x^2 - x + \frac{1}{4} = \left(x - \frac{1}{2}\right)^2 \ge 0.$$

Therefore, $x(1-x) = x - x^2 \le \frac{1}{4}$.

(2) If $x = 0$, the claim is obvious. Thus, we may assume that $x \in (0, 1]$. Suppose first that

$$0 \le y \le \frac{1}{4x}.$$

Then, $xy \le \frac{1}{4}$. Otherwise, $y > \frac{1}{4x}$ so that

$$1 - y < 1 - \frac{1}{4x}.$$

In this case,

$$\begin{aligned}
(1-x)(1-y) \leq (1-x)\left(1-\frac{1}{4x}\right) &= 1-x-\frac{1}{4x}+\frac{1}{4} \\
&= \frac{1}{x}\left[x(1-x)\right]-\frac{1}{4x}+\frac{1}{4} \\
&\leq \frac{1}{x}\cdot\frac{1}{4}-\frac{1}{4x}+\frac{1}{4}=\frac{1}{4}
\end{aligned}$$

This completes the proof. □

4 Fourth Tutorial

We begin with a rapid review of the definitions. Let X be a subset of $\mathbb{R}$ (a priori, possibly empty). A real number u is said to be an *upper bound* for X if $x \leq u$ for all $x \in X$. If X has an upper bound, we say that X is bounded from above. We call a point $s \in \mathbb{R}$ a *least upper bound* for X if both of the following hold:

(I) s is an upper bound for X;

(II) if u is any upper bound for X, then $s \leq u$.

A given subset of $\mathbb{R}$ may or may not have a least upper bound. If one exists, however, it must be unique. To see this, let s and s' be least upper bounds for a set X. In particular, they are both upper bounds. Thus, we must have $s \leq s'$ and $s' \leq s$ by (II) above. Of course, this forces $s = s'$.

Completeness Property of $\mathbb{R}$. Let X be a non-empty subset of $\mathbb{R}$ that is bounded from above. Then X has a least upper bound, which we denote by $\sup X$.

In your assignment, you have proven that a non-empty subset of $\mathbb{R}$ that is bounded from below has an *infimum*, i.e. a greatest lower bound. If X is such a set, we denote this quantity by $\inf X$. As before, $\inf X$ can be shown to be unique. With this in mind, we can state a converse to the completeness property of $\mathbb{R}$.

Proposition 4.1. *Let X be a subset of $\mathbb{R}$.*

1. *If X has a least upper bound (i.e. a supremum), then X is non-empty and bounded from above.*

2. *If X has a greatest lower bound (i.e. an infimum), then is non-empty and bounded from below.*

Proof. We only prove the first claim, and leave the second as an exercise. Suppose that X has a supremum s in $\mathbb{R}$. Then, s is an upper bound for X. By definition, X must be bounded from above. All that remains is to check that X is non-empty. For this, it is enough to show that the empty set has no least upper bound.

Consider the empty set $\varnothing$ and suppose (for a contradiction) that $\varnothing$ has a least upper bound u in $\mathbb{R}$. Then, $x \leq u$ for all $x \in \varnothing$. But, we also have $x \leq u - 1$ for all $x \in \varnothing$. This means that $u - 1 < u$ is an upper bound for $\varnothing$, contradicting the choice of u as the least upper bound. $\square$

Let X be a non-empty set that is bounded from above and let s denote its supremum in $\mathbb{R}$. Let $\varepsilon > 0$ be given and note that $s - \varepsilon < s$ cannot be an upper bound for X (since s is the *least* upper bound for X). Thus, there must exist $x_\varepsilon \in X$ such that $s - \varepsilon < x_\varepsilon$. Since s is an upper bound for X, this gives

$$s - \varepsilon < x_\varepsilon \leq s. \tag{4.1}$$

Conversely, let s be an upper bound for X and suppose that for each $\varepsilon > 0$ one can find $x_\varepsilon \in X$ such that (4.1) holds true. We claim that $s = \sup X$. As s is an upper bound for X, it is enough to show that $s \leq u$ for all upper bounds u of X. If u is an upper bound for X with $u < s$, then $\varepsilon := (s-u) > 0$ whence there exists $x_\varepsilon \in X$ with

$$s - \varepsilon < x_\varepsilon \leq s.$$

But, $s - \varepsilon = s - (s - u) = u$. This gives $u < x_\varepsilon$ which contradicts the fact that u is an upper bound for X. Thus, we must have $s \leq u$. Since u was an arbitrary upper bound, we have $s = \sup X$. This gives the following result from Bartle.

Theorem 4.2. *Let X be a non-empty set and s an upper bound for X in $\mathbb{R}$. The following statements are equivalent.*

1. $s = \sup X$;

2. *for every $\varepsilon > 0$ there exists $x_\varepsilon \in X$ such that $s - \varepsilon < x_\varepsilon \leq s$.*

For the infimum, a similar argument to what we used above yields an analogous result.

Theorem 4.3. *Let X be a non-empty set and v an lower bound for X in $\mathbb{R}$. The following statements are equivalent.*

1. $v = \inf X$;

2. *for every* $\varepsilon > 0$ *there exists* $x_\varepsilon \in X$ *such that* $v \leq x_\varepsilon < v + \varepsilon$.

It's probably best if we do at least one worked example.

Problem 1. *Consider the set*

$$X = \left\{ \frac{1}{n} - \frac{1}{m} : n, m \in \mathbb{N} \right\}.$$

Compute $\sup X$ *and* $\inf X$.

Solution. We first check that X is bounded (bounded from above and below). By the triangle inequality, for all $n, m \in \mathbb{N}$:

$$\left| \frac{1}{n} - \frac{1}{m} \right| \leq \frac{1}{n} + \frac{1}{m} \leq 1 + 1 = 2.$$

Hence, X is indeed bounded. By the completeness of $\mathbb{R}$ (and your homework problems), both $\sup X$ and $\inf X$ exist in $\mathbb{R}$. Do note that

$$\frac{1}{n} - \frac{1}{m} \leq \frac{1}{n} \leq 1,$$

for all $n, m \in \mathbb{N}$. Hence, 1 is an upper bound for X. Informally, we should try to look for the largest elements in X and see what they approach. Here, we should take $n = 1$ and try to make $\frac{1}{m}$ as small as possible. In doing so, $1 - \frac{1}{m}$ seems to approach 1. With this in mind, a good idea is to try and show that $\sup X = 1$.

Since we have already shown that 1 is an upper bound, it remains to check that 1 is the least upper bound. Let $\varepsilon > 0$ be given, we want to show that there exists $x_\varepsilon \in X$ such that

$$1 - \varepsilon < x_\varepsilon \leq 1.$$

Choose $m \in \mathbb{N}$ such that $\frac{1}{m} < \varepsilon$ (how?) and note that

$$x_\varepsilon := 1 - \frac{1}{m} \in X.$$

Then, $-\varepsilon < -\frac{1}{m}$ whence

$$1 - \varepsilon < 1 - \frac{1}{m} = x_\varepsilon \leq 1.$$

By Theorem 4.2, we see that $\sup X = 1$. For $\inf X$, we will use a trick from your homework assignment. Observe that

$$X = \left\{ \frac{1}{n} - \frac{1}{m} : n, m \in \mathbb{N} \right\} = \left\{ \frac{1}{m} - \frac{1}{n} : n, m \in \mathbb{N} \right\} = -X.$$

Hence, by Assignment 3 we get $\inf X = \inf(-X) = -\sup X = -1$. $\square$

Given a subset X of $\mathbb{R}$, we define $\max X$ to be the largest element of X, whenever it exists. Similarly, we define $\min X$ as the smallest element of X, whenever it exists. Note that, when these exist, we have $\min X$, $\max X \in X$.

Proposition 4.4. *Let X be a subset of $\mathbb{R}$. Then $\max X$ exists if and only if $\sup X$ exists and is an element of X. Similarly, $\min X$ exists if and only if $\inf X$ exists and is an element of X.*

Proof. We will only prove the first statement, leaving the second as a potential exercise. Suppose that $\max X$ exists and denote it by m. By definition of $\max X$, we have $m \in X$. Especially, $X \neq \varnothing$. Then, $x \leq m$ for all $x \in X$ whence X is bounded from above. Therefore, $\sup X$ exists. Since $m \in X$, we get $m \leq \sup X$. On the other hand, m is an upper bound for X. By definition of $\sup X$ as the least upper bound, we also have $\sup X \leq m$. It follows that $\sup X = m$ whence $\sup X \in X$.

Conversely, suppose that $\sup X$ exists and is an element of X. By definition, $x \leq \sup X$ for all $x \in X$. Since $\sup X \in X$, this means that $\sup X$ is the maximal element of X. Thus, $\max X$ exists. $\square$

Remark 4.1. This proof tells us that $\max X = \sup X$, whenever $\max X$ exists. Similarly, $\inf X = \min X$ when the latter exists.

4.1 The Existence of $\sqrt{2}$

We will elaborate on the argument used in Theorem 2.4.7 of Bartle's *Introduction to Real Analysis*. We claim that there exists a real number x such that $x^2 = 2$. To show this, we consider the set

$$X = \left\{ s \in [0, \infty) : s^2 < 2 \right\}.$$

Clearly, X is non-empty for $1^2 < 2$ implies $1 \in X$. Also, if $s > 2$ then $s^2 > 4$ implies $s \notin X$. Hence, $s \leq 2$ for all $s \in X$. This means that X is a non-empty bounded subset of $\mathbb{R}$. It therefore has a supremum x in $\mathbb{R}$. Since $1 \in X$ and x is an upper bound for X, we immediately have

$$\boxed{0 < 1 \leq x.}$$

We will show that both cases $x^2 < 2$ and $x^2 > 2$ are impossible. Clearly, this will force $x^2 = 2$ whence x is a square-root of 2.

First, suppose that $x^2 < 2$. For every $n \in \mathbb{N}$ we have $\frac{1}{n^2} \leq \frac{1}{n}$ so that

$$\left(x + \frac{1}{n}\right)^2 = x^2 + \frac{2x}{n} + \frac{1}{n^2} \leq x^2 + \frac{2x+1}{n}. \tag{4.2}$$

By the assumption that $x^2 < 2$, we see that

$$\frac{2 - x^2}{2x + 1} > 0.$$

Thus, we can choose $n \in \mathbb{N}$ such that

$$0 < \frac{1}{n} < \frac{2 - x^2}{2x + 1}.$$

This implies that

$$\frac{2x + 1}{n} < 2 - x^2.$$

Returning to (4.2) yields

$$\left(x + \frac{1}{n}\right)^2 = x^2 + \frac{2x}{n} + \frac{1}{n^2} \leq x^2 + \frac{2x+1}{n} < x^2 + 2 - x^2 = 2. \tag{4.3}$$

Because $x > 0$, we get that $x + \frac{1}{n}$ is an element of X. Since x is an upper bound for X we must have $x + \frac{1}{n} \leq x$, which is a contradiction. Therefore, the case $x^2 < 2$ is impossible.

Next, we show that the case $x^2 > 2$ is impossible. Since we know that $x \geq 1$ and $x^2 > 2$, we actually have $x > 1$. Given $m \in \mathbb{N}$ we calculate

$$\left(x - \frac{1}{m}\right)^2 = x^2 - \frac{2x}{m} + \frac{1}{m^2} > x^2 - \frac{2x}{m}. \tag{4.4}$$

Since $x^2 > 2$, $\frac{x^2-2}{2x}$ is positive. We can therefore choose $m \in \mathbb{N}$ such that

$$0 < \frac{1}{m} < \frac{x^2-2}{2x} \iff 0 < \frac{2x}{m} < x^2 - 2.$$

Of course, this implies

$$2 - x^2 < -\frac{2x}{m}.$$

Using this choice of m in (4.4), we obtain

$$\left(x - \frac{1}{m}\right)^2 > x^2 - \frac{2x}{m} > x^2 + 2 - x^2 = 2.$$

Since $x > 1$, we have $x - \frac{1}{m} > 0$. If $s > x - \frac{1}{m}$ then

$$s^2 > \left(x - \frac{1}{m}\right)^2 > 2.$$

Hence, $s \leq x - \frac{1}{m}$ for all $s \in X$. This means that $x - \frac{1}{m}$ is an upper bound for X, contradicting the least upper bound property of x.

4.2 The Density of the Irrationals

Recall that a non-empty set $A \subseteq \mathbb{R}$ is said to be dense in $\mathbb{R}$ if, for any $x < y$ in $\mathbb{R}$, there exists $a \in A$ such that

$$x < a < y.$$

You have seen in class that the set of rational numbers $\mathbb{Q}$ is dense in $\mathbb{R}$. Here, we show that the irrationals are *also* dense in $\mathbb{R}$. First, let us recall the proof that $\mathbb{R}$ is uncountable. Typically, one applies the diagonal argument to show that the closed interval $[0,1]$ is uncountable. From there, it follows that $\mathbb{R} \supset [0,1]$ is also uncountable. Note that $(0,1)$ cannot be countable, for then $[0,1] = (0,1) \cup \{0,1\}$ would also be.

Next we fix two real numbers a, b with $a < b$ and consider the function

$$f : (0,1) \to (a,b), \quad x \mapsto (b-a)x + a$$

which is a bijection (check this yourself!). Therefore, every open interval (a,b) (with $a, b \in \mathbb{R}$) is uncountable.

Proposition 4.5. *The set of irrational numbers is dense in $\mathbb{R}$.*

Proof. Let $x, y \in \mathbb{R}$ be such that $x < y$. We want to show that there exists an irrational number $\xi \in (x, y)$. Let $\mathbb{Q}_x^y$ denote the set of rational numbers in the interval (x, y), i.e.

$$\mathbb{Q}_x^y := (x, y) \cap \mathbb{Q}.$$

As a subset of a countable set, $\mathbb{Q}_x^y$ is countable. Since (x, y) is uncountable, the set $(x, y) \setminus \mathbb{Q}_x^y$ cannot be countable.[11] Put otherwise, $(x, y) \setminus \mathbb{Q}_x^y$ must be uncountable and, in particular, non-empty. Let ξ be any element of $(x, y) \setminus \mathbb{Q}_x^y$. Then,

$$x < \xi < y \quad \text{and} \quad \xi \notin (x, y) \cap \mathbb{Q}.$$

Thus, ξ is an irrational number satisfying $x < \xi < y$. □

Theorem 4.6 (Characterization of Intervals)**.** *Let $S \subseteq \mathbb{R}$ contain at least two points and assume that S satisfies the property*

$$\textit{if} \quad x, y \in S \quad \textit{and} \quad x < y \quad \textit{then} \quad [x, y] \subseteq S.$$

Then, S is an interval.

Proof. There are 4 distinct cases to consider.

(i) S is bounded;

(ii) S is bounded from above but not from below;

(iii) S is bounded from below but not from above;

(iv) S is unbounded from above and below.

We begin with (i). Define $a := \inf S$ and $b := \sup S$ and observe that $S \subseteq [a, b]$. We now claim that $(a, b) \subseteq S$. Let $z \in (a, b)$ be given. Then, z is not a lower bound for S. We can therefore choose a point $x \in S$ with $x < z$. Similarly, z is not an upper bound for S and there must exist $y \in S$ with $z < y$. But then, $z \in [x, y] \subseteq S$. Since z was arbitrary, we obtain $(a, b) \subseteq S$. If $a, b \in S$, then $[a, b] = S$. If $a \in S$ but $b \notin S$, then $[a, b) = S$. Similarly, if $a \notin S$ but $b \in S$ we have $(a, b] = S$. In either case, S is an interval.

For (ii) we will use a similar argument. Define $b := \sup S$ and observe that $S \subseteq (-\infty, b]$. We claim that $(-\infty, b) \subseteq S$. If $z \in (-\infty, b)$ then $z < b$ whence z

[11]If it were, then as above, we could write $(x, y) = \mathbb{Q}_x^y \cup \left[(x, y) \setminus \mathbb{Q}_x^y\right]$ as the union of two countable sets, making (x, y) countable.

is not an upper bound for S. Thus, there exists $y \in S$ with $z < y$. On the other hand, z cannot be a lower bound for S (since S is unbounded from below). This means that we can choose $x \in S$ with $x < z < y$. But then, $z \in [x, y] \subseteq S$. It follows that $(-\infty, b) \subseteq S$. If $b \in S$, then $(-\infty, b] = S$. Otherwise, $(-\infty, b) = S$. In either case, S is an interval.

For (iii) we will "piggy back" off (ii). Suppose that S is bounded from below but not from above. Then, $-S$ is bounded from above but not from below. By (ii), $-S$ must be an interval of the form $(-\infty, b)$ or $(-\infty, b]$. But then, $S = -(-S)$ is of the form

$$(-b, \infty) \quad \text{or} \quad [-b, \infty).$$

Now we handle (iv). Suppose that S is unbounded from above and below. We claim that $(-\infty, \infty) \subseteq S$. Let $z \in (-\infty, \infty)$ be given and note that z is neither a lower bound nor an upper bound for S. Thus, we can find $x, y \in S$ such that $x < z < y$. That is, $z \in [x, y] \subseteq S$. This yields $\mathbb{R} = (-\infty, \infty) \subseteq S \subseteq \mathbb{R}$. Since $S = \mathbb{R}$, we see that S is an interval. □

4.3 Algebraic Properties of the Supremum/Infimum

Problem 2. *Let A and B be non-empty subsets of $\mathbb{R}$ that are bounded above. Suppose both A and B only contain non-negative elements. Show that*

$$\sup(A \cdot B) = \sup A \cdot \sup B.$$

Here, $A \cdot B = \{ab : a \in A,\ b \in B\}$.

Proof. Clearly, $A \cdot B \neq \emptyset$. Since $0 \leq a \leq \sup A$ and $0 \leq b \leq \sup B$, we have

$$0 \leq ab \leq \sup A \cdot \sup B$$

for all $ab \in A \cdot B$. Hence, $A \cdot B$ is bounded from above. This means that $A \cdot B$ has a least upper bound. The above tells us that $A \cdot B$ is bounded above by $\sup A \cdot \sup B$; it only remains to check that this is the least upper bound.

First, we handle the case where $\sup A = 0$ or $\sup B = 0$. Since $0 \leq a \leq \sup A$ and $0 \leq b \leq \sup B$ for all $a \in A$ and $b \in B$, we would have either $A = \{0\}$ or $B = \{0\}$. In either case, $A \cdot B = \{0\}$ whence $\sup(A \cdot B)$ is indeed equal to $0 = \sup A \cdot \sup B$.

Now we check the case where $\sup A$ and $\sup B$ are non-zero (and hence strictly larger than 0). Let u be an upper bound for $A \cdot B$. Then,

$$0 \leq ab \leq u, \quad \forall a \in A,\ b \in B.$$

Let $b \in B$ with $b \neq 0$. Then,

$$0 \leq a \leq \frac{u}{b}, \quad \forall a \in A.$$

This forces $0 \leq \sup A \leq \frac{u}{b}$ whence $0 \leq b \sup A \leq u$. Since $b \neq 0$ was arbitrary, this actually holds for all $b \in B \setminus \{0\}$. Since it is trivially true for $b = 0$, we have

$$0 \leq b \sup A \leq u, \quad \forall b \in B.$$

But this yields

$$0 \leq b \leq \frac{u}{\sup A}$$

for all $b \in B$ so that $\sup B \leq \frac{u}{\sup A}$. Of course, this gives $\sup A \sup B \leq u$. Since u was any upper bound of $A \cdot B$, we conclude that $\sup(A \cdot B) = \sup A \cdot \sup B$. □

4.4 More About Functions

Let X be a non-empty set and $f : X \to \mathbb{R}$ a function. We say that f is bounded from above if $f(X)$ is bounded from above in $\mathbb{R}$. Similarly, f is said to be bounded from below if $f(X)$ is bounded from below in $\mathbb{R}$. If f is bounded from above, we define

$$\sup_{x \in X} f := \sup f(X) = \sup \{y : y = f(x) \text{ for some } x \in X\}.$$

If f is bounded from below, we have the analogous definition:

$$\inf_{x \in X} f := \inf f(X) = \inf \{y : y = f(x) \text{ for some } x \in X\}.$$

Problem 3. *Let X and Y be non-empty sets. Let $f : X \to Y$ be any function and $g : Y \to \mathbb{R}$ bounded from above. Show that $g \circ f$ is bounded from above and that*

$$\sup_{x \in X} (g \circ f) \leq \sup_{y \in Y} g. \tag{4.5}$$

In addition, we have equality when f is surjective.

Proof. The function $g \circ f$ is bounded from above whenever g is bounded from above since $(g \circ f)(X) \subseteq g(Y)$. Therefore,

$$\sup_{x \in X} (g \circ f) = \sup\, (g \circ f)(X)$$

exists. To prove the desired inequality, it is enough to show that $\sup_{y\in Y} g$ is an upper bound for the set $(g\circ f)(X)$. Clearly, any element of $(g\circ f)(X)$ takes the form $g(f(x))$ for some $x\in X$. Rather, it takes the form $g(y)$ for some element $y=f(x)\in Y$. Thence,

$$(g\circ f)(x) = g(f(x)) = g(y) \le \sup_{y\in Y} g.$$

We have thus proven (4.5). Now, suppose that f is surjective. This means that $g(Y)=g(f(X))$ whence

$$\sup_{y\in Y} g = \sup g(Y) = \sup g(f(X)) = \sup\,(g\circ f)(X) = \sup_{x\in X}\,(g\circ f).$$

See the remark below for an example in which equality fails when f is not surjective. □

Remark 4.2. Equality may fail in (4.5) if f is not surjective. Let $X=Y=[0,1]$ and define functions $f:X\to Y$, $g:Y\to\mathbb{R}$ according to the formulas

$$f(x)=0, \quad g(y)=y.$$

Then, $(g\circ f)(X)=\{0\}$ so that

$$\sup_{x\in X}\,(g\circ f) = 0.$$

On the other hand, $g(Y)=[0,1]$. By an earlier result (Proposition 4.4), we have $\sup g(Y)=\max g(Y)=1>0$. Therefore,

$$\sup_{x\in X}\,(g\circ f) < \sup_{y\in Y} g. \tag{4.6}$$

5 Fifth Tutorial

We first recall some notions about open and closed subsets of $\mathbb{R}$. A set $U\subseteq\mathbb{R}$ is called **open** if, for every $x\in U$, there exists $\varepsilon>0$ such that

$$V_\varepsilon(x) := (x-\varepsilon, x+\varepsilon) \subseteq U.$$

In symbolic terms, U is said to be open if and only if:

$$(\forall x\in U)\,(\exists\varepsilon>0 \text{ such that } V_\varepsilon(x)\subseteq U)\,.$$

In other words, a set U is open if for every point $x \in U$ there exists an open interval $(x-\varepsilon, x+\varepsilon)$ contained in U.[12] Since an open interval $(x-\varepsilon, x+\varepsilon)$ contains all points "near x", the definition for an open set can be interpreted as follows:

> *A set $U \subseteq X$ is called open if, for every $x \in U$, the set U also contains all points near x.*[13]

A set $F \subseteq \mathbb{R}$ is said to be **closed** if $F^{\mathrm{c}} = \mathbb{R} \setminus F$ is open in $\mathbb{R}$. Let us now summarize some properties of open sets that you will have proven in the lectures.

- If $\{U_\alpha\}_{\alpha \in I}$ is an indexed family of open sets, then $\bigcup_{\alpha \in I} U_\alpha$ is also open.
- If U_1 and U_2 are open subsets of $\mathbb{R}$, then so is $U_1 \cap U_2$. By induction, finite intersections of open sets are thus open.
- Both $\mathbb{R}$ and $\varnothing$ are open in $\mathbb{R}$.

The set $\mathfrak{T}$ of all open subsets of $\mathbb{R}$ is therefore closed under unions, finite intersections, and contains both $\varnothing$ and $\mathbb{R}$. This makes $\mathfrak{T}$ into what we call a *topology.* Since $\mathbb{R}$ has a topology (i.e. open subsets), it is called a *topological space.*

Example 5.1. Let us show that infinite intersections of open sets need not be open. For every $n \in \mathbb{N}$, let U_n be the open interval:

$$U_n := \left(-\frac{1}{n}, \frac{1}{n}\right).$$

Now consider the infinite family $\{U_n\}_{n=1}^{\infty}$, what can be said about $\bigcap_{n=1}^{\infty} U_n$? Well, after a moment's consideration, it is not difficult to see that

$$\bigcap_{n=1}^{\infty} U_n = \bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}. \tag{5.1}$$

To prove this, we first note the obvious inclusion $\{0\} \subseteq U_n$, for each $n \in \mathbb{N}$. Since this holds for all n, we have $\{0\} \subseteq \bigcap_{n=1}^{\infty} U_n$. Conversely, let $x \in \bigcap_{n=1}^{\infty} U_n$ be given; we will show that $x \neq 0$ is impossible. If $x \neq 0$, then either $x < 0$ or $x > 0$.

[12]Note that the empty set $\varnothing$ is vacuously open by this definition as well. If $\varnothing$ were *not* open, then there must exist $x \in \varnothing$ such that $V_\varepsilon(x) \nsubseteq \varnothing$ for all $\varepsilon > 0$. Since this would imply $x \in \varnothing$, we have a contradiction.

[13]Please note that this is not rigorous, and is only intended to provide intuition.

1. Case $x > 0$. By the Archimedean property, there exists $n \in \mathbb{N}$ so large that $\frac{1}{n} < x$. Thus, $x \notin U_n$.

2. Case $x < 0$. Then, $-x > 0$. As above, we can find $n \in \mathbb{N}$ such that $n > -\frac{1}{x}$. This implies that $x < -\frac{1}{n}$ whence $x \notin U_n$.

In either case, we get $x \notin U_n$ for some $n \in \mathbb{N}$. Since this contradicts the assumption that $x \in \bigcap_{n=1}^{\infty} U_n$, we conclude that $x \neq 0$ is impossible. Thus, $x \in \bigcap_{n=1}^{\infty} U_n$ implies $x = 0$, i.e. $\bigcap_{n=1}^{\infty} U_n \subseteq \{0\}$. To summarize, we have proven (5.1).

But $\{0\}$ is not open! If it were, then for $0 \in \{0\}$ we could find $\varepsilon > 0$ such that $(0 - \varepsilon, 0 + \varepsilon) \subseteq \{0\}$; which is clearly absurd. Hence $\bigcap_{n=1}^{\infty} U_n$ is an infinite intersection of open sets that is not open.

Example 5.2. You have proven (using De Morgan's laws) that any intersection of closed sets is necessarily closed. You have also seen that finite unions of closed sets are closed. What about infinite unions of closed sets? First, note that any singleton $\{x\} \subset \mathbb{R}$ is closed. Indeed,

$$\mathbb{R} \setminus \{x\} = (-\infty, x) \cup (x, \infty)$$

is a union of open sets, and hence open. Consider the union $\bigcup_{x \in (0,1)} \{x\} = (0, 1)$ of infinitely many closed sets. We claim that $(0, 1)$ is not closed in $\mathbb{R}$. If it were, then $\mathbb{R} \setminus (0, 1) = (-\infty, 0] \cup [1, \infty)$ would be open in $\mathbb{R}$. But, for any $\varepsilon > 0$,

$$V_\varepsilon(0) = (-\varepsilon, \varepsilon) \nsubseteq (-\infty, 0] \cup [1, \infty).$$

This means that $\mathbb{R} \setminus (0, 1)$ is not open whence $(0, 1)$ is not closed.

By way of another example, we show that $\mathbb{N}$ is closed in $\mathbb{R}$. Indeed, $\mathbb{R} \setminus \mathbb{N}$ can be written as the union of open sets:

$$\mathbb{R} \setminus \mathbb{N} = (-\infty, 1) \cup \bigcup_{n=1}^{\infty} (n, n + 1).$$

Thus, $\mathbb{R} \setminus \mathbb{N}$ is open. A more interesting example is the following.

Example 5.3. We show that $\mathbb{Q}$ is neither open nor closed in $\mathbb{R}$. Suppose for a contradiction that $\mathbb{Q}$ is open in $\mathbb{R}$. Then, for any $q \in \mathbb{Q}$ there exists $\varepsilon > 0$ such that $V_\varepsilon(q) = (q - \varepsilon, q + \varepsilon) \subseteq \mathbb{Q}$. Recall from the previous tutorial that the irrationals are dense in $\mathbb{R}$. Thus, every open interval in $\mathbb{R}$ contains an irrational number. In particular, so does $V_\varepsilon(q)$. It follows that $V_\varepsilon(q)$ cannot be contained in $\mathbb{Q}$, which is a contradiction. We conclude that $\mathbb{Q}$ is *not* open.

To see that $\mathbb{Q}$ is not closed, we also argue by contradiction. If $\mathbb{Q}$ were closed, then the set of irrationals $\mathbb{R} \setminus \mathbb{Q}$ would be open in $\mathbb{R}$. Then, for any irrational number ξ, we could find $\varepsilon > 0$ such that $V_\varepsilon(\xi) = (\xi - \varepsilon, \xi + \varepsilon) \subseteq \mathbb{R} \setminus \mathbb{Q}$. Since any open interval contains a rational number ($\mathbb{Q}$ is dense in $\mathbb{R}$), this is impossible. Hence, $\mathbb{Q}$ cannot be closed.

5.1 About Intervals

Proposition 5.1. *Let $U \subseteq \mathbb{R}$. Then U is open if and only if it can be written as a union of open intervals.*

Proof. Since open intervals are open, and unions of open sets are always open, the direction " $\Longleftarrow$ " is immediate. Conversely, let U be an open subset of $\mathbb{R}$. If $U = \varnothing$, then U can be written as the *empty* union of open intervals. Otherwise, U is non-empty. For every $x \in U$, we can find $\varepsilon_x > 0$ such that $V_{\varepsilon_x}(x) \subseteq U$. Therefore,

$$\bigcup_{x \in U} V_{\varepsilon_x}(x) \subseteq U.$$

On the other hand, $x \in U$ implies $x \in V_{\varepsilon_x}(x)$ from which we get

$$\bigcup_{x \in U} V_{\varepsilon_x}(x) \supseteq U.$$

We conclude that

$$\bigcup_{x \in U} V_{\varepsilon_x}(x) = U.$$

Since every $V_{\varepsilon_x}(x)$ is itself an open interval, we see that U can be written as the union of open intervals. ◻

In the previous tutorial we proved the characterization of intervals theorem, which we recall below for the sake of convenience.

Theorem 5.2 (Characterization of Intervals)**.** *Let $S \subseteq \mathbb{R}$ contain at least two points and assume that S satisfies the property*

$$\textit{if} \quad x, y \in S \quad \textit{and} \quad x < y \quad \textit{then} \quad [x, y] \subseteq S.$$

Then, S is an interval.

Using this theorem, we will be able to deduce a very intuitive result that, a priori, seems *very* tricky to prove rigorously.

Proposition 5.3. *Let J be an index set and $\left\{I_j\right\}_{j\in J}$ a family of intervals such that $\bigcap_{j\in J} I_j \neq \varnothing$. Then, $I := \bigcup_{j\in I} I_j$ is an interval in $\mathbb{R}$.*

Proof. As mentioned, we will be using the characterization of intervals theorem. Since every interval contains at least two points, so must I. Let now $x, y \in I$ with $x < y$ be given. By Theorem 4.6, it is enough to show that $[x, y] \subseteq I$. Let $a \in \bigcap_{j\in J} I_j$; choose indices j_1 and $j_2 \in J$ such that $x \in I_{j_1}$ and $y \in I_{j_2}$. We now distinguish three cases.

1. Case $a \le x < y$. Since $a, y \in I_{j_2}$, we must have $[a, y] \subseteq I_{j_2} \subseteq I$. In particular, $[x, y] \subseteq I$.

2. Case $x < a < y$. Since $x \in I_{j_1}$ and $a \in I_{j_1}$, we get that $[x, a] \subseteq I_{j_1} \subseteq I$. Similarly, we see that $[a, y] \subseteq I$. Therefore, $[x, y] = [x, a] \cup [a, y] \subseteq I$.

3. Case $x < y \le a$. Since $x \in I_{j_1}$ and $a \in I_{j_1}$, it follows from the properties of intervals that $[x, a] \subseteq I_{j_1} \subseteq I$. However, because $y \le a$, we have

$$[x, y] \subseteq [x, a] \subseteq I.$$

In either case we have $[x, y] \subseteq I$. By the characterization of intervals theorem, I is an interval in $\mathbb{R}$. With this, the proof is complete. □

5.2 The Boundary of a Set

Let X be a subset of $\mathbb{R}$. The *boundary* of X is defined to be the set

$$\partial X := \left\{x \in \mathbb{R} : \forall \varepsilon > 0,\ V_\varepsilon(x) \cap X \neq \varnothing \text{ and } V_\varepsilon(x) \cap X^{\mathrm{c}} \neq \varnothing\right\}.$$

In other words, ∂X consists of all points $x \in X$ such that every ε-neighbourhood of x intersects both X and X^{c}. This definition is a lot to unpack so let us take a step back. Consider the set $(0, 1) \subset \mathbb{R}$. The point $0 \notin (0, 1)$ belongs to $\partial(0, 1)$ since every ε-neighbourhood $(-\varepsilon, \varepsilon)$ of 0 contains negative numbers (points not in $(0, 1)$) and points between 0 and 1 (i.e. points in $(0, 1)$). Similarly, $1 \in \partial S$.

In the tutorial, I will draw a picture that I hope will make of this clearer. If this doesn't help, come see me or send an me an email; I will then update these notes with more examples. For now, we will stick to more "proofy" problems.

Proposition 5.4. *For any subset $X \subseteq \mathbb{R}$, we have the $\partial X = \partial(\mathbb{R} \setminus X)$. Therefore, the boundary of a set is complement symmetric.*

Proof. We could go through showing that $\partial X \subseteq \partial(\mathbb{R} \setminus X)$ and $\partial(\mathbb{R} \setminus X) \subseteq \partial X$. Or, we could be clever and realize that the definition of ∂X is symmetric in X and $\mathbb{R} \setminus X = X^{\mathsf{c}}$. Indeed, the proof follows from the following observation:

$$\begin{aligned}\partial X &= \{x \in \mathbb{R} : \forall \varepsilon > 0,\ V_\varepsilon(x) \cap X \neq \varnothing \text{ and } V_\varepsilon(x) \cap X^{\mathsf{c}} \neq \varnothing\} \\ &= \{x \in \mathbb{R} : \forall \varepsilon > 0,\ V_\varepsilon(x) \cap X^{\mathsf{c}} \neq \varnothing \text{ and } V_\varepsilon(x) \cap X \neq \varnothing\} \\ &= \partial(X^{\mathsf{c}}) \\ &= \partial(\mathbb{R} \setminus X).\end{aligned}$$

□

Proposition 5.5. *Let $\varnothing \subsetneq X \subsetneq \mathbb{R}$ be given. Then, $\partial X \neq \varnothing$.*

Proof. Since X is non-empty, we can choose $x \in X$. Since $X \neq \mathbb{R}$, there exists $y \in \mathbb{R}$ with $y \notin X$. For the moment, let us make the assumption:

$$\boxed{\text{Assumption} : x < y}$$

Then, $[x, y] \cap X$ is a non-empty and bounded subset of $\mathbb{R}$. Let s be the supremum of $[x, y] \cap X$ in $\mathbb{R}$. Clearly, $x \leq s \leq y$. We now show that $s \in \partial X$. Given $\varepsilon > 0$, we know that $s - \varepsilon$ is not an upper bound of $[x, y] \cap X$. Therefore, we can find an element $z \in [x, y] \cap X$ such that

$$s - \varepsilon < z \leq s.$$

Then, $z \in V_\varepsilon(s) \cap X$. In particular, $V_\varepsilon(s) \cap X \neq \varnothing$. We now consider two sub-cases:

(i) If $s \notin X$ then $V_\varepsilon(s) \ni s$ intersects $\mathbb{R} \setminus X$ at s.

(ii) Suppose that $s \in X$ so that $s \in [x, y] \cap X$. Since $y \notin X$ and $s \leq y$, we must have $s < y$. In this case, we obtain

$$(s, \min\{s + \varepsilon, y\}) \subseteq (s, s + \varepsilon) \subseteq V_\varepsilon(s).$$

But, since s is an upper bound for $[x, y] \cap X$, the interval $(s, \min\{s + \varepsilon, y\})$ is contained in $\mathbb{R} \setminus X$. Indeed, for any $w \in (s, \min\{s + \varepsilon, y\})$ we have $x \leq s < w \leq y$ whence $w \in [x, y]$. If $w \in X$, then $w \in [x, y] \cap X$ would contradict the upper bound property of s.

In either case, we have $V_\varepsilon(s)\cap X \neq \varnothing$ and $V_\varepsilon(s)\cap(\mathbb{R}\setminus X) \neq \varnothing$. Thus, $s \in \partial X$. In the case $y < x$, we argue similarly but using instead the infimum of $[y, x] \cap X$. ◻

If X is a set, the definition of ∂X says that for each $x \in \partial X$ and every $\varepsilon > 0$ the ε-neighbourhood $V_\varepsilon(x)$ intersects both X and X^c. Thus, by shrinking ε as necessary, this seems to suggest that one can approximate all $x \in \partial X$ by points in X *and* also by elements of X^c. This is confirmed by the following theorem:

Theorem 5.6. *Let X be a set. Then, $z \in \partial X$ if and only if for each $n \in \mathbb{N}$ there exists $x \in X$ and $y \in X^c$ such that*

$$|z - x| < \frac{1}{n} \quad \textit{and} \quad |z - y| < \frac{1}{n}. \tag{5.2}$$

Proof. First assume that $z \in \partial X$ and let $n \in \mathbb{N}$ be given. By definition of ∂X, $V_{1/n}(z) \cap X \neq \varnothing$ and $V_{1/n}(z) \cap X^c \neq \varnothing$. Thus, we can pick $x \in V_{1/n}(z) \cap X$ and $y \in V_{1/n}(z) \cap X^c$. Since this implies $|z - x| < \frac{1}{n}$ and $|z - y| < \frac{1}{n}$, the first implication has been proven.

Conversely, fix z and assume that for each $n \in \mathbb{N}$ one can find $x \in X$ and $y \in X^c$ such that (5.2) holds. Let now $\varepsilon > 0$ be given. By the Archimedean Property, there exists $n \in \mathbb{N}$ such that $n > \frac{1}{\varepsilon}$. Or, equivalently, such that $\frac{1}{n} < \varepsilon$. Now, applying the assumption for this n, (5.2) provides the existence of $x \in X$ and $y \in X^c$ such that $x \in V_{1/n}(z)$ and $y \in V_{1/n}(z)$. Using the choice of n, we find that

$$x \in V_{1/n}(z) \subseteq V_\varepsilon(z) \quad \text{and} \quad y \in V_{1/n}(z) \subseteq V_\varepsilon(z).$$

Thus, $V_\varepsilon(z) \cap X$ and $V_\varepsilon(z) \cap X^c$ are both non-empty. Hence, $z \in \partial X$. ◻

We now a recall a problem you will prove in your fourth assignment:

Lemma 5.7. *Let $X \subseteq \mathbb{R}$ be closed. Then, $\partial X \subseteq X$.*

A subset X of $\mathbb{R}$ is said to be **clopen** if it is both open and closed. With the help of this lemma, we can completely characterize the clopen subsets of $\mathbb{R}$.

Proposition 5.8. *A subset X of $\mathbb{R}$ is clopen if and only if $X = \mathbb{R}$ or $X = \varnothing$.*

Proof. It was shown in class that $\varnothing$ and $\mathbb{R}$ are clopen. Thus, we only need to show the “ $\Longrightarrow$ ” implication. To this end, let X be a non-empty clopen subset of $\mathbb{R}$. We claim that $X = \mathbb{R}$. Since X is closed, our previous lemma forces $\partial X \subseteq X$.

We now show that ∂X must be empty. If this is true, then Proposition 5.5 would imply that $X = \mathbb{R}$, since $X \neq \varnothing$. Suppose that ∂X is non-empty and choose $x \in \partial X$. As mentioned above, $\partial X \subseteq X$ implies $x \in X$. Since X is open, there exists an ε-neighbourhood $V_\varepsilon(x) \subseteq X$. Thus, $V_\varepsilon(x) \cap (\mathbb{R} \setminus X) = \varnothing$ for this particular ε. Of course, this contradicts the assumption that $x \in \partial X$. Our only option is therefore $\partial X = \varnothing$ which forces (by Proposition 5.5) $X = \mathbb{R}$. □

From this, we can deduce a very nice topological property.

Corollary 5.9. *$\mathbb{R}$ cannot be written as the union of two non-empty open subsets X and Y of $\mathbb{R}$ such that $X \cap Y = \varnothing$.*

Proof. Let X and Y be open subsets of $\mathbb{R}$, with $X \cap Y = \varnothing$, such that $\mathbb{R} = X \cup Y$. Clearly, $X^{\mathsf{c}} = Y$ and $Y^{\mathsf{c}} = X$. Since Y is open, this tells us that X^{c} is open. Hence, X is closed. Similarly, Y can be seen to be closed. But, this would mean that X and Y are clopen. By the last proposition, we must have $X = \mathbb{R}$ or $Y = \mathbb{R}$. Without loss of generality, suppose that $X = \mathbb{R}$. Finally, observe that

$$Y = Y \cap \mathbb{R} = Y \cap X = \varnothing.$$

This shows that X and Y cannot be non-empty disjoint open subsets of $\mathbb{R}$. With this, the proof is complete. □

Finally, we consider the question of when a given set arises as the boundary of an open set. As it turns out, this question has an elegant answer.

Theorem 5.10. *Let $X \subseteq \mathbb{R}$. Then, X is the boundary of an open set if and only if X is closed and contains no non-trivial open intervals.*

Proof. First let us assume that X is the boundary of some open set. That is, there exists an open set $U \subseteq \mathbb{R}$ such that $\partial U = X$. Since you have proven in your assignment that the boundary of a set is always closed, we see that $X = \partial U$ is closed. Now, we must show that X contains no non-empty open intervals. Assume that $I \subseteq X$ is a non-empty interval and let $x \in I$ be given. Since I is open, there exists $\varepsilon > 0$ such that $V_\varepsilon(x) \subseteq I \subseteq X$. Therefore, $V_\varepsilon(x) \subseteq \partial U = X$. In particular, $x \in \partial U$ so that

$$V_\varepsilon(x) \cap U \neq \varnothing \quad \text{and} \quad V_\varepsilon(x) \cap U^{\mathsf{c}} \neq \varnothing.$$

However, because U is open, $V_\varepsilon(x) \cap U \subseteq \partial U \cap U = \varnothing$. Clearly, this is a contradiction.

Conversely, let us assume that X is closed and contains no non-empty open intervals; we will construct an open set $U \subseteq \mathbb{R}$ such that $\partial U = X$. To this end, let $U := X^{\mathsf{c}}$. Since X is closed, U is an open subset of $\mathbb{R}$. We must now show that $\partial U = X$. Here, there are two inclusions to demonstrate:

- We claim that $X \subseteq \partial U$. Given $x \in X$ and $\varepsilon > 0$, we know that $V_\varepsilon(x) \cap X \neq \emptyset$. Since $X = U^c$, this means that $V_\varepsilon(x) \cap U^c$ is non-empty. Now, since X contains no non-empty open intervals, $V_\varepsilon(x) \not\subseteq X$. Thus,

$$V_\varepsilon(x) \cap X^c = V_\varepsilon(x) \cap U$$

  is non-empty as well. Since $\varepsilon > 0$ was arbitrary, this implies that $x \in \partial U$.

- We now show that $\partial U \subseteq X$. Otherwise, $\partial U \cap X^c = \partial U \cap U$ is non-empty. However, this is impossible by assignment 4 since U is open.

With this, the proof is complete. □

6 *Sixth Tutorial – The Cantor Set*

We devote this tutorial to the construction of an infamous topological structure known as the *Cantor set*. It is this set which helped give rise to the subject known today as point-set topology. The Cantor set is very special in the sense that it defies intuition. For one, the Cantor set is a *closed uncountable set* that does not contain **any** open interval. Moreover, the Cantor set is both totally disconnected and nowhere dense (these terms will be defined rigorously in a few moments). Despite this, the Cantor set does not contain any *isolated points* (again, we will give precise meaning to this term shortly). In short, the Cantor set is a prime example of how weak our ingrained topological intuition can be.

6.1 *Constructing the Cantor Set*

We begin with the closed and bounded interval $[0, 1]$, which has length equal to 1. Let us put $A_0 := [0, 1]$. We can now subdivide A_0 into three **disjoint bounded** intervals, each of length $\frac{1}{3}$, according to the following rule:

$$A_0 := \left[0, \frac{1}{3}\right] \cup \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left[\frac{2}{3}, 1\right].$$

Note that the three intervals given above are pairwise disjoint (i.e. none of these intervals have common points). The idea is now to *remove* what we will call the **middle open third interval**. In the case of A_0, this "middle open third interval" will simply be the open interval

$$\left(\frac{1}{3}, \frac{2}{3}\right).$$

Deleting this interval from A_0, we obtain a *new* subset of $[0, 1]$ given by

$$A_1 := \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Note that A_1 is actually a subset of A_0. Moreover, since finite unions of closed sets are always closed, A_1 will be both closed and bounded. Now, A_1 consists of two intervals, each having length $\frac{1}{3}$. We will now remove the "middle open third interval" from each of these two sub-intervals. To do this, we decompose $[0, \frac{1}{3}]$ into thirds as follows:

$$\left[0, \frac{1}{3}\right] = \left[0, \frac{1}{9}\right] \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left[\frac{2}{9}, \frac{1}{3}\right].$$

Thus, when removing the "open middle third" interval from $\left[0, \frac{1}{3}\right]$, we would be left with

$$\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right]. \tag{6.1}$$

Similarly, we break the second interval in A_1 into thirds:

$$\left[\frac{2}{3}, 1\right] = \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup \left[\frac{8}{9}, 1\right].$$

The "open middle third" interval corresponding to the interval above is therefore $\left(\frac{7}{9}, \frac{8}{9}\right)$. Deleting this from $\left[\frac{2}{3}, 1\right]$ leaves us with the following:

$$\left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]. \tag{6.2}$$

Thus, (6.1)-(6.2) consist of the sets leftover after deleting the "open middle third" intervals from the two subintervals of A_1, respectively. We then define

$$A_2 := \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

Again, A_2 is the disjoint union of finitely many closed intervals in $[0, 1]$. Hence, A_2 is closed and bounded. Note that every subinterval in the definition of A_2 has length equal to $\frac{1}{3^2}$. We now apply the last steps recursively. Given A_n, for some $n \in \mathbb{N}$, we remove the "open middle third" interval from every (disjoint) closed subinterval given in the definition of A_n. After removing an "open middle third" interval from a closed interval $[a, b]$, we are always left with two disjoint closed intervals. Therefore, upon removing the "open middle third" subintervals from each of the subintervals of A_n, we will be left with a finite family of disjoint closed intervals, each contained in A_n. We then define A_{n+1} to be the union of these closed intervals. The following properties will hold for this A_{n+1}.

- A_{n+1} is closed. Indeed, A_{n+1} is by definition the finite union of disjoint closed intervals.

- $A_{n+1} \subseteq A_n$. Certainly, every interval in the definition of A_{n+1} was obtained by removing a "middle open third" interval from some subinterval of A_n. In particular, every interval that makes up A_{n+1} is a subset of A_n. Since A_{n+1} is merely the union of these intervals, we get $A_{n+1} \subseteq A_n$.

- $0 \in A_n$ for every $n \in \mathbb{N}$. Indeed, this follows by observing that the endpoints of intervals are never removed when deleting the "open middle third" interval. Similarly, $1 \in A_n$.

Therefore, for every $n \geq 0$, we obtain a sequence of closed and bounded sets, denoted by $\{A_n\}_{n\in\mathbb{N}_0}$. Note that, by construction, $0 \in \bigcap_{n\in\mathbb{N}_0} A_n$. In particular, this intersection is non-empty.

Definition 6.1 (Cantor Set)**.** The Cantor set $\mathfrak{C}$ is the non-empty set

$$\mathfrak{C} := \bigcap_{n\in\mathbb{N}_0} A_0. \tag{6.3}$$

Note that $\mathfrak{C} \subseteq A_0 \subseteq [0,1]$. Moreover, since A_n is closed for each $n \in \mathbb{N}$, it is immediate from the intersection property of closed sets that $\mathfrak{C}$ is closed.

Since the Cantor set is very difficult to picture, I have included a depiction from Wikipedia below.

Figure 1: Six iterations of the Cantor set

6.2 The Cantor Set is Uncountable

Given the title of this subsection, we can all guess what we will prove next.

Proposition 6.1. *The Cantor set* $\mathfrak{C}$ *defined in* (6.3) *is uncountable.*

Proof. Suppose, by way of contradiction, that $\mathfrak{C}$ is countable and enumerate its elements:

$$\mathfrak{C} := \{x_1, x_2, \ldots, x_n, \ldots\}\,.$$

By construction, $\mathfrak{C} = \bigcap_{n\in\mathbb{N}_0} A_n \subseteq A_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. Since these two intervals are disjoint, the point x_1 only belongs to one of these two intervals. Let I_1 denote the interval **not** containing this point x_1. Dividing this subinterval I_1 by removing its "open middle third" interval, we obtain two disjoint closed subintervals of A_2 (each of length $\frac{1}{3^2}$). Again, since these two intervals are disjoint and $x_2 \in A_2$, we see that x_2 cannot belong to both of these subintervals. Let I_2 be any one of these two subintervals **not** containing x_2. Since I_2 was obtained by deleting points from I_1, we clearly have $I_2 \subseteq I_1 \subseteq [0, 1]$.

We proceed recursively, obtaining a nested sequence

$$I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n \cdots$$

of closed and bounded intervals, with $x_n \notin I_n$ for every $n \in \mathbb{N}$. By the *nested interval property* of $\mathbb{R}$, we know that $\bigcap_{n\in\mathbb{N}} I_n$ is non-empty. Let now $x \in \bigcap_{n\in\mathbb{N}} I_n$ so that $x \in I_n$ for all $n \in \mathbb{N}$. By construction, $I_n \subseteq A_n$ whence $x \in \bigcap_{n\in\mathbb{N}_0} A_n = \mathfrak{C}$. Thus, x belongs to the Cantor set $\mathfrak{C}$ and must be present in our enumeration. On the other hand, $x_n \notin I_n$ for every index n. Since $x \in I_n$, this forces $x \neq x_n$ for all $n \in \mathbb{N}$. We have therefore found an element of $\mathfrak{C}$ not present in our enumeration. Hence $\mathfrak{C}$ cannot be countable. □

6.3 The Topology of $\mathfrak{C}$

We now discuss the surprising topological aspects of the Cantor set. We have already seen that $\mathfrak{C}$ is closed and bounded (making it an example of what is known as a *compact* set). Now, we claim that $\mathfrak{C}$ does not contain any open interval (and hence does not contain any non-trivial interval).

Proposition 6.2. *The Cantor set $\mathfrak{C}$ does not contain any non-trivial interval.*

Proof. Here we have to look back at our construction of the Cantor set. At the n^{th} stage of the construction, A_n is the *disjoint* union of 2^n closed intervals in $[0, 1]$, each having length 3^{-n}. Thus, any interval contained in A_n must have length no larger than 3^{-n}. Hence, if I is a non-trivial interval contained in $\mathfrak{C}$, it must satisfy

$$|I| \leq \inf_{n\geq 1} 3^{-n}.$$

Here, $|I|$ simply denotes the length of I. By Bernoulli's inequality,

$$3^n > n \quad \text{for all } n \geq 1.$$

Thus,

$$0 \leq \inf_{n \geq 1} 3^{-n} \leq \inf_{n \geq 1} \frac{1}{n} = 0.$$

We then see that $|I| = 0$, which is absurd for any non-trivial interval. We conclude that $\mathfrak{C}$ does not contain any non-trivial interval. $\square$

Corollary 6.3. *The Cantor set $\mathfrak{C}$ has empty interior. This makes $\mathfrak{C}$ and example of a nowhere dense set.*[14]

Proof. Let $\mathring{\mathfrak{C}}$ denote the interior of $\mathfrak{C}$. Then, $\mathring{\mathfrak{C}}$ is a subset of $\mathfrak{C}$ by definition. If $\mathring{\mathfrak{C}}$ were non-empty, then for any point $x \in \mathring{\mathfrak{C}}$ we could find $\varepsilon > 0$ such that $V_\varepsilon(x) \subseteq \mathring{\mathfrak{C}} \subseteq \mathfrak{C}$. Of course, this means that $\mathfrak{C}$ would contain an open interval. This contradiction shows that $\mathfrak{C}$ has empty interior. $\square$

Remark 6.1. Note that $0 = \inf \mathfrak{C}$ and $1 = \sup \mathfrak{C}$.[15] Since $0, 1 \in \mathfrak{C}$, we see that $\mathfrak{C}$ contains both of its "endpoints". But this does not imply that $\partial\mathfrak{C} = \{0, 1\}$! Indeed, as we shall see $\partial\mathfrak{C}$, contains much more than the two endpoints 0 and 1 of $\mathfrak{C}$.

We claim that $\partial\mathfrak{C} = \mathfrak{C}$. Since $\mathfrak{C}$ is closed, the inclusion $\partial\mathfrak{C} \subseteq \mathfrak{C}$ is immediate. Conversely, let $x \in \mathfrak{C}$ be given and fix $\varepsilon > 0$. Clearly, $V_\varepsilon(x) \cap \mathfrak{C} \neq \varnothing$ since $x \in \mathfrak{C}$. However, because $\mathfrak{C}$ does not contain any open interval, we have $V_\varepsilon(x) \nsubseteq \mathfrak{C}$. This implies that

$$V_\varepsilon(x) \cap \mathfrak{C} \neq \varnothing \quad \text{and} \quad V_\varepsilon(x) \cap \mathfrak{C}^{\mathsf{c}} \neq \varnothing.$$

Since $\varepsilon > 0$ was arbitrary, it follows that $x \in \partial\mathfrak{C}$. Thus, $\mathfrak{C} \subseteq \partial\mathfrak{C}$ and $\partial\mathfrak{C} = \mathfrak{C}$. In particular, $\partial\mathfrak{C}$ is uncountable. Therefore, the Cantor set is an example of a closed set with boundary points that are **not** endpoints. In fact, every point in $\mathfrak{C}$ is a boundary point of $\mathfrak{C}$.

Definition 6.2. A subset X of $\mathbb{R}$ is said to be *totally disconnected* if, for any $x, y \in X$ with $x < y$, there exists $z \in \mathbb{R} \setminus X$ such that $x < z < y$.

One example of a totally disconnected set is $\mathbb{Q}$. Indeed, by the density of the irrationals, for any $x < y$ in $\mathbb{Q}$ one can always find an irrational number z with $z \in (x, y)$.

---

[14] A set $A \subseteq \mathbb{R}$ is said to be nowhere dense if the interior of $\overline{A}$ is empty.

[15] This is because $\mathfrak{C} \subseteq [0, 1]$ and $0, 1 \in \mathfrak{C}$.

Proposition 6.4. *The Cantor set $\mathfrak{C}$ is totally disconnected.*

Proof. Let $x, y \in \mathfrak{C}$ be given and assume that $x < y$. We must show that there exists some $z \notin \mathfrak{C}$ with $x < z < y$. Since $\mathfrak{C}$ does not contain any non-trivial interval, we cannot have $(x, y) \subseteq \mathfrak{C}$. Namely, $(x, y) \cap \mathfrak{C}^c \neq \varnothing$. Then, any $z \in (x, y) \cap \mathfrak{C}^c$ is satisfies $x < z < y$ and $z \notin \mathfrak{C}$. □

Definition 6.3. Let X be a non-empty subset of $\mathbb{R}$. We say that a point $x \in X$ is an *isolated point* of X if there exists $\varepsilon > 0$ such that $V_\varepsilon(x) \cap (X \setminus \{x\}) = \varnothing$.

In more familiar language, $x \in X$ is called an isolated point of X if there exists an ε-neighbourhood of x that does not contain points in X other than x.

Proposition 6.5. *The Cantor set $\mathfrak{C}$ has no isolated points.*

Proof. Fix $x \in \mathfrak{C}$ and let $\varepsilon > 0$ be given; we must show that $V_\varepsilon(x)$ intersects $\mathfrak{C}$ at a point other than x. Since $x \in \mathfrak{C}$, we must have $x \in A_n$ for all $n \geq 1$. Arguing as before, we can find $N \in \mathbb{N}$ so large that $3^{-N} < \varepsilon$. Consider this A_N, and recall that every subinterval of A_N has length equal to $3^{-N} < \varepsilon$. Let $I = [a, b]$ be the closed subinterval of A_N containing the point x. For any $y \in I$:

$$|x - y| \leq |b - a| .$$ [16]

Since I has length $3^{-N} < \varepsilon$, it follows that

$$|x - y| \leq 3^{-N} < \varepsilon$$

whence $y \in V_\varepsilon(x)$. Or, rather, $I = [a, b] \subseteq V_\varepsilon(x)$. Now, I is a subinterval from the construction of A_N whence $I \subseteq A_n$ for all $n \leq N$. On the other hand, when removing the "open middle third" in the construction, the end points of the intervals are never removed. Therefore the end points a, b of I will belong to A_n for all $n \in \mathbb{N}$. This simply means that

$$a, b \in \mathfrak{C} = \bigcap_{n \in \mathbb{N}_0} A_n.$$

Since $a, b \in V_\varepsilon(x)$, it follows that $V_\varepsilon(x)$ intersects $\mathfrak{C}$ at a point other than x. □

[16] First note that $x \leq b$ and $y \geq a$. Therefore, $x - y \leq b - a$. Similarly, one can show that $y - x \leq b - a$ whence $-(b - a) \leq x - y$.

Remark 6.2. In the previous proof, we have used that if x is an endpoint of one of the disjoint intervals that make up A_n for some $n \in \mathbb{N}$, then $x \in \mathfrak{C}$. It is therefore natural to ask whether the converse is true as well. That is, suppose $x \in \mathfrak{C}$. Can we conclude that x is an endpoint of one of the disjoint intervals making up A_n for some $n \in \mathbb{N}$? Unfortunately, this is not the case. To see that not all of points of $\mathfrak{C}$ are of this form, recall $\mathfrak{C}$ is uncountable. On the other hand, the collection of endpoints of intervals making up A_n for each $n \in \mathbb{N}$ is countable.

7 Seventh Tutorial

We begin by reiterating the definition of a limit for sequences. Recall that a sequence is a function $x : \mathbb{N} \to \mathbb{R}$, where we denote by x_n the value $x(n)$. It has become convention to denote a sequence by either (x_n) or $\{x_n\}_{n\in\mathbb{N}}$; I prefer the former. Intuitively, we like to picture a sequence as an "infinite list" of numbers that may have repetitions. For this reason, it is important not to confuse a sequence with a subset of the real numbers, as sets do not allow for repetitions.

Definition 7.1. Let (x_n) be a sequence and fix $x \in \mathbb{R}$. We say that x_n *converges* to x (as $n \to \infty$) if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $|x_n - x| < \varepsilon$ for all $n \geq N$. In symbolic terms:

$$(\forall \varepsilon > 0)\,(\exists N \in \mathbb{N})\,(n \geq N \implies |x_n - x| < \varepsilon)\,.$$

If this is the case, then we write $x_n \to x$, $x_n \xrightarrow{n\to\infty} x$, or

$$\lim x_n = x.$$

To help solidify the definition of convergence, we give a few explicit examples of convergent sequences. Although in practice one does not argue directly from the definition, it is important to fully understand the logic behind this definition.

Example 7.1. Show that

$$\lim \frac{2n}{n+1} = 2.$$

Proof. Let $\varepsilon > 0$ be given, we must find $N \in \mathbb{N}$ such that

$$\left|\frac{2n}{n+1} - 2\right| < \varepsilon, \quad \forall n \geq N.$$

First, we calculate for all $n \in \mathbb{N}$:

$$\left|\frac{2n}{n+1} - 2\right| = \left|\frac{2n - 2(n+1)}{n+1}\right| = \frac{2}{n+1} \leq \frac{2}{n}.$$

By the Archimedean property, we can find $N \in \mathbb{N}$ such that $N > \frac{2}{\varepsilon}$. Then, for all $n \geq N$ there holds

$$\frac{2}{n} \leq \frac{2}{N} < 2 \cdot \frac{\varepsilon}{2} = \varepsilon.$$

Hence, if $n \geq N$, we see that

$$\left|\frac{2n}{n+1} - 2\right| \leq \frac{2}{n} < \varepsilon,$$

which is what had to be shown. □

Example 7.2. Prove that

$$\lim \frac{3n+1}{2n+5} = \frac{3}{2}.$$

Proof. Let $\varepsilon > 0$ be given, we seek $N \in \mathbb{N}$ such that

$$\left|\frac{3n+1}{2n+5} - \frac{3}{2}\right| < \varepsilon$$

for all $n \geq N$. Naturally, for each $n \in \mathbb{N}$, we have the following estimate:

$$\begin{aligned}
\left|\frac{3n+1}{2n+5} - \frac{3}{2}\right| = \left|\frac{2(3n+1) - 3(2n+5)}{2(2n+5)}\right| &= \left|\frac{2-15}{2(2n+5)}\right| \\
&= \frac{13}{2(2n+5)} \\
&\leq \frac{13}{2n} \\
&\leq \frac{13}{n}.
\end{aligned}$$

Choose $N \in \mathbb{N}$ so large that $\frac{1}{N} < \frac{\varepsilon}{13}$ (why can we do this?); if $n \geq N$ the above implies that

$$\left|\frac{3n+1}{2n+5} - \frac{3}{2}\right| \leq \frac{13}{n} \leq \frac{13}{N} < 13 \cdot \frac{\varepsilon}{13} = \varepsilon.$$

Put otherwise, we have

$$\left|\frac{3n+1}{2n+5} - \frac{3}{2}\right| < \varepsilon, \quad \forall n \geq N.$$

By definition of the limit, this means that

$$\frac{3n+1}{2n+5} \xrightarrow{n\to\infty} \frac{3}{2}.$$

□

Remark 7.1. We would like to point out our general strategy for computing limits via the definition. If we wish to show that a sequence (x_n) converges to $x \in \mathbb{R}$, we start by fixing $\varepsilon > 0$. Our goal is then to show that there exists $N \in \mathbb{N}$ such that $|x_n - x| < \varepsilon$ whenever $n \geq N$. To *find* such an N, we try to estimate $|x_n - x|$ for **all** large enough n. This is done in the hopes of bounding the (often complex) expression $|x_n - x|$ by some term, depending on n, that is easier to estimate.

Example 7.3. Prove that

$$\lim \frac{n^2-1}{2n^2+3} = \frac{1}{2}.$$

Proof. Following the outline in the remark, let $\varepsilon > 0$ and consider the expression:

$$\begin{aligned}\left|\frac{n^2-1}{2n^2+3} - \frac{1}{2}\right| &= \left|\frac{2(n^2-1)-(2n^2+3)}{2(2n^2+3)}\right| \\ &= \left|\frac{5}{2(2n^2+3)}\right| \\ &\leq \frac{5}{4n^2} \\ &\leq \frac{2}{n^2}.\end{aligned}$$

Now, it will be easier to find our candidate for N. Let $N \in \mathbb{N}$ be such that

$$N > \sqrt{\frac{2}{\varepsilon}},$$

which we know to exist by the Archimedean property. If $n \geq N$, then our estimate above tells us that

$$\left|\frac{n^2-1}{2n^2+3} - \frac{1}{2}\right| \leq \frac{2}{n^2} \leq \frac{2}{N^2} < 2 \cdot \frac{\varepsilon}{2} = \varepsilon.$$

This is what had to be shown. □

We give one final example.

Example 7.4. Prove that
$$\lim \frac{\sqrt{n}}{n+1} = 0.$$

Proof. Let $\varepsilon > 0$ be given. For every $n \in \mathbb{N}$, we have
$$\left|\frac{\sqrt{n}}{n+1} - 0\right| = \frac{\sqrt{n}}{n+1} \leq \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}.$$

Let now $N \in \mathbb{N}$ be such that $N > \frac{1}{\varepsilon^2}$. Then, $\frac{1}{\sqrt{N}} < \varepsilon$ so that, for all $n \geq N$,
$$\left|\frac{\sqrt{n}}{n+1} - 0\right| \leq \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{N}} < \varepsilon.$$

□

7.1 Other Properties and a Return to Topology

We note that the convergence of a sequence (x_n) does imply that convergence of the sequence $(|x_n|)$. Indeed, we prove this below.

Proposition 7.1. *Let (x_n) be a sequence of real numbers converging to $x \in \mathbb{R}$. Then, $|x_n|$ converges to $|x|$, as $n \to \infty$.*

Proof. Let $\varepsilon > 0$ be given, since $\lim x_n = x$, we can find $N \in \mathbb{N}$ so large that $|x_n - x| < \varepsilon$ for all $n \geq N$. By the reverse triangle inequality, for all $n \geq N$ there then holds
$$||x_n| - |x|| \leq |x_n - x| < \varepsilon.$$

Hence, $\lim |x_n| = |x|$. □

The converse need not hold in general. Indeed, consider the sequence (x_n) given by $x_n := (-1)^n$. This sequence does **not** converge to any real number x, but the sequence given by $|x_n|$ is always equal to 1. Since constant sequences converge, this means that $(|x_n|)$ is convergent. Despite this, a partial converse to the above does hold.

Proposition 7.2. *Let (x_n) be a sequence of real numbers. If $(|x_n|)$ converges to* 0 *as $n \to \infty$, then so does (x_n).*

Proof. Given $\varepsilon > 0$, we can find $N \in \mathbb{N}$ so large that $||x_n| - 0| < \varepsilon$, whenever $n \geq N$. But this means that $|x_n| < \varepsilon$ for all $n \geq N$. This proves the claim. □

Theorem 7.3. *If $a > 1$ is a real number, then*

$$\lim\left(\frac{1}{a^n}\right) = 0.$$

Moreover, if $a \in (-1, 1)$, then $\lim a^n = 0$.

Proof. First suppose that $a > 1$. Then, $a = 1 + t$ for some $t > 0$. By Bernoulli's inequality, we know that

$$a^n = (1+t)^n \geq 1 + nt, \quad \forall n \geq 1.$$

Therefore, for $n \in \mathbb{N}$, we have the estimate:

$$\left|\frac{1}{a^n}\right| = \frac{1}{a^n} \leq \frac{1}{1+nt} \leq \frac{1}{nt}.$$

Choose $N \in \mathbb{N}$ such that $N > \frac{1}{t\varepsilon} > 0$. For all $n \geq N$, we have

$$\left|\frac{1}{a^n} - 0\right| \leq \frac{1}{nt} \leq \frac{1}{Nt} < \frac{t\varepsilon}{t} = \varepsilon.$$

This shows that $\frac{1}{a^n} \to 0$, as $n \to \infty$. Now, we handle the case where $a \in (0, 1)$. Consider $b := \frac{1}{a} > 1$; for every $n \in \mathbb{N}$ we have

$$a^n = \frac{1}{b^n}.$$

By the first part, it follows that

$$\lim a^n = \lim\left(\frac{1}{b^n}\right) = 0.$$

If $a = 0$, then obviously $\lim a^n = 0$. Thus, it only remains to handle the case $a \in (-1, 0)$. In this case, note that $|a| \in (0, 1)$ and

$$|a^n| = |a|^n, \quad \forall n \in \mathbb{N}.$$

By the previous case (the case of $a \in (0, 1)$), we see that

$$\lim |a^n| = \lim (|a|^n) = 0.$$

Invoking Proposition 7.2, we see that $\lim a^n = 0$ in this case as well. With this, the proof is complete. □

Before moving onto more quantitative results, we should take a moment to discuss the relationship between open sets and limits. As it turns out, the $\varepsilon - N$ definition of the limit for sequences can be reformulated entirely in terms of open sets.

Theorem 7.4 (Open set characterization of the limit)**.** *Let (x_n) be a sequence and fix a point $x \in \mathbb{R}$. The following statements are equivalent:*

(1) $\lim x_n = x$*, i.e. x_n converges to x as $n \to \infty$;*

(2) for each open set U containing x, there exists $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$.[17]

Proof. We first show that (1) implies (2). Assume that $x_n \to x$ and let $U \ni x$ be an open set. There exists $\varepsilon > 0$ such that $V_\varepsilon(x) \subseteq U$. For this ε, using that $\lim x_n = x$, we can find $N \in \mathbb{N}$ such that

$$|x_n - x| < \varepsilon, \quad \forall n \geq N.$$

Equivalently, $x_n \in V_\varepsilon(x) \subseteq U$ for all $n \geq N$. This verifies that (2) holds true.

Conversely, assume that criterion (2) is true; we must show that x_n converges to x. Given $\varepsilon > 0$, it is clear that $V_\varepsilon(x)$ is an open set containing x. Therefore, by (2) with $U = V_\varepsilon(x)$, there exists $N \in \mathbb{N}$ such that $x_n \in V_\varepsilon(x)$ for all $n \geq N$. However, this is equivalent to writing

$$|x_n - x| < \varepsilon \quad \forall n \geq N.$$

Since $\varepsilon > 0$ was arbitrary, this means that $\lim x_n = x$. □

Remark 7.2. Let (x_n) be a sequence converging to a point x. Then, intuitively speaking, x_n can be made arbitrarily close to x by making n very large. Now, given an open set U containing the point x, the theorem above says that x_n must belong to U for all n large. This agrees with the intuitive statement that open sets contain all points "near" every one of its elements.

Sequences can also be used to characterized closed sets. We formalize this notion below.

Theorem 7.5. *Let $F \subseteq \mathbb{R}$ be a set. The following statements are equivalent:*

(1) F is closed;

[17] Put otherwise, the open set U contains the point x_n for all $n \geq N$.

(2) *F contains the limits of all its convergent sequences. That is, if (x_n) is any convergent sequence with $x_n \in F$ for all $n \in \mathbb{N}$, then $\lim x_n \in F$.*

Proof. You will prove (1) $\implies$ (2) in your assignment. Thus, we need only show that (2) implies (1). For this, we argue by contradiction. Assume that (2) holds but that F is not closed. Then, F^c is not open. That is, there exists a point $x \in F^c$ such that $V_\varepsilon(x) \nsubseteq F^c$ for each $\varepsilon > 0$. In particular, for each $n \geq 1$, we have $V_{1/n}(x) \cap F \neq \varnothing$. Thus, given $n \in \mathbb{N}$, we may select a point $x_n \in V_{1/n}(x) \cap F$. This gives us a sequence (x_n) in F such that

$$|x_n - x| < \frac{1}{n}$$

for all $n \in \mathbb{N}$. By the Squeeze theorem, this implies that $x_n \to x$ as $n \to \infty$. However, as (x_n) is a sequence in F by construction, our assumption implies that $\lim x_n = x \in F$. As we have chosen $x \in F^c$, this is a contradiction. □

7.2 About the Ratio Test

We begin with the following theorem.

Theorem 7.6. *Let (x_n) be a sequence of* ***positive*** *real numbers such that*

$$L := \lim \left(\frac{x_{n+1}}{x_n} \right)$$

exists. If $L < 1$, then (x_n) converges to 0.

Proof. Fix $r \in (L, 1)$ and define $\varepsilon_0 := r - L > 0$. Choose $N \in \mathbb{N}$ such that

$$\left| \frac{x_{n+1}}{x_n} - L \right| < \varepsilon_0, \quad \forall n \geq N.$$

Since $x_n > 0$ for all n, we know that $L \geq 0$. If $n \geq N$, then the above implies that

$$\frac{x_{n+1}}{x_n} = \left| \frac{x_{n+1}}{x_n} \right| \leq \left| \frac{x_{n+1}}{x_n} - L \right| + |L| < L + \varepsilon_0 = r.$$

Thus, for all $n \geq N$ there holds

$$0 < x_{n+1} < x_n r < x_{n-1} r^2 < \cdots < x_N r^{n-N+1} =: C r^n,$$

for some constant $C > 0$ that does not depend on n. By Theorem 7.3, we know that $\lim r^n = 0$. In particular, we know that $\lim(Cr^n) = 0$. Thus, given $\varepsilon > 0$, we can find $N' \in \mathbb{N}$ so large that

$$Cr^n = |Cr^n| < \varepsilon, \quad \forall n \geq N'.$$

Let $K := \max(N, N')$. If $n \geq K + 1$, then

$$|x_n| \leq Cr^{n-1} < \varepsilon,$$

since $n - 1 \geq N'$. Thus, the theorem is proven. □

Remark 7.3. In the context of the previous result, when $L > 1$ we can instead say that the sequence (x_n) is *divergent*. Indeed, assuming that $\lim \frac{x_{n+1}}{x_n} = L > 1$, there exists $N \in \mathbb{N}$ such that

$$\left|\frac{x_{n+1}}{x_n} - L\right| < L - 1$$

for all $n \geq N$. Then, for all such $n \in \mathbb{N}$, we find that

$$L - \frac{x_{n+1}}{x_n} < L - 1 \iff 1 < \frac{x_{n+1}}{x_n},$$

whence $x_{n+1} > x_n$ for all $n \geq N$. In particular, for any $n \geq N$, we see that

$$x_n > x_{n-1} > x_{n-2} > \cdots > x_N > 0.$$

So, if x_n converges to some $x \in \mathbb{R}$, we must have $\lim x_n = x \geq x_N > 0$. However, it would then follow from the limit laws that

$$\lim \frac{x_{n+1}}{x_n} = \frac{\lim x_{n+1}}{\lim x_n} = \frac{x}{x} = 1$$

which contradicts our assumption that $L > 1$.

In light of Proposition 7.2, the positivity assumption is not required. Indeed, we only require that the sequence (x_n) be non-zero.

Corollary 7.7 (Ratio Test for Sequences)**.** *Let (x_n) be a sequence of non-zero real numbers and assume that*

$$L = \lim \frac{|x_{n+1}|}{|x_n|}$$

exists. If $L < 1$, then (x_n) converges to 0*. If $L > 1$, then (x_n) does not converge.*

Proof. If $L < 1$ then Theorem 7.6 ensures that $|x_n| \to 0$ as $n \to \infty$. Of course, citing Proposition 7.2, we infer that $x_n \to 0$ as well. If $L > 1$, our remark tells us that $(|x_n|)$ does not converge. This implies that (x_n) cannot converge. Indeed, if (x_n) were convergent, then so would be $(|x_n|)$ by Proposition 7.1. □

Remark 7.4. If $L = 1$ then nothing can be said about the convergence of (x_n). Indeed, the constant sequence (x_n) given by $x_n = 1$ is convergent and satisfies

$$\lim \frac{|x_{n+1}|}{|x_n|} = 1.$$

On the other hand, the alternating sequence $x_n := (-1)^n$ is divergent but *also* satisfies the above. Similarly, the sequence $x_n = n$ is unbounded (and therefore divergent) but once again satisfies the above.

7.2.1 A First Look at Infinite Series

Let (x_n) be a sequence of real numbers. For every $N \in \mathbb{N}$, let us define

$$S_N := \sum_{n=1}^{N} x_n.$$

This gives us a *new* sequence (S_N) in $\mathbb{R}$. We then say the infinite series $\sum_{n=1}^{\infty} x_n$ converges if there exists a real number x such that

$$\lim_{N\to\infty} S_N = \lim_{N\to\infty} \sum_{n=0}^{N} x_n = x.$$

In this case, we will denote this limit by the symbol $\sum_{n=1}^{\infty} x_n$. The series $\sum_{n=1}^{\infty} x_n$ is said to be *absolutely convergent* if the series

$$\sum_{n=1}^{\infty} |x_n|$$

is convergent. It is a fact that every absolutely convergent series is convergent.

Theorem 7.8. *Let (x_n) be a sequence of real numbers and assume that the series $\sum_{n=1}^{\infty} |x_n|$ is convergent. That is, let $\sum_{n=1}^{\infty} x_n$ be absolutely convergent. Then, $\sum_{n=1}^{\infty} x_n$ converges in $\mathbb{R}$.*

Proof. We postpone the proof; we will return to this result once we have studied Cauchy sequences. □

Proposition 7.9. *Let (x_n) be a sequence of non-zero real numbers such that*

$$\lim \left| \frac{x_{n+1}}{x_n} \right| = L$$

exists. If $L < 1$, then the series $\sum_{n=1}^{\infty} x_n$ is absolutely convergent.

Proof. It is automatic that $L \geq 0$. Let us fix $L < r < 1$ and define $\varepsilon_0 := r - L > 0$. There exists $N \in \mathbb{N}$ so large that

$$\left| \left| \frac{x_{n+1}}{x_n} \right| - L \right| < \varepsilon_0, \quad \forall n \geq N.$$

Therefore, by the triangle inequality,

$$\begin{aligned} \left| \frac{x_{n+1}}{x_n} \right| = \left| \left| \frac{x_{n+1}}{x_n} \right| - L + L \right| &\leq \left| \left| \frac{x_{n+1}}{x_n} \right| - L \right| + L \\ &< \varepsilon_0 + L = r \end{aligned}$$

for all $n \geq N$. That is,

$$|x_{n+1}| \leq r\,|x_n| \quad \forall n \geq N.$$

It follows that $|x_{n+k}| \leq r^k\,|x_n|$ for all $k \geq 1$ and $n \geq N$. But then, for every $K \geq N + 1$, we see that

$$\begin{aligned} \sum_{n=N+1}^{K} |x_n| \leq \sum_{k=1}^{K} |x_{N+k}| \leq \sum_{k=1}^{K} r^k\,|x_N| = |x_N| \sum_{k=1}^{k} r^k \\ = |x_N|\, r \frac{1 - r^k}{1 - r} \\ \leq |x_N|\, \frac{r}{1 - r}. \end{aligned}$$

Therefore, the sequence

$$(S_K)_{k=1}^{\infty} = \left(\sum_{k=1}^{K} |x_k| \right)$$

is bounded. This implies (by the monotone convergence theorem - which you will all see soon) that the series $\sum_{n=1}^{\infty} |x_n|$ converges. □

7.3 Some Additional Limit Proofs

We will use the remaining time to discuss additional "less obvious" limits. For these, we will simplify our argument by using the *squeeze* and limit theorems. First, we show that

$$\lim \frac{2^n}{n!} = 0. \tag{7.1}$$

To do this, we will show that

$$0 < \frac{2^n}{n!} \leq 2\left(\frac{2}{3}\right)^{n-2}, \quad \forall n \geq 2.$$

If this indeed true, then the squeeze theorem would imply that

$$\lim \frac{2^n}{n!} = 0.$$

Thus, we are reduced to verifying the aforementioned inequality. For this, we shall argue by induction. As a base case, let us check the inequality for $n = 2$:

$$0 < \frac{2^2}{2!} = 2 \quad \text{and} \quad 2\left(\frac{2}{3}\right)^0 = 2.$$

Our base case is thence satisfied. Now, for the inductive step. Assume that

$$0 < \frac{2^n}{n!} \leq 2\left(\frac{2}{3}\right)^{n-2}$$

for some $n \geq 2$. Clearly, by the inductive hypothesis:

$$\begin{aligned} 0 < \frac{2^{n+1}}{(n+1)!} = \frac{2}{(n+1)} \cdot \frac{2^n}{n!} &\leq \frac{2 \cdot 2}{(n+1)}\left(\frac{2}{3}\right)^{n-2} \\ &\leq 2\left(\frac{2}{3}\right)^{n-2+1} \\ &= 2\left(\frac{2}{3}\right)^{(n+1)-2}. \end{aligned}$$

Here, we have used that $n \geq 2$ implies $n + 1 \geq 3$ and, consequently, $\frac{1}{n+1} \leq \frac{1}{3}$. By our earlier remarks (where we discussed the application of the squeeze theorem), we see that (7.1) holds.

Next, we discuss the following limit:

$$\lim \frac{n!}{n^n} = 0. \tag{7.2}$$

Again, this can be solved relatively easily by the use of the squeeze theorem. Given $n \in \mathbb{N}$, observe that

$$0 < \frac{n!}{n^n} = \frac{n \cdot (n-1) \cdots 2 \cdot 1}{\underbrace{n \cdot n \cdot \cdots \cdot n}_{n \text{ times}}} = \prod_{k=1}^{n} \frac{k}{n} \leq \frac{1}{n}.$$

This last inequality uses the fact that $\frac{k}{n} \leq 1$, for all $k = 2, \ldots, n$. Thus, for every $n \in \mathbb{N}$, one has

$$0 \leq \frac{n!}{n^n} \leq \frac{1}{n}.$$

Since

$$\frac{1}{n} \xrightarrow{n \to \infty} 0,$$

the Squeeze Theorem shows that (7.2) is true.

8 Eighth Tutorial

Let (x_n) be a sequence and suppose that x_n converges to $x \in \mathbb{R}$. It was seen in the lectures that any subsequence (x_{n_k}) of (x_n) must *also* converge to this point x. Moreover, let us recall that

$$e := \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n. \tag{8.1}$$

Note that this is our *definition* of e! We still do not know how to define e^x for all real numbers x! Similarly, we cannot yet make sense of the logarithm. Next, we warm up by practicing with subsequences and limit-laws.

Example 8.1. We show that

$$\lim \left(1 - \frac{1}{n}\right)^n = \frac{1}{e}.$$

To see this, note that for $n \geq 2$:

$$\left(1 - \frac{1}{n}\right)^{-n} = \left(\frac{n-1}{n}\right)^{-n} = \left(\frac{n}{n-1}\right)^n.$$

Or, equivalently,

$$\left(1-\frac{1}{n}\right)^{-n} = \left(1+\frac{1}{n-1}\right)^{n} = \left(1+\frac{1}{n-1}\right)^{n-1}\left(1+\frac{1}{n-1}\right).$$

Therefore,

$$\begin{aligned}\lim\left(1-\frac{1}{n}\right)^{-n} &= \lim\left(1+\frac{1}{n-1}\right)^{n-1}\lim\left(1+\frac{1}{n-1}\right)\\ &= \lim\left(1+\frac{1}{n-1}\right)^{n-1}\\ &= e,\end{aligned}$$

where we have used (8.1) in this last step. Now, we write

$$\lim\left(1-\frac{1}{n}\right)^{n} = \lim\frac{1}{\left(1-\frac{1}{n}\right)^{-n}} = \frac{1}{\lim\left(1-\frac{1}{n}\right)^{-n}} = \frac{1}{e}.$$

Example 8.2. Similarly, we will show that

$$\lim\left(1+\frac{1}{2n}\right)^{n} = \sqrt{e}.$$

Here, we will instead make use of subsequences. For each $n \in \mathbb{N}$, we can write

$$\left(1+\frac{1}{2n}\right)^{n} = \left[\left(1+\frac{1}{2n}\right)^{2n}\right]^{\frac{1}{2}}. \tag{8.2}$$

Now,

$$\left(1+\frac{1}{2n}\right)^{2n}$$

is a subsequence of $\left(1+\frac{1}{n}\right)^{n}$, which we know to be convergent with limit e. This tells us that the sequence given above converges:

$$\lim\left(1+\frac{1}{2n}\right)^{2n} = \lim\left(1+\frac{1}{n}\right)^{n} = e.$$

Using this in (8.2), the limit laws tell us that

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

Example 8.3. Find the flaw in the following "proof" that

$$\lim \left(1 + \frac{2}{n}\right)^n = e^2.$$

"Proof". For every $n \in \mathbb{N}$ we can write

$$\left(1 + \frac{2}{n}\right)^n = \left[\left(1 + \frac{1}{n/2}\right)^{n/2}\right]^2.$$

Since we know that

$$\lim \left(1 + \frac{1}{n}\right)^n = e,$$

we get that

$$\lim \left(1 + \frac{1}{n/2}\right)^{n/2} = e.$$

Thus,

$$\lim \left(1 + \frac{2}{n}\right)^n = \lim \left[\left(1 + \frac{1}{n/2}\right)^{n/2}\right]^2 = e^2.$$

□

What exactly goes wrong with the proof above? Well, everything is correct up until the statement

$$\lim \left(1 + \frac{1}{n/2}\right)^{n/2} = e.$$

This does **not** follow from the fact that

$$\lim \left(1 + \frac{1}{n}\right)^n = e,$$

Indeed, although it seems to be at a first glance,

$$\left(1 + \frac{1}{n/2}\right)^{n/2}$$

is not a subsequence of

$$\left(1 + \frac{1}{n}\right)^n.$$

This is because $n/2$ need not be a natural number for every $n \in \mathbb{N}$. Having said this, it is only natural to give a correct proof of this result.

Proposition 8.1. *One has* $\lim \left(1+\frac{2}{n}\right)^n = e^2$.

Proof. By the limit laws and (8.1), it is easy to see that

$$e^2 = \lim \left(1+\frac{1}{n}\right)^{2n} = \lim \left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n. \tag{8.3}$$

Next, we observe that for each $n \in \mathbb{N}$

$$\left(1+\frac{2}{n}\right)^n \leq \left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n \leq \left(1+\frac{2}{n}\right)^n \left(1+\frac{1}{n^2}\right)^n. \tag{8.4}$$

Here, we have used the fact that

$$\left(1+\frac{2}{n}+\frac{1}{n^2}\right) \leq \left(1+\frac{2}{n}\right)\left(1+\frac{1}{n^2}\right), \quad \forall n \in \mathbb{N}.$$

Now, (8.4) shows that

$$\left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n \left(1+\frac{1}{n^2}\right)^{-n} \leq \left(1+\frac{2}{n}\right)^n, \quad \forall n \in \mathbb{N}.$$

Furthermore, a direct calculation ensures that

$$\left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n \left(1+\frac{1}{n^2}\right)^{-n} = \left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n \left(1-\frac{1}{n^2+1}\right)^n$$

for all $n \in \mathbb{N}$. Combining this with the previous inequality, we find that

$$\left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n \left(1-\frac{1}{n^2+1}\right)^n \leq \left(1+\frac{2}{n}\right)^n \leq \left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n.$$

Finally, an application of Bernoulli's inequality to $\left(1-\frac{1}{n^2+1}\right)^n$ gives the inequality

$$\left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n \left(1-\frac{n}{n^2+1}\right) \leq \left(1+\frac{2}{n}\right)^n \leq \left(1+\frac{2}{n}+\frac{1}{n^2}\right)^n$$

for all $n \in \mathbb{N}$. Note also that

$$\lim \left(1-\frac{n}{n^2+1}\right) = 1.$$

Thus, by the limit laws, (8.3), and the Squeeze Theorem, we infer that

$$e^2 = \lim \left(1+\frac{2}{n}\right)^n.$$

$\square$

8.1 Approximating square roots

Let $a > 0$ be given; we will construct a sequence in $\mathbb{R}$ converging to $\sqrt{a}$. In fact, as we shall soon see, this sequence will not require that we know the numerical value of $\sqrt{a}$. As a first step, pick **any** $x_1 > 0$ from $\mathbb{R}$. For $n \geq 1$, we define

$$x_{n+1} := \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) > 0 \tag{8.5}$$

This gives rise to a recursive sequence (x_n), with $x_n > 0$ for every $n \in \mathbb{N}$. As a first step, we will show that $x_n^2 \geq a$ for all $n \geq 2$. For any n, note that x_n is a root of the following equation:

$$x^2 - 2x_{n+1}x + a = 0.$$

In particular, the quadratic equation above has a real root. It follows from the quadratic formula that the discriminant

$$4x_{n+1}^2 - 4a$$

must be non-negative, i.e. $4x_{n+1}^2 - 4a \geq 0$. Thus, $x_{n+1}^2 \geq a$ whence $x_{n+1} \geq \sqrt{a}$.

As a second step, we show that the sequence (x_n) is decreasing for all $n \geq 2$. Namely, we show that $x_{n+1} \leq x_n$, whenever $n \geq 2$. To see this, we simply calculate

$$\begin{aligned} x_n - x_{n+1} = x_n - \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) &= \frac{x_n}{2} - \frac{a}{2x_n} \\ &= \frac{1}{2}\left(\frac{x_n^2 - a}{x_n}\right) \geq 0. \end{aligned}$$

Here, we have used that $x_n^2 \geq a$ whenever $n \geq 2$. Thus, $x_{n+1} \geq x_n$ whenever $n \geq 2$. Combined, we have shown that (x_n) is an eventually decreasing sequence with $x_n \geq \sqrt{a}$ for $n \geq 2$. By the monotone convergence theorem, we know that (x_n) will converge to some $x \geq \sqrt{a}$. In fact, using the limit laws with (8.5) shows that x will satisfy

$$x = \frac{1}{2}\left(x + \frac{a}{x}\right) \geq 0.$$

Rather,

$$x = \frac{a}{x} \geq 0.$$

Thus, we have $x^2 = a$ as was required. This recursive procedure is in general quite fruitful and simple. For this reason, we give one final example before moving on to more general results.

Example 8.4. Let $x_1 \geq 2$ be given and define for $n \geq 1$:

$$x_{n+1} := 1 + \sqrt{x_n - 1}. \tag{8.6}$$

We will show that the recursive sequence (x_n) converges and calculate its limit. First, let us show (by induction) that $x_n \geq 2$ for all $n \in \mathbb{N}$. The base case $n = 1$ is trivial. Assuming that $x_n \geq 2$, we then have

$$x_{n+1} = 1 + \sqrt{x_n - 1} \geq 1 + \sqrt{2 - 1} = 2.$$

Here, we have used that the function $f(x) := \sqrt{x}$ is monotone increasing on $[0, \infty)$.[18] We conclude that $x_n \geq 2$ for all $n \in \mathbb{N}$. In particular, $x_n - 1 \geq 1$ for all $n \in \mathbb{N}$. Thus, given $n \in \mathbb{N}$, we have

$$x_{n+1} = 1 + \sqrt{x_n - 1} \leq 1 + (x_n - 1) = x_n.$$

In this last step, we have used that $\sqrt{t} \leq t$ for all $t \geq 1$.[19] Hence, (x_n) is a monotone decreasing sequence that is bounded below by 2. It therefore converges to some point $x \geq 2$. Using the limit laws and (8.6), this point x must satisfy

$$x = 1 + \sqrt{x - 1}.$$

It follows that $(x - 1) = \sqrt{x - 1}$, where $x - 1 \geq 1$. Hence, $\sqrt{x - 1} = 1$ whence $x = 2$. We conclude that $\lim x_n = 2$.

[18] To see why this is true, suppose that $0 \leq x < y$. Then, we calculate

$$\sqrt{y} - \sqrt{x} = \frac{(\sqrt{y} - \sqrt{x})(\sqrt{y} + \sqrt{x})}{\sqrt{y} + \sqrt{x}} = \frac{y - x}{\sqrt{x} + \sqrt{y}} \geq 0.$$

Thus, $\sqrt{x} \leq \sqrt{y}$ whenever $0 \leq x < y$. Since this also holds for $0 \leq x = y$, we see that $f(x) = \sqrt{x}$ is monotone increasing on $[0, \infty)$.

[19] To justify this inequality, let $t \geq 1$ and note that

$$t - \sqrt{t} = \sqrt{t}(\sqrt{t} - 1) \geq \sqrt{t} - 1 \geq 0.$$

8.2 The Bolzano-Weierstrass Theorem

In this section, we prove one of the most fundamental and useful theorems in real analysis. Although the result heavily hinges upon our cumulative efforts thus far, the theorem will easily follow from one simple observation.

Definition 8.1. Let (x_n) be a sequence of real numbers. The N^{th} term x_N is called a *peak* of the sequence (x_n) if, for every $n \geq N$, there holds $x_n \leq x_N$.

In light of this definition, we have a relatively simple proof of our key lemma.

Lemma 8.2. *Any sequence of real numbers has a monotone subsequence.*

Proof. Let (x_n) be a sequence of real numbers. We distinguish the only two possible cases.

1. Suppose (x_n) has *infinitely many* peaks. Then, we can find a subsequence (x_{n_k}) of (x_n) consisting entirely of peaks. For every index k, x_{n_k} is a peak of the sequence (x_n). In particular, x_{n_1} is a peak of the sequence (x_n). Since $n_2 > n_1$, we must then have $x_{n_1} \geq x_{n_2}$. Similarly, since x_{n_2} is a peak and $n_3 > n_2$, there holds
$$x_{n_2} \geq x_{n_3}$$
and so forth. In general, we have
$$x_{n_1} \geq x_{n_2} \geq \cdots \geq x_{n_k} \geq \cdots$$
That is, (x_{n_k}) is a monotone subsequence of (x_n).

2. Otherwise, (x_n) has only finitely many peaks. Thus, we can find $N \in \mathbb{N}$ so large that x_n is not a peak of the sequence (x_n) for all $n \geq N$. Define $x_{n_1} := x_N$. Since x_{n_1} is not a peak of the sequence (x_n), there is some x_j in our sequence, with $j \geq n_1$, such that $x_j > x_{n_1}$. In particular, $j \neq n_1$. Put then
$$x_{n_2} := x_j$$
so that $x_{n_1} < x_{n_2}$ and $N \leq n_1 < n_2$. Again, x_{n_2} is not a peak of the sequence (x_n). By a similar argument, we can find $j > n_2$ such that $x_j > x_{n_2}$. Define then $x_{n_3} := x_j$ so that
$$x_{n_1} < x_{n_2} < x_{n_3}.$$
Continuing in this way, we construct a monotone increasing subsequence (x_{n_k}) of (x_n).

In either case, we have found a monotone subsequence. □

Armed with this lemma, we can easily establish the following powerhouse of a theorem.

Theorem 8.3 (Bolzano-Weierstrass)**.** *Any bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

Remark 8.1. The theorem above continues to hold in $\mathbb{R}^m$ and $\mathbb{C}^n$ for all $m, n \in \mathbb{N}$. Such theorems are known as *compactness* theorems and are extremely powerful.

Proof of Theorem. By our lemma, any sequence has a monotone subsequence. If (x_n) is bounded, then any monotone subsequence of (x_n) is also bounded. By the monotone convergence theorem, such a subsequence converges. □

8.3 More about the Supremum

Proposition 8.4. *Let $E \subseteq \mathbb{R}$ be non-empty and bounded from above. Fix a point $y \in \mathbb{R}$; the following statements are equivalent:*

(1) y is the limit of two sequences (x_n) and (y_n) such that, for every $n \in \mathbb{N}$, $x_n \in E$ and y_n is an upper bound of E.

(2) $y = \sup E$.

Proof. Suppose that (1) holds. For every $n \in \mathbb{N}$, y_n is an upper bound for E and $x_n \in E$. Therefore, we must have

$$x_n \leq \sup E \leq y_n$$

for each $n \in \mathbb{N}$. Letting $n \to \infty$, the Squeeze Theorem implies that $y \leq \sup E \leq y$.

Conversely, let $y := \sup E$. For every $n \in \mathbb{N}$, the quantity $y - \frac{1}{n}$ is no longer an upper bound for E (y is the least upper bound). Thus, there exists $x_n \in E$ such that

$$y - \frac{1}{n} < x_n \leq y.$$

This gives us a sequence (x_n) in E with $\lim x_n = y$. Now, take (y_n) to be the constant sequence $y_n := y$. □

9 Ninth Tutorial

Since this is the last tutorial before the second midterm exam, we will spend some ftime reviewing topics in point-set topology and sequences. As for the more recent material covered (e.g. Cauchy sequences), we will focus on solving some of the ungraded homework questions.

9.1 Accumulation Points

Proposition 9.1. *Let (x_n) be a bounded sequence in $\mathbb{R}$. Then,* $\lim x_n$ *exists if and only if the sequence (x_n) has exactly one accumulation point.*

Proof. One direction is easy: if (x_n) converges to $x \in \mathbb{R}$, then so must any subsequence of (x_n). In particular, x can be the only accumulation point of (x_n).

Conversely, suppose that the sequence (x_n) has only one accumulation point, say, x. Then, every convergent subsequence (x_{n_k}) of (x_n) must converge to this point x. By way of contradiction, suppose that (x_n) does *not* converge to this point x. Then, there is some $\varepsilon_0 > 0$ such that, for every $N \in \mathbb{N}$, there exists $n \geq N$ with the property that $|x_n - x| \geq \varepsilon_0$.[20] In particular, we can find a subsequence (x_{n_k}) of (x_n) such that $\left|x_{n_k} - x\right| \geq \varepsilon_0$, for all $k \in \mathbb{N}$. Note that this subsequence is also a bounded sequence. By the Bolzano-Weierstrass theorem, we can find a subsequence $(x_{n_{k_j}})$ of our subsequence (x_{n_k}) that converges. Since $(x_{n_{k_j}})$ is then a convergent subsequence of (x_n), it must converge to x. But this is impossible for, as a subsequence of (x_{n_k}),

$$\left|x_{n_{k_j}} - x\right| \geq \varepsilon_0, \quad \text{for all } j \in \mathbb{N}.$$

This proves our claim. □

Remark 9.1. The boundedness assumption in the previous proposition is necessary. Indeed, consider the sequence (x_n) in $\mathbb{R}$ given by

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

Clearly, $\lim x_{2n} = \lim 1 = 1$ whence 1 is an accumulation point of (x_n). Let now (x_{n_k}) be a convergent subsequence of (x_n). In particular, (x_{n_k}) is bounded.

[20] To see this, just negate the definition of convergence to x.

Now, (x_{n_k}) either contains a subsequence of (x_{2n}), or a subsequence of (x_{2n+1}). However, since any subsequence of (x_{2n+1}) is unbounded, (x_{n_k}) must contain a subsequence of (x_{2n}). In particular, (x_{n_k}) contains a subsequence converging to 1. Because (x_{n_k}) is assumed to be convergent, all subsequences converge and have the same limit as (x_{n_k}). This tells us that $\lim x_{n_k} = 1$. In short, (x_n) has a single accumulations point, i.e. 1. However, (x_n) is unbounded and clearly cannot converge.

Proposition 9.2. *Let (x_n) be a bounded sequence in $\mathbb{R}$. Then, (x_n) converges if and only if*

$$\liminf x_n = \limsup x_n.$$

In either case,

$$\lim x_n = \liminf x_n = \limsup x_n.$$

Proof. Let (x_n) be a convergent sequence, with limit $x \in \mathbb{R}$. By the previous proposition, (x_n) has but a single accumulation point, denoted a. Hence,

$$\liminf x_n = \limsup x_n = a.$$

But, (x_n) is a convergent subsequence of (x_n). Therefore, $\lim x_n = x$ must be an accumulation point of (x_n). Since a is the only such point, we conclude that $x = a$. Especially,

$$\lim x_n = \liminf x_n = \limsup x_n.$$

Conversely, suppose that $\liminf x_n = \limsup x_n$. Denote by $\mathrm{Acc}(x_n)$ the set of all accumulation points of (x_n). Since (x_n) is bounded, is has a convergent subsequence (why?). Thus, $\mathrm{Acc}(x_n)$ is non-empty. If it contains two distinct elements $a_1 < a_2$, then we would have

$$\liminf x_n \leq a_1 < a_2 \leq \limsup x_n.$$

which would be a contradiction. Hence, $\mathrm{Acc}(x_n)$ consists of a single point a. In particular,

$$a = \liminf x_n = \limsup x_n.$$

Once again invoking the previous proposition, we see that $\lim x_n$ exists. As before, (x_n) is a convergent subsequence of (x_n) and therefore $\lim x_n$ is an accumulation point of (x_n). It follows that

$$\lim x_n = a = \liminf x_n = \limsup x_n.$$

□

This convergence criterion easily leads to the following:

Proposition 9.3. *Let (x_n) be a bounded sequence and fix $x \in \mathbb{R}$. If every subsequence of (x_n) has a subsequence converging to x, then $x_n \to x$ as $n \to \infty$.*

Proof. In light of the previous result, it is enough to show that x is the only accumulation point of (x_n). Clearly, since (x_n) is a subsequence of itself, there is a subsequence (x_{n_k}) converging to x. Hence, x is an accumulation point of (x_n). Now, if (x_{n_k}) is any convergent subsequence of (x_n), it must converge to x. Indeed, this is because (x_{n_k}) has (by assumption) a subsequence converging to x. We infer that every convergent subsequence of (x_n) converges to x. It follows that

$$\limsup x_n = \liminf x_n = x$$

and the proof is complete. □

Example 9.1. Let (x_n) be an unbounded sequence. We claim that there exists a subsequence (x_{n_k}) such that

$$\lim \frac{1}{x_{n_k}} = 0.$$

Since (x_n) is unbounded, we know that for every $M > 0$, one can find some $n \in \mathbb{N}$ such that $|x_n| > M$. In fact, we can find *infinitely many* n such that $|x_n| > M$. Otherwise, we would have

$$|x_n| \leq M$$

for all but finitely many $n \in \mathbb{N}$. If this were the case, then the sequence (x_n) would be bounded. Now, given $k \geq 1$, we can find $n_k \in \mathbb{N}$ such that $\left|x_{n_k}\right| > k$. Having chosen x_{n_k}, we can guarantee the existence of $n_{k+1} \in \mathbb{N}$ such that

$$\left|x_{n_{k+1}}\right| > k + 1 \quad \text{and} \quad n_{k+1} > n_k.$$

In this way, we construct a subsequence (x_{n_k}) such that $\left|x_{n_k}\right| > k$, for every $k \in \mathbb{N}$. It follows that,

$$\left|\frac{1}{x_{n_k}} - 0\right| = \frac{1}{\left|x_{n_k}\right|} < \frac{1}{k}, \quad \forall k \geq 1.$$

Hence, $\lim x_{n_k} = 0$.

We conclude this section with a more difficult example.

Example 9.2. Let (x_n) be a bounded sequence and let

$$\liminf x_n = l \leq L = \limsup x_n.$$

Show that if $x_{n+1} - x_n \to 0$ then all points in $[l, L]$ are accumulation points of (x_n).

Solution. Suppose there exists $x \in [l, L]$ which is not an accumulation point of (x_n). Since we know that both l and L are accumulation points, we must have $l < x < L$. As x is not an accumulation point of the sequence, there exist $\varepsilon > 0$ and $N \in \mathbb{N}$ such that

$$x_n \notin (x - \varepsilon, x + \varepsilon) \quad \forall n \geq N \tag{9.1}$$

Furthermore, since $x_{n+1} - x_n \to 0$, we may suppose without loss of generality that N is sufficiently large so that

$$|x_{n+1} - x_n| < \varepsilon \quad \forall n \geq N. \tag{9.2}$$

Let now $\delta = L - x > 0$. Since L is an accumulation point, we may find $M \geq N$ such that

$$x_M \in (L - \delta, L + \delta) = (x, L + \delta).$$

Combining the above with equation (9.1), we see that $x_M \geq x + \varepsilon$. We now show by induction that $x_n \geq x + \varepsilon$ for all $n \geq M$. Indeed, if $x_n \geq x + \varepsilon$ for some $n \geq M \geq N$, then equation (9.1) implies that either $x_{n+1} \geq x + \varepsilon$ or $x_{n+1} \leq x - \varepsilon$. In the latter case, we have

$$|x_{n+1} - x_n| \geq x_n - x_{n+1} \geq x + \varepsilon - (x - \varepsilon) \geq 2\varepsilon.$$

Since this contradicts equation (9.2), we conclude that $x_{n+1} \geq x + \varepsilon$. We have deduced that

$$x_n \geq x + \varepsilon > l + \varepsilon$$

for all $n \geq M$. But this contradicts the fact that l is an accumulation point. □

9.2 Cauchy Sequences and Applications

Example 9.3. Let (x_n) be a sequence of real numbers and let $r \in (0, 1)$. In addition, suppose that $|x_{n+1} - x_n| < r^n$, for all $n \in \mathbb{N}$. We show that (x_n) is convergent. Based on the given information, it is not clear what $\lim x_n$ should be. Thus, our only real option is to check that (x_n) is Cauchy.[21]

[21] Since $\mathbb{R}$ is complete, a sequence in $\mathbb{R}$ converges if and only if it is Cauchy. The importance of this fact should not be underestimated.

To this end, let $\varepsilon > 0$ be given, and choose $N \in \mathbb{N}$ so large that

$$\frac{r^N}{1-r} < \varepsilon.$$

This can be done because $r \in (0,1)$ implies $\lim r^n = 0$. Now, given any natural numbers $m > n \geq N$,

$$\begin{aligned}
|x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} + \cdots - x_{n+1} + x_{n+1} - x_n| \\
&\leq \sum_{k=n}^{m-1} |x_{k+1} - x_k| \\
&\leq \sum_{k=n}^{m-1} r^k \\
&= \sum_{k=0}^{m-1} r^k - \sum_{k=0}^{n-1} r^k \\
&= \frac{1-r^m}{1-r} - \frac{1-r^n}{1-r} \\
&= \frac{r^n - r^m}{1-r} \\
&\leq \frac{r^n}{1-r} \\
&\leq \frac{r^N}{1-r}.
\end{aligned}$$

Hence, $|x_m - x_n| < \varepsilon$ whenever $m \geq n \geq N$. This shows that (x_n) is Cauchy.

Problem 4. *Let (x_n) be a sequence with $x_n \geq 0$, for all $n \in \mathbb{N}$. Suppose that $\lim (-1)^n x_n$ exists. Prove that (x_n) converges.*

Proof. Let x denote the limit of the sequence $(-1)^n x_n$. Then, all subsequences of $(-1)^n (x_n)$ converge to this point x. In particular,

$$x = \lim (-1)^{2n} x_{2n} = \lim x_{2n} \geq 0$$

and, since $x_n \geq 0$ for all $n \in \mathbb{N}$,

$$x = \lim (-1)^{2n+1} x_{2n+1} = \lim (-x_{2n+1}) \leq 0.$$

Therefore, the only possibility is to have $x = 0$. That is, $\lim (-1)^n x_n = 0$. Let now $\varepsilon > 0$ be given. Using this last property, we can find $N \in \mathbb{N}$ such that

$$|x_n| = \left|(-1)^n x_n - 0\right| < \varepsilon$$

for all $n \geq N$. This means that $\lim x_n = 0$. $\square$

9.3 Absolutely Convergent Series

Let us now return to the notion of *absolute convergence*, which was quickly introduced in §7.2.1. For the sake of clarity, we will reiterate everything here. Let (x_n) be a sequence of real numbers and define, for each $N \in \mathbb{N}$, the partial sum

$$S_N := \sum_{n=1}^{N} x_n.$$

Note that (S_N) is itself a sequence of real numbers indexed by the variable N. We then say that the series $\sum_{n=1}^{\infty} x_n$ converges if the sequence of partial sums (S_N) converges. If this is the case, we *define*

$$\sum_{n=1}^{\infty} x_n := \lim_{N\to\infty} S_N.$$

On the other hand, the sequence $\sum_{n=1}^{\infty} x_n$ is said to converge *absolutely* when the series $\sum_{n=1}^{\infty} |x_n|$ converges. By the monotone convergence theorem, this is equivalent to saying that the partial sums $\sum_{n=1}^{N} |x_n|$ are bounded:

Lemma 9.4. *Let (x_n) be a sequence of real numbers and define*

$$S_N := \sum_{n=1}^{N} |x_n|, \quad \forall N \in \mathbb{N}.$$

Then, the series $\sum_{n=1}^{\infty} x_n$ is absolutely convergent if and only if there exists $M > 0$ such that $|S_N| \leq M$ for all $N \in \mathbb{N}$. That is, $\sum_{n=1}^{\infty} x_n$ converges absolutely if and only if (S_N) is bounded.

Proof. By definition, if $\sum_{n=1}^{\infty} |x_n|$ converges, then the sequence (S_N) defined above is convergent with $\lim S_N = \sum_{n=1}^{\infty} |x_n|$. Since convergent sequences are bounded, it follows that (S_N) is bounded. Conversely, assume that there exists $M > 0$ such that

$$|S_N| = S_N = \sum_{n=1}^{\infty} |x_n| \leq M$$

for all $N \in \mathbb{N}$. Since $S_N \leq S_{N+1}$ for all $N \in \mathbb{N}$, the sequence (S_N) is therefore a bounded monotone increasing sequence. By the monotone convergence theorem, (S_N) is convergent. By definition, this implies the absolute convergence of $\sum_{n=1}^{\infty} x_n$. □

Theorem 7.8 from §7.2.1 states that every absolutely convergent series is also convergent. Equipped with the notion of a Cauchy sequence, we can finally supply the proof of this important result.

Proof of Theorem 7.8. Let (x_n) be a sequence of real numbers such that $\sum_{n=1}^{\infty} x_n$ converges absolutely. That is, assume that $\sum_{n=1}^{\infty} |x_n|$ is convergent. Define for each $N \in \mathbb{N}$ the quantities

$$S_N := \sum_{n=1}^{N} x_n \quad \text{and} \quad R_N := \sum_{n=1}^{N} |x_n| \,.$$

Since $\sum_{n=1}^{\infty} |x_n|$ converges, the sequence (R_N) is Cauchy in $\mathbb{R}$. So, for each $\varepsilon > 0$, we can find $K \in \mathbb{N}$ such that

$$|R_N - R_M| = \sum_{n=M+1}^{N} |x_n| < \varepsilon$$

whenever $N \geq M \geq N$. In particular, by the triangle inequality, we obtain

$$|S_N - S_M| = \left| \sum_{n=M+1}^{N} x_n \right| \leq \sum_{n=M+1}^{N} |x_n| < \varepsilon$$

for all $N \geq M \geq K$. This shows that the sequence of partial sums (S_N) is Cauchy and thus convergent. It follows that $\sum_{n=1}^{\infty} x_n$ converges. □

In order to show that absolute convergence implies convergence, we used that every Cauchy sequence in $\mathbb{R}$ is convergent. As it turns out, the convergence of every Cauchy sequence is equivalent to every absolutely convergent series being convergent. Namely, if we assume that every absolutely convergent series converges, we can show that every Cauchy sequence must converge.

Theorem 9.5. *The following statements are equivalent:*

(1) *Every Cauchy sequence in $\mathbb{R}$ is convergent.*

(2) *Every absolutely convergent series is convergent.*

We omit the proof of this theorem as it does not pertain directly to the course material.

9.4 Divergence of the Harmonic Series

We now consider the infinite series

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

which is often called the *harmonic series*. As mentioned previously, the infinite series above can be formally defined by the following

$$\lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n} =: \lim_{N\to\infty} S_N.$$

provided the limit exists. Here, we are setting $S_N := \sum_{n=1}^{N} \frac{1}{n}$ for each $N \geq 1$. Then, the infinite series $\sum_{n=1}^{\infty} \frac{1}{n}$ exists if and only if the sequence (S_N) converges. Thus, to show that the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent, it is enough to show that the sequence (S_N) is not Cauchy.

Let now $N, M \in \mathbb{N}$ be given. If $M > N$, an easy calculation tells us that

$$\begin{aligned} |S_N - S_M| = S_M - S_N = \sum_{n=1}^{M} \frac{1}{n} - \sum_{n=1}^{N} \frac{1}{n} &= \sum_{n=N+1}^{M} \frac{1}{n} \\ &\geq \sum_{n=N+1}^{M} \frac{1}{M} \\ &= \frac{M-N}{M} \\ &= 1 - \frac{N}{M}. \end{aligned}$$

In particular, taking $M = 2N$ gives

$$|S_{2N} - S_N| \geq 1 - \frac{1}{2} = \frac{1}{2}, \quad \forall N \in \mathbb{N}.$$

This shows that (S_N) cannot be Cauchy, and thus does not converge.

10 Tenth Tutorial

This tutorial is all about limits of functions. Before giving any problems and examples, let us first reiterate the $\varepsilon - \delta$ definition of the limit. To properly make sense of the behaviour of a function "near a point x", it is critical that the function be defined at *some* points near x. This is precisely what the concept of a cluster point formalizes:

Definition 10.1. Let $A \subseteq \mathbb{R}$. A point $a \in \mathbb{R}$ is called a *cluster point* of A if, for each $\delta > 0$, $V_\delta(a) \cap (A \setminus \{a\}) \neq \emptyset$. Or, as we shall see later in Corollary 10.7, a is a cluster point of A if and only if

$$a \in \overline{A \setminus \{a\}}.$$

Plainly put, this guarantees that there are "enough points" in A near a to define a meaninful limit for a function f defined on A.

Definition 10.2. Let $A \subseteq \mathbb{R}$ and a be a cluster point of A. Let $f : A \to \mathbb{R}$ be a function and fix $L \in \mathbb{R}$. We say that $f(x)$ *converges* to L as $x \to a$, written

$$\lim_{x \to a} f(x) = L,$$

provided for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{whenever } x \in A \text{ and } 0 < |x - a| < \delta.$$

Or, symbolically, if

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in A)(0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon).$$

Note that the value of $f(a)$ (if it even exists) is of no importance as the requirement that $0 < |x - a| < \delta$ excludes the possibility that $x = a$.

Remark 10.1. The $\varepsilon - \delta$ definition of the limit bears a significant resemblance to the $\varepsilon - N$ definition for sequences. Indeed, the ε serves as an error term to measure how close to the limit L the value $f(x)$ is. However, instead of looking at very large natural numbers, we care about values of x close to a. So, in the $\varepsilon - \delta$ definition, the δ-term measures how close to a the variable x is. Hence, the $\varepsilon - \delta$ definition formalizes the following intuitive statement:

> *If the distance between $f(x)$ and L can be made arbitrarily small by making x sufficiently close to a, then* $\lim_{x \to c} f(x) = L$.

10.1 Examples Using the $\varepsilon - \delta$ Definition

To help us better understand the $\varepsilon - \delta$ characterization of the limit, we work out two intuitive examples explicitly.

Example 10.1. Using the $\varepsilon - \delta$ definition of the limit, prove that

$$\lim_{x \to c} f(x) = \sqrt{c}, \quad \forall c \geq 0,$$

where $f : [0, \infty) \to \mathbb{R}$ is given by $f(x) := \sqrt{x}$.

Proof. Let $\varepsilon > 0$ be given, we must show that there is some $\delta > 0$ such that

$$\left| f(x) - \sqrt{c} \right| < \varepsilon$$

for all $x \in [0, \infty)$ with $0 < |x - c| < \delta$. We distinguish two possible cases:

1. Case $c = 0$. Here, we really want to show that $\sqrt{x} \to 0$ as $x \to 0$. Choose $\delta := \varepsilon^2 > 0$. Then, if $0 < x < \delta = \varepsilon^2$, we calculate

   $$|f(x) - 0| = \left| \sqrt{x} \right| = \sqrt{x} < \sqrt{\delta} < \sqrt{\varepsilon^2} = \varepsilon.$$

   This shows that $\sqrt{x} \to 0$ as $x \to 0$.

2. We now handle the case $c > 0$. For any $x \in [0, \infty)$ with $x \neq c$, we calculate

   $$\begin{aligned} \left| \sqrt{x} - \sqrt{c} \right| &= \left| \frac{(\sqrt{x} - \sqrt{c})(\sqrt{x} + \sqrt{c})}{\sqrt{x} + \sqrt{c}} \right| \\ &= \frac{|x - c|}{\sqrt{x} + \sqrt{c}} \\ &\leq \frac{|x - c|}{\sqrt{c}}. \end{aligned}$$

   Now, let

   $$\delta := \sqrt{c}\varepsilon > 0.$$

   If $x \geq 0$ and $0 < |x - c| < \delta$ we find that

   $$\left| \sqrt{x} - \sqrt{c} \right| \leq \frac{|x - c|}{\sqrt{c}} < \frac{\delta}{\sqrt{c}} = \varepsilon.$$

This completes the proof. □

Let us consider another example.

Example 10.2. Prove that, for every $c > 0$, one has

$$\lim_{x \to c} \frac{1}{\sqrt{x}} = \frac{1}{\sqrt{c}}.$$

Of course, we are are viewing $\frac{1}{\sqrt{x}}$ as a function with domain equal to $(0, \infty)$.

Proof. Given $\varepsilon > 0$ and $x > 0$ with $x \neq c$, we have the estimate

$$\begin{aligned} \left| \frac{1}{\sqrt{x}} - \frac{1}{\sqrt{c}} \right| = \frac{\left| \sqrt{x} - \sqrt{c} \right|}{\sqrt{x}\sqrt{c}} &= \frac{|x - c|}{\sqrt{x}\sqrt{c}\left(\sqrt{x} + \sqrt{c}\right)} \\ &\leq \frac{|x - c|}{x\sqrt{c}}. \end{aligned}$$

Let now

$$\delta := \min\left\{ \frac{c^{3/2}}{2}\varepsilon, \frac{c}{2} \right\} > 0.$$

If $0 < |x - c| < \delta$, then $x \neq c$ and

$$-\delta < x - c < \delta.$$

But then, $\delta \leq \frac{c}{2}$ would imply

$$x > c - \delta \geq \frac{c}{2}$$

so that

$$\left| \frac{1}{\sqrt{x}} - \frac{1}{\sqrt{c}} \right| \leq \frac{|x - c|}{x\sqrt{c}} < \frac{\delta}{\sqrt{c}} \frac{2}{c} = \frac{2\delta}{c^{3/2}} \leq \varepsilon.$$

This shows that

$$\frac{1}{\sqrt{x}} \to \frac{1}{\sqrt{c}} \quad \text{as } x \to c.$$

□

10.2 Cluster Points and the Closure

We now discuss the link between closed sets and cluster points. More precisely, we study the relationship between the cluster points of a set A and the closure of A.

Proposition 10.1. *Let A be a subset of $\mathbb{R}$. Denote by A' the set of all cluster points of A. Then, $\bar{A} = A \cup A'$.*

Proof. First, if $x \in A'$ then $x \in \overline{A \setminus \{x\}} \subseteq \bar{A}$. Hence, we see that $A' \subseteq \bar{A}$. Since there also holds $A \subseteq \bar{A}$, it follows that

$$A \cup A' \subseteq A.$$

Conversely, let $x \in \bar{A}$. If $x \in A$ then we have $x \in A \cup A'$. Otherwise, $x \notin A$ whence $A = A \setminus \{x\}$. In particular,

$$\bar{A} = \overline{A \setminus \{x\}}.$$

Since $x \in \bar{A}$, we see that

$$x \in \overline{A \setminus \{x\}}$$

whence $x \in A'$. This shows that $\bar{A} \subseteq A \cup A'$ and the proof is complete. □

This gives us an easy corollary that helps to get a handle on this concept.

Corollary 10.2. *A subset of $\mathbb{R}$ is closed if and only if it contains all of its cluster points.*

Proof. By the previous proposition, it is easy to see that $\bar{A} = A$ if and only if A contains all of its cluster points. □

Let us think back to our discussion of the Cantor set $\mathfrak{C}$. There, we had introduced the notion of an *isolated point*. For the sake of completeness, let us reiterate this definition here.

Definition 10.3. Let A be a subset of $\mathbb{R}$. A point $x \in A$ is said to be an *isolated point* of A if there exists $\varepsilon > 0$ such that $V_\varepsilon(x) \cap (A \setminus \{x\})$ is empty.

Remark 10.2. This is quite the opposite of a cluster point. Indeed, a cluster point x of A is a point in $\mathbb{R}$ such that we can always find elements of A arbitrarily close to x. On the other hand, when dealing with an isolated point x, if we "zoom in" enough, there are no points in A that are near x.

10.3 The Relationship with Accumulation Points

Let (x_n) be a sequence in $\mathbb{R}$. The sequence (x_n) is said to be *injective* if, for any $n, m \in \mathbb{N}$, having $x_n = x_m$ implies $n = m$. In other words, the sequence (x_n) is called injective if it is an injective function $\mathbb{N} \to \mathbb{R}$.

Lemma 10.3. *Let A be an infinite subset of $\mathbb{R}$ and let x be a cluster point of A. For each $\varepsilon > 0$, the intersection*

$$V_\varepsilon(x) \cap (A \setminus \{x\})$$

is infinite.

Proof. Given $\varepsilon > 0$, we at least know that the intersection above is non-empty (because x is a cluster point of A). Assume for a contradiction that it is finite. Then,

$$V_\varepsilon(x) \cap (A \setminus \{x\}) = \{x_1, \ldots, x_n\} \tag{$*$}$$

for some $x_1, \ldots, x_n \in \mathbb{R}$. Clearly, every x_i is different from x. Let now

$$\alpha := \min_{1 \leq i \leq n} |x_i - x| > 0.$$

Also, $\alpha < \varepsilon$. Since x is a cluster point of A, we know that $V_\alpha(x) \cap (A \setminus \{x\})$ is non-empty. Choose any y from this intersection and note that

$$y \in V_\alpha(x) \cap (A \setminus \{x\}) \implies y \in V_\varepsilon(x) \cap (A \setminus \{x\}).$$

However, $|y - x| < \alpha < |x - x_i|$, for all $i = 1, \ldots, n$. This means that $y \neq x_i$, for all $i = 1, \ldots, n$. Since $y \in V_\varepsilon(x) \cap (A \setminus \{x\})$, this contradicts ($*$). $\square$

Proposition 10.4. *Let (x_n) be an injective sequence in $\mathbb{R}$. A point $x \in \mathbb{R}$ is an accumulation point of (x_n) if and only if it is also a cluster point of the set*

$$S := \{x_n : n \in \mathbb{N}\} \subset \mathbb{R}.$$

Proof. First, let $x \in \mathbb{R}$ be an accumulation point of (x_n). That is, one can find a subsequence (x_{n_k}) of (x_n) converging to x. Given $\varepsilon > 0$, one can always find $N \in \mathbb{N}$ so large that $\left|x_{n_k} - x\right| < \varepsilon$ whenever $n_k \geq N$. This is the same as saying that $x_{n_k} \in V_\varepsilon(x)$, for every $n_k \geq N$. In particular, $V_\varepsilon(x) \cap \{x_n : n \in \mathbb{N}\}$ is non-empty. Since our sequence is injective, this implies that

$$V_\varepsilon(x) \cap (S \setminus \{x\}) = \varnothing.$$

As $\varepsilon > 0$ was arbitrary, we conclude that x is a cluster point of S.

Conversely, let x be a cluster point of the set S. Because (x_n) is injective, S must be infinite. For a given $k \in \mathbb{N}$, we know that

$$V_{\frac{1}{k}}(x) \cap (S \setminus \{x\}) \subseteq S \neq \emptyset.$$

In fact, by the previous lemma, there will be infinitely many elements in this intersection. This allows us to construct a subsequence (x_{n_k}) of (x_n) such that

$$\left|x_{n_k} - x\right| < \frac{1}{k}, \quad \forall k \in \mathbb{N}.$$

By the squeeze theorem, we conclude that $\lim x_{n_k} = x$. Especially, x is an accumulation point of (x_n). □

Remark 10.3. The previous claim fails if we do not require (x_n) to be injective. Indeed, consider the sequence $(x_n) := (-1)^n$. Obviously,

$$\lim x_{2n} = 1 \quad \text{and} \quad \lim x_{2n+1} = -1.$$

Therefore, ± 1 are accumulation points of the sequence (x_n). Since any subsequence of (x_n) contains a subsequence of (x_{2n}) or (x_{2n+1}), we see that these are the only accumulation points of the sequence. That is,

$$\operatorname{Acc}(x_n) = \{\pm 1\}.$$

However, the set

$$\{x_n : n \in \mathbb{N}\} = \{1, -1\}$$

has **no** cluster points. This is confirmed by the following result.

Proposition 10.5. *A finite subset of* $\mathbb{R}$ *has no cluster points.*

Proof. Let $A \subseteq \mathbb{R}$ be finite. If A is empty, then the claim is obvious. Otherwise, we can write A in the following way:

$$A = \{x_1, \ldots, x_n\}.$$

Now, assume that $x \in \mathbb{R}$ is a cluster point of x. Then, $x \in \overline{A \setminus \{x\}}$. However, since finite sets are close, this implies that

$$x \in \overline{A \setminus \{x\}} = A \setminus \{x\}.$$

This contradiction shows that A cannot have any cluster points. □

As mentioned in Definition 10.1, we will now verify that $x \in \mathbb{R}$ is a cluster point of a set A if and only if

$$x \in \overline{A \setminus \{x\}}$$

To establish this, it is enough to prove the following theorem.

Theorem 10.6. *Let A be a subset of $\mathbb{R}$ and fix a point x. The following statements are equivalent.*

(1) $x \in \bar{A}$;

(2) every open set U containing x intersects A, i.e. $U \cap A \neq \varnothing$;

(3) for every $\varepsilon > 0$, one has $V_\varepsilon(x) \cap A \neq \varnothing$.

Proof. We first show that (1) is equivalent to (2). Let $x \in \bar{A}$ and assume that there is some open set $U \ni x$ with $U \cap A = \varnothing$. Then, $A \subseteq U^c$ whence $\bar{A} \subseteq U^c$. But, this would imply $x \in U \cap U^c$ which is absurd. Conversely, suppose that every open set $U \ni x$ has non-empty intersection with A. We must prove that $x \in \bar{A}$. If $x \notin \bar{A}$, then we can find a closed set $F \supseteq A$ **not** containing the point x. In this case, F^c is an open set containing the point x. By hypothesis, we must have

$$\varnothing \neq F^c \cap A \subseteq F \cap F^c,$$

which is a contradiction. We have thus established the equivalence (1) $\iff$ (2).

Let us now show that (2) is equivalent to (3), whence the theorem would follow. Since every $V_\varepsilon(x)$ is open, it is obvious that (2) implies (3). Conversely, suppose that (3) holds and let U be an open set containing the point x. Choose $\varepsilon > 0$ such that $V_\varepsilon(x) \subseteq U$. By assumption, $V_\varepsilon(x) \cap A \neq \varnothing$. This implies that

$$U \cap A \supseteq V_\varepsilon(x) \cap A \neq \varnothing.$$

This completes the proof. □

Corollary 10.7. *Let $A \subseteq \mathbb{R}$ and fix $a \in \mathbb{R}$. Then, a is a cluster point of A if and only if $a \in \overline{A \setminus \{a\}}$.*

11 Eleventh Tutorial

Let's first spend a minute discussing the definition of the limit for sequences. Let (x_n) be a sequence and fix a point $x \in \mathbb{R}$. We have $\lim x_n = x$ if and only if

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n \geq N \implies |x_n - x| < \varepsilon).$$

If $\lim x_n \neq x$, then negating the above says that

$$(\exists \varepsilon > 0)(\forall N \in \mathbb{N})(\exists n \geq N \implies |x_n - x| \geq \varepsilon).$$

This means that there exist *infinitely many* $n \in \mathbb{N}$ with the property that

$$|x_n - x| \geq \varepsilon.$$

We cannot say that there exists $N \in \mathbb{N}$ such that $|x_n - x| \geq \varepsilon$ for all $n \geq N$. Indeed, this would imply that $|x_n - x| \geq \varepsilon$ for all *but finitely many* n, which is entirely different from the statement that $|x_n - x| \geq \varepsilon$ for infinitely many $n \in \mathbb{N}$.

11.1 More Examples of Function Limits

Example 11.1. We show that

$$\lim_{x \to a} \frac{1}{x^2} = \frac{1}{a^2}$$

for all $a \neq 0$. Obviously, $f(x) = \frac{1}{x^2}$ is considered as a function with domain $\mathbb{R} \setminus \{0\}$. Let now $\varepsilon > 0$ be given. If $\delta > 0$ and $x \neq 0$ satisfies $0 < |x - a| < \delta$, then we have the estimate

$$\left|\frac{1}{x^2} - \frac{1}{a^2}\right| = \frac{\left|x^2 - a^2\right|}{x^2 a^2} = \frac{|x - a|\,|x + a|}{x^2 a^2} < \delta \frac{|x| + |a|}{x^2 a^2}.$$

If in addition $\delta \leq |a|$, then

$$|x| \leq |x - a| + |a| < \delta + |a| \leq 2\,|a|\,.$$

Hence,

$$\left|\frac{1}{x^2} - \frac{1}{a^2}\right| < \delta \frac{|x| + |a|}{x^2 a^2} \leq \delta \frac{3\,|a|}{x^2 a^2} = \frac{3\delta}{x^2\,|a|}.$$

If we assume further that $\delta \leq \frac{|a|}{2}$, then

$$||x| - |a|| \leq |x - a| < \delta$$

implies $|a| - \delta < |x| < |a| + \delta$ whence $|x| > |a| - \frac{|a|}{2} = \frac{|a|}{2}$. Rather, we obtain

$$\frac{1}{x^2} < \frac{4}{a^2}$$

so that
$$\left|\frac{1}{x^2} - \frac{1}{a^2}\right| < \frac{3\delta}{x^2\,|a|} \le \frac{12\delta}{|a|^3}.$$
Finally, if
$$\delta \le \frac{\varepsilon\,|a|^3}{12},$$
then
$$\left|\frac{1}{x^2} - \frac{1}{a^2}\right| < \frac{3\delta}{x^2\,|a|} \le \frac{12\delta}{|a|^3} < \varepsilon.$$
Therefore, we define
$$\delta := \min\left\{\frac{|a|}{2}, \frac{\varepsilon\,|a|^3}{12}, |a|\right\} = \min\left\{\frac{|a|}{2}, \frac{\varepsilon\,|a|^3}{12}\right\}.$$
By our earlier calculations, for every $x \neq 0$ with $0 < |x - a| < \delta$, one has
$$\left|\frac{1}{x^2} - \frac{1}{a^2}\right| < \varepsilon.$$
This completes the proof.

Example 11.2. Given $a \neq -1$, we claim that
$$\lim_{x \to a} \frac{x}{1+x} = \frac{a}{1+a}.$$
Again, here we are viewing $\frac{x}{1+x}$ as a function $\mathbb{R} \setminus \{-1\} \to \mathbb{R}$.

Proof. Let $\varepsilon > 0$ be given and fix $a \neq -1$. Henceforth, we shall always assume that $x \neq -1$. If for some $\delta > 0$ one has $0 < |x - a| < \delta$, then
$$\left|\frac{1}{1+x} - \frac{1}{1+a}\right| = \frac{|x-a|}{|1+x|\,|1+a|} < \frac{\delta}{|1+x|\,|1+a|}.$$
Next, we have to "rid ourselves" of the $|1+x|$ term. To this end, suppose additionally that
$$\delta \le \frac{|1+a|}{2}.$$
Then, by the reverse triangle inequality,
$$|1+x| \ge |1+a| - |x-a| > |1+a| - \delta \ge \frac{|1+a|}{2} > 0.$$

Therefore,

$$\left|\frac{1}{1+x} - \frac{1}{1+a}\right| < \frac{\delta}{|1+x|\,|1+a|} \leq \frac{2\delta}{(1+a)^2}.$$

Thus, if

$$\delta \leq \varepsilon \frac{(1+a)^2}{2}$$

then

$$\left|\frac{1}{1+x} - \frac{1}{1+a}\right| < \frac{2\delta}{(1+a)^2} \leq \varepsilon.$$

Finally, we define

$$\delta := \min\left\{\frac{|1+a|}{2}, \varepsilon\frac{(1+a)^2}{2}\right\}.$$

By the calculations we carried out above, we conclude that

$$\left|\frac{1}{1+x} - \frac{1}{1+a}\right| < \varepsilon$$

for all $x \neq -1$ with $0 < |x - a| < \delta$. □

Remark 11.1. Using the sequential criterion for the limit, we can easily conclude that the limit

$$\lim_{x \to -1} \frac{x}{1+x}$$

does not exist. Indeed, consider the sequence (x_n) in $\mathbb{R}$ defined by

$$x_n := -1 - \frac{1}{n}.$$

Clearly, $x_n \neq -1$ for all $n \in \mathbb{N}$ and $\lim x_n = -1$. Furthermore, for each $n \in \mathbb{N}$ we calculate:

$$f(x_n) = \frac{x_n}{1+x_n} = \frac{-1-\frac{1}{n}}{-\frac{1}{n}} = n\left(1 + \frac{1}{n}\right) \geq n.$$

Hence, we get that $\lim f(x_n)$ does not exist in $\mathbb{R}$. This shows that $\lim\limits_{x \to -1} \dfrac{x}{1+x}$ does not exist either.

11.2 Limits at Infinity

Let $f : \mathbb{R} \to \mathbb{R}$ be a function; we would like to make the statement

$$\lim_{x \to \pm\infty} f(x) = \ell$$

meaningful and rigorous. Ideally, we want to capture what it means for a function f to "approach" a value ℓ "at infinity". Naturally, we achieve this by tweaking the $\varepsilon - \delta$ definition for the limit. In fact, I think our definition will more closely resemble the $\varepsilon - N$ definition for sequences.

Definition 11.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We say that $f(x)$ converges to $\ell \in \mathbb{R}$ as $x \to \infty$, written

$$\lim_{x \to \infty} f(x) = \ell,$$

if for every $\varepsilon > 0$, there is some $M > 0$ such that $|f(x) - \ell| < \varepsilon$ whenever $x > M$. In more symbolic terms,

$$(\forall \varepsilon > 0)(\exists M > 0)(\forall x > M \implies |f(x) - \ell| < \varepsilon).$$

Similarly, we can define limits at $-\infty$.

Definition 11.2. We say that $f(x)$ converges to $\ell \in \mathbb{R}$ as $x \to -\infty$, denoted

$$\lim_{x \to -\infty} f(x) = \ell,$$

if for every $\varepsilon > 0$, there is some $N < 0$ such that $|f(x) - \ell| < \varepsilon$ whenever $x < N$. Or, symbolically,

$$(\forall \varepsilon > 0)(\exists N < 0)(\forall x < N \implies |f(x) - \ell| < \varepsilon).$$

We will write

$$\ell = \lim_{|x| \to \infty} f(x)$$

to say that

$$\ell = \lim_{x \to \infty} f(x) = \lim_{x \to -\infty} f(x).$$

Example 11.3. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) := \frac{1}{1+x^2}$. Then,

$$\lim_{|x| \to \infty} f(x) = 0.$$

Proof. Let $\varepsilon > 0$ be given and let $M \in \mathbb{N}$ be so large that

$$M^2 > \frac{1}{\varepsilon} - 1.$$[22]

Clearly,

$$\frac{1}{1+M^2} < \varepsilon.$$

If $x > M$, then $x^2 > M^2$ and $1 + x^2 > 1 + M^2$. Therefore,

$$\left|\frac{1}{1+x^2}\right| = \frac{1}{1+x^2} < \frac{1}{1+M^2} < \varepsilon.$$

This shows that

$$\lim_{x\to\infty} f(x) = 0.$$

Similarly, one can show that $\lim_{x\to-\infty} f(x) = 0$. ◻

11.3 Continuity of Functions

We have now come to the topic of continuity. For the moment, let us fix a set $A \subseteq \mathbb{R}$ and a point $a \in A$. We say that a function $f : A \to \mathbb{R}$ is *continuous* at the point a if for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - f(a)| < \varepsilon$$

whenever $x \in A$ satisfies $|x - a| < \delta$. Equivalently, if

$$(\forall \varepsilon > 0)\,(\exists \delta > 0)(\forall x \in A)(|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon). \tag{11.1}$$

Note that in contrast to the $\varepsilon - \delta$ definition of the limit, we do not require that $|x - a|$ be positive. That is, we allow for $x = a$. It is a routine exercise to show that the continuity of f at a implies the continuity of $|f|$ at a. Indeed, the proof is immediate from the reverse triangle inequality:

$$||f(x)| - |f(a)|| \leq |f(x) - f(a)|\,.$$

This is highly reminiscent of a result we had for sequences, and as such it is natural to check whether the converse holds.

[22]Note that such and M exists by the Archimedean property.

Example 11.4. Consider the function $f : [0,1] \to \mathbb{R}$ given by the rule

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

Clearly, $|f(x)| = |\pm 1| = 1$ for all $x \in [0,1]$. Hence, $|f|$ is continuous. However, f will turn out to be discontinuous at every point $x \in [0,1]$. To see this, we fix a point $x \in [0,1]$ and distinguish the only two possible cases.

(1) Suppose that $x \in \mathbb{Q}$ so that $f(x) = 1$. By their density, we can select a sequence of irrational numbers (x_n) from $[0,1]$ such that $\lim x_n = x$. Since every $x_n \notin \mathbb{Q}$, we see that

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} (-1) = -1 \neq 1 = f(x).$$

(2) Suppose that $x \notin \mathbb{Q}$, whence $f(x) = -1$. As above, choose a sequence (x_n) in $\mathbb{Q} \cap [0,1]$ that converges to x as $n \to \infty$. Since every x_n is rational, we see that

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} 1 = 1 \neq -1 = f(x).$$

In either case, we see (by the sequential criterion for continuity) that the function f is nowhere continuous on $[0,1]$.

Let us now see how the continuity of a functions helps to locally preserve the “sign” or “magnitude” of a function.

Proposition 11.1. *Let $A \subseteq \mathbb{R}$ and assume that $f : A \to \mathbb{R}$ is continuous at a point $a \in A$. If $f(a) > 0$, there exists a δ-neighbourhood $V_\delta(a)$ such that $f(x) > 0$ for all $x \in A \cap V_\delta(a)$.*

Remark 11.2. Loosely speaking, if $f : A \to \mathbb{R}$ is continuous at a and $f(a) > 0$, the result above says that $f(x)$ is positive at all points in A “near” a.

Proof. Let us define

$$\varepsilon := \frac{f(a)}{2} > 0.$$

Since f is continuous at a, we can find $\delta > 0$ such that

$$|f(x) - f(a)| < \varepsilon = \frac{f(a)}{2}$$

for all $x \in A$ with $|x - a| < \delta$. That is, the above holds for all $x \in V_\delta(x) \cap A$. However, for any such x, we have

$$f(a) - f(x) \leq |f(x) - f(a)| < \frac{f(a)}{2}$$

whence it follows that

$$f(a) - \frac{f(a)}{2} < f(x).$$

Or, rather, we find that

$$f(x) > \frac{f(a)}{2} > 0$$

for all $x \in V_\delta(x) \cap A$. □

Corollary 11.2. *Let $A \subseteq \mathbb{R}$ and assume that $f : A \to \mathbb{R}$ is continuous at a point $a \in A$. If $f(a) < 0$, there exists a δ-neighbourhood $V_\delta(a)$ such that $f(x) < 0$ for all $x \in A \cap V_\delta(a)$.*

Proof. Apply the previous result to the function $g := -f$. □

The argument used in the proof of this proposition can be generalized to obtain a result that applies to $|f|$.

Proposition 11.3. *Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$ be a function and assume that f is continuous at a point $a \in A$. If $f(a) \neq 0$, then one can find $\delta > 0$ such that*

$$|f(x)| > 0$$

for all $x \in V_\delta(a) \cap A$.

Proof. If f is continuous at a, then so is $|f|$. Also, if $f(a) \neq 0$, then $|f(a)| > 0$. Therefore, it suffices to apply Proposition 11.1 to $|f|$. □

11.3.1 Thomae's Function

Throughout this subsection, we denote by I the open interval $(0, 1)$. Thomae's function is defined to be the mapping given by the rule

$$f : I \to \mathbb{R}, \quad x \mapsto \begin{cases} 0 & \text{if } x \notin \mathbb{Q}, \\ \frac{1}{q} & \text{if } x \in \mathbb{Q} \text{ and } x = \frac{p}{q} \text{ with } p, q \in \mathbb{N} \text{ and } \gcd(p, q) = 1. \end{cases} \tag{11.2}$$

Note that $f(x) \neq 0$ for all $x \in \mathbb{Q} \cap I$.

We assert that f is continuous at all irrational numbers and discontinuous at every rational number. This would show that the function f defined above has only countably many discontinuities on the interval $(0, 1)$. In fact, f would be discontinuous only on a countably infinite set.

Proposition 11.4. *Thomae's function f defined in (11.2) is continuous everywhere on I except at those point in $I \cap \mathbb{Q}$.*

Proof. We first show that f is discontinuous at every $x \in \mathbb{Q} \cap I$. Let (ξ_n) be a sequence of irrational numbers in $(0, 1)$ that converge to x as $n \to \infty$.[23] By definition, $f(\xi_n) = 0$ for all $n \in \mathbb{N}$. However,

$$\lim f(\xi_n) = 0 \neq f(x)$$

which shows that f is not continuous at x by the sequential characterization of the limit.

To prove that f is continuous at the irrational points of I, we follow the argument used in Assignment 10 (see MyCourses). □

11.3.2 Continuity of the Maximum Operator

Let $f, g : \mathbb{R} \to \mathbb{R}$ be functions and assume that these are continuous at $a \in \mathbb{R}$. Is it true that the function $x \mapsto \max\{f(x), g(x)\}$ is also continuous at a? To help answer this question, let us first justify two elementary identities. Given $x, y \in \mathbb{R}$, we claim that

$$\max\{x, y\} = \frac{x + y + |x - y|}{2}. \tag{11.3}$$

To check this, we distinguish two cases:

1. If $x \geq y$ then

$$\frac{x + y + |x - y|}{2} = \frac{x + y + x - y}{2} = \frac{2x}{2} = x = \max\{x, y\}.$$

[23] Since I is open, there exists $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subseteq I$. Therefore, we have $\left(x - \frac{\varepsilon}{n}, x + \frac{\varepsilon}{n}\right) \subseteq (x - \varepsilon, x + \varepsilon) \subseteq I$ for all $n \geq 1$. Since the irrationals are dense in $\mathbb{R}$, there exists for each $n \in \mathbb{N}$ and irrational number $\xi_n \in \left(x - \frac{\varepsilon}{n}, x + \frac{\varepsilon}{n}\right)$. Since $|x - \xi_n| < \frac{\varepsilon}{n} \to 0$ as $n \to \infty$, the Squeeze Theorem ensures that $\xi_n \to x$ as $n \to \infty$.

2. If instead $x < y$ then

$$\frac{x+y+|x-y|}{2} = \frac{x+y+y-x}{2} = \frac{2y}{2} = y = \max\{x,y\}.$$

Hence, we see that (11.3) holds. In a similar way, one can check that

$$\min\{x,y\} = \frac{x+y-|x-y|}{2}.$$

So, in particular, for each x

$$\max\{f(x),g(x)\} = \frac{f(x)+g(x)+|f(x)-g(x)|}{2}, \tag{11.4}$$

$$\min\{f(x),g(x)\} = \frac{f(x)+g(x)-|f(x)-g(x)|}{2}. \tag{11.5}$$

This shows that $\max\{f(x),g(x)\}$ and $\min\{f(x),g(x)\}$ can both be expressed as the linear combination of continuous functions, and hence will be continuous at a as well.

11.4 Open Mappings

A function $f:\mathbb{R}\to\mathbb{R}$ is said to be an *open map* if $f(O)$ is open in $\mathbb{R}$ whenever O is open in $\mathbb{R}$. That is, if f sends open sets to open sets. It is **not** true that all continuous maps are open, as we shall see below.

Example 11.5. Consider the function $f:\mathbb{R}\to\mathbb{R}$ given by $x\mapsto x^2$. Clearly, f is continuous on all of $\mathbb{R}$. However, it is not an open map. To see this, note that f maps the open set $(-1,1)$ onto the set $[0,1)$, which is not open!

To make this phenomenon more concrete, we fix an open interval $(a,b)\subseteq\mathbb{R}$. If $a\ge 0$, then $f((a,b)) = (a^2,b^2)$ which is again open in $\mathbb{R}$. Similarly, if $a<b\le 0$ then

$$f((a,b)) = (b^2,a^2)$$

which is an open subset of $\mathbb{R}$. So far, f has taken open sets to open sets. The problem we encountered earlier arises when $a<0<b$, in which case

$$f((a,b)) = \left[0, \max\{|a|,|b|\}^2\right)$$

which is not open in $\mathbb{R}$.

Although every continuous function need not be an open map, the following characterization of continuous functions does hold true.

Theorem 11.5. *A function $f : \mathbb{R} \to \mathbb{R}$ is continuous if and only if $f^{-1}(V)$ is open for every open set $V \subseteq \mathbb{R}$.*

Proof. First suppose that f is continuous and fix an open set $V \subseteq \mathbb{R}$. We claim that $f^{-1}(V)$ is open. If $f^{-1}(V)$ is empty, then we are done. Otherwise, let us take $x \in f^{-1}(V)$ and consider $y = f(x) \in V$. Since V is open, we can find $\varepsilon > 0$ such that $V_\varepsilon(y) \subseteq V$. For this $\varepsilon > 0$, there is $\delta > 0$ with the property that

$$|f(z) - f(x)| = |f(z) - y| < \varepsilon$$

for all $z \in \mathbb{R}$ with $|z - x| < \delta$. That is, $f(z) \in V_\varepsilon(y) \subseteq V$ for all $z \in V_\delta(x)$. It follows that $V_\delta(x) \subseteq f^{-1}(V)$ whence $f^{-1}(V)$ is open.

Conversely, fix a point $c \in \mathbb{R}$ and let $\varepsilon > 0$. Define $y := f(c)$ and consider the ε-neighbourhood $V_\varepsilon(y)$. By hypothesis, we know that $f^{-1}(V_\varepsilon(y))$ is open in $\mathbb{R}$. Since $c \in f^{-1}(V_\varepsilon(y))$, we can therefore find $\delta > 0$ such that $V_\delta(c) \subseteq f^{-1}(V_\varepsilon(y))$. In particular, this means that

$$f(x) \in V_\varepsilon(y), \quad \text{whenever } |x - c| < \delta.$$

Of course, this is the same as saying that $|f(x) - f(c)| < \varepsilon$ for all $x \in \mathbb{R}$ with $|x - c| < \delta$. Hence, f is continuous at c. □

In a similar vein, we say that f is a *closed map* if it sends closed sets to closed sets. That is, if $\Sigma \subseteq \mathbb{R}$ is closed, then so is $f(\Sigma)$.

Example 11.6. Consider once more the function $f(x) = x^2$. If $[a, b] \subset \mathbb{R}$ is a closed interval, then $f([a, b])$ is also closed in $\mathbb{R}$. To see this, we consider three cases as above. First if $a \geq 0$, then $f([a, b]) = [a^2, b^2]$ which is again a closed interval in $\mathbb{R}$. Similarly, if $a < b \leq 0$ then

$$f([a, b]) = [b^2, a^2]$$

Finally, if $a < 0 < b$ then

$$f([a, b]) = \left[0, \max\{|a|, |b|\}^2\right]$$

which is also closed in $\mathbb{R}$.

11.5 Compactness and Open Covers

Let us consider a set $K \subseteq \mathbb{R}$. An *open cover* of K is a family $\mathcal{U} := \{U_\alpha\}_{\alpha \in I}$ of **open sets** such that $K \subseteq \bigcup_{\alpha \in I} U_\alpha$. That is, an open cover of K is simply a family of open sets whose union *contains* K. Using the definition of an open cover, we can define the concept of *compactness*. This very general notion captures what it means for a "space" to behave like a finite set (albeit in a loose sense). Although its definition is purely topological, continuous functions on compact sets will enjoy very nice analytic properties.

Definition 11.3. A subset $K \subseteq \mathbb{R}$ is said to be *compact* whenever **every** open cover of K has a finite subcover. More precisely, a set K is compact if, for any open cover $\mathcal{U} = \{U_\alpha\}_{\alpha \in I}$ of K, there exist finitely many $U_{\alpha_1}, \ldots, U_{\alpha_n} \in \mathcal{U}$ such that $K \subseteq \bigcup_{j=1}^{n} U_{\alpha_j}$.

Next, we provide some examples (mostly without proof[24]) of sets that are compact.

- The empty set $\varnothing$ is compact since it is a subset of every set. Hence, given any open cover $\mathcal{U}$ of $\varnothing$, we have $\varnothing \subseteq U$ for each $U \in \mathcal{U}$. That is, every open set $U \in \mathcal{U}$ is trivially a finite subcover of $\varnothing$.

- The closed interval $[0, 1]$ is compact. In fact, every interval $[a, b]$ with $a < b$ is compact.

- If $S \subset \mathbb{R}$ is bounded, then ∂S is compact.

- Every finite set is compact. Indeed, let $X = \{x_1, \ldots, x_n\}$ be any finite set and let $\{U_\alpha\}_{\alpha \in I}$ be an open cover of X. That is, each U_α is open and $X \subseteq \bigcup_{\alpha \in I} U_\alpha$. Now, for each $j = 1, \ldots, n$, there exists $\alpha_j \in I$ such that $x_j \in U_{\alpha_j}$. Hence, $\{x_1, \ldots, x_n\} = X \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$. We infer that X is compact.

- The Cantor set $\mathfrak{C}$ defined in §6 is compact.

To help solidify the concept of compactness, let us give two explicit examples of an open cover which has no finite subcover (to see why some sets might not be compact).

[24] The compactness of these sets will follow at once when we prove the Heine-Borel theorem.

Example 11.7. Consider the set $A = [0, \infty)$, which is unbounded. We claim that A has an open cover that does **not** have a finite subcover. In this particular case, such an open cover is easy to construct. Given $n \geq -1$, we define $U_n := (n, n+2)$, which is clearly open in $\mathbb{R}$. Moreover, $(-1, \infty) = \bigcup_{n \geq -1} U_n$. Thus, the family

$$\mathcal{U} := \{U_n : n \geq -1,\ n \in \mathbb{Z}\}$$

is an open cover of $[0, \infty)$ that has no finite subcover.[25]

One last trickier example is in order.

Example 11.8. We claim that $[0, 1)$ is not compact. Let us try to find an open cover of $[0, 1)$ having no finite subcover. Since $[0, 1)$ is bounded, the "boundedness trick" we used in the last example will fail here. Instead, let us try to cook up a slightly more contrived open cover. First, note that given $x \in [0, 1) \subseteq (-1, 1)$, we can find $\varepsilon_x > 0$ such that $V_{\varepsilon_x}(x) \subseteq (-1, 1)$. In fact, by replacing ε_x with

$$\frac{\varepsilon_x}{2} < \varepsilon_x,$$

we can make sure that there always exists $y \notin V_{\varepsilon_x}(x)$ with $0 < y < 1$. Then, the family

$$\{V_{\varepsilon_x}(x) : x \in [0, 1)\}$$

certainly forms an open cover of $[0, 1)$. However, given finitely many $x_1, \ldots, x_n$ in $[0, 1)$, the union $\bigcup_{j=1}^{n} V_{\varepsilon_{x_j}}(x_j)$ cannot cover all of $[0, 1)$. Indeed, this is because every $V_{\varepsilon_{x_j}}(x_j)$ is of the form (a_j, b_j), with $b_j < 1$. That is,

$$\bigcup_{j=1}^{n} V_{\varepsilon_{x_j}}(x_j) \subseteq \left(\min_{1 \leq j \leq n} a_j, \max_{1 \leq j \leq n} b_j \right)$$

where $\max_{1 \leq j \leq n} b_j < 1$.

The next result is a fundamental property of compact sets that continues to hold in more general settings. For now, however, it will be more than enough to have the result for $\mathbb{R}$.

Lemma 11.6. *If $K \subseteq \mathbb{R}$ is compact, then it is also bounded.*

[25]The union of finitely many members of $\mathcal{U}$ always gives a bounded set, and thus no finite subcollection of $\mathcal{U}$ can cover $[0, \infty)$.

Proof. For each $n \in \mathbb{N}$ define $U_n := (-n, n)$. Then, U_n is open and

$$K \subseteq \mathbb{R} = \bigcup_{n \in \mathbb{N}} U_n.$$

Thus, $\{U_n\}_{n \in \mathbb{N}}$ is an open cover of K. Since K is compact, we can find finitely many $U_{n_1}, \ldots, U_{n_k}$ such that $K \subseteq U_{n_1} \cup \cdots \cup U_{n_k}$. Let $N := \max(n_1, \ldots, n_k)$ and note that

$$K \subseteq U_{n_1} \cup \cdots \cup U_{n_k} = U_N = (-N, N).$$

Since K is a subset of a bounded set, the proof is complete. □

12 Twelfth/Last Tutorial

Recall the (topological/generalized) definition of a compact set: a set $K \subseteq \mathbb{R}$ is said to be *compact* if every open cover of K has a finite subcover. As seen in the lectures, the Heine-Borel theorem provides a complete characterization of the compact subsets of $\mathbb{R}$.

Theorem 12.1 (Heine-Borel)**.** *A subset $K \subseteq \mathbb{R}$ is compact (according to Definition 11.3) if and only if it is both closed and bounded.*

At the end of the previous tutorial, we showed that any compact set is automatically bounded (see Lemma 11.6). In fact, our proof of this fact was elementary in the sense that it did not rely on anything other than the very definition of compactness. Arguing in a similar way, it is not too hard to show using only elementary arguments that every compact set is closed. Although we already know this from class, it is still a good exercise to understand a *different* proof.

Theorem 12.2. *Every compact subset of $\mathbb{R}$ is both closed and bounded.*

Proof. By Lemma 11.6, we already know that every compact set is bounded. Now let $K \subseteq \mathbb{R}$ be compact; we must show that K is closed. To establish this, we will prove that K^c is open. Fix a point $x \in K^c$ and define

$$U_n := \left\{ y \in \mathbb{R} : |x - y| > \frac{1}{n} \right\} = \mathbb{R} \setminus \left[x - \frac{1}{n}, x + \frac{1}{n} \right]$$

for each $n \in \mathbb{N}$. Clearly, U_n is open and $U_n \subseteq U_{n+1}$ for all $n \geq 1$. Finally, observe that

$$\bigcup_{n=1}^{\infty} U_n = \mathbb{R} \setminus \{x\}.$$

Since $x \notin K$, it follows that $K \subseteq \mathbb{R} \setminus \{x\} = \bigcup_{n\in\mathbb{N}} U_n$. That is, $\{U_n\}_{n\in\mathbb{N}}$ is an open cover of K. By definition of compactness, there exist finitely many $U_{n_1}, \dots, U_{n_k}$ such that

$$K \subseteq U_{n_1} \cup \cdots \cup U_{n_k}.$$

As in the proof of Lemma 11.6, we define $N := \max(n_1, \dots, n_k)$ and observe that

$$K \subseteq U_{n_1} \cup \cdots \cup U_{n_k} \subseteq U_N.$$

Thus, $U_N^{\mathsf{c}} \subseteq K^{\mathsf{c}}$. However, $U_N^{\mathsf{c}} = \left[x - \frac{1}{N}, x + \frac{1}{N}\right]$ whence

$$\left(x - \frac{1}{N}, x + \frac{1}{N}\right) \subseteq U_N^{\mathsf{c}} \subseteq K^{\mathsf{c}}.$$

Put otherwise, we have $V_{1/N}(x) \subseteq K^{\mathsf{c}}$. As $x \in K^{\mathsf{c}}$ was arbitrary, we infer that K^{c} is open. This completes the proof. □

12.1 Uniform Continuity

Given a subset $A \subseteq \mathbb{R}$ and a function $f : A \to \mathbb{R}$, we say that f is *uniformly continuous* on A if for each $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in A$ satisfy $|x - y| < \delta$. In symbolic terms:

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x, y \in A)(|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon).$$

At a first glance this may seem to be the same as saying that f is continuous on A, but this is not the case. By way of clarification, let us compare our definition of *uniform continuity* to our $\varepsilon - \delta$ definition of *continuity* in (11.1).

- Continuity is a *local property* and uniform continuity is a *global property.* We know what it means for the function f to be continuous at an individual point $a \in A$, but we have only defined uniform continuity for the entire set A. Namely, the continuity of f at a single point a depends only on the behaviour of f in a neighbourhood of a. On the other hand, uniform continuity depends on how f behaves on the entire set A.

- We say that f is continuous at a point $y \in A$ if

  $$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in A)(|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon).$$

  Thus, we say that f is continuous on the *entire set* A if

  $$(\forall y \in A)(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in A)(|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon).$$

In this expression, the variable y appears before the "$\exists\delta$", which means that δ may depend on the point y. However, this is not the case in the definition of uniform continuity. For f to be uniformly continuous on A, the $\delta > 0$ cannot depend on x or y.

Let us also recall the following fundamental result from class.

Theorem 12.3 (Uniform Continuity Theorem)**.** *Let $K \subset \mathbb{R}$ be compact and let $f : K \to \mathbb{R}$ be continuous. Then, f is uniformly continuous on A.*

If K is not compact, this implication may fail. However, there are still many nice continuous functions that turn out to be uniformly continuous on unbounded intervals. One example is the function

$$f : (0, \infty) \to \mathbb{R}, \quad x \mapsto \sqrt{x}.$$

Example 12.1. We claim that the function f described above is uniformly continuous on $[0, \infty)$. First, let us establish the following useful inequality:

$$\left|\sqrt{x} - \sqrt{y}\right| \leq \sqrt{|x - y|}, \quad \forall x, y \geq 0. \tag{12.1}$$

By symmetry, it is enough to show that the above holds whenever $0 \leq y \leq x$. For such x, y, the inequality in (12.1) reduces to proving that

$$\begin{aligned}
\sqrt{x} - \sqrt{y} \leq \sqrt{x - y} &\iff \left(\sqrt{x} - \sqrt{y}\right)^2 \leq (x - y) \\
&\iff x - 2\sqrt{xy} + y \leq x - y \\
&\iff -2\sqrt{xy} + y \leq -y \\
&\iff 2y \leq 2\sqrt{xy} \\
&\iff y \leq \sqrt{xy}.
\end{aligned}$$

This last inequality holds true because $y \leq x$ implies

$$y = \sqrt{y^2} = \sqrt{y \cdot y} \leq \sqrt{x \cdot y}.$$

Thus, we have proven (12.1). It now becomes trivial to show that $\sqrt{x}$ is uniformly continuous on $[0, \infty)$. To see this, let $\varepsilon > 0$ be given and take $\delta := \varepsilon^2 > 0$. If $x, y \in [0, \infty)$ are such that $|x - y| < \delta$, then (12.1) tells us that

$$|f(x) - f(y)| = \left|\sqrt{x} - \sqrt{y}\right| \leq \sqrt{|x - y|} < \sqrt{\delta} = \varepsilon.$$

Note our δ only depends on ε and **not** on x or y.

12.1.1 A Return to Cauchy Sequences

Let $f : A \to \mathbb{R}$ be a uniformly continuous function and let (x_n) be a sequence in A. If (x_n) is Cauchy, then the induced sequence $(f(x_n))$ turns out to be *also* Cauchy. In this sense, we say that f takes Cauchy sequences to Cauchy sequences. This is **not** a property satisfied by functions that are merely continuous. Indeed, consider the function

$$f : (0,1) \to \mathbb{R}, \quad x \mapsto \frac{1}{x}.$$

Clearly, f is continuous on its entire domain. Now, the sequence (x_n) in $(0,1)$ defined by $x_n := \frac{1}{n}$ converges to 0 and must therefore be Cauchy. However, we have $f(x_n) = n$ for each $n \in \mathbb{N}$. Hence, the sequence $(f(x_n))$ is not Cauchy.

Proposition 12.4. *Let $A \subseteq \mathbb{R}$ be non-empty and let $f : A \to \mathbb{R}$ be uniformly continuous. If (x_n) is a Cauchy sequence in A, then $(f(x_n))$ is also Cauchy.*

Proof. Let $\varepsilon > 0$ be given. Since f is uniformly continuous, there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $x, y \in A$ satisfying $|x - y| < \delta$. Now, using that (x_n) is Cauchy in A, there exists $N \in \mathbb{N}$ such that $|x_n - x_m| < \delta$ whenever $n, m \geq N$. Consequently,

$$|f(x_n) - f(x_m)| < \varepsilon$$

for all $n, m \geq N$. As $\varepsilon > 0$ was arbitrary, it follows that $(f(x_n))$ is Cauchy. □

Next we ask whether the converse is true. Namely, if a function maps Cauchy sequences to Cauchy sequences, must it be uniformly continuous? To help answer this question, we consider a continuous function on an interval that is not uniformly continuous.

Lemma 12.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) := x^2$. The function f is not uniformly continuous on $\mathbb{R}$, but is uniformly continuous on any compact interval $[-a, a]$.*

Proof. The only non-trivial claim is that f is not uniformly continuous on all of $\mathbb{R}$. To see this, we argue by contradiction. If f is uniformly continuous on $\mathbb{R}$, then for $\varepsilon := 1$ we can find $\delta > 0$ such that

$$\left|x^2 - y^2\right| < 1$$

whenever $|x - y| < \delta$. Given $x \in \mathbb{R}$ we consider the point

$$y := x + \frac{\delta}{2}.$$

It is clear that $|x - y| < \delta$, and we must therefore have

$$\left| x^2 - \left(x + \frac{\delta}{2} \right)^2 \right| < 1.$$

However, the expression above reduces to the following:

$$\left| x\delta + \frac{\delta^2}{4} \right| < 1.$$

Since $x \in \mathbb{R}$ can be made very large, this is impossible. $\square$

Having exhibited a continuous function $\mathbb{R} \to \mathbb{R}$ that is not uniformly continuous, we can answer our original question.

Example 12.2. We claim that the function $f : \mathbb{R} \to \mathbb{R}$ given by $x \mapsto x^2$ takes Cauchy sequences in $\mathbb{R}$ to Cauchy sequences. To see this, let (x_n) be a Cauchy sequence. Since Cauchy sequences are bounded, the sequence (x_n) is contained within some compact interval $[-a, a]$. That is, (x_n) is a Cauchy sequence in $[-a, a]$. By continuity on all of $\mathbb{R}$, the function f is uniformly continuous on $[-a, a]$. Since uniformly continuous functions take Cauchy sequences to Cauchy sequences, we see that $(f(x_n))$ is Cauchy in $\mathbb{R}$.

Thus, we have found a continuous map $\mathbb{R} \to \mathbb{R}$ that takes Cauchy sequences to Cauchy sequences, even though it is *not* uniformly continuous on $\mathbb{R}$. Moving away from examples, let us now prove some more powerful/general statements.

Proposition 12.6. *Let $A \subseteq \mathbb{R}$ be non-empty and let $f, g : A \to \mathbb{R}$ be uniformly continuous on A.*

(1) $f + g$ is uniformly continuous as a function $A \to \mathbb{R}$.

(2) The product fg may not be uniformly continuous on A.

Proof. We begin with the first claim. Let $\varepsilon > 0$ be given and let $\delta_1 > 0$ be such that

$$|f(x) - f(y)| < \frac{\varepsilon}{2}$$

whenever $x, y \in A$ satisfy $|x - y| < \delta_1$. Similarly, because g is uniformly continuous on A, we can find $\delta_2 > 0$ with the property that $|x - y| < \delta_2$ (with $x, y \in A$) implies

$$|g(x) - g(y)| < \frac{\varepsilon}{2}.$$

Taking $\delta := \min\{\delta_1, \delta_2\} > 0$, we see that for any $x, y \in A$ with $|x - y| < \delta$ we have:

$$\begin{aligned}
|(f+g)(x) - (f+g)(y)| &= |f(x) + g(x) - f(y) - g(y)| \\
&\leq |f(x) - f(y)| + |f(y) - g(y)| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon.
\end{aligned}$$

This means that $f+g$ is uniformly continuous on A. For the second part, consider the functions $f(x) = g(x) = x$ on all of $\mathbb{R}$. Clearly, both f and g are uniformly continuous but their product $fg = x^2$ is *not* (by our previous lemma). □

12.2 Uniformly Continuous Functions Are Bounded

Given a non-compact *bounded* subset A of $\mathbb{R}$, a continuous function $f : A \to \mathbb{R}$ need not be bounded. Indeed, this is the case for the continuous map

$$f : (0, 1] \to \mathbb{R}, \quad x \mapsto \frac{1}{x}.$$

Thus, continuous functions need not (in general) be bounded. However, it is true that uniformly continuous functions on bounded domains are bounded. We prove this below.

Theorem 12.7. *Let $A \subset \mathbb{R}$ be bounded and suppose that $f : A \to \mathbb{R}$ is uniformly continuous on A. Then, f is bounded on A.*

Proof. We argue by contradiction. If f were unbounded on A, we could find a sequence (x_n) in A such that $|f(x_n)| > n$ for each $n \in \mathbb{N}$. Since A is bounded, the sequence (x_n) is bounded. Thus, we can find a subsequence (x_{n_k}) that converges in $\mathbb{R}$. Equivalently, (x_{n_k}) is a Cauchy sequence in A such that

$$\left|f(x_{n_k})\right| > n_k \geq k$$

for all $k \in \mathbb{N}$. However, since f is uniformly continuous on A, we see that $(f(x_{n_k}))$ is Cauchy in $\mathbb{R}$. In particular, the sequence $(f(x_{n_k}))_k$ is bounded. This contradicts the fact that

$$\left|f(x_{n_k})\right| > n_k \geq k.$$

Hence, we conclude that f is bounded on A. □

12.3 Lipschitz Continuity

Lipschitz continuity is an additional continuity condition one can impose on a function that is stronger than uniform continuity.

Definition 12.1. Let A be a non-empty subset of $\mathbb{R}$. A function $f : A \to \mathbb{R}$ is said to be *Lipschitz continuous* on A if there exists $L > 0$ such that

$$|f(x) - f(y)| \leq L |x - y|$$

for all $x, y \in A$. The set of all Lipschitz continuous functions $A \to \mathbb{R}$ is often denoted by $\mathrm{Lip}(A; \mathbb{R})$ or $\mathrm{Lip}(A)$.

It is automatic that Lipschitz continuous functions are uniformly continuous on their domains. Indeed, given $f \in \mathrm{Lip}(A)$ and $\varepsilon > 0$, we define $\delta = \frac{\varepsilon}{L}$ and notice that if $x, y \in A$ satisfy $|x - y| < \delta$ then

$$|f(x) - f(y)| \leq L |x - y| < L\delta = \varepsilon.$$

One can extend the concept of a Lipschitz function to that of a Hölder continuous function. We make this precise below.

Definition 12.2. Let $A \subseteq \mathbb{R}$ be non-empty and let $f : A \to \mathbb{R}$ be a function. We say that f is Hölder continuous on A with exponent $\alpha > 0$ if there exists $M > 0$ such that

$$|f(x) - f(y)| \leq M |x - y|^{\alpha}$$

for all $x, y \in A$.

Note that $f \in \mathrm{Lip}(A; \mathbb{R})$ if and only if f is Hölder continuous on A with exponent $\alpha = 1$.

Exercise 12.1. Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$ be Hölder continuous on A with exponent $\alpha > 0$. Prove that f is uniformly continuous on A.

Let us now give some examples of Lipschitz/Hölder continuous functions.

(i) By the reverse triangle inequality, the function $f(x) := |x|$ is Lipschitz on $\mathbb{R}$. Indeed, note that

$$||x| - |y|| \leq |x - y|$$

for all $x, y \in \mathbb{R}$.

(ii) In (5.1) we showed that

$$\left|\sqrt{x} - \sqrt{y}\right| \leq \sqrt{|x-y|} = |x-y|^{1/2}, \quad \forall x, y \in [0, \infty).$$

Thus, we were really proving that $f(x) := \sqrt{x}$ is Hölder continuous on $[0, \infty)$ with exponent $\alpha = \frac{1}{2}$.

(iii) The function $f(x) = \sqrt{x}$ is *not* Lipschitz on $[0, 1]$. To see this, we argue by contradiction. If it were Lipschitz on $[0, 1]$, we could find $L > 0$ such that

$$\left|\sqrt{x} - \sqrt{y}\right| \leq L\,|x-y|, \quad \forall x, y \in [0, 1].$$

Thus,

$$\sup_{\substack{x,y\in[0,1]\\ x\neq y}} \frac{\left|\sqrt{x} - \sqrt{y}\right|}{|x-y|} \leq L < \infty.$$

In particular (taking $y = 0$),

$$\sup_{\substack{x\in[0,1]\\ x\neq 0}} \frac{\sqrt{x}}{x} = \sup_{\substack{x\in[0,1]\\ x\neq 0}} \frac{1}{\sqrt{x}} < \infty.$$

Since this is absurd, we have a contradiction.

12.4 The Lebesgue Number Lemma (Optional)

If time permits, let us conclude by exploring a more advanced and less intuitive property that compact sets enjoy. Despite being somewhat strange at first glance, the usefulness of this result should not be underestimated.

Theorem 12.8 (Lebesgue Number Lemma)**.** *Let $K \subseteq \mathbb{R}$ be compact and let*

$$\mathcal{U} = \{U_\alpha\}_{\alpha\in I}$$

be an open cover of K. There exists $\delta > 0$ such that, for any $x \in K$, there is some $U_\alpha \in \mathcal{U}$ with $V_\delta(x) \subseteq U_\alpha$. In symbolic terms:

$$(\exists\delta > 0)(\forall x \in K)(\exists\alpha \in I \text{ with } V_\delta(x) \subseteq U_\alpha).$$

In other words, every δ-neighbourhood centered at a point $x \in K$ is entirely contained within a single member of $\mathcal{U}$. This δ is called a ***Lebesgue number****.*

Proof. For every $x \in K$, we can find $U_{\alpha_x} \in \mathcal{U}$ containing the point x. Since U_{α_x} is open, there is some corresponding ε_x such that $V_{\varepsilon_x}(x) \subseteq U_{\alpha_x}$. Next, we consider the family of open sets

$$\mathcal{A} := \{V_{\varepsilon_x/2}(x) : x \in K\}$$

which certainly forms an open covering of K. Since K is compact, it has a finite subcover. Namely, we can find finitely many points $x_1, \dots, x_n \in K$ such that

$$K \subseteq \bigcup_{j=1}^{n} V_{\varepsilon_j/2}(x_j), \qquad \varepsilon_j = \varepsilon_{x_j}.$$

Now, define

$$\delta := \frac{1}{2} \min_{1 \le j \le n} \varepsilon_j > 0.$$

Let $x \in K$ be given and consider $B(x, \delta)$. Since

$$K \subseteq \bigcup_{j=1}^{n} V_{\varepsilon_j/2}(x_j),$$

we can find $j = 1, \dots, n$ with $x \in V_{\varepsilon_j/2}(x_j)$. Then, by the triangle inequality, it follows that

$$V_\delta(x) \subseteq V_{\varepsilon_j}(x_j) \subseteq U_{\alpha_j}.$$

□