Notes on real analysis

Rodolfo J Amancio

February 21, 2023

Contents

Chapter 1

Preface

First, let me be clear. I am not a mathematician. These notes are not intended as a manual, however I like to teach, explain science, and I am firm believer that the best way to learn something is to teach each. Richard Feynman famously said the best way to learn is to follow these steps:

1. Study: arguably the easiest part, whether you like to take notes on paper, tablet or your computer. Whether you like to sit on a table or on the couch. However, as many people who have come this far know, reading, taking notes or making exercises only take you so far.

2. Teach: this is where the fun begins, as you try to explain something you know (or think you know) to someone, you start being more aware of your limitations, of the gaps in the proofs you cannot explain, the unexpected questions that may appear lead you astray. Even without an audience, this is a nice thing to do as it forces you to be clear and think about how to explain something in a clear yet rigorous way.

3. Fill the gaps: now it is time to come back to studying, reading more, exploring new books, papers or what else. Once you have discovered your limitations on the previous step, you are once again in the position to learn and study, but now you know where to look.

4. Simplify: one of the greatest sins we commit is to get stuck with fancy proofs, delude ourselves in the beauty of math. Make it so that people will understand and enjoy what they are reading or listening.

In this document, I have written my learnings from studying Real Analysis. I hope to learn more while writing it.

Chapter 2

The real numbers

During high school math we are often given a simplified definition of the real numbers, one it may take a while to fully grasp how awkward it is: "The real numbers is the set which contains the rational numbers and the irrational numbers". Taking alone it may seem a reasonable statement. In fact, it is true. However, if we start with the natural numbers there is a very concise and clear way of writing it:

$$\mathbb{N} = \{1, 2, 3, ...\} \tag{2.1}$$

Taking one step further, the integers follow quite naturally:

$$\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\} \tag{2.2}$$

And even for the rationals, we can clearly write:

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z} \text{ and } q \neq 0 \right\} \tag{2.3}$$

Now, for the real numbers things are not so clear. So, we are stuck with our initial understanding of the big set which includes the rational and irrational numbers. And as the saying goes "The devil in the details". Calculus, with all of its beautiful and powerful tools: limits, derivatives and integrals, relies fundamentally on real numbers. In order to rigorously work with its results, investigate the extremes and challenging cases, and proving its results depends on a formal definition for the real numbers.

So, let's start by looking more carefully at this similarly weird creature.

2.1 Irrational numbers

Before we proceed, let's take a minute to appreciate why we need irrational numbers. The following result will play an important role to distinguish the "holes" of the rational numbers when compared with the reals. We begin with a theorem.

Theorem 1. There is no such number whose square root is 2

Proof. As stated before, a rational number is one that can be written in the form p/q, with $q \neq 0$. Our approach here is what is called proof by contradiction. We will assume the opposite of what we want to prove, once we arrive at some absurd result we will conclude our initial assumption was wrong. Therefore, assume $\exists p, q \in \mathbb{Z} : (p/q)^2 = 2$, additionally, we take p and q with no common factors, such that the fraction p/q is written in its simplest form.
If this is true, we can rearrange the relation into: $p^2 = 2q^2$. Which implies p^2 is an even number, since the square of any odd number is odd, p must also be even, *i.e.* $p = 2r$.
Now, replacing p on the previous equation yields $4r^2 = 2q^2 \Rightarrow q^2 = 2r^2$, which implies q^2 and so is q.
This directly contradicts our initial assumption, since p and q are both even from the result above. Hence, our initial assumption must be wrong, and we conclude $\nexists p, q \in \mathbb{Z} : (p/q)^2 = 2$. □

In order to deal with irrational numbers, the set of real numbers is the natural extension necessary. Before we deal with it in a rigorous way, will start with the necessary tools to help us on this journey.

2.2 Preliminaries

This section aims to define some basic definitions and results that will help us to deal with real numbers, and the other topics of interest.

2.2.1 Set theory

Definition 1 (Set)**.** A set is a collection of objects, called elements or members. An empty set is a set with no elements, denoted $\varnothing$.

Usually, we write a set A as $A = \{a_1, a_2, ...\}$ where $a_1, a_2, ...$ are the elements of the set. Some important notations are:

- $a \in A$: meaning a is an element of A
- $a \notin A$: meaning a is not an element of A
- $\forall$: meaning 'for all'. For example, in mathematical notation the expression 'for all a which is an element of A' would be $\forall a \in B$
- $\exists$: meaning 'there exists'. The opposite would be $\nexists$
- $\Rightarrow$: implies
- $\Leftrightarrow$: if and only if

On a side note, the terms 'implies' and 'if and only if' have fundamental differences which will lead to different approaches in demonstrations and results. For instance, let's say proposition P_1 implies proposition P_2. In mathematical notation $P_1 \Rightarrow P_2$. This means that if P_1 is true, so is P_2, but it does not say anything about the opposite direction. That is, if P_2 is true, not necessarily P_1

is true. On the other hand when the relation is P_1 is true if, and only if, P_2 is true. Or, $P_1 \Leftrightarrow P_2$ then the relation works both ways: if P_1 is true, so is P_2 and if P_2 is true, so is P_1. During proofs, when we have a $\Leftrightarrow$ relation, the result must be proven in both directions.

Definition 2 (Subset)**.** A is a subset of B if, every element of A is also an element of B. Notation: $A \subseteq B$. Equivalently, if B is a superset of A, it is denoted $B \supseteq A$.

Informally, we understand that two sets are equal if every element of one set is also an element of the other, and vice-versa. On mathematical notation:

Definition 3 (Equal sets)**.** Two sets, A and B, are equal if $A \subseteq B$ and $B \subseteq A$. Hence, $A = B$.

Definition 4 (Proper subset)**.** A set A is a proper subset of B if $A \subseteq B$ and $A \neq B$. Notation: $A \subsetneq B$.

Now, tow (or more) sets can be combined by operations. We define:

- Union: $A \cup B = \{x : x \in A \text{ or } x \in B\}$
- Intersection: $A \cap B = \{x : x \in A \text{ and } x \in B\}$
- Difference: $A \setminus B = \{x : x \in A \text{ and } x \notin B\}$
- Complement: $A^C = \{x \notin A\}$

Definition 5 (Disjoint sets)**.** Two sets, A and B, are disjoint if $A \cap B = \varnothing$.

Some important from set theory are the so-called De Morgan's laws:

Theorem 2 (De Morgan's Laws)**.** If A, B and C are sets, then:

- $(B \cup C)^C = B^C \cap C^C$
- $(B \cap C)^C = B^C \cup C^C$
- $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
- $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

Following the proof of the first result is shown, the other results can be derived similarly.

Proof. Two sets X and Y are equal if $X \subseteq Y$ and $Y \subseteq X$. Our goal is to show that $(B \cup C)^C \subseteq B^C \cap C^C$ and $(B \cup C)^C \supseteq B^C \cap C^C$.
Let $x \in (B \cup C)^C$. It follows that $x \notin (B \cup C)$ then $x \notin B$ and $x \notin C$. So, $x \in B^C \cap C^C$ and we have $(B \cup C)^C \subseteq B^C \cap C^C$.
From the opposite direction, let $x \in B^C \cap C^C$. Then $x \in B^C$ and $x \in C^C$ which means $x \notin B$ and $x \notin C$. So $x \notin (B \cup C) \Rightarrow x \in (B \cup C)^C$. So $B^C \cap C^C \subseteq (B \cup C)^C$. Since

$(B \cup C)^C \subseteq B^C \cap C^C$ and $B^C \cap C^C \subseteq (B \cup C)^C$, we have $(B \cup C)^C \subseteq B^C \cap C^C$. $\square$

Fields

Definition 6 (Field). A set F is a field if satisfies the following properties:

- For addition
  1. If $x, y \in F \Rightarrow x + y \in F$
  2. Commutativity: $\forall x, y \in F : x + y = y + x$
  3. Associativity: $\forall x, y, z \in F : (x + y) + z = x + (y + z)$
  4. Additive identity: $\exists 0 \in F : 0 + x = x, \forall x \in F$
  5. Additive inverse: $\exists - x \in F : x + (-x) = 0, \forall x \in F$
- For multiplication
  1. If $x, y \in F \Rightarrow x \cdot y \in F$
  2. Commutativity: $\forall x, y \in F : x \cdot y = y \cdot x$
  3. Associativity: $\forall x, y, z \in F : (x \cdot y)z = x(y \cdot z)$
  4. Multiplicative identity: $\exists 1 \in F : x \cdot 1 = x, \forall x \in F$
  5. Multiplicative inverse: $\exists x^{-1} \in F : x \cdot x^{-1} = 1, \forall x \in F$

Theorem 3. If F is a field, $\forall x \in F : x \cdot 0 = 0$.

Proof. If $x \in F$ then $0x \in F$ so $0 = 0x + (-0x) = 0x + 0x + (-0x) = 0x$. $\square$

Definition 7 (Ordered field). An ordered field F is a field which satisfies $\forall x, y, z \in F$:

1. If $x < y$ then $x + z < y + z$
2. If $x > 0$ and $y > 0$ then $xy > 0$

Bounds

Definition 8 (Bounds). Let $A \subseteq B$. Then,

1. If $\exists u \in B : u \geq a, \forall a \in A$ then A is bounded above and u is an upper bound for A.
2. If $\exists l \in B : l \leq a, \forall a \in A$ then A is bounded below and l is a lower bound for A.

Example. Consider the set $B = \mathbb{R}$ and $A = [0, 1]$. Then, $2, 2.5, \pi$ are all upper bounds for A. Similarly, $-1, 0, -\pi$ are all lower bounds for A.

Definition 9 (Supremum). Let $A \subseteq B$ with A bounded above. Then s is the least upper bound (or supremum) if:

1. s is an upper bound for A, and
2. If u is another upper bound for A then $s \leq u$.

Mathematically, we write $s = \sup A$.

Definition 10 (Infimum). Let $A \subseteq B$, with A bounded below. Then, i is the greatest lower bound (or infimum) of A if:

1. i is a lower bound for A, and
2. If l is another lower bound for A then $i \geq l$.

Mathematically, we write $i = \inf A$.

Example. Consider $B = \mathbb{R}$ and $A = (0, 1) \subseteq A$. Then $1, \pi, 10$ are all upper bounds for A but 1 is the least upper bound (or infimum). On the other hand, $-10, -1, 0$ are all lower bounds for A, but only 0 is the greatest lower bound (or infimum) of A.

The previous example shows an important characteristic of the supremum (or infimum). In this case $0 \notin A$ and $1 \notin B$. We can also define:

Definition 11 (Maximum). Let A be a set bounded above, then M is the maximum of A if $M \in A$ and $M \geq a, \forall a \in A$.

Definition 12 (Minimum). Let A be a set bounded below, then m is the minimum of A if $m \in A$ and $m \leq a, \forall a \in A$.

Remark. Notice that a set may have an infimum and not a minimum, as the previous example, since $0 \notin A$. The same result is valid for the supremum and maximum. On the other hand, if a set A has a maximum, then it necessarily has a supremum. An equivalent result holds for the infimum and minimum.

2.2.2 Function

The formal definition of function is the following:

Definition 13 (Function). Given a set A and a set B, a function is a mapping rule which takes as an argument an element $a \in A$ and associates it with an element of B. We write $f : A \to B$. $f(a)$ is used to express the element of B, $f(a) \in B$, associated with the element $a \in A$. A is called the domain of the function, while B is its codmoain. The image of f is not necessarily equal to B, but refers to $\{b \in B : b = f(a) \text{ for some } a \in A\} \subseteq B$.

It is worth noting how this definition liberates math from the usual 'formula' understanding of a function. In particular, this definition is closer to Dirichlet's definition, and it allows math to deal with more interesting and complex functions, such as:

Example (Dirichlet's function).

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

This broader definition of function will lead to interesting results and test some limits in maths. But more on that latter.

Classification

Definition 14 (Injective function)**.** The function $f : A \to B$ is called 1-1 or injective if $f(a_1) = f(a_2) \Rightarrow a_1 = a_2$. Equivalently, $a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$.

Definition 15 (Surjective function)**.** The function $f : A \to B$ is called onto or surjective if $\forall b \in B, \exists a \in A$ such that $f(a) = b$.

Definition 16 (Bijective function)**.** A function that is both injective and surjective is called bijective.

Composition and inverse

Definition 17 (Composite function)**.** If $f : A \to B$ and $g : B \to C$, then $f \circ g : A \to C$ is defined by $(f \circ g)(x) = g(f(x))$.

Definition 18 (Inverse function)**.** Consider $f : A \to B$ a bijective function. Then the inverse function $f^{-1} : B \to A$ is defined by: if $b \in B$ then $f^{-1}(b) \in A$ is the unique element $f^{-1}(b)$ such that $f(f^{-1}(b)) = b$.

2.2.3 The absolute function

The absolute function plays an important role in the proofs and arguments that are to come. First, it is defined as:

Definition 19 (Absolute function)**.** The absolute function $f(x) : \mathbb{R} \to \mathbb{R}_+$ is defined as:

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0 \end{cases} \tag{2.4}$$

It leads to a very important result, called the triangle inequality:

Theorem 4 (Triangle inequality)**.** $\forall a, b \in \mathbb{R}$, $|a + b| \leq |a| + |b|$

Proof. Let $x, y \in \mathbb{R}$. Then, $x + y \leq |x| + |y|$ and

$$(-x) + (-y) \leq |-x| + |-y| = |x| + |y|$$

Hence, $-(|x| + |y|) \leq x + y \leq |x| + |y|$ and we obtain

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

□

2.2.4 Induction

The natural numbers have a property which leads to very important applications. This can be enunciated as:

Well ordering property of $\mathbb{N}$

If $S \subseteq \mathbb{N}$ and $S \neq \varnothing$. Then, $\exists x \in S$ such that $x \leq y, \forall y \in S$.

An important tool that arises from it is called 'Induction'. We can state it as:

Proof by induction

Let $P(n)$ be a statement depending on $n \in \mathbb{N}$. Assume:

1. Base case: $P(1)$ is true
2. Inductive case: If $P(m)$ is true, so is $P(m + 1)$.

From it, we conclude $P(n)$ is true for all $n \in \mathbb{N}$.

Example. Prove that

$$1 + c + c^2 + ... + c^n = \frac{1 - c^{n+1}}{1 - c}, \forall c \neq 1, \forall n \in \mathbb{N}$$

using induction.

Following the algorithm presented before:

1. Base case.
$$1 + c = \frac{1 - c^2}{1 - c} = \frac{(1 - c)(1 + c)}{1 - c} = 1 + c$$
As expected.

2. Inductive case. Assume
$$1 + c + c^2 + ... + c^m = \frac{1 - c^{m+1}}{1 - c}$$
is true. Now, for $m + 1$:
$$\begin{aligned} 1 + c + c^2 + ... + c^{m+1} &= (1 + c + c^2 + ... + c^m) + c^{m+1} \\ &= \frac{1 - c^{m+1}}{1 - c} + c^{m+1} \\ &= \frac{1 - c^{m+1} + c^{m+1} + c^{m+2}}{1 - c} \\ &= \frac{1 - c^{m+2}}{1 - c} \end{aligned}$$
Hence, the relation still holds.

2.3 Defining $\mathbb{R}$

2.3.1 The incompleteness of $\mathbb{Q}$

Now, let's revisit our initial problem, namely $\sqrt{2} \notin \mathbb{Q}$. First, we start with a theorem:

Theorem 5. The set $E = \{x \in \mathbb{Q} : 0 < x < \sqrt{2}\}$ is bounded above and does not have a supremum in $\mathbb{Q}$.

Proof. First, consider $q \in \mathbb{Q}$ then $q^2 < 2 < 4 \Rightarrow q^2 - 4 < 0 \Rightarrow (q - 2)(q + 2) < 0$. Since $q > 0$ we have $q - 2 < 0 \Rightarrow q < 2$. Hence, 2 is an upper bound for E.
Next, to show that $\nexists \sup E \in \mathbb{Q}$ we begin by assuming $x = \sup E \in Q$.
Assume, for contradiction, $x^2 < 2$. Define
$$h = \min\left\{\frac{1}{2}, \frac{2 - x^2}{2(2x + 1)}\right\} < 1$$
Then, $h > 0$. Now we prove $h + x \in E$. Computing $(x + h)^2 = x^2 + 2xh + h^2 < x^2 + 2xh + h$

since $h < 1$. So

$$\begin{aligned}(x+h)^2 < x^2 + (2x+1)h &= x^2 + (2x+1)\frac{2-x^2}{2(2x+1)} \\ &= x^2 + 2 - x^2 \\ &= 2\end{aligned}$$

Therefore $(x+h)^2 < 2$ which implies $x + h \in E$ and $x + h > x$ so $x \neq \sup E$ which is a contradiction. Therefore, $x^2 \geq 2$.
Now, assume for contradiction $x > 2$. Then, define

$$h = \frac{x^2 - 2}{2x}$$

Note that $x^2 > 2 \Rightarrow h > 0 \Rightarrow x - h < x$. Now we prove $x - h$ is an upper bound for E. Compute $(x-h)^2 = x^2 - 2xh + h^2 = x^2 - (x^2 - 2) + h^2 = 2 + h^2 > 2$. Let $q \in E$, *i.e.* $0 < q < \sqrt{2}$. Then, $q^2 < 2 < (x-h)^2 \Rightarrow 0 < (x-h)^2 - q^2 \Rightarrow 0 < (x-h+q)(x-h-q)$ and

$$0 < \left(\frac{x^2+2}{2x} + q\right)(x - h - q).$$

Since $q > 0$ and $(x^2+2)/(2x) > 0$ then $0 < x - h - q \Rightarrow q < x - h$. Thus, $\forall q \in E, q < x - h \Rightarrow x - h$ is an upper bound for E. Since $x = \sup E \Rightarrow x \leq x + h \Rightarrow h \leq 0$, which is a contradiction.
Thus, $x^2 = 2$ and $x > 1$.
For contradiction, assume $\exists m, n \in \mathbb{N}$ such that $m > n, x = m/n$. Then, $\exists n \in \mathbb{N}$ such that $nx \in \mathbb{N}$. Let $S = \{k \in \mathbb{N} : kx \in \mathbb{N}\}$, note that $n \in S \Rightarrow S \neq \varnothing$. By the well-ordering of $\mathbb{N}$, S has the least element $k_0 \in S$. Define $k_1 = k_0 x - k_0 \in \mathbb{Z}$. Since $x > 1, k_1 = k_0(x-1) > 0 \Rightarrow k_1 \in \mathbb{N}$. Since $x^2 = 2 \Rightarrow 4 - x^2 > 0 \Rightarrow (2-x)(2+x) > 0 \Rightarrow 2 - x > 0 \Rightarrow x < 2$. Then $k_1 = k_0(x-1) < k_0(2-x) = k_0$. Thus, $k_1 \in \mathbb{N}$ and $k_1 < k_0$. Computing $xk_1 = x(xk_0 - k_0) = x^2 k_0 - xk_0 = 2k_0 - xk_0 = k_0 + (k_0 - xk_0) = k_0 - k_1 \in \mathbb{N}$. Thus, $k_1 \in S$ and $k_1 < k_0$ which means k_0 is not the least element in S and $\sup E$ does not exist in $\mathbb{Q}$. $\square$

2.3.2 The definition of $\mathbb{R}$

First, in order to define $\mathbb{R}$ the previous result about the lack of an upper bound for $E = \{q \in \mathbb{Q} : 0 < q < \sqrt{2}\}$ allows us to introduce a definition.

Definition 20 (Least upper bound property)**.** An ordered set S has the least upper bound property if every nonempty and bounded above subset $E \subseteq S$ has a supremum in S.

The previous definition could be stated about a 'Greatest upper bound property'. Clearly $\mathbb{Q}$ does not have the Least upper bound property as the previous subsection has shown.

Now, for the real numbers,

Theorem 6 (Existence of $\mathbb{R}$)**.** There exists a unique ordered field which contains $\mathbb{Q}$ and has the least upper bound property. This field is denoted by $\mathbb{R}$.

Theorem 7. There exists a unique $r \in \mathbb{R}$ such that $r > 0$ and $r^2 = 2$.

Proof. First, let $\tilde{E} = \{x \in \mathbb{R} : 0 < x < \sqrt{2}\}$. Then, $\tilde{E}$ is bounded above. Take $r = \sup \tilde{E}$. The same proof as before show $r > 1$ and $r^2 = 2$. We now prove r is unique. Suppose $\tilde{r} \in \mathbb{R}, \tilde{r} > 0$ and $\tilde{r}^2 = 2$. Then, $0 = \tilde{r}^2 - r^2 = (\tilde{r} - r)(\tilde{r} + r) \Rightarrow 0 = \tilde{r} - r \Rightarrow r = \tilde{r}$. □

2.3.3 The density of $\mathbb{Q}$ in $\mathbb{R}$

The set $\mathbb{Q}$ contains $\mathbb{N}$, and $\mathbb{R}$ contains $\mathbb{Q}$. The following theorems shows how $\mathbb{N}$ and $\mathbb{Q}$ sit inside $\mathbb{R}$:

Theorem 8 (Archimedean property)**.** If $x, y \in \mathbb{R}$ and $x > 0$ then $\exists n \in \mathbb{N}$ such that $nx < y$.

Proof. Suppose $x, y \in \mathbb{R}$ and $x > 0$. We need to show $\exists n \in \mathbb{N}$ such that $nx < y$, *i.e.* $x < y/n$. Assume for contradiction $\forall n \in \mathbb{N} : n \leq y/x$. Then $\mathbb{N} \subseteq \mathbb{R}$ is bounded above, hence it has a supremum, by the least upper bound property of $\mathbb{R}$ with value $a \in \mathbb{R}$. Since a is the supremum of $\mathbb{N}$ then $a - 1$ is not an upper bound for $\mathbb{N}$. Therefore, $\exists m \in \mathbb{N}$ such that $a - 1 < m \Rightarrow a < m + 1$ which implies a is not an upper bound for $\mathbb{N}$ contradiction our initial claim. □

Theorem 9 (Densit of $\mathbb{Q}$ in $\mathbb{R}$)**.** If $x, y \in \mathbb{R}$ and $x < y$ then $\exists q \in \mathbb{Q}$ such that $x < r < y$.

Proof. Let $x, y \in \mathbb{R}$ and $x < y$, then:

1. If $x < 0 < y$ we have $r = 0 \in \mathbb{Q}$.
2. If $0 \leq x < y$ then by the Archimedean property, $\exists n \in \mathbb{N}$ such that $n(y - x) > 1$ and $\exists l \in \mathbb{N}$ such that $l > nx$. Thus, $S = \{k \in \mathbb{N} : k > nx\} \neq \varnothing$. By the well ordering property of $\mathbb{N}$, S has the least element m.
   Since $m \in S \Rightarrow nx < m$. Since m is the least element of S, $m - 1 \notin S \Rightarrow m - 1 \leq nx \Rightarrow m \leq nx + 1$. Thus, $nx < m < nx + 1 \Rightarrow x < m/n < y$. So, $r = m/n \in \mathbb{Q}$ is the solution.
3. If $x < y \leq 0$, then $0 \leq -y < x$. Then by the previous result $\exists \tilde{r} \in \mathbb{Q}$ such that $-y < \tilde{r} < -x$, or equivalently $x < -\tilde{r} < y$ and $r = -\tilde{r}$ is the solution.

□

2.4 Cardinality

Now, we turn our attention to cardinality, which is an approach to compare the size of sets.

Definition 21 (Cardinality). Two sets, A and B, have the same cardinality if there exists a bijective function $f : A \to B$.

Notation

- If A and B have the same cardinality we write $|A| = |B|$
- If $|A| = |\{1, 2, 3, ..., n\}|$ we write $|A| = n$
- If there exists an injective function $f : A \to B$ we write $|A| \leq |B|$
- If $|A| \leq |B|$ and $|A| \neq |B|$ then $|A| < |B|$

Theorem 10 (Cantor-Schorer-Bernstein). If $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$

2.4.1 Countable, uncountable and countably infinite sets

Definition 22 (Countably infinite). If $|A| = |\mathbb{N}|$ then A is countably infinite.

Definition 23 (Countable sets). If A is countably infinite or finite, then A is countable.

Definition 24 (Uncountable set). If A is neither countably infinite nor finite, then A is uncountable.

Since both cardinality and countability have been introduced it is time to appreciate some very interesting results.

> There are twice as many numbers as numbers
>
> *Richard Feynman*

Theorem 11. The set of positive even numbers is countable, *i.e.* $|\{2 \times n : n \in \mathbb{N}\}| = |\mathbb{N}|$. And so are the odd numbers, $|\{2 \times n - 1 : n \in \mathbb{N}\}|$.

Proof. Let $f : \mathbb{N} \to \{2 \times n : n \in \mathbb{N}\}$. So, $f(n) = 2n, \forall n \in \mathbb{N}$.
First, if $f(n_1) = f(n_2)$ then $2n_1 = 2n_2$, hence $n_1 = n_1$ and f is injective.
Second, let $m \in \{2 \times k : k \in \mathbb{N}\}$. Then, $\exists n \in \mathbb{N}$ such that $m = 2n$, the function $f(n) = 2n = m \Rightarrow n = m/2$. Therefore, f is also surjective, and by result there exists a bijective function $f : \mathbb{N} \to \{2 \times n : n \in \mathbb{N}\}$, and we conclude $|N| = |\{2 \times n : n \in \mathbb{N}\}|$.
The proof for the odd numbers is similar. $\square$

Theorem 12 (Countability of $\mathbb{Z}$)**.** The set of integers is countable, *i.e.* $|\mathbb{N}| = |\mathbb{Z}|$.

Proof. Define $f : \mathbb{N} \to \mathbb{Z}$ as

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

$f(n)$ is both injective and surjective. Hence, $|\mathbb{Z}| = |\mathbb{N}|$. □

Theorem 13 (Countability of $\mathbb{Q}$)**.** $|\mathbb{Q}| = |\mathbb{N}|$

Proof. Set $A_1 = \{0\}$, for $n \geq 2$ define $A_n = \{\pm p/q : p, q \in \mathbb{N}$ and are in the lowest terms with $p+q = n\}$. For example, $A_1 = \{0\}$, $A_2 = \{1/1, -1/1\}$, $A_3 = \{1/2, -1/2, 2/1, -2/1\}$, and so on. Since each rational number appear in only one A_n and every rational number can be represented by the relation above, $|\mathbb{Q}| = |\mathbb{N}|$. □

Theorem 14. If $A \subseteq B$ and B is countable, then A is either countable or finite.

Theorem 15.

1. If $A_1, A_2, ..., A_m$ are each countable sets, then $\bigcup_{n=1}^{m} A_n$ is countable
2. If A_n is a countable $\forall n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} A_n$ is countable

2.4.2 Cantor's theorem

Cantor's diagonalization method

Theorem 16. The open interval $(0, 1) = \{x \in \mathbb{R} : 0 < x < 1\}$ is uncountable.

Proof. For contradiction, assume there exists $f : \mathbb{N} \to (0, 1)$ bijective. Next, for each $m \in \mathbb{N}, \exists f(m) \in (0, 1)$ we write the decimal representation $f(m) = 0.a_{m,1}a_{m,2}\dots$. Next, define $b \in (0, 1)$ such that $b = 0.b_1b_2...$ with each digit following

$$b_n = \begin{cases} 2 \text{ if } a_{n,n} \neq 2 \\ 3 \text{ if } a_{n,n} = 2 \end{cases} \tag{2.5}$$

Since $b_n \neq a_{n,n} \forall n \in \mathbb{N}$ there exists $b \in (0, 1)$ such that $\nexists n \in \mathbb{N} : f(n) = b$ and $f(m)$ is not surjective. □

Corollary. The set of real numbers, $\mathbb{R}$, is uncountable.

Power sets and Cantor's theorem

Given a set A, the power set $\mathcal{P}(A)$ refers to the collection of all subsets of A.

Theorem 17 (Cantor's theorem)**.** Given a set A, there is no function $f : A \to \mathcal{P}(A)$ which is onto.

Proof. Assume $f : A \to \mathcal{P}(A)$ is bijective. We prove f cannot be surjective by finding a subset $B \subseteq A$ that is not equal to $f(a)$ for any $a \in A$. Take $B = \{a \in A : a \notin f(a)\}$. If f is surjective, then $B = f(a')$ for some $a' \in A$. However:

1. If $a' \in B$ then $a' \notin f(a')$. However, since $B = f(a')$ this implies $a' \notin B$
2. Else, if $a' \notin B$ then $a' \in f(a') = B$, which is again a contradiction. So, there is no function $f : A \to P(A)$ which is onto.

□

Theorem 18. $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})| < |\mathcal{P}(\mathcal{P}(\mathbb{N}))| < ...$

Informally, there exists an infinite number of infinitudes.

Theorem 19. If $|A| = n$ then $\mathcal{P}(A) = 2^n$.

Corollary. $\forall n \in \mathbb{N} \cup \{0\}, n < 2^n$.

2.5 Epilogue

Cardinality allows us to create an equivalence relation between sets. In this sense, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ are grouped together and are called countable sets. On the other hand, $\mathbb{R}, (a, b), P(\mathbb{N})$ are uncountable, and belong to a separate group.

Because of the importance of the countable sets, it is usual to denote $\aleph_0 = |\mathbb{N}|$. In terms of cardinal numbers, if $|X| < \aleph_0$ then X is finite. In this way, $\aleph_0$ is the smallest infinite cardinal number. The cardinality of $\mathbb{R}$ also deserves its special designation $\boldsymbol{c} = |\mathbb{R}| = |(0, 1)|$. Hence, $\aleph_0 < \boldsymbol{c}$.

From this point, one possible question to ask is: "is there a set $A \subseteq \mathbb{R} : \aleph_0 < |A| < \boldsymbol{c}$"?

Cantor believed there was no such set, leading to the "continuum hypothesis" i.e. $\nexists A \subseteq \mathbb{R} : \aleph_o < |A| < \boldsymbol{c}$. In 1940, Kurt Gödel showed there was no way to disprove this hypothesis from the axioms of set theory. Latter, in 1964 Paul Cohen showed it was also impossible to prove this conjecture. Hence, the problem of the continuum hypothesis is undecidable.

Chapter 3

Sequences and series

3.1 The starting problem

Basically a series is a sum of infinite terms. On the following example, some problems will appear as we try to manipulate the series as standard mathematical entities. Consider, for instance:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ... \tag{3.1}$$

We can consider the partial sum, s_n, *i.e.* the sum of the n first terms of the series. In this case we would obtain: $s_1 = 1$, $s_2 = 1/2$, $s_3 = 5/6$,... and so on. Interestingly, the odd sums decrease ($s_1 > s_3 > s_5 > ...$), while the even sums increase ($s_2 < s_4 < s_6 < ...$). It gives the idea that (s_n), the sequence of partial sums, converges to some number S. And we may feel tempted to write:

$$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...$$

However, the use of standard mathematical notation $(+, -, =)$ for series can be misleading. Take the previous equation, multiply it for $1/2$ and add it to itself. We would get:

$$\frac{3}{2}S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...$$

Which seems to be a contradiction to our initial claim. In a certain sense, addition in this infinite setting is not commutative.

Another example is the series:

$$\sum_{n=1}^{\infty} (-1)^n = -1 + 1 - 1 + 1 - 1 + ... \tag{3.2}$$

Depending on how we group the terms we would find different results:

$$(-1 + 1) + (-1 + 1) + (-1 + 1) + ... = 0$$

On the other hand,

$$-1 + (1 - 1) + (1 - 1) + (1 - 1) + ... = 1$$

In order to deal with the tricks hidden in infinite series, we begin by discussing sequences.

3.2 Sequences

3.2.1 Convergent sequences

Definition 25 (Sequence)**.** A sequence is a function, f, whose domain is $\mathbb{N}$. In this way, $f : \mathbb{N} \to \mathbb{R}$. Hence, $f(n)$ is the n-th term of the sequence. Notation: usually, a sequence is presented in the form (x_n), or $(x_n)_{n=1}^{\infty}$, or $x_1, x_2, x_3, \ldots$.

Definition 26 (Convergence of a sequence)**.** A sequence (x_n) converges to x if $\forall \varepsilon > 0, \exists N \in N$ such that $|x_n - x| < \varepsilon, \forall n \geq N$. There are a few different ways to denote convergence, such as $(x_n) \to x$, $\lim\limits_{n\to\infty} x_n = x$ or $x_n \to x$.

The negation of the convergence of a sequence would be:

Definition 27. A sequence (x_n) does not converge to x if $\exists \varepsilon_0 > 0$ such that $\exists m \in \mathbb{N}$ such that $|x_n - x| \geq \varepsilon, \forall n \geq m$.

Example.

$$\lim_{n\to\infty} \frac{1}{n^2 + 30n + 1} = 0$$

Solution:
We need to find $N \in \mathbb{N}$ such that

$$\frac{1}{n^2 + 30n + 1} < \varepsilon, \forall n \geq N$$

But

$$\frac{1}{n^2 + 30n + 1} \leq \frac{1}{n^2 + 30n} \leq \frac{1}{30n} \leq \frac{1}{n}$$

Hence, if $1/n < \varepsilon$ the initial inequality is immediately satisfied.
Let $\varepsilon > 0$, set $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Then, for all $n \geq N$:

$$\left|\frac{1}{n^2 + 30n + 1} - 0\right| = \frac{1}{n^2 + 30n + 1} \leq \frac{1}{30n} \leq \frac{1}{n} \leq \frac{1}{N} < \varepsilon$$

Definition 28 (Bounded sequences)**.** A sequence (x_n) is bounded if there exists a number $M > 0$ such that $|x_n| < M, \forall n \in \mathbb{N}$.

Theorem 20. If (x_n) is convergent, then (x_n) is bounded.

Proof. Suppose $(x_n) \to x$, then $\exists N \in \mathbb{N}$ such that $\forall n \geq N, |x_n - x| < \varepsilon, \forall \varepsilon > 0$. Regardless if x is positive or negative, we can write $|x_n| < |x| + \varepsilon$. Define $M = \max(||x_1|, |x_2|, ..., |x_{N-1}|, |x| + \varepsilon)$. Then, $|x_n| \leq M, \forall n \in \mathbb{N}$. □

Definition 29. A sequence (x_n) is:

1. Monotone increasing, if $x_n \leq x_{n+1}, \forall n \in \mathbb{N}$,
2. Monotone decreasing, if $x_n \geq x_{n+1}, \forall n \in \mathbb{N}$,
3. If it is either monotone increasing or decreasing, then it is called monotone.

Theorem 21. A monotonic sequence is convergent if, and only if, it is bounded.

Proof. Suppose (x_n) is a monotonic increasing sequence. Then,

1. $(\Rightarrow)$ follows from the previous theorem.
2. $(\Leftarrow)$. Suppose (x_n) is bounded. Then, $(x_n : n \in \mathbb{N}) \subseteq \mathbb{R}$ is a bounded set. Let $x = \sup(x_n : n \in \mathbb{R})$. We claim
$$x = \lim_{n \to \infty} x_n$$
Let $\varepsilon > 0$. Since $x - \varepsilon$ is not an upper bound for $(x_n : n \in \mathbb{N})$, $\exists M_0 \in \mathbb{N}$ such that $x_n - \varepsilon < x_{M_0} < x$. Choose $M = M_0$, then $\forall n \geq M, x - \varepsilon < x_{M_0} < x_n \leq x + \varepsilon$, or $x - \varepsilon < x_M < x + \varepsilon$.

□

Theorem 22 (Algebraic limit theorem)**.** Let $(a_n) \to a$ and $(b_n) tob$. Then,

1. $(ca_n) \to ca, \forall c \in \mathbb{R}$
2. $(a_n + b_n) \to a + b$
3. $(a_n b_n) \to ab$
4. $(a_n / b_n) \to a/b$, given $b \neq 0$

Proof. Let's take each item individually:

1. First, note $|ca_n - ca| = |c||a_n - a|$. Hence, for $\varepsilon > 0$ we have $|ca_n - ca| < \varepsilon \Leftrightarrow |a_n - a| < \varepsilon/|c|$. Since $(a_n) \to a$ then $\exists N \in \mathbb{N}$ such that $|a_n - a| < \varepsilon/|c|$, so $|ca_n - ca| = |c||a_n - a| < |c|\varepsilon/|c|, \forall n \geq N$.
2. From the triangle inequality, $|(a_n + b_n) - (a - b)| \leq |a_n - a| + |b_n - b|$. Set $N_1 \in \mathbb{N}$ such that $|a_n - a| < \varepsilon/2, \forall n \geq N_1$ with $\varepsilon > 0$. And set $N_2 \in \mathbb{N}$ such that $|b_n - b| < \varepsilon/2, \forall n \geq N_2$.

Then, for $N = \max(N_1, N_2)$ we obtain: $|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

3. First, $|a_nb_n - ab| = |a_nb_n - ab_n + ab_n - ab| \leq |a_nb_n - ab_n| + |ab_n - ab| = |b_n||a_n - a| + |a||b_n - b|$. Take $N_1 \in \mathbb{N}$ such that $|b_n - b| < \varepsilon/(2|a|), \forall n \geq N_1$ with $\varepsilon > 0$. Since every convergent sequence is bounded, take $M > 0$ so that $|b_n| < M, \forall n \in \mathbb{N}$. Then, set $N_2 \in \mathbb{N}$ such that $|a_n - a| < \varepsilon/(2M), \forall n \geq N_2$. Finally, for $N = \max(N_1, N_2)$ we obtain $|a_nb_n - ab| \leq |b_n||a_n - a| + |a||b_n - b| < M\varepsilon/(2M) + |a|\varepsilon/(2|a|) = \varepsilon$.

4. $(a_n/b_n) \to a/b$ follows from the previous result by noting $((1/b_n)) \to 1/b$, provided $b \neq 0$. So,
$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b||b_n|}$$
$|b_n - b|$ can be made arbitrarily small. On the other hand, considering $\varepsilon_0 = |b|/2$, define $N_1 \in \mathbb{N}$ such that $|b_n - b| < |b|/2, \forall n \geq N_1$, hence $|b_n| > |b|/2, \forall n \geq N_1$. Now, set $N_2 \in \mathbb{N}$ such that $|b_n - b| < \varepsilon|b|^2/2$. Taking $N = \max(N_1, N_2)$ leads to
$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = |b - b_n|\frac{1}{|b||b_n|} < \frac{\varepsilon|b|^2}{2}\frac{1}{|b|\frac{|b|}{2}} = \varepsilon, \forall n \geq N$$

□

Theorem 23 (Order limit theorem)**.** Assume $(a_n) \to a$ and $(b_n) \to b$, then:

1. If $a_n \geq 0, \forall n \in \mathbb{N} \Rightarrow a \geq 0$,

2. If $a_n \leq b_n, \forall n \in \mathbb{N} \Rightarrow a \leq b$,

3. If there exists $c \in \mathbb{R}$ such that $c \leq b_n, \forall n \in \mathbb{N}$, then $c \leq b$. Similarly, if $a_n \leq c, \forall n \in \mathbb{N}$ then $a \leq c$.

Proof. Proof for each statement:

1. Assume $a < 0$. Consider $\varepsilon = |a|$, since $(a_n) \to a$ we have $|a_n - a| < |a|, \forall n \geq N$. In particular, $|a_N - a| < |a|$ hence $a_N < 0$ which is a contradiction. Therefore, $a \geq 0$.

2. From the algebraic theorem, $(b_n - a_n) \to a - b$. Since $b_n - a_n \geq 0$ from the previous result, we get $b - a \geq 0$.

3. Take $a_n = c, \forall n \in \mathbb{N}$. From the previous theorem, if $c \leq b_n$ then $b - c \geq 0$. Hence, $b \geq c$.

□

3.2.2 Operations involving convergent sequences

In order to find out if a sequence converges, and to what value, there are a few tools at our disposal. We begin with the very popular Squeeze theorem, sometimes refered here as ST.

Theorem 24 (Squeeze theorem)**.** Let $(a_n), (b_n), (x_n)$ be sequences such that $\forall n \in \mathbb{N}$, $a_n \leq x_n \leq b_n$. Suppose (a_n) and (b_n) both converge and

$$\lim_{n\to\infty} a_n = x = \lim_{n\to\infty} b_n$$

So, $(x_n) \to x$.

Proof. Let $\varepsilon > 0$. Since $(a_n) \to x$, $\exists M_0 \in \mathbb{N}$ such that $\forall n \geq M_0, |a_n - x| < \varepsilon \Rightarrow x - \varepsilon < a_n$.
Since $(b_n) \to x$, then $\exists M_1 \in \mathbb{N}$ such that $\forall n \geq M_1, |b_n - x| < \varepsilon \Rightarrow b_n < x + \varepsilon$.
Choose $M = \max(M_0, M_1)$. Then, if $n \geq M, x - \varepsilon < a_n \leq x_n \leq b_n < x - \varepsilon \Rightarrow |x_n - x| > \varepsilon$. □

The limit of a function can be also expressed in another form, which can be pretty useful.

Theorem 25. Another way to check if $(x_n) \to x$ would be

$$\lim_{n\to\infty} x_n = x \iff \lim_{n\to\infty} |x_n - x| = 0$$

Example. Show that:

$$\lim_{n\to\infty} \frac{n^2}{n^2+n+1} = 1$$

Solution:
We have

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

Thus,

$$0 \leq \left|\frac{n^2}{n^2+n+1} - 1\right| \leq \frac{1}{n} \Longrightarrow \lim_{n\to\infty} \left|\frac{n^2}{n^2+n+1} - 1\right| = 0$$

by the Squeeze theorem.

3.2.3 Some special sequences

Theorem 26. If (x_n) is a convergent sequence such that $\forall n \in \mathbb{N}, x_n \geq 0$, then $(\sqrt{x_n})$ is convergent and

$$\lim_{n\to\infty} \sqrt{x_n} = \sqrt{\lim_{n\to\infty} x_n} \tag{3.3}$$

Proof. Let $x = \lim\limits_{n\to infty} x_n$, then:

1. If $x = 0$. Let $\varepsilon > 0$ then, since $(x_n) \to 0$, $\exists N_0 \in \mathbb{N}$ such that $x_n = |x_n - 0|\varepsilon^2, \forall n \geq N_0$. Choose $N = N_0$, then $\sqrt{x_n} - \sqrt{0} = \sqrt{x_n} < \sqrt{\varepsilon^2} = \varepsilon$
2. If $x > 0$. Then,

$$\begin{aligned}|\sqrt{x_n} - \sqrt{x}| &= \left|\frac{\sqrt{x_n} - \sqrt{x}}{\sqrt{x_n} + \sqrt{x}}(\sqrt{x_n} + \sqrt{x})\right| && (3.4)\\ &= \frac{1}{\sqrt{x_n} - \sqrt{x}}|x_n - x| && (3.5)\\ &\leq \frac{1}{\sqrt{x}}|x_n - x|, \forall n \in \mathbb{N} && (3.6)\end{aligned}$$

And,

$$0 \leq |\sqrt{x_n} - \sqrt{x}| \leq \frac{1}{\sqrt{x}}|x_n - x|, \forall n \in \mathbb{N}$$

So, by the squeeze theorem, $\lim_{n\to\infty} |\sqrt{x_n} - \sqrt{x}| = 0$

□

Theorem 27. If (x_n) is convergent and $\lim_{n\to\infty} x_n = x$, then $(|x_n|)$ is convergent and $\lim_{n\to\infty} |x_n| = |x|$.

Proof. Note that $|x| = \sqrt{x^2}, \forall x \in \mathbb{R}$. Then,

$$\lim_{n\to\infty} |x_n| = \lim_{n\to\infty} \sqrt{x_n^2} = \sqrt{x^2} = x$$

□

Theorem 28. If $x \in (0, 1)$ then $\lim_{n\to\infty} c^n = 0$. If $c > 1$ then (c^n) is unbounded.

Proof. For each case:

1. If $0 < c < 1$. Note that $0 < c^{n+1} < c^n < 1, \forall n \in \mathbb{N}$. This can be shown by induction:
   - Base case, consider $0 < c^2 < c < 1$, since $0 < c < 1$.
   - Inductive case, consider $0 < c^{m+1} < c^m < 1$ to be true. Multiplying the former inequality by c we obtain $0 < c^{m+2} < c^{m+1}$.

   Thus, (c^n) is monotone decreasing sequence and is bounded below, which implies (c^n) is convergent. Set $L = \lim_{n\to\infty} c^n$. Take $\varepsilon > 0$, then $\exists N \in \mathbb{N}$ such that

$$
\begin{aligned}
(1-c)|L| = |L-cL| &= |L-c^{M+1}+c^{M+1}-cL| \\
&\leq |L-c^{M+1}| + c|c^M - L| \\
&< (1-c)\frac{\varepsilon}{2} + c(1-c)\frac{\varepsilon}{2} \\
&< (1-c)\varepsilon, \forall n \geq N
\end{aligned}
$$

Hence, $|L| < \varepsilon, \forall \varepsilon > 0 \longrightarrow L = 0$.

2. For $c > 1$. Note that $\forall B \geq 0, \exists n \in \mathbb{N}$ such that $c^n > B$. For $n \in \mathbb{N}$ such that $n > B/(c-1)$ then $c^n = (1+(1-c))^n \geq 1 + n(c-1) \geq n(c-1) > B$. Hence, (c^n) is unbounded $\forall c > 1$ and the sequence does not converge.

$\square$

Theorem 29. If $p > 0$ then $\lim_{n\to\infty} n^{-p} = 0$

Proof. Let $\varepsilon > 0$. Take $N > (1/\varepsilon)^{1/p}$, then

$$\left|\frac{1}{n^p} - 0\right| = \frac{1}{|n^p|} \leq \frac{1}{N^p} < \varepsilon$$

$\square$

Theorem 30. If $p > 0$ then $\lim_{n\to\infty} p^{1/n} = 1$.

3.2.4 Subsequences and Bolzano–Weierstrass theorem

Definition 30 (Subsequence)**.** Let (x_n) be a sequence of real numbers, and (n_k) be a strictly increasing sequence of natural numbers. Then, $(x_{n_k})_{k=1}^{\infty}$ is called a subsequence of (x_n).

Theorem 31. If (x_n) converges to x then any subsequence of (x_n) will converge to x.

Proof. Suppose $(x_n) \to x$. Let $\varepsilon > 0$, then $\exists M_0 \in \mathbb{N}$ such that $\forall n \geq M_0, |x_n - x| < \varepsilon$. Choose, $M = M_0$. If $k \geq M$, then $n_k \geq k \geq M = M_0$, hence for $\varepsilon > 0, \exists M \in \mathbb{N}$ such that $|x_{n_k} - x| < \varepsilon, \forall n_k \geq M$. $\square$

From the decision of subsequence we may ask: "does a bounded sequence have a convergent subsequence?". The answer is yes, before we show it, we need to define some specific limits.

Definition 31 (Limsup/liminf). Let (x_n) be a bounded sequence. If the limit exists, we can define:

- Limit superior: $\limsup_{n\to\infty} x_n = \lim_{n\to\infty} (\sup\{x_k : k \geq n\})$
- Limit inferior: $\liminf_{n\to\infty} x_n = \lim_{n\to\infty} (\inf\{x_k : k \geq n\})$

Now we proceed to show an interesting result: these limits always exist.

Theorem 32. Let (x_n) be a bounded sequence, and

- $a_n = \sup\{x_k : k \geq n\}$
- $b_n = \inf\{x_k : k \geq n\}$

Then, the following statements are true:

1. (a_n) is monotone decreasing and bounded,
2. (b_n) is monotone increasing and bounded,
3. $\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$

Proof. Proving each of the results:

1. First, $\{x_k : k \geq n+1\} \subseteq \{x_k : k \geq n\}, \forall n \in N$, so $a_{n+1} = \sup\{x_k : k \geq n+1\} \leq \sup\{x_k : k \geq n\} = a_n$.
2. Similarly, $b_{n+1} \geq b_n, \forall n \in \mathbb{N}$. Since (x_n) is bounded, $\exists M \geq 0$ such that $-B \leq x_n \leq B, \forall n \in \mathbb{N}$. So, $-B \leq b_n \leq a_n \leq B$.
3. By the previous result, $b_n \leq a_n, \forall n \in \mathbb{N} \implies \liminf_{n\to\infty} x_n = \lim_{n\to\infty} b_n \leq \lim_{n\to\infty} a_n = \limsup_{n\to\infty} x_n$

$\square$

Example. Consider the sequence (x_n) with $x_n = (-1)^n$. Calculate the limit superior and limit inferior.
Solution:
First, notice that $\{(-1)^k : k \geq n\} = \{-1, 1\}, \forall n \in \mathbb{N}$. Hence, the supremum is always 1 and the infimum is always -1. So,

$$\limsup_{n\to\infty} = 1$$

$$\liminf_{n\to\infty} = -1$$

Theorem 33 (Bolzano-Weierstrass). Every bounbed sequence has a convergent subsequence.

Proof. This result follows from the previous theorem. □

Theorem 34. Let (x_n) be a bounded subsequence. Then, (x_n) converges if, and only if:

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

Proof. Proving each direction separately:

- $(\Longrightarrow)$: let $x = \lim_{n\to\infty} x_n$ then every subsequence converges to x, so $\liminf_{n\to\infty} x_n = x$ and $\limsup_{n\to\infty} x_n = x$. Hence, $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.
- $(\Longleftarrow)$: suppose $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$. Then, $\inf\{x_k : k \geq n\} \leq x_k \leq \sup\{x_k : k \geq n\}, \forall n \in \mathbb{N}$. By the squeeze theorem we obtain: $\lim_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.

□

3.2.5 Cauchy sequences

Definition 32 (Cauchy sequence)**.** A sequence (x_n) is Cauchy if $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $|x_n - x_k| < \varepsilon, \forall n, k \geq N$.

Example. Show $x_n = 1/n$ is Cauchy.
Solution:
Let $\varepsilon > 0$, take $N \in \mathbb{N}$ such that $1/N < \varepsilon/2$. Then,

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

Theorem 35. If (x_n) is Cauchy, then (x_n) is bounded.

Proof. From the definition of a Cauchy sequence, take $\varepsilon = 1$ without loss of generality, then $\exists N \in \mathbb{N}$ such that $|x_n - x_k| < 1, \forall n, k \geq N$. So, $|x_n| \leq |x_n - x_N| + |x_N| < |x_M| + 1$. Let $M = |x_1| + |x_1| + ... + |x_M| + 1$, so $|x_n| \leq M, \forall n \in \mathbb{N}$. □

Theorem 36. IF (x_n) is Cauchy and a subsequence (x_{n_k}) converges, then (x_n) converges.

Proof. Suppose (x_{n_k}) is a subsequence of (x_n) such that $\lim_{n\to\infty} x_n = x$. Let $\varepsilon > 0$, then $\exists N_1 \in \mathbb{N}$ such that $|x_{n_k} - x| < \varepsilon/2, \forall k \geq N_1$. Since (x_n) is Cauchy, $\exists N_2 \in \mathbb{N}$ such that $|x_n - x_m| < \varepsilon/2, \forall n, m \geq M_1$.
Take $N = \max\{N_1, N_2\}$. Then, $|x_n - x| \leq |x_n - x_{n_N}| + |x_{n_M} - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Theorem 37. A sequence (x_n) is Cauchy if, and only if, it is convergent.

Proof. Proving each direction of the theorem:

- $(\Longrightarrow)$: if (x_n) is Cauchy then it is bounded. So, by the Bolzano-Weierstrass theorem, there exists a convergent subsequence. From the previous theorem, if a sequence is Cauchy and has a convergent subsequence, then it is convergent. So, (x_n) is convergent.
- $(\Longleftarrow)$: Suppose (x_n) is convergent and $(x_n) \to x$. Let $\varepsilon > 0$, then $\exists N_0 \in \mathbb{N}$ such that $|x_n - x| < \varepsilon/2, \forall n \geq N_0$. Choose $N = N_0$, then $|x_n - x_k| \leq |x_n - x| + |x_k - x| <$

□

3.3 Series

3.3.1 Convergent series

As pointed out by David Bressoud in " A radical approach to real analysis", the infinite summation is in itself an oxymoron. That is, the sum is the process of adding up, or reaching the totality. On the other hand, infinite means never-ending. Weird things can happen as we deal with series. The formal treatment of real analysis aims to safeguard us against danger.

Definition 33 (Series convergence)**.** Given the sequence, (x_n), the series associated with it is the summation of its terms, denoted as $\sum\limits_{n=1}^{\infty} x_n = \sum x_n$. The series converges the sequence of partial sums

$$\left(s_m = \sum_{n=1}^{m} x_n\right)_{m=1}^{\infty} \tag{3.7}$$

converges. If $\lim\limits_{m\to\infty} s_m = s$ we write $\sum x_n = s$ and treat $\sum x_n$ as a number.

Example. Prove that the series $\sum_{n=1}^{\infty} 1/(n(n+1))$ converges.
Solution:
First, note that

$$\begin{aligned} s_m = \sum_{n=1}^{\infty} \frac{1}{n(n+1)} &= \sum_{n=1}^{\infty} \frac{1}{n} - \frac{1}{n+1} \\ &= \left(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{m}\right) - \left(\frac{1}{2} + \frac{1}{3} + ... + \frac{1}{m+1}\right) \\ &= 1 - \frac{1}{m+1} \end{aligned}$$

Thus, $s_m = 1 - \frac{1}{m+1} \to 1$ and the series converges.

Theorem 38 (Geometric series convergence)**.** If $|x| < 1$ then $\sum_{n=0}^{\infty} r^n$ converges and

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-x} \tag{3.8}$$

Proof. First, note that

$$s_m = \sum_{n=0}^{m} x^n = \frac{1-x^{m+1}}{1-x}, \forall m \in \mathbb{N}$$

which can be proven by induction (not the point here). Since $|x| < 1$ we have $\lim_{m\to\infty} |x|^{m+1} = 0$ and

$$\lim_{n\to\infty} s_m = \frac{1-0}{1-x} = \frac{1}{1-x}$$

□

Theorem 39. Consider the sequence (x_n) and let $N \in \mathbb{N}$. Then, $\sum_{n=1}^{\infty} x_n$ converges if and only if $\sum_{n=N}^{\infty} x_n$ converges.

Proof. The partial sums satisfy

$$\sum_{n=1}^{N} x_n = \sum_{n=1}^{M} x_n + \sum_{n=M}^{N} x_n, \forall N \in \mathbb{N} \text{ and } 1 \leq M \leq N$$

□

Definition 34. The series $\sum x_n$ is Cauchy if the sequence of partial sums $(s_m = \sum_{n=1}^{m} x_n)$ is Cauchy.

Theorem 40. The series $\sum x_n$ is Cauchy if, and only if, $\sum x_n$ is convergent.

Proof. This follows from the previous theorem which states that a sequence is Cauchy if, and only if, it is convergent. □

Theorem 41. The series $\sum x_n$ is Cauchy if, and only if, $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $\forall m \geq N$ and $l \geq m$,

$$\left| \sum_{n=m+1}^{l} x_n \right| < \varepsilon$$

Proof. ($\Longrightarrow$) If $\sum x_n$ is Cauchy, let $\varepsilon > 0$ then $\exists N_0 \in \mathbb{N}$ such that $|s_m - s_l| < \varepsilon, \forall m, l \geq N_0$, where s_m is the partial sum of the first m terms. Hence, for $N = N_0$ if $m \geq N$ and $l > m$ we obtain:

$$|s_l - s_m| = \left| \sum_{n=m+1}^{l} x_n \right| < \varepsilon$$

($\Longleftarrow$) To-do $\square$

Theorem 42. If $\sum x_n$ converges, then $(x_n) \to 0$.

Proof. If $\sum x_n$ converges then it is Cauchy, hence for $\varepsilon > 0, \exists N_0$ such that $\forall l > m \geq N_0$,

$$\left| \sum_{n=m+1}^{l} x_n \right| < \varepsilon$$

For $N = N_0 + 1$, then $m \geq N$ implies $m - 1 \geq N_0$. For $l = m$:

$$|s_m| = \left| \sum_{n=m}^{m} x_n \right| < \varepsilon$$

And since ε can be made arbitrarily small $(x_n) \to 0$. $\square$

Theorem 43. If $|x| \geq 1$ then $\sum_{n=0}^{\infty} x^n$ diverges.

Proof. If $|x| \geq 1$ then $\lim_{n\to\infty} x_n \neq 0$ and $\sum x^n$ diverges from the previous theorem. $\square$

Corollary. The series $\sum_{n=0}^{\infty} \alpha \cdot x^n$ converges if, and only if, $|x| < 1$.

First, let's revisit one of the last results, namely that for a convergent series $\sum x_n$ we have $(x_n) \to x$. Is the converse true? That is, if $(x_n) \to 0$ can we conclude $\sum x_n$ converges? The answer is no, we will show it with a counter-example.

Theorem 44 (Divergence of the harmonic series)**.** The harmonic series $\sum_{n=1}^{\infty} 1/n$ does not converge.

Proof. We will show there exists a subsequence $s_m = \sum_{n=1}^{m} 1/n$ which is unbounded, hence the series diverges. Consider:

$$\begin{aligned}
s_{2^l} &= 1 + \left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + ... + \frac{1}{8}\right) + ... + \left(\frac{1}{2^{l-1}+1} + \frac{1}{2^l}\right) \\
&= 1 + \sum_{\lambda=1}^{l} \sum_{n=2^{\lambda-1}+1}^{2^\lambda} \frac{1}{n} \\
&\geq 1 + \sum_{\lambda=1}^{l} \sum_{n=2^{\lambda-1}+1}^{2^\lambda} \frac{1}{2^\lambda} \\
&= 1 + \sum_{\lambda=1}^{l} \frac{1}{2^\lambda}(2^\lambda - (2^{\lambda-1}+1) + 1) \\
&= 1 + \sum_{\lambda=1}^{l} \frac{2^{\lambda-1}}{2^\lambda} \\
&= 1 + \frac{l}{2}
\end{aligned}$$

Hence, $(s_{2^l})_{l=1}^{\infty}$ is unbounded and by consequence (s_{2^l}) does not converge. □

3.3.2 Properties of series

Theorem 45 (Algebraic limit theorem for series)**.** Consider two convergent series: $\sum_{n=1}^{\infty} x_n = X$ and $\sum_{n=1}^{\infty} y_n = Y$. Then:

1. $\sum_{n=1}^{\infty} cx_n = cX, \forall c \in \mathbb{R}$,
2. $\sum_{n=1}^{\infty}(x_n + y_n) = X + Y$

Proof. Proving each of the statements:

1. First, note that the sequence of partial sums takes the form $t_m = cx_1 + cx_2 + ... + cx_m$, which is equivalent to $t_m = cs_m$ where $s_m = x_1 + x_2 + ... + x_m$. By the algebraic limit theorem, if the sequence $(s_m) \to x$ then $(t_m) = (cs_m) \to cX$.

2. Equivalently, we have $s_m = x_1 + x_2 + ... + x_3$ and $t_m = y_1 + y_2 + ... + y_m$. And, $u_m = (x_1 + y_1) + (x_2 + y_2) + ... + (x_m + y_m)$, which is equivalent to $u_m = s_m + t_m$. Since $(s_m) \to X$ and $(t_m) \to Y$, then by the algebraic limit theorem $(u_m) = (s_m + t_m) \to X + Y$.

□

Theorem 46. If $x_n \geq 0, \forall n \in \mathbb{N}$ then $\sum x_n$ converges if, and only if, the sequence of partial sums, (s_m), is bounded.

Proof. If $x_n \geq 0, \forall n \in \mathbb{N}$ then

$$s_{m+1} = \sum_{n=1}^{m+1} x_n = \sum_{n=1}^{m} + x_{m+1} = s_m + x_{m+1} \geq s_m$$

Which implies (s_m) is a monotone increasing sequence, which converges if, and only if, it is bounded. $\square$

Definition 35 (Absolute convergence)**.** The series $\sum x_n$ converges absolutely if $\sum |x_n|$ converges

Theorem 47. If $\sum x_n$ converges absolutely then $\sum x_n$ converges.

Proof. If $\sum |x_n|$ converges then from the previous theorem (s_m) is bounded, since $|x_n| \geq 0, \forall n \in \mathbb{N}$ and $s_m = \sum_{n=1}^{m} |x_n|$. From the previous result $s_{m+1} \geq s_m$. Since $\sum |x_n|$ converges by the initial hypothesis that means $\sum |x_n|$ is Cauchy. So, $\forall \varepsilon > 0, \exists N_0 \in \mathbb{N}$ such that:

$$\sum_{n=m+1}^{l} |x_n| < \varepsilon, \forall l > m \geq N_0$$

Take $N = N_0$. Then,

$$\left| \sum_{n=m+1}^{l} x_n \right| \leq \sum_{n=m+1}^{l} |x_n| < \varepsilon$$

Hence, $\sum x_n$ is Cauchy and converges. $\square$

From the previous theorem, we may be tempted to ask " If a series converges, does it mean it also converges absolutely?". We will show this is not the case by a counter-example:

Example. Consider the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

This series converges, as we will see in a while. Now, show it does not converge absolutely.
Solution:
Note that the sum of absolute values is exactly the harmonic series, which we have already seen does not converge.

3.3.3 Convergence tests

Theorem 48 (Comparison test)**.** Suppose $0 \leq x_n \leq y_n, \forall n \in \mathbb{N}$, then:

1. if $\sum y_n$ converges, then $\sum x_n$ converges,

2. if $\sum x_n$ diverges, then $\sum y_n$ diverges.

Proof. Proving each statement:

1. If $\sum y_n$ converges, then the sequence of partial, $\left(\sum_{n=1}^m y_n\right)$, sums is bounded. So, $\exists M \geq 0$ such that:
$$\sum_{n=1}^{m} y_n \leq B, \forall m \in \mathbb{N} \tag{3.9}$$
Hence, $0 \leq x_n \leq y_n, \forall n \in \mathbb{N} \Longrightarrow 0 \leq \sum_{n=1}^m x_n \leq \sum_{n=1}^m y_n \leq B, \forall m \in \mathbb{N}$, which implies $\left(\sum_{n=1}^m x_n\right)$ is bounded and therefore it converges.

2. if $\sum x_n$ diverges, then $\left(\sum_{n=1}^m x_n\right)$ is unbounded. Let $M \geq 0$, then:
$$\sum_{n=1}^{m} x_n \geq M, \forall m \in \mathbb{N} \tag{3.10}$$
Therefore, $\sum_{n=1}^m y_n \geq \sum_{n=1}^m x_n \geq M$. which implies $\left(\sum_{n=1}^m y_n\right)$ is unbounded and therefore diverges.

□

Theorem 49 (Ratio test)**.** Consider the series $\sum x_n$ and $x_n \neq 0, \forall n \in \mathbb{N}$. Suppose
$$L = \lim_{n\to\infty} \frac{|x_{n+1}|}{|x_n|} \tag{3.11}$$
exists. Then,

1. if $L < 1$ then $\sum x_n$ converges absolutely
2. if $L > 1$ then $\sum x_n$ diverges
3. if $L = 1$ no assertion can be made

Proof. Proving the first two statements:

1. For $L < 1$, take $\alpha \in (L, 1)$. Then, $\exists N_0 \in \mathbb{N}$ such that $|x_{n+1}|/|x_n| < \alpha, \forall n \geq N_0$. Equivalently $|x_{n+1}| \leq \alpha |x_n|, \forall n \geq N_0$. Which leads to $|x_n| \leq \alpha |x_{n+1}| \leq \alpha^2 |x_{n+2}| \leq ... \leq$

$\alpha^{n-N_0}|x_{n+1}|^{n-N_0}, \forall n \geq N_0$. Now, for $m \in \mathbb{N}$:

$$\begin{aligned}\sum_{n=1}^{m} |x_n| &= \sum_{n=1}^{N_0-1} |x_n| + \sum_{n=N_0}^{m} |x_n| \\ &\leq \sum_{n=1}^{N_0-1} |x_n| + |x_{N_0}| \sum_{n=N_0}^{m} \alpha^{n-N_0} \\ &\leq \sum_{n=1}^{N_0-1} |x_n| + |x_{N_0}| \sum_{l=0}^{\infty} \alpha^{l} \\ &= \sum_{n=1}^{N_0-1} |x_n| + \frac{|x_{N_0}|}{1-\alpha}\end{aligned}$$

Therefore, $(\sum_{n=1}^{m} |x_n|)_{m=1}^{\infty}$ is bounded and $\sum |x_n|$ converges.

2. For $L > 1$, take $\alpha \in (1, L)$. Then, $\exists N_0 \in \mathbb{N}$ such that $|x_{n+1}|/|x_n| \geq \alpha > 1, \forall n \geq N_0$. Which means $|x_{n+1}| \geq |x_n|, \forall n \geq N_0$. Hence, $\lim_{n\to\infty} |x_n| \neq 0$ and $\sum x_n$ diverges.

□

Example. Consider the series

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

Use the ratio test to verify it converges absolutely.
Solution:
We begin by noticing:

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

So, for the limit:

$$\lim_{n\to\infty} \left|\frac{\frac{(-1)^{n+1}}{(n+1)^2+1}}{\frac{(-1)^n}{(n)^2+1}}\right| < \lim_{n\to\infty} \frac{n^2}{(n+1)^2} = 1$$

Since the limit is less then 1, the series converges absolutely by the ration test.

Theorem 50 (Root test)**.** Consider the series $\sum x_n$ and suppose the limit

$$L = \lim_{n\to\infty} |x_n|^{1/n}$$

exists. Then,

1. if $L < 1$ the series converges absolutely,

2. if $L > 1$ the series diverges,

3. if $L = 1$ no assertion can be made.

Proof. Proving the first two assertions:

1. Take $r \in (L, 1)$. Since $(|x_n|^{1/n} \to L), \exists N \in \mathbb{N}$ such that $|x_n|^{1/n} < r, \forall n \geq N$ which is equivalent to $|x_n| \leq r^n, \forall n \geq N$. So,

$$\begin{aligned}
\sum_{n=1}^{n} |x_n| &= \sum_{n=1}^{N-1} |x_n| + \sum_{n=N}^{m} |x_n| \\
&\leq \sum_{n=1}^{N-1} |x_n| + \sum_{n=N}^{m} r^n \\
&\leq \sum_{n=1}^{N-1} |x_n| + \sum_{n=1}^{\infty} r^n \\
&= \sum_{n=1}^{N-1} |x_n| + \frac{1}{1-r}
\end{aligned}$$

   Thus, the sequence of partial sums of absolute values is bounded and the series converges absolutely.

2. Since $(|x_n|^{1/n}) \to L > 1, \exists N \in \mathbb{N}$ such that $|x_n|^{1/n} > 1, \forall n \geq N$, which is equivalent to $|x_n| > 1, \forall n \geq N$ and $\lim_{n\to\infty} x_n \neq 0$ and the series diverges.

$\square$

Theorem 51 (Alternating series test)**.** Consider the sequence (x_n) to be monotone decreasing, with $x_n \to x$. Then $\sum(-1)^n x_n$ converges.

Proof. Considering (x_n) is monotone decreasing, then for the partial sums:

$$\begin{aligned}
s_{2k} &= \sum_{n=1}^{2k} (-1)^n x_n \\
&= (x_2 - x_1) + (x_4 - x_3) + ... + (x_{2k} - x_{2k-1}) \\
&\geq (x_2 - x_1) + (x_4 - x_3) + ... + (x_{2k} - x_{2k-1}) + (x_{2k+2} - x_{2k+1}) \\
&= s_{2(k+1)}
\end{aligned}$$

Hence, (s_{2k}) is monotone decreasing. And, $s_{2k} = -x_1 + (x_2 - x_3) + (x_4 - x_5) + ... + (x_{2k-2} - x_{2k-1}) + x_{2k} \geq -x_1$. So (s_{2k}) is monotone decreasing and bounded below. Therefore it converges.
Take $s = \lim_{k\to\infty} s_{2k}$ and $\varepsilon > 0$. Since $s_{2k} \to s, \exists N_1 \in \mathbb{N}$ such that $|s_{2k} - s| < \varepsilon/2, \forall k \geq N_1$. On the other hand, since $x_n \to 0, \exists N_2 \in \mathbb{N}$ such that $|x_n| < \varepsilon/2, \forall n \geq N_2$. Choose

$N = \max(2N_1+1, N_2)$, then $|s_m - s| = |s_{2m/2} - s| < \varepsilon/2 < \varepsilon, \forall m \in\geq N$, with m even and $m/2 \geq N_1 + 1/2 \geq M_0$.
If m is even, take $k = (m-1)/2$ so $m = 2k+1$. Then, $|s_m - s| = |s_{m-1} + x_m - s| \leq |s_{2k} - s + x_m| \leq |s_{2k} - s| + |x_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Thus, $s_m \to s$ and $\sum(-1)^n x_n$ converges. $\square$

Remark. From the previous theorem, it is clear that $\sum(-1)^n/n$ converges since $(1/n)$ is monotone decreasing, however as it was shown before it does not converge absolutely as the sum of the absolute terms of the series is exactly the harmonic series, which does not converge.

Theorem 52. Suppose $\sum x_n$ converges absolutely to X. Consider $f : \mathbb{N} \to \mathbb{N}$ a bijective function, then $\sum x_{f(n)}$ converges absolutely and $\sum x_{f(n)}$ converges to X. In other words "If a series converges absolutely, then any rearrangement of its terms also converges to the same limit".

Proof. To show $\sum |x_{f(n)}|$ converges it suffices to show $\sum_{n=1}^m |x_{f(n)}|$ is bounded. Since $\sum x_n$ converges, then it is also bounded hence $\exists M \in \mathbb{N}$ such that

$$\sum_{n=1}^{l} |x_n| \leq M, \forall l \in \mathbb{N}$$

Let $m \in \mathbb{N}$, then $f((1,2,...,m))$ is a finite subset of $\mathbb{N}$, so $\exists l \in \mathbb{N}$ such that $f((1,2,...,m)) \subseteq (1,2,...,l)$. And,

$$\sum_{n=1}^{n} |x_{f(n)}| = \sum_{n\in f((1,2,...,m))} |x_n| \leq \sum_{n=1}^{l} |x_n| \leq M$$

So, $\sum |x_{f(n)}|$ converges. Let $\varepsilon > 0$ then $\exists N_0 \in \mathbb{N}$ such that:

$$\left| \sum_{n=1}^{m} x_n - X \right| < \varepsilon/2, \forall m \geq N_0$$

Since $\sum |x_n|$ converges, $\exists N_1 \in \mathbb{N}$ such that:

$$\sum_{n=m+1}^{m} |x_n| < \varepsilon/2, \text{ with } l > m > N_1$$

Take $N_2 = \max(N_0, N_1)$, then:

$$\left| \sum_{n=1}^{m} x_n - X \right| < \varepsilon/2 \text{ and } \sum_{n=m+1}^{m} |x_n| < \varepsilon/2 \text{ and } l > m \geq N_2$$

Since $f^{-1}((1,...,N_2))$ is finite, set $N \in \mathbb{N}$ such that $(1,...,N_2) \subseteq (1,...,N)$, then:

$$\begin{aligned}
\left|\sum_{n'=1}^{m'} x_{f(n')} - X\right| &= \left|\sum_{n \in f((1,...,m'))} x_n - X\right| \\
&= \left|\sum_{n=1}^{N} x_n - X + \sum_{n \in f((1,...,m'))\setminus(1,...,N)} x_n\right| \\
&\leq \left|\sum_{n=1}^{N} x_n - X\right| + \sum_{n=N+1}^{\max(f((1,...,m'))}|x_n| \\
&\leq \left|\sum_{n=1}^{N} x_n - X\right| + \sum_{n=N+1}^{l}|x_n| \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon
\end{aligned}$$

$\square$

Chapter 4

The derivative

Before we begin dealing with derivative, we must build some of the tools needed to define it and evaluate its existence. We begin by evaluating limits on a function, which lead us to the cornerstone concept of continuity.

4.1 Continuity

In order to expand our concepts of limits from sequences to functions on real numbers, we begin by defining the points on the real line we will be able to evaluate the future concepts.

4.1.1 Limits of function

Definition 36 (Cluster point)**.** Let $S \subseteq \mathbb{R}$. Then, $x \in \mathbb{R}$ is a cluster point of S if $\forall \delta > 0, (x - \delta, x + \delta) \cap S \setminus \{x\} \neq \varnothing$

This notion can be further clarified by a few examples:

Examples

1. $S = \{1/n : n \in \mathbb{N}\}$: 0 is a cluster point of S, since $1/n$ can be made arbitrary small, so $(0 - \delta, 0 + \delta) \cap S \setminus \{0\} \neq \varnothing, \forall \delta > 0$.

2. $S = (0, 1)$ then $[0, 1]$ is the set of cluster points of S.

3. $S = \mathbb{Q}$ then $\mathbb{R}$ is the set of cluster points.

Theorem 53. Let $S \subseteq \mathbb{R}$, then x is a cluster point of S if, and only if, there exists a sequence $\{x_n\}$ of elements in $S \setminus \{x\}$ such that $x_n \to x$.

Definition 37 (Function convergence)**.** Consider $S \subseteq \mathbb{R}$, c a cluster point of S, and $f : S \to \mathbb{R}$. Then, $f(x)$ converges to $L \in \mathbb{R}$ at c if $\forall \varepsilon > 0, \exists \delta > 0$ such that if $x \in S$ and $0 < |x - c| < \delta$, then $|f(x) - L| < \varepsilon$.

We can write $f(x) \to L$ as $x \to c$ or $\lim_{x\to c} f(x) = L$.

Theorem 54 (Uniqueness of the limit of a function)**.** Let c be a cluster point of $S \subseteq \mathbb{R}$ and $f : S \to \mathbb{R}$. If $f(x) \to L_1$ and $f(x) \to L_2$ as $x \to c$ then $L_1 = L_2$.

Proof. Take $\varepsilon > 0$, since $f(x) \to L_1$ and $f(x) \to L_2, \exists \delta_1, \delta_2 \in \mathbb{R}$ such that if $x \in S$ with $0 < |x-c| < \delta_1$ and $0 < |x-c| < \delta_2$ we have $|f(x) - L_1| < \varepsilon/2$ and $|f(x) - L_2| < \varepsilon/2$. Choose $\delta = \min\{\delta_1, \delta_2\}$. Since c is a cluster point of $S, \exists x_0 \in S$ such that $0 < |x_0 - c| < \delta \longrightarrow$ $|L_1 - L_2| = |L_1 - f(x_0) + f(x_0) - L_2| \leq |L_1 - f(x_0)| + |f(x_0) - L_2| < \varepsilon$. □

Theorem 55. Consider c a cluster point in $S \subseteq \mathbb{R}$, and $f : S \to \mathbb{R}$. Then, the following statements are equivalent:

- $\lim_{x\to c} f(x) = L$ and,
- for every sequence $\{x_n\}$ in $S \setminus \{x\}$ such that $x_n \to c$, then $f(x_n) \to L$.

Proof. Proving each direction of the theorem individually:

1. Suppose $f(x) \to L$ as $x \to c$, then consider $\{x_n\}$ in $S \setminus \{x\}$ such that $x_n \to c$. Let $\varepsilon > 0, \exists \delta > 0$ such that $|f(x)-L| < \varepsilon$ if $x \in S$ and $0 < |x-c| < \delta$. Since $x_n \to c, \exists N \in \mathbb{N}$ such that $0 < |x_n - c| < \delta, \forall n \geq N$, since $|f(x) - L|\varepsilon, \forall 0 < |x-c| < \delta$ then $f(x_n) \to L$.
2. Assuming the second part is false, for contradiction, $\exists \varepsilon_0 > 0$ such that $\forall \delta > 0, \exists x \in S$ such that $0 < |x-c| < \delta$ and $|f(x) - L| \geq \varepsilon_0$. Then, $\forall n \in \mathbb{N}, \exists x_n \in S$ such that $0 < |x_n - c| < 1/n$ and $|f(x_n) - L| \geq \varepsilon_0$. By the squeeze theorem we conclude $x_n \to c$ and $0 = \lim_{n\to\infty} |f(x_n) - L| \geq \varepsilon_0 > 0$ which is a contradiction.

□

Theorem 56.

$$\lim_{x\to 0} \sin(1/x) \text{ does not exist}$$

Proof. Let $x_n = \frac{2}{(2n-1)\pi}$. Then, $x_n \neq 0$ and $x_n \to 0$. Now,

$$\sin(1/x_n) = \sin\frac{(2n-1)\pi}{2} = (-1)^{n+1}$$

□

Theorem 57.

$$\lim_{x\to 0} x\sin(1/x) = 0$$

Proof. Suppose $x_n \neq 0$ and $x_n \to 0$. Then,

$$0 \leq |x_n \sin(1/x_n)| = |x_n||\sin(1/x_n)| \leq |x_n|$$

By the squeeze theorem, $\lim_{n\to\infty} |x_n \sin(1/x_n)| = 0$ □

Theorem 58. Consider c a cluster point in $S \subseteq \mathbb{R}$ and $f, g : S \to \mathbb{R}$, with $f(x) \leq g(x), \forall x \in S$. Suppose $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ both exist. Then,

$$\lim_{x\to c} f(x) \leq \lim_{x\to c} g(x)$$

Proof. Define $L_1 = \lim_{x\to c} f(x)$ and $L_2 = \lim_{x\to c} g(x)$, and $\{x_n\}$ to be a sequence in $S \setminus \{c\}$ with $x_n \to c$. Then, $f(x_n) \leq g(x_n), \forall n \in \mathbb{N}$. So,

$$L_1 = \lim_{n\to\infty} f(x_n) \leq \lim_{n\to\infty} g(x_n) = L_2$$

□

Definition 38 (Convergence from the left)**.** Consider c to be a cluster point of $S \cap (-\infty, c)$ with $S \subseteq \mathbb{R}$. Then, we say $f(x)$ converges to L from the left (or as $x \to c^-$) if $\forall \varepsilon > 0, \exists \delta > 0$ such that if $x \in S$ and $c - \delta < x < c$ we obtain $|f(x) - L| < \varepsilon$.
We denote it by $L = \lim_{x\to c^-} f(x)$.

Definition 39 (Convergence from the right)**.** Consider c to be a cluster point of $S \cap (c, \infty)$ with $S \subseteq \mathbb{R}$. Then, we say $f(x)$ converges to L from the right (or as $x \to c^+$) if $\forall \varepsilon > 0, \exists \delta > 0$ such that if $x \in S$ and $c < x < c + \delta$ we obtain $|f(x) - L| < \varepsilon$.
We denote it by $L = \lim_{x\to c^+} f(x)$.

Example Consider

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

Then, $\lim_{x\to 0^-} f(x) = 0$ and $\lim_{x\to 0^+} f(x) = 0$, despite $f(0)$ being undefined.

Theorem 59. Consider c a cluster point of $S \cap (-\infty, c)$ and $S \cap (c, \infty)$, with $S \subseteq \mathbb{R}$. Then, c is a cluster point of S. Or equivalently:

$$\lim_{x\to c} f(x) = L \longleftrightarrow \lim_{x\to c^-} f(x) = \lim_{x\to c^+} f(x) = L$$

4.1.2 Continuity of a function

As shown in a past example, it is possible that $\lim_{x\to c} f(x) \neq f(c)$. In other words, it is possible that a limit of a function as $x \to c$ (or $x \to c^-$, $x \to c^+$) differs from $f(c)$. Continuity links the two concepts.

Definition 40 (Continuous function)**.** Consider $c \in S \subseteq \mathbb{R}$ a cluster point. We say f is continuous at c if $\forall \varepsilon > 0, \exists \delta > 0$ such that if $x \in S$ with $|x - c| < \delta$ then $|f(x) - f(c)| < \varepsilon$. If f is continuous at all points of $U \subseteq S$ then f is continuous on U.

Theorem 60. Consider $c \in S \subseteq \mathbb{R}$, and $f : S \to \mathbb{R}$, then:

1. if c is not a cluster point of f, then f is continuous at c,
2. if c is a cluster point of f, then f is continuous at c if, and only if, $\lim_{x\to c} f(x) = f(c)$,
3. f is continuous at c, and only if, for all sequence $\{x_n\}$ in S with $x_n \to c$ then $f(x_n) \to f(c)$.

Proof. Proving each statement:

1. Consider $\varepsilon > 0$, since c is not a cluster point of S then $\exists \delta > 0$ such that $(c-\delta, c+\delta) \cap S = \{c\}$, so if $x \in S$ and $|x - c| < \delta$ then $x = c$ and $|f(x) - f(c)| < \varepsilon$.
2. Proving each direction of the statement:
   - ($\Longleftarrow$) If $\lim_{x\to c} f(x) = f(c)$ then $\forall \varepsilon > 0, \exists \delta > 0$ such that if $x \in S$ and $|x - c| < \delta$ then $|f(x) - f(c)| < \varepsilon$.
   - ($\Longrightarrow$) If f is continuous at c then $\forall \varepsilon > 0, \exists \delta > 0$ such that if $x \in S$ with $|x - c| < \delta$ then $|f(x) - f(c)| < \varepsilon$.
3. Proving each direction of the statement:
   - ($\Longrightarrow$), let $\{x_n\}$ in S with $x_n \to c$. Take $\varepsilon > 0$, since f is continuous at c, $\exists \delta > 0$ such that if $|x - c| < \delta$ with $x \in S$ then $|f(x_n) - f(c)| < \varepsilon$. Since $x_n \to c, \exists N \in \mathbb{N}$ such that $|x_n - c| < \delta, \forall n \geq N$. So, $|x_n - c| < \delta \longrightarrow |f(x_n) - f(c)| < \varepsilon$.
   - ($\Longleftarrow$) For contradiction, assume f is not continuous at c, then $\exists \varepsilon_0 > 0$ such that $\forall \delta > 0, \exists x \in S$ such $|x - c| < \delta$ and $|f(x) - f(c)| \geq \varepsilon_0$. Thus, $\forall n \in N, \exists x_n \in S$ such that $|x_n - c| < 1/n$ and $|f(x_n) - f(c)| \geq \varepsilon_0$. Thus, by the squeeze theorem $|x_n - c| \to 0$ and $x_n \to c$ which implies $f(x_n) \to x$ which is a contradiction.

$\square$

Definition 41 (Bounded function)**.** A function $f : S \to \mathbb{R}$ is bounded if $\exists M \geq 0$ such that $|f(x)| \leq M, \forall x \in S$.

Theorem 61. If $f : [a,b] \to \mathbb{R}$ is continuous at $[a,b]$ then it is bounded.

Proof. For contradiction, assume f is continuous but not bounded. Then, $\forall n \in \mathbb{N}, \exists x_n \in [a,b]$ such that $|f(x)| \geq n$. By the Bolzano-Weierstrass theorem, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ with $x \in \mathbb{R}$ and $x_{n_k} \to x$. Since, $x_{n_k} \in [a,b], \forall k \in \mathbb{N}$ then $x \in [a,b]$. Given f is continuous

$$f(x) = \lim_{k\to\infty} f(x_{n_k}) \implies |f(x)| = \lim_{k\to\infty} |f(x_{n_k})|$$

Hence, $\{|f(x_{n_k})\}$ is bounded and so is $\{n_k\}$ since $n_k \leq |f(x_{n_k})|$, from the definition of a subsequence $k \leq x_k, \forall k \in N$, contradicting the initial claim. ☐

Theorem 62 (Min-max theorem or Extreme value theorem)**.** Consider $f : [a,b] \to \mathbb{R}$. If f is continuous on $[a,b]$ then it achieves an absolute maximum and absolute minimum on $[a,b]$.

Proof. For the absolute maximum, if f is continuous then f is bounded. Thus, $E = \{f(x) : x \in [a,b]\}$ is bounded. Let $L = \sup E$ then,

- L is an upper bound for E
- There exists a sequence $\{f(x_n)\}$ with $x_n \in [a,b]$ such that $f(x_n) \to L$

By the Bolzano-Weierstrass theorem, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ and $d \in [a,b]$ such that $x_n \to d$ as $k \to \infty$. Hence,

$$f(d) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = L$$

by the continuity of f on $[a,b]$. So, f achieves an absolute maximum at d. The proof for the absolute minimum follows similarly. ☐

Theorem 63. Consider $f : [a,b] \to \mathbb{R}$. If $f(a) < 0$ and $f(b) > 0$, then $\exists c \in (a,b)$ such that $f(c) = 0$

Proof. Let $a_1 = a$ and $b_1 = b$. Define a_n, b_n as follows:

- If $f((a_{n-1} + b_{n-1})/2) \geq 0$ then $a_n = a_{n-1}$ and $b_n = (a_{n-1} + b_{n-1})/2$,
- If $f((a_{n-1} + b_{n-1})/2 < 0$ then $a_n = (a_{n-1} + b_{n-1})/2$ and $b_n = b_{n-1}$.

In this way, we obtain:

1. $a \leq a_n \leq a_{n+1} \leq b_{n+1} \leq b_n \leq b, \forall n \in N$,
2. $b_{n+1} - a_{n+1} = (b_n - a_n)/2, \forall n \in \mathbb{N}$,
3. $f(a_n) \leq 0$ and $f(b_n) \geq 0, \forall n \in N$.

From 1., $\{a_n\}$ and $\{b_n\}$ are bounded and monotone increasing and decreasing respectfully. Thus, $\exists c, d \in [a,b]$ such that $a_n \to c$ and $b_n \to d$. By 2.,

$$b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2} = \frac{1}{4}(b_{n-2} - a_{n-2}) = \ldots = \frac{1}{2^{n-1}}(b-a)$$

And,

$$d - c = \lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty}\frac{1}{2^{n-1}}(b-a) = 0 \Longrightarrow d = c$$

So, $a_n \to c$ and $b_n \to c$. By 3., $f(c) = \lim_{n\to\infty} f(a_n) \leq 0$ and $f(c) = \lim_{n\to\infty} f(b_n) \geq 0$. Therefore, $f(c) = 0$. $\square$

Theorem 64 (Bolzano intermediate value theorem)**.** Consider $f : [a,b] \to \mathbb{R}$ continuous. If $f(a) < f(b)$ with $y \in (f(a), f(b)), \exists c \in (a,b)$ such that $f(c) = y$. Else, if $f(b) < f(a)$ with $y \in (f(b), f(a))$ then $\exists c \in (a,b)$ such that $f(c) = y$.

Proof. Suppose $f(a) < f(b)$ with $y \in (f(a), f(b))$. Define $g(x) = f(x) - y$. Then, $f(x) = c \Longleftrightarrow g(x) = 0$ and $g : [a,b] \to \mathbb{R}$ is continuous, more importantly $g(a) = f(a) - y < 0$ and $g(b) = f(b) - y > 0$, then by the previous theorem $\exists c \in (a,b)$ such that $g(c) = y$ which is equivalent to $f(c) = y$. The proof for $f(b) < f(a)$ follows similarly. $\square$

Theorem 65. Consider $f : [a,b] \to \mathbb{R}$ to be continuous. Take $c \in [a,b]$ to be where f achieves a minimum value in $[a,b]$ and $d \in [a,b]$ to be where f achieves a maximum value in $[a,b]$. Then, $f([a,b]) = [f(c), f(d)]$. Putting it in words: every value between the maximum and minimum is achieved.

Proof. It is clear that $f([a,b]) \subseteq [f(c), f(d)]$. By the intermediate value theorem applied to $f : [c,d] \to \mathbb{R}$, we obtain $[f(c), f(d)] \subseteq f([c,d]) \subset f([a,b])$. Therefore, $f([a,b]) = [f(c), f(d)]$. $\square$

4.1.3 Uniform continuity

Definition 42 (Uniform continuity)**.** Consider $f : S \to \mathbb{R}$. Then f is continuous on S if $\forall \varepsilon > 0, \exists \delta = \delta(\varepsilon) > 0$ such that $|x - c| < \delta$ implies $|f(x) - f(c)| < \varepsilon, \forall x \in S$.

Theorem 66. Consider $f : [a,b] \to \mathbb{R}$, then f is continuous if, and only if, f is uniformly continuous.

Proof. Proving each direction of the statement:

- ($\Longrightarrow$): Suppose f is continuous and assume for contradiction that f is not uniformly continuous. Then, $\exists \varepsilon_0 > 0$ such that $\forall n \in \mathbb{N}, \exists x_n, c_n \in [a,b]$ such that $|x_n - c| < 1/n$ and $|f(x_n) - f(c_n)| \geq \varepsilon_0$.
  By the Bolzano-Weierstrass theorem, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ and $x \in$

$[a, b]$ such that $x_{n_k} \to x$. Similarly, there also exists a subsequence $\{c_{n_k}\}$ of $\{c_n\}$ and $c \in [a, b]$ such that $c_{n_k} \to c$. And also, the subsequence $\{x_{n_{k_k}}\}$ of $\{x_{n_k}\}$ satisfies $x_{n_{k_j}} \to x$. Then, $|x - c| = \lim_{j\to\infty} |x_{n_{k_j}} - c_{n_{k_j}}| \leq \lim_{j\to\infty} 1/n_{k_j} - 0$.
Thus, $x = c$. But since f is continuous at c, $0 = |f(c) - f(c)| = \lim_{j\to\infty} |f(x_{n_{k_j}}) - f(c_{n_{k_j}})| \geq \varepsilon_0$ which is a contradiction to the initial claim.

□

4.2 Differentiation

4.2.1 Definition and properties

Definition 43 (Derivative)**.** Let I be an interval with $f : I \to \mathbb{R}$ and $c \in I$. Then, f is differentiable at c if the limit

$$\lim_{x\to c} \frac{f(x) - f(c)}{x - c} \tag{4.1}$$

exists, in this case we write:

$$f'(c) = \lim_{x\to c} \frac{f(x) - f(c)}{x - c} \tag{4.2}$$

Furthermore, if f is differentiable $\forall c \in I$ then we write the derivative as f', or $f'(x)$ or $\frac{\mathrm{d}f}{\mathrm{d}x}$.

Example For all $n \in \mathbb{N}$, the derivative of the power function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \alpha x^n$ is given by $f'(c) = \alpha n c^{n-1}, \forall c \in \mathbb{R}$.

Proof. First, note that $\forall n \in \mathbb{N}$:

$$(x - c) \sum_{j=0}^{n-1} x^{n-1-j} c^j = \sum_{j=0}^{n-1} x^{n-j} c^j - \sum_{j=0}^{n-1} x^{n-1-j} c^{j+1}$$

Defining $l = j + 1$:

$$\begin{aligned}
(x - c) \sum_{j=0}^{n-1} x^{n-1-j} c^j &= \sum_{l=1}^{n} x^{n-j} c^j - \sum_{j=0}^{n-1} x^{n-l} c^l \\
&= x^{n-0} c^0 - x^{n-n} c^n \\
&= x^n - c^n
\end{aligned}$$

Therefore,

$$f'(c) = \lim_{x\to c} \frac{\alpha x^n - \alpha c^n}{x - c} = \alpha \lim_{x\to c} \sum_{j=0}^{n-1} x^{n-1-j} c^j = \alpha \sum_{j=0}^{n-1} c^{n-1-j} c^j = \alpha n c^{n-1}$$

□

Theorem 67. If the function $f : I \to \mathbb{R}$ is differentiable at $c \in I$, then it is also continuous at c.

Proof. Since every point of I is also a cluster point, then f is continuous at $c \in I$ if, and only if, $\lim_{x\to c} f(x) = f(c)$. Now,

$$\begin{aligned}\lim_{x\to c} f(x) &= \lim_{x\to c}(f(x) - f(c) + f(c)) \\ &= \lim_{x\to c}\left((x-c)\frac{f(x)-f(c)}{x-c} + f(c)\right) \\ &= 0 \cdot f'(c) + f(c) = f(c)\end{aligned}$$

□

4.2.2 Weierstrass' function

Continuity seems to be a prerequisite for the differentiation of a function. However, we may be tempted to take it as a sufficient condition, which unfortunately is not the case. This behaviour can be seen by an example.

Example Consider $f(x) = |x|$. Then, f is not differentiable at 0, even though it is continuous at 0.

Proof. Consider a sequence $\{x_n\}$ such that $x_n \to 0$ and

$$\lim_{n\to\infty} \frac{f(x_n) - f(0)}{x_n - 0}$$

does not exist. Let $x_n = (-1)^n/n$. Then, $x_n \to 0$ and

$$\frac{f(x_n) - f(0)}{x_n - 0} = \frac{|(-1)^n/n|}{(-1)^n/n} = (-1)^n$$

Hence, $\lim_{n\to\infty}(-1)^n$ does not exist. □

It is clear that a function may be continuous, and yet non-differentiable at some point. However, if $f : \mathbb{R} \to \mathbb{R}$ is continuous at $\mathbb{R}$, is there a point $c \in \mathbb{R}$ such that f is differentiable at c. The answer is no. There exists a function everywhere continuous and nowhere differentiable, called the Weierstrass' function.

In order to prove such function exists and fulfils the description above it is necessary to gather some tools, the following theorems are presented with this goal.

Theorem 68. For the cosine function, it is true that:

1. $\forall x, y \in \mathbb{R}, |\cos x - \cos y| \leq |x - y|$
2. For $c \in \mathbb{R}$ and $k \in \mathbb{N}$, $\exists y \in (c + \pi/k, c + 3\pi/k)$ such that $|\cos(kc) - \cos(ky)| \leq 1$.

Theorem 69. For $a, b, c \in \mathbb{R}$, $|a+b+c| \leq |a| - |b| - |c|$.

Proof. This follows from the triangle inequality:

$$|a| = |a+b+c+(-b)+(-c)| \leq |a+b+c| + |b+c| \leq |a+b+c| + |b| + |c|$$

□

Theorem 70. Consider the function:

$$f(x) = \sum_{k=0}^{\infty} \frac{\cos(160^k x)}{4^k} \tag{4.3}$$

Then,

1. $\forall x \in \mathbb{R}, f(x)$ is absolutely convergent,
2. $f(x)$ is bounded and continuous.

Proof. Proving each statement:

1. First, note that
$$\left| \frac{\cos(160^k x)}{4^k} \right| \leq 4^{-k}, \forall k \in \mathbb{N}$$
Hence, the the comparison test,
$$\sum_{k=0}^{\infty} \left| \frac{\cos(160^k x)}{4^k} \right|$$
converges.

2. We begin by noticing:
$$|f(x)| \leq \sum_{k=0}^{\infty} \frac{|\cos(160^k)|}{4^k} \leq \sum_{k=0}^{\infty} \frac{1}{4^k} = \frac{4}{3}$$
Hence, f is bounded.
Next, suppose $c \in \mathbb{R}$ and $x_n \to c$. Note that $\{|f(x_n) - f(c)|\}$ is bounded. Thus,
$$\lim_{n\to\infty} |f(x_n) - f(c)| = 0 \iff \limsup_{n\to\infty} |f(x_n) - f(c)| = 0$$
It is necessary to show $\limsup_{n\to\infty} |f(x_n) - f(c)| \leq \varepsilon, \forall \varepsilon > 0$. Choose $N_0 \in \mathbb{N}$ such that

$\sum_{k=N_0+1}^{\infty} 4^{-k} < \varepsilon/2$. Then,

$$\begin{aligned}
&\limsup_{n\to\infty} |f(x_n) - f(c)| \\
&= \limsup_{n\to\infty} \left| \sum_{k=0}^{N_0} \frac{\cos(160^k x_n)}{4^k} - \frac{\cos(160^k c)}{4^k} + \sum_{k=N_0+1}^{\infty} \frac{\cos(160^k x_n)}{4^k} - \frac{\cos(160^k c)}{4^k} \right| \\
&\leq \limsup_{n\to\infty} \sum_{k=0}^{N_0} 4^{-k} |\cos(160^k x) - \cos(160^k c) + \sum_{k=N_0+1}^{\infty} 4^{-k} |\cos(160^k x) - \cos(160^k c) \\
&\leq \limsup_{n\to\infty} \left(\sum_{k=0}^{N_0} 4^{-k} \right) |x_n - c| + \varepsilon = \varepsilon
\end{aligned}$$

□

Theorem 71 (Weierstrass function)**.** The function:

$$f(x) = \sum_{k=0}^{\infty} \frac{\cos(160^k x)}{4^k} \tag{4.4}$$

is nowhere differentiable.

Proof. Consider $c \in \mathbb{R}$, our goal is to find a sequence $\{x_n\}$ with $x_n \to c$ such that

$$\left\{ \frac{f(x_n) - f(c)}{x_n - c} \right\}$$

is unbounded. From one of the previous theorem, $\forall n \in \mathbb{N}, \exists x_n$ such that $\pi/160^n < x_n - c < 3\pi/160^n$ and $|\cos(160^n c) - \cos(160^n x_n)| \geq 1$. So, $x_n \neq 0, \forall n \in \mathbb{N}$ and $|x_n - c| \leq 3\pi/160^n \to 0$. Define:

$$f_k(x) = \frac{\cos(160^k x)}{4^k}$$

So, $f(x) = \sum f_k(x)$. Thus, define:

$$\begin{aligned}
f(c) - f(x_n) &= f_n(c) - f_n(x_n) + \sum_{k=0}^{n-1} (f_k(c) - f_k(x_n)) + \sum_{k=n}^{\infty} (f_k(c) - f_k(x_n)) \\
&= a_n + b_n + c_n
\end{aligned}$$

Then, $|a_n| = 4^{-n}|\cos(160^k x_n) - \cos(160^k c)| \geq 4^{-n}$. And,

$$\begin{aligned}
|b_n| &\leq \sum_{k=0}^{n-1} 4^{-k}|\cos(160^k c) - \cos(160^k x_n)|) \\
&\leq \sum_{k=0}^{n-1} 4^{-k}160^k|x_n - c| \\
&\leq \frac{3\pi}{160^n}\sum_{k=0}^{n-1} 40^k \\
&= \frac{3\pi}{160^n}\frac{40^n - 1}{39} \leq \frac{4^{-n+1}}{13}
\end{aligned}$$

And,

$$\begin{aligned}
|c_n| &\leq \sum_{k=n+1}^{\infty} 4^{-k}(|\cos(160^k c)| + |\cos(160^k x_n)|) \\
&\leq 2\sum_{k=n+1}^{\infty} 4^{-k} \\
&= 2 \cdot 4^{-n+1}\frac{4}{3} = 4^{-n}\frac{2}{3}
\end{aligned}$$

Combining the former inequalities, we obtain:

$$|f(c) - f(x_n)| \geq 4^{-n}\left(1 - \frac{4}{13} - \frac{2}{3}\right) = 4^{-n}\frac{1}{39}$$

Therefore,

$$\frac{|f(c) - f(x_n)|}{|c - x_n|} \geq \frac{160^n}{3\pi}4^{-n}\frac{1}{39} = \frac{40^n}{117\pi}$$

Thus, the sequence is unbound and therefore does not converge for any $x \in \mathbb{R}$ and the derivative does not exist. □

4.2.3 Differentiation rules and theorems

Theorem 72 (Chain rule)**.** Consider $f : A \to B$ and $g : B \to \mathbb{R}$, with f differentiable at $c \in A$ and g differentiable at $f(c) \in B$. Then, $(g \circ f)'(c) = g'(f(c))f'(c)$.

Proof. Let $h(x) = (g \circ f)(x)$ and $d = f(c)$. Define:

$$u(y) = \begin{cases} \frac{g(y)-g(d)}{y-d} & \text{if } y \neq d \\ g'(d) & \text{if } y = d \end{cases}$$

and,

$$v(x) = \begin{cases} \frac{f(x)-f(c)}{x-c} & \text{if } x \neq c \\ f'(c) & \text{if } x = c \end{cases}$$

Then,

$$\lim_{y \to d} u(y) = \lim_{y \to d} \frac{g(y) - g(d)}{y - d} = g'(d) = u(d)$$

and,

$$\lim_{x \to c} v(x) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = d'(c) = v(c)$$

Which shows $u(y)$ and $v(x)$ are continuous. Now, $g(y) - g(d) = u(y)(y - d)$ and $f(x) - f(c) = v(x)(x - c)$. Then, $h(x) - h(c) = g(f(x)) - g(f(c)) = g(f(x)) - g(d) = u(f(x))(f(x) - f(c)) = u(f(x))v(x)(x - c)$. So,

$$\begin{aligned} \lim_{x \to c} \frac{h(x) - h(c)}{x - c} &= \lim_{x \to c} u(f(x))v(x) \\ &= u(f(c))v(c) \\ &= f'(g(c))g'(c) \end{aligned}$$

□

Theorem 73. Consider $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$, both differentiable at $c \in I$. Then,

1. $(\alpha f)'(c) = \alpha f'(c), \forall \alpha \in \mathbb{R}$,
2. $(f + g)'(c) = f'(c) + g'(c)$,
3. $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$,
4. $(f/g)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{[g(c)]^2}$, provided $g(c) \neq 0$.

Proof. Proving each statement:

1. This result follows directly from the definition:

   $$(\alpha f)'(c) = \lim_{x \to c} \frac{(\alpha f)(x) - (\alpha f)(c)}{x - c} = \alpha \frac{f(x) - f(c)}{x - c} = \alpha f'(c)$$

   Considering the algebraic properties of limits.

2. Again, from the definition:

   $$\begin{aligned} (f + g)'(c) &= \lim_{x \to c} \frac{(f + g)(x) - (f + g)(c)}{x - c} = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} + \lim_{x \to c} \frac{g(x) - g(c)}{x - c} \\ &= f'(c) + g'(c) \end{aligned}$$

3. First, note:

$$\frac{f(x)g(x)-f(c)g(c)}{x-c} = \frac{f(x)-f(c)}{x-c}g(x) + f(c)\frac{g(x)-g(c)}{x-c}$$

Then, taking $\lim_{x\to c}$, we obtain: $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$.

4. Consider $h(x) = 1/g(x)$. By the chain rule $h'(c) = -g'(c)/[g(c)]^2$. Then,

$$\begin{aligned}(f/g)'(c) = (fh)'(c) = f'(c)h(c) + f(c)h'(c) &= [f'(c)/g(c) - f(c)[g(c)]^2/g'(c)] \\ &= [f'(c)g(c) - f(c)g'(c)]/[g(c)]^2\end{aligned}$$

□

Theorem 74. Consider $f : [a,b] \to \mathbb{R}$ to be a bijective function, and define f^{-1}, the inverse function such that $f^{-1}(y) = x$. If $f(x)$ is differentiable on $[a,b]$ and $f(x) \neq 0, \forall x \in [a,b]$ then:

$$(f^{-1})'(y) = \frac{1}{f'(x)} \text{ where } y = f(x)$$

Proof. Since $f^{-1}(y) = x$ then $y = f(x)$. Taking the derivative with respect to y on both sides:

$$1 = f'(x)\frac{\mathrm{d}x}{\mathrm{d}y}$$

Thus,

$$\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{f'(x)}$$

But $x = f^{-1}(y)$, so:

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

□

4.2.4 Mean value theorem

Definition 44 (Relative maximum/minimum)**.** Consider $S \subseteq \mathbb{R}$ and $f : S \to \mathbb{R}$. Then, f has a relative maximum at $c \in S$ if $\exists \delta > 0$ such that $\forall x \in S : |x - c| < \delta$ then $f(x) \leq f(c)$. The definition of minimum follows analogously.

Theorem 75. If $f : [a,b] \to \mathbb{R}$, f has a relative min or max at $c \in (a,b)$ and f is differentiable at c, then: $f'(c) = 0$.

Proof. If f has a relative maximum at $c \in (a,b)$, then $\exists \delta > 0$ such that $f(c) \leq f(x), \forall xin(c-$

$\delta, c+\delta)$, with $x \in [a,b]$. Let

$$x_n = c - \frac{\delta}{2n} \in (c-\delta, c)$$

Then, $x_n \to c$, so:

$$f'(c) = \lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c} \geq 0$$

Now, define:

$$y_n = c + \frac{\delta}{2n} \in (c, c+\delta)$$

Then, $y_n \to c$, and

$$f'(c) = \lim_{n\to\infty} \frac{(f(y_n) - f(c))}{y_n - c} \leq 0$$

Therefore, $f'(c) = 0$. $\square$

Theorem 76 (Rolle's theorem)**.** Consider $f : [a,b] \to \mathbb{R}$ and f differentiable in (a,b), additionally if $f(a) = f(b)$, then $\exists c \in (a,b)$ such that $f'(c) = 0$.

Proof. Let $Y = f(a) = f(b)$. Since f is continuous $\exists c_1, c_2 \in [a,b]$ to be a relative maximum and a relative minimum, respectfully. Then if $f(c_1) > Y$ then $c_1 \in (a,b)$ and $f'(c_1) = 0$. Similarly, if $f(c_2) < Y$ then $c_2 \in (a,b)$ and $f'(c_2) = 0$. If $f(c_1) \leq Y \leq f(c_2)$ then $f(x) = Y, \forall x \in [a,b]$, so $f'(c) - 0$ for any $c \in (a,b)$. $\square$

Theorem 77 (Mean value theorem)**.** Consider $f : [a,b] \to \mathbb{R}$ to be continuous and differentiable in (a,b). Then, $\exists c \in (a,b)$ such that $f(b) - f(a) = f'(c)(b-a)$.

Proof. Define

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

Then, $g(a) = g(b) = 0$. Thus, by the Rolle's theorem, $\exists c \in (a,b)$ such that $g'(c) = 0$. Hence,

$$g'(c) = 0 = f'(c) - \frac{f(b) - f(a)}{b-a}$$

$\square$

Corollary. If $f : [a,b] \to \mathbb{R}$ is differentiable on $[a,b]$ and $f'(x) = 0, \forall x \in (a,b)$ then $f(x) = K$ for some constant $K \in \mathbb{R}$.

Proof. Take $x_1, x_2 \in [a,b]$ and assume $x_1 < x_2$. Applying the Mean value theorem on $[x_1, x_2]$:

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

for some $c \in [x_1, x_2]$. On the other hand $f'(x) = 0 \forall x \in [a, b]$, so $f'(c) = 0$ which implies $f(x_2) = f(x_1)$. Define $K = f(x_1) = f(x_2)$. Since x_1, x_2 are arbitrary, $f(x) = K, \forall x \in [a, b]$. □

Corollary. If $f : S_1 \to \mathbb{R}$ and $g : S_2 \to \mathbb{R}$ are differentiable on $[a, b] \subseteq S_1 \cup S_2$ and $f'(x) = g'(x), \forall x \in [a, b]$ then $f(x) = g(x) + K$ with $K \in \mathbb{R}$ constant.

Proof. Define $h(x) = f(x) - g(x)$. Then $h'(x) = f'(x) - g'(x) = 0 \forall x \in [a, b]$. So, from the previous theorem $h(x) = K$ which implies $f(x) = g(x) + K$. □

Theorem 78. If $f : I \to \mathbb{R}$ is differentiable and $f'(x) = 0, \forall x \in I$ then f is constant.

Proof. Let $a, b \in I$ with $a < b$. Then, f is continuous on $[a, b]$ and differentiable on (a, b). So, by the previous theorem, $\exists c \in (a, b)$ such that $f(b) - f(a) = (b - a)f'(c) = 0$. Hence, $f(b) = f(a), \forall a, b \in I$ such that $a < b$. □

Theorem 79. Consider $f : I \to \mathbb{R}$ differentiable, then:

1. f is increasing if, and only if, $f'(x) \geq 0, \forall x \in I$, and
2. f is decreasing if, and only if, $f'(x) \leq 0, \forall x \in I$.

Proof. Proving f is increasing if $f'(x) \geq 0$.

- ($\Longleftarrow$): Suppose $f'(x) \geq 0, \forall x \in I$. Then, let $a, b \in I$ with $a < b$. By the mean value theorem, $\exists c \in (a, b)$ such that $f(b) - f(a) = (b - a)f'(c) \geq 0 \Longrightarrow f(a) \leq f(b)$.

- ($\Longrightarrow$): Suppose f is increasing. Let $c \in$ and $\{x_n\}$ be a sequence in I such that $x_n \to c$, with $x_n < c, \forall n \in \mathbb{N}$. Then, $f(x_n) - f(c) \leq 0, \forall n \in N$, and by consequence:

  $$f'(c) = \lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c} \geq 0$$

  On the other hand, if $\{x_n\}$ is such that $x_n \to c$ and $x_n > c, \forall n \in \mathbb{N}$ then: $f(x_n) - f(c) \geq 0, \forall n \in \mathbb{N}$, and:

  $$f'(c) = \lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c} \geq 0$$

  In either case, $f'(c) \geq 0$.

The proof for the decreasing function follows from taking $-f$ increasing, which is equivalent to f decreasing. □

Theorem 80 (Generalised mean value theorem)**.** If $f : S_1 \to \mathbb{R}$ and $g : S_2 \to \mathbb{R}$ are continuous on $[a, b] \subseteq S_1 \cap S_2$ and differentiable on (a, b) then, there exists a point $c \in (a, b)$ such that:

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c)$$

If $g'(c) \neq 0, \forall c \in (a, b)$ then:

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

Proof. Consider $h(x) = [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x)$, then from the Mean value theorem $\exists c \in (a, b)$ such that:

$$h'(c) = \frac{h(b) - h(a)}{b - a}$$

From the definition of $h(x)$, substituting $h(a)$ and $h(b)$:

$$\begin{aligned} h'(c) &= \frac{[f(b) - f(a)]g(b) - [g(b) - g(a)]f(b) - [f(b) - f(a)]g(a) + [g(b) - g(a)]f(a)}{b - a} \\ &= \frac{f(b)g(b) - f(a)g(b) - g(b)f(b) + g(a)f(b) - f(b)g(a) + f(a)g(a) + g(b)f(a) - g(a)f(a)}{b - a} \\ &= \frac{0}{b - a} \\ &= 0 \end{aligned}$$

On the other hand,

$$h'(c) = [f(b) - f(a)]g'(c) - [g(b) - g(a)]f'(c)$$

So,

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c)$$

If $g'(c) \neq 0, \forall c \in (a, b)$ then:

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

□

Definition 45 (Higher-order derivatives)**.** Consider $f : I \to \mathbb{R}$. Then, f is n times differentiable on $J \subseteq I$ if f', f'', ..., $f^{(n)}$ exist at every point of J. The n-th derivative of f is denoted by $f^{(n)}$.

Theorem 81 (Second derivative test)**.** Suppose $f : (a, b) \to \mathbb{R}$ has two continuous derivatives. If $x_0 \in (a, b)$ is such that $f'(x_0) = 0$ then:

- If $f''(x_0) > 0$, then f has a relative minimum at x_0,
- If $f''(x_0) < 0$, then f has a relative maximum at x_0,
- If $f''(x_0) = 0$, then f is an inflection point.

Proof. Proving only the first result: if f'' is continuous at x_0 and $\lim_{c \to x_0} f''(c) = f''(x_0) > 0$. Then, $\exists \delta > 0$ such that $f''(c) > 0, \forall c \in (x_0 - \delta, x_0 + \delta)$. Take $x \in (x_0 - \delta, x_0 + \delta)$, then, by Taylor's theorem, $\exists c \in (x, x_0)$ and $c \in (x_0 - \delta, x_0 + \delta)$ by consequence, such that:

$$f(x) = f(x_0) + \frac{f''(c)}{2}(x - x_0)^2 \geq f(x_0)$$

With $f(x) > f(x_0)$ if $x \neq x_0$. □

4.2.5 L'Hospital's Rules

From the algebraic limit theorem,

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{\lim_{x \to c} f(x)}{\lim_{x \to c} g(x)}$$

provided $\lim_{x \to c} g(x) \neq 0$. It is not difficult to argue that if the numerator tends to any number different from zero, while the denominator tends to zero the quotient explodes to infinity (positive or negative). However, there may be cases where both the numerator and denominator tend to zero, or alternatively where both terms tend to infinity. L'Hospital's rules provide an important tool for dealing with these cases.

Theorem 82 (L'Hospital's rule: $0/0$ case)**.** Consider $f : S_1 \to \mathbb{R}$ and $g : S_2 \to \mathbb{R}$ continuous with both functions differentiable on $A \subseteq S_1 \cap S_2$ with the possible exception of $a \in A$. If $f(a) = g(a) = 0$ and $g'(x) \neq 0, \forall x \neq a$, then:

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)} = L$$

provided $\lim_{x \to a} f'(x)/g'(x)$ exists.

Proof. From the Generalized mean value theorem:

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

Taking the limit of $b \to a$ we have:

1. Since $c \in (a, b)$ then $a < c < b$. So, $a - b < c - b < 0$. $a - b \to 0$ as $b \to a$. From the Squeeze theorem, $c \to b$ as $b \to a$. So, for the left-hand side of the Generalized mean value theorem stated above:
$$\frac{f'(b)}{g'(b)}$$

2. For the right-hand side, since $f(a) = f(b) = 0$:

$$\frac{f(b)}{g(b)}$$

So,

$$\lim_{b\to a}\frac{f'(b)}{g'(b)} = \lim_{b\to a}\frac{f(b)}{g(b)}$$

□

Definition 46. Given a function $f : A \to \mathbb{R}$ and a point $a \in A$, then we denote $\lim_{x\to a} f(x) = +\infty$ if $\forall N > 0, \exists\delta > 0$ such that $f(x) > N, \forall x \in (a-\delta, a+\delta)$. Similarly, we can define $\lim_{x\to a} = -\infty$.

Theorem 83 (L'Hospital's rule: ∞/∞ case)**.** Consider $f : S_1 \to \mathbb{R}$ and $g : S_2 \to \mathbb{R}$ to be differentiable on $(a,b) \in S_1 \cap S_2$. If $g'(x) \neq 0, \forall x \in (a,b)$, $\lim_{x\to c} f(x) = \infty$ or $(-\infty)$ and $\lim_{x\to c} g(x) = \infty$ or $(-\infty)$, for $c \in (a,b)$, then:

$$\lim_{x\to c}\frac{f(x)}{g(x)} = \lim_{x\to c}\frac{f'(x)}{g'(x)} = L$$

provided $\lim_{x\to c} f'(x)/g'(x)$ exist.

Proof. Let $\varepsilon > 0$, since $\lim_{x\to c} f'(x)/g'(x) = L$, there exists δ_1 such that:

$$\left|\frac{f'(x)}{g'(x)} - L\right| < \frac{\varepsilon}{2}, \forall x \in (c-\delta_1, c+\delta_1)$$

Applying the Generalized mean theorem for $[x, c+\delta_1]$:

$$\frac{f(c+\delta_1) - f(x)}{g(c+\delta_1) - g(x)} = \frac{f'(d)}{g'(d)}$$

for $d \in (x, c+\delta_1)$. So,

$$L - \frac{\varepsilon}{2} < \frac{f(c+\delta_1) - f(x)}{g(c+\delta_1) - g(x)} < L + \frac{\varepsilon}{2}, \forall x \in (c, c+\delta_1)$$

In order to isolate $f(x)/g(x)$ in the previous equation it is necessary to multiply the inequality by $(g(c+\delta_1) - g(x))/g(x)$. However, first this last term must be show to be strictly positive. We begin by noticing,

$$\frac{g(c+\delta_1) - g(x)}{g(x)} = \frac{g(c+\delta_1)}{g(x)} - \frac{g(x)}{g(x)} > 0 \Longrightarrow \frac{g(c+\delta_1)}{g(x)} > 1$$

So, it is necessary x such that $g(c+\delta_1) > g(x)$. Define δ_2 such that $g(c+\delta_1) > g(x), \forall x \in (c, c+\delta_2)$. We obtain:

$$\left(L - \frac{\varepsilon}{2}\right)\left(\frac{g(c+\delta_1)}{g(x)} - 1\right) < \frac{f(c+\delta_1) - f(x)}{g(x)} < \left(L + \frac{\varepsilon}{2}\right)\left(\frac{g(c+\delta_1)}{g(x)} - 1\right)$$

which leads to:

$$L\frac{g(c+\delta_1)}{g(x)} - \frac{\varepsilon}{2}\frac{g(c+\delta_1)}{g(x)} - L + \frac{\varepsilon}{2} < \frac{f(c+\delta_1) - f(x)}{g(x)} < L\frac{g(c+\delta_1)}{g(x)} + \frac{\varepsilon}{2}\frac{g(c+\delta_1)}{g(x)} - L - \frac{\varepsilon}{2}$$

$$L + \frac{\varepsilon}{2} - \frac{Lg(c+\delta_1) + \frac{\varepsilon}{2} - f(c+\delta_1)}{g(x)} < \frac{f(x)}{g(x)} < L - \frac{\varepsilon}{2} - \frac{Lg(c+\delta_1) - \frac{\varepsilon}{2} - f(c+\delta_1)}{g(x)}$$

Define δ_3 such that:

$$\frac{Lg(c+\delta_1) + \frac{\varepsilon}{2} - f(c+\delta_1)}{g(x)} < \varepsilon$$

And

$$\frac{Lg(c+\delta_1) - \frac{\varepsilon}{2} - f(c+\delta_1)}{g(x)} < \varepsilon$$

for all $x \in (c, c+\delta_3)$. Then, for $\delta = \min\{\delta_1, \delta_2, \delta_3\}$,

$$\left|\frac{f(x)}{g(x)} - L\right| < \varepsilon$$

□

4.2.6 Taylor's theorem

Theorem 84 (Taylor's theorem)**.** Suppose $f : [a, b] \to \mathbb{R}$ is continuous and has n continuous derivatives on $[a, b]$ such that $f^{(n+1)}$ exists on (a, b). Given $x, x_0 \in [a, b]$, there exists $c \in (a, b)$ such that:

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1} \tag{4.5}$$

Denote the large sum as $P_n(x)$ and the remainder as $R_n(x)$.

Proof. Let $x, x_0 \in [a, b]$. If $x = x_0$ then any c satisfies the theorem. Suppose $x \neq x_0$. Define:

$$M_{x,x_0} = \frac{f(x) - P_n(x)}{(x - x_0)^{n+1}}$$

Then, $f(x) = P_n(x) + M_{x,x_0}(x - x_0)^{n+1}$. And $f^{(k)}(x_0) = P^{(k)}(x_0), \forall 0 \leq k \leq n$. Let $g(s) =$

$f(x) - P_n(s) - M_{x,x_0}(s-x_0)^{n+1}$, then:

$$\begin{aligned} g(x_0) &= f(x_0) - P_n(x_0) - M_{x,x_0}(x-x_0)^{n+1} = 0 \\ g'(x_0) &= f'(x_0) - P_n'(x_0) - M_{x,x_0}(n+1)(x-x_0)^{n} = 0 \\ g''(x_0) &= f''(x_0) - P_n''(x_0) - M_{x,x_0}(n+1)n(x-x_0)^{n-1} = 0 \\ &\dots \\ g^{(n)}(x_0) &= f^{(n)}(x_0) - P_n^{(n)}(x_0) - M_{x,x_0}(n+1)!(x-x_0) = 0 \end{aligned}$$

Since $g(x) = 0$ and $g(x_0) = 0$, by the Mean value theorem, there exists $x_1 \in (x_0, x)$ such that $g'(x_1) = 0$. Thus, $g'(x_0) = 0$ and $g'(x_1) = 0$. By consequence, there exists $x_2 \in (x_0, x_1)$ such that $g''(x_2) = 0$. Preceding similarly, we find $x_n \in (x_0, x_{n-1})$ such that $g^{(n)}(x_n) = 0$. Finally, $g^{(n)}(x_0) = 0$ and $g^{(n)}(x_n) = 0$ which implies $\exists c \in (x_0, x_n)$ such that $g^{(n+1)}(c) = 0$. So,

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}s^{n+1}} M_{x,x_0}(s-x_0)^{n+1} = M_{x,x_0}(n+1)!$$

Additionally, $P_n^{(n+1)}(c) = 0$ since $P_n(x)$ is a polynomial with degree n. Hence,

$$0 = g^{(n+1)}(c) = f^{(n+1)}(c) - M_{x,x_0}(n+1)! \implies M_{x,x_0} = \frac{f^{(n+1)}(c)}{(n+1)!}$$

Thus,

$$f(x) = P_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$$

□

4.2.7 Smoothness Classes

Definition 47 (Smoothness classes). Consider $f : S \to \mathbb{R}$ and k a non-negative integer. If the function f is k-th order differentiable with $f^{(k)}(x)$ continuous on S, then f is of class $\mathcal{C}^k(S)$. If f is infinitely differentiable with all derivatives continuous on S, then f is called smooth, and belongs to the class $\mathcal{C}^\infty(S)$.

Remark. A continuous function belongs to the class $\mathcal{C}^0$.

Each class $\mathcal{C}^r(S)$ can be understood as a "set of functions of class $\mathcal{C}^r(S)$". In this way,

$$\mathcal{C}^\infty(S) = \bigcap_{r\in\mathbb{N}} \mathcal{C}^r(S) \subseteq \cdots \subseteq \mathcal{C}^1(S) \subseteq \mathcal{C}^0(S)$$

Where each inclusion $\mathcal{C}^i(S) \subseteq \mathcal{C}^{i-1}(S)$ is proper.

Chapter 5

Integration

5.1 The Riemann Integral

The Riemann integral is the first rigorous theory of 'area', and it is the inverse of differentiation. However, it is not a complete theory of 'area', the problems it fails to solve are addressed Lebesque integration.

First, we will define and calculate integrals of continuous function on the interval $[a, b]$. In order to avoid writing it every time, we define:

Definition 48. Define: $C([a, b]) = \{f : [a, b] \to \mathbb{R} : f \text{ is continuous on } [a, b]\}$.

The Riemann integral will be defined over a partition of an interval, we define:

Definition 49 (Partition)**.** A partition $\underline{x}$ of $[a, b]$ is a set $\underline{x} = \{a = x_0 < x_1 < ... < x_n = b\}$. The norm of $\underline{x}$ is denoted by $||\underline{x}||$ and defined by $||\underline{x}|| := \max\{x_1 - x_0, x_2 - x_1, ..., x_n - x_{n-1}\}$.

Definition 50 (Tag)**.** If $\underline{x}$ is a partition, the tag of $\underline{x}$ is defined as a finite set $\underline{\xi} = \{\xi_1, \xi_2, ..., \xi_n\}$ such that $a = x_0 \leq \xi_1 \leq x_1 \leq \xi_2 \leq x_2 \leq ... \leq x_{n-1} \leq \xi_n \leq x_n = b$. The pair $(\underline{x}, \underline{\xi})$ is called tagged partition.

Definition 51 (Riemann sum)**.** Consider $f \in C([a, b])$. Then, the Riemann sum corresponding to the tagged partition $(\underline{x}, \underline{\xi})$ is the number:

$$S_f(\underline{x}, \underline{\xi}) := \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) \tag{5.1}$$

The Riemann sum can be understood as an approximation for the area under the graph of the function f. As $||\underline{x}|| \to 0$ we expect this number to converge to A, which we interpret as the area under the curve of f on the interval $[a, b]$.

Theorem 85 (Riemann Integral). Consider $f \in C([a,b])$. Then, there exists a unique number denoted $\int_a^b f(x)\mathrm{d}x \in \mathbb{R}$ with the property that all sequences of tagged partition $\{(\underline{x}^r, \underline{\xi}^r)\}$ such that $||\underline{x}^r|| \to 0$ then:

$$\lim_{r\to\infty} S_f(\underline{x}^r, \underline{\xi}^r) = \int_a^b f(x)\mathrm{d}x \tag{5.2}$$

Next, the derivation of this theorem will be shown. The uniqueness of the integral follows immediately from the uniqueness of the limit, and need not be proven here. In order to prove it, a few tools are presented ahead:

Definition 52 (Modulus of continuity). Consider $f \in C([a,b])$, and $\eta > 0$. Then, the modulus of continuity $\omega_f(\eta)$ is defined as: $\omega_f(\eta) = \sup\{|f(x) - f(y)| : |x-y| \leq \eta\}$.

Theorem 86 (Theorem I).

$$\lim_{\eta\to 0} \omega_f(\eta) = 0, \forall f \in C([a,b]) \tag{5.3}$$

Putting it into words: for any continuous function $f \in C([a,b])$, for all $\varepsilon > 0$, there exists $\delta > 0$ such that $\forall \eta < \delta, \omega_f(\eta) < \varepsilon$.

Proof. Consider $\varepsilon > 0$. Since $f \in C([a,b])$ then f is uniformly continuous on $[a,b]$. Thus, $\exists\delta > 0$ such that $|f(x) - f(y)| < \varepsilon/2, \forall |x-y| < \delta$. Take $\eta < \delta$, then if $|x-y| \leq \eta < \delta$ we have $|f(x) - f(y)| < \varepsilon/2$. So, $\varepsilon/2$ is an upper bound for $\{|f(x) - f(y)| : |x-y| \leq \eta\}$. Therefore, $\omega_f(\eta) \leq \varepsilon/2 < \varepsilon$. □

Theorem 87 (Theorem II). If $(\underline{x}, \underline{\xi})$ and $(\underline{x}', \underline{\xi}')$ are tagged partitions of $[a,b]$ such that $\underline{x} \subseteq \underline{x}'$, then if $f \in C([a,b])$ we have $|S_f(\underline{x}, \underline{\xi} - S_f(\underline{x}', \underline{\xi}') \leq \omega_f(||\underline{x}||)(b-a)$.

Definition 53 (Refinement). If $(\underline{x}, \underline{\xi})$ and $(\underline{x}', \underline{\xi}')$ are tagged partitions of $[a,b]$ such that $\underline{x} \subseteq \underline{x}'$, then $\underline{x}$ is a refinement of $\underline{x}'$.

Proof. For $k = 1, ..., n.$ let:

$$\begin{aligned} \underline{y}(k) &= \{x_{k-1} = x'_l, x'_{l+1}, ..., x'_m = x_k\} \\ \underline{\eta}(k) &= \{\xi'_{l+1}, \xi'_{l+2}, ..., \xi'_m\} \end{aligned}$$

Then,

$$\begin{aligned} |f(\xi_k)(x_k - x_{k-1}) - S_f(\underline{y}(k), \underline{\eta}(k))| &= \left| f(\xi_k) - \sum_{j=l+1}^{m} f(\xi'_j)(x'_j - x'_{j-1}) \right| \\ &= \left| \sum_{j=l+1}^{m} (f(\xi_k) - f(\xi'_j))(x'_j - x'_{j-1}) \right| \end{aligned}$$

Since $\sum_{j=1}^{m} x'_j - x'_{j-1} = x_m - x'_l = x_k - x_{k-1}$:

$$\begin{aligned}|f(\xi_k)(x_k - x_{k-1}) - S_f(\underline{y}(k), \underline{\eta}(k))| &\le \sum_{j=l+1}^{m} |f(\xi_k) - f(\xi'_j)|(x'_j - x'_{j-1}) \\ &\le \sum_{j=l+1}^{m} \omega_f(|x_k - x_{k-1}|)(x'_j - x'_{j-1}) \\ &\le \omega_f(||\underline{x}||)(x_k - x_{k-1})\end{aligned}$$

Therefore,

$$\begin{aligned}|S_f(\underline{x}, \underline{\xi}) - S_f(\underline{x}', \underline{\xi}')| &= \left|\sum_{k=1}^{m}(f(\xi_k)(x_k - x_{k-1}) - S_j(\underline{y}(k), \underline{\eta}(k)))\right| \\ &\le \sum_{k=1}^{m} |f(\xi_k)(x_k - x_{k-1}) - S_f(\underline{y}(k), \underline{\eta}(k))| \\ &\le \omega_f(||\underline{x}||) \sum_{k=1}^{n} x_k - x_{k-1} \\ &= \omega_f(||\underline{x}||)(b-a)\end{aligned}$$

□

Theorem 88 (Theorem III). If $(\underline{x}, \underline{\xi})$ and $(\underline{x}', \underline{\xi}')$ are two tagged partitions of $[a, b]$, and $f \in C([a, b])$, then $|S_f(\underline{x}, \underline{\xi}) - S_f(\underline{x}', \underline{\xi}')| \le (\omega_f(||\underline{x}||) + \omega_f(||\underline{x}'||))(b-a)$.

Proof. Let $\underline{x}'' = \underline{x} \cup \underline{x}'$ and $\underline{\xi}''$ a tag of $\underline{x}''$. Then, by Theorem II, $|S_f(\underline{x}, \underline{\xi}) - S_f(\underline{x}', \underline{\xi}')| \le |S_f(\underline{x}, \underline{\xi}) - S_f(\underline{x}'', \underline{\xi}'')| - |S_f(\underline{x}', \underline{\xi}') - S_f(\underline{x}'', \underline{\xi}'')| \le \omega_f(||\underline{x}||)(b-a) + \omega_f(||\underline{x}'||)(b-a)$ □

Finally, we can return to the derivation of the Riemann integral, as desired.

Proof. Let $\{\underline{y}(r), \underline{\zeta}(r)\}$ be a sequence of tagged partitions with $||\underline{y}(r)|| \to 0$ as $r \to \infty$. First, we prove $\{\underline{y}(r), \underline{\zeta}(r)\}$ is Cauchy: Let $\varepsilon > 0$. By theorem I, $\exists \delta > 0$ such that $\forall \eta < \delta$,

$$\omega_f(\eta) < \frac{\varepsilon}{2(b-a)}$$

Since $||\underline{y}(r)|| \to 0, \exists N_0 \in N$ such that $||\underline{y}(r)|| < \delta, \forall r \ge N_0$. For $N = N_0$ with $r, s \ge N$, $|S_f(\underline{y}(r), \underline{\zeta}(r)) - S_f(\underline{y}(s), \underline{\zeta}(r))| \le (\omega_f(||\underline{y}(r)||) + \omega(||\underline{y}(s)||))(b-a)$, by Theorem III. Hence,

$$|S_f(\underline{y}(r), \underline{\zeta}(r)) - S_f(\underline{y}(s), \underline{\zeta}(r))| < \left(\frac{\varepsilon}{2(b-a)} + \frac{\varepsilon}{2(b-a)}\right)(b-a) = \varepsilon$$

So, since the sequence is Cauhcy, define $L = \lim_{r\to\infty} S_f(\underline{y}(r), \underline{\zeta}(r))$ which exists.
Next, we prove that $\lim_{r\to\infty} S_f(\underline{y}(r), \underline{\xi}(r)) = L$ for any sequence of partitions $\{(\underline{x}(r), \underline{\xi}(r))\}$

with $||\underline{x}(r)|| \to 0$.
By the triangle inequatlity and Theorem III,

$$\begin{aligned}|S_f(\underline{x}(r), \underline{\xi}(r)) - L| &\leq |S_f(\underline{x}(r), \underline{\xi}(r)) - S_f(\underline{y}(r), \underline{\zeta}(r)| + |S_f(\underline{y}(r), \underline{\zeta}(r)) - L| \\ &\leq (\omega_f(||\underline{x}(r)||) + \omega_f(||\underline{y}(r)||))(b-a) + |S_f(\underline{y}(r), \underline{\zeta}(r)) - L| \\ &\to 0\end{aligned}$$

Thus, by the Squeeze theorem, $|S_f(\underline{x}(r), \underline{\xi}(r)) - L| \to 0$ as $r \to \infty$. □

5.2 Properties of the Riemann integral

Theorem 89. If $f, g \in C([a,b])$ and $\alpha \in \mathbb{R}$, then

1. Linearity
$$\int_a^b (\alpha f)\mathrm{d}x = \alpha \int_a^b f\mathrm{d}x \tag{5.4}$$
2. Additivity
$$\int_a^b (f+g)\mathrm{d}x = \int_a^b f\mathrm{d}x + \int_a^b g\mathrm{d}x \tag{5.5}$$

Proof. Let $\{\underline{x}(r), \underline{\xi}(r)\}$ be a sequence of tagged partitions such that $||\underline{x}(r)|| \to 0$. Then,

1. $$S_{\alpha f}(\underline{x}(r), \underline{\xi}(r)) = \alpha S_f(\underline{x}(r), \underline{\xi}(r))$$
   So,
   $$\begin{aligned}\int_a^b \alpha f\mathrm{d}x &= \lim_{r\to\infty} S_{\alpha f}(\underline{x}(r), \underline{\xi}(r)) \\ &= \lim_{r\to\infty} \alpha S_f(\underline{x}(r), \underline{\xi}(r)) \\ &= \alpha \int_a^b f\mathrm{d}x\end{aligned}$$
2. $$S_{f+g}(\underline{x}(r), \underline{\xi}(r)) = S_f(\underline{x}(r), \underline{\xi}(r)) + S_g(\underline{x}(r), \underline{\xi}(r))$$
   So,
   $$\begin{aligned}\int_a^b (f+g)\mathrm{d}x &= \lim_{r\to\infty} S_{f+g}(\underline{x}(r), \underline{\xi}(r)) \\ &= \lim_{r\to\infty} [S_f(\underline{x}(r), \underline{\xi}(r)) + S_g(\underline{x}(r), \underline{\xi}(r))] \\ &= \int_a^b f\mathrm{d}x + \int_a^b g\mathrm{d}x\end{aligned}$$

$\square$

Theorem 90. Consider $f \in C([a,b])$, and

$$\begin{aligned} m_f &= \inf\{f(x) : a \leq x \leq b\} \in \mathbb{R} \\ M_f &= \sup\{f(x) : a \leq x \leq b\} \in \mathbb{R} \end{aligned}$$

Then,

$$m_f(b-a) \leq \int_a^b f(x)\mathrm{d}x \leq M_f(b-a)$$

Proof. Consider $\{(\underline{x}(r), \underline{\xi}(r)\}$ a sequence of tagged partitions such that $||\underline{x}(r)|| \to 0$ as $r \to \infty$. Then,

$$S_f(\underline{x}(r), \underline{\xi}(r)) = \sum_{k=1}^{n} f(\xi_k(r))(x_k(r) - x_{k-1}(r)) \geq m_f \sum_{k=1}^{n}(x_k(r) - x_{k-1}(r)) = m_f(b-a)$$

And,

$$S_f(\underline{x}(r), \underline{\xi}(r)) = \sum_{k=1}^{n} f(\xi_k(r))(x_k(r) - x_{k-1}(r)) \leq M_f \sum_{k=1}^{n}(x_k(r) - x_{k-1}(r)) = M_f(b-a)$$

Therefore,

$$m_f(b-a) \leq S_f(\underline{x}(r), \underline{\xi}(r)) \leq M_f(b-a), \forall r$$

Taking $\lim_{r\to\infty}$:

$$m_f(b-a) \leq \int_a^b f(x)\mathrm{d}x \leq M_f(b-a)$$

$\square$

Theorem 91. Consider $f, g \in C([a,b])$ with $f(x) \leq g(x), \forall x \in [a,b]$, then:

$$\int_a^b f(x)\mathrm{d}x \leq \int_a^b f(x)\mathrm{d}x$$

Proof. Consider $\{\underline{x}(r), \underline{\xi}(r)\}$ a sequence of tagged partitions, such that $||\underline{x}(r)|| \to 0$ with

$r \to \infty$. Then,

$$\begin{aligned} S_f(\underline{x}(r), \underline{\xi}(r)) &= \sum_{j=1}^{n} f(\xi_j(r)(x_j(r) - x_{j-1}(r)) \\ &\leq \sum_{j=1}^{n} g(\xi_j(r))(x_j(r) - x_{j-1}(r)) \\ &= S_g(\underline{x}(r), \underline{\xi}(r)) \end{aligned}$$

Taking the limit as $r \to \infty$

$$\int_a^b f(x)\mathrm{d}x \leq \int_a^b g(x)\mathrm{d}x$$

□

Theorem 92. Consider $f \in C([a,b])$, then:

$$\left| \int_a^b f(x)\mathrm{d}x \right| \leq \int_a^b |f(x)|\mathrm{d}x$$

Proof. Notice $\pm f(x) \leq |f(x)|, \forall x \in [a,b]$, so:

$$\pm \int_a^b f(x)\mathrm{d}x \leq \int_a^b f(x)\mathrm{d}x \Longrightarrow -\int_a^b f(x)\mathrm{d}x \leq \int_a^b f(x)\mathrm{d}x \leq \int_a^b |f(x)|\mathrm{d}x$$

□

5.3 Fundamental theorem of calculus

Theorem 93 (Fundamental theorem of calculus)**.** Consider $f \in C([a,b])$.

1. If $F : [a,b] \to \mathbb{R}$ is differentiable and $F'(x) = f(x)$, then

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$$

2. The function

$$G(x) := \int_a^x f(t)\mathrm{d}t$$

is differentiable on $[a,b]$ and

$$\begin{cases} G'(x) = f(x) \\ G(a) = 0 \end{cases}$$

Proof. Proving each statement:

1. Consider $\{\underline{x}(r)\}$ a sequence of partitions with $||\underline{x}|| \to 0$ as $r \to \infty$. Then, by the Mean Value Theorem, $\exists \xi_j(r) \in [x_{j-1}(r), x_j(r)], \forall r, j$ such that $F(x_j(r)) - F(x_{j-1}(r)) = F'(\xi_j(r))(x_j(r) - x_{j-1}(r)) = f(\xi_j(r))(x_j(r) - x_{j-1}(r))$. Thus,

$$\begin{aligned}
\int_a^b f(x)\mathrm{d}x &= \lim_{r\to\infty} \sum_{j=1}^{n(r)} f(\xi_j(r))(x_j(r) - x_{j-1}(r)) \\
&= \lim_{r\to\infty} \sum_{j=1}^{n(r)} [F(x_j(r)) - F(x_{j-1}(r))] \\
&= \lim_{r\to\infty} F(b) - F(a) \\
&= F(b) - F(a)
\end{aligned}$$

2. Consider $c \in [a, b]$ and let $\varepsilon > 0$, since f is continuous at c, $\exists \delta > 0$ such that $|f(t) - f(c)| < \varepsilon/2, \forall t \in (c - \delta, c + \delta)$.
   Suppose $0 < x - c < \delta$, if $t \in [c, x]$, then $|t - c| \le |x - c| < \delta$. Thus,

$$\begin{aligned}
\left|\frac{1}{x-c}\int_c^x f(t)\mathrm{d}t - f(c)\right| &= \left|\frac{1}{x-c}\int_c^x f(t)\mathrm{d}t - \frac{1}{x-c}\int_c^x f(c)\mathrm{d}t\right| \\
&= \frac{1}{x-c}\left|\int_c^x (f(t) - f(c))\mathrm{d}t\right| \\
&\le \frac{1}{x-c}\int_c^x |f(t) - f(c)|\mathrm{d}t \\
&\le \frac{1}{x-c}\int_c^x \frac{\varepsilon}{2}\mathrm{d}t \\
&= \frac{1}{x-c}\frac{\varepsilon}{2}(x-c) \\
&= \frac{\varepsilon}{2}
\end{aligned}$$

   A similar result can be found for $0 < c - x < \delta$. Thus,

$$\left|\frac{1}{x-c}\left(\int_a^x f(t)\mathrm{d}t - \int_a^c f(t)\mathrm{d}t\right) - f(c)\right| \le \frac{\varepsilon}{2} < \varepsilon, \forall x \in (c - \delta, c + \delta)$$

   So,

$$G'(c) = \lim_{x\to c} \frac{1}{x-c}\left(\int_a^x f(t)\mathrm{d}t - \int_a^c f(t)\mathrm{d}t\right) = f(c), \forall c \in [a, b]$$

   Hence, $G'(x) = f(x)$.

$\square$

5.4 Integration methods

Theorem 94 (Integration by parts)**.** Suppose $f, g \in C([a,b])$ and $f', g' \in C([a,b])$. Then,

$$\int_a^b f'(x)g(x)\mathrm{d}x = (f(b)g(b) - f(a)g(a)) - \int_a^b f(x)g'(x)\mathrm{d}x$$

Proof. From the derivation properties we have: $(f(x)g(x))' = f'(x)g(x) - f(x)g'(x)$. Therefore, by the Fundamental theorem of calculus:

$$(f(b)g(b) - f(a)g(a)) = \int_a^b f'(x)g(x)\mathrm{d} + \int_a^b f(x)g'(x)\mathrm{d}x$$

□

Lemma 1 (Riemman-Lebesgue)**.** Consider $f \in C([-\pi,\pi])$, $f' \in C([a,b])$ and f even periodic with period 2π, *i.e.* $f(-\pi) = f(\pi)$. For $n \in \mathbb{N} \cup \{0\}$, define:

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\mathrm{d}x \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\mathrm{d}x \end{aligned}$$

Then,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$$

Definition 54 (Fourier coefficients)**.** The terms a_n and b_n in the previous lemma are called *Fourier coefficients* of f.

Returning to the proof of the Riemman-Lebesgue lemma:

Proof. Using integration by parts,

$$\begin{aligned} |b_n| &= \frac{1}{\pi}\left|\int_{-\pi}^{\pi} f(x)\cos(nx)\mathrm{d}x\right| \\ &= \frac{1}{\pi}\left|\int_{-\pi}^{\pi}\left(\frac{1}{n}\sin(nx)\right)' f(x)\mathrm{d}x\right| \\ &= \left|\frac{1}{n}(f(\pi)\sin(n\pi) - f(-\pi)\sin(n(-\pi))) - \frac{1}{n}\int_{-\pi}^{\pi}\sin(nx)f'(x)\mathrm{d}x\right| \end{aligned}$$

Notice that $\sin(n\pi) = \sin(n(-\pi)) = 0, \forall n \in \mathbb{N}$. Hence,

$$\begin{aligned} |b_n| &\leq \frac{1}{n}\int_{-\pi}^{\pi} |\sin(nx)||f'(x)|\mathrm{d}x \\ &\leq \frac{1}{n}\int_{-\pi}^{\pi} |f'| \to 0 \end{aligned}$$

So, $|b_n| \to 0$. A similar argument holds for a_n. □

Theorem 95 (Change of variables)**.** Consider $\phi : [a, b] \to [c, d]$ to be continuously differentiable with $\phi > 0$ on $[a, b], \phi(a) = c$, and $\phi(b) = d$. Then,

$$\int_c^d f(u)\mathrm{d}u = \int_a^b f(\phi(x))\phi'(x)\mathrm{d}x$$

Proof. Consider $F : [a, b] \to \mathbb{R}$ such that $F'(x) = f(x)$, then $F(\phi(x))' = f(\phi(x))$. Hence, by the Fundamental theorem of calculus,

$$\begin{aligned} \int_a^b f(\phi(x))\phi'(x)\mathrm{d}x &= \int_a^b F(\phi(x))'\mathrm{d}x \\ &= F(\phi(b)) - F(\phi(a)) \\ &= F(d) - F(c) \end{aligned}$$

Furthermore,

$$\int_c^d f(u)\mathrm{d}u = \int_c^d F(u)'\mathrm{d}u = F(d) - F(c)$$

□

Chapter 6

Sequences and series of functions

6.1 Pointwise and uniform convergence

Previously, we have defined and dealt with the convergence of a sequence of numbers. So, we must begin with an understanding of what it means to a sequence of functions to converge.

Definition 55 (Pointwise convergence). Define $f : S \to \mathbb{R}$ and $f_n : S \to \mathbb{R}$ with $n \in \mathbb{N}$. The sequence of functions (f_n) is said to *converge pointwise* to f if:

$$\lim_{n\to\infty} f_n(x) = f(x), \forall x \in S$$

Example.

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 \text{ if } x \in [0,1) \\ 1 \text{ if } x = 1 \end{cases}$$

Thus $\{f_n(x)\}$ converges pointwise to the function above, hence a sequence of continuous function may not converge to a continuous function.

Example. Consider $f_n(x) = \sum_{m=0}^{n} x^m$ for $x \in (-1,1)$. Then,

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \sum_{m=0}^{n} x_m = \frac{1}{1-x}$$

Definition 56 (Uniform convergence). For $n \in \mathbb{N}$, define $f_n : S \to \mathbb{R}$ and $f : S \to \mathbb{R}$. Then, f_n *converges uniformly* to f if $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon, \forall n \geq N, \forall x \in S$.

Theorem 96. Consider $f_n : S \to \mathbb{R}$ and $f : S \to \mathbb{R}$. Then, if f_n converges to f uniformly it also converges to f pointwise.

Proof. Consider $c \in S$ and let $\varepsilon > 0$. Then, if $f_n \to f$ uniformly, $\exists N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon, \forall n \geq N, \forall x \in S$. So, $\lim_{n\to\infty} f_n(c) = f(c), \forall c \in S$. Therefore, f_n converges to f pointwise. $\square$

Theorem 97 (Cauchy criterion for uniform convergence)**.** Consider $S \subseteq \mathbb{R}$, non-empty. Let $f_n : S \to \mathbb{R}$ be a sequence of functions. Then, f_n converges to f uniformly on S if $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that:

$$|f_n(x) - f_m(x)| < \varepsilon, \forall n, m \geq N, \forall x \in S$$

Proof. Suppose f_n converges uniformly to f on S, as $n \to \infty$. Then, $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that:

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2}, \forall n \geq N, \forall x \in S$$

On the other hand,

$$\begin{aligned}
|f_n(x) - f_m(x)| &= |f_n(x) - f(x) + f(x) - f_m(x)| \\
&\leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon
\end{aligned}$$

If $|f_n(x) - f_m(x)| < \varepsilon, \forall \varepsilon > 0, \forall x \in S$, then $(f_n(x))_n$ is Cauchy. So,

$$f(x) := \lim_{n\to\infty} f_n(x), \forall x \in S$$

from the Cauchy theorem of sequences. Additionally,

$$\lim_{m\to\infty} |f_n(x) - f_m(x)| = |f_n(x) - f(x)| < \frac{\varepsilon}{2} < \varepsilon, \forall n \geq N, \forall x \in S$$

So, from the definition, $f_n \to f$ uniformly. $\square$

Theorem 98 (Weierstrass M-test)**.** Define $f_n : S \to \mathbb{R}$ and suppose $\exists M_n > 0, \forall n \in \mathbb{N}$ such that:

- $|f_n(x)| < M_n, \forall x \in S$,
- $\sum_{n=1}^{\infty} M_n$ converges

Then,

1. $\sum_{n=1}^{\infty} f_n(x)$ converge absolutely for all $x \in S$,
2. Define $f(x) = \sum_{n=1}^{\infty} f_n(x), \forall x \in S$, then $\sum_{n=1}^{\infty} f_n$ converges uniformly to f on S.

Proof. Proving each statement:

1. The first item implies the sequence $\{|f_i(x)|\}_i$ is bounded for all $x \in S$ by M_i. So,

$$\sum_{i=1}^{n} |f_i(x)| \leq \sum_{i=1}^{n} M_i$$

   By the comparison test, the sequence $\{\sum_{i=1}^{n} |f_i(x)|\}_n$ converges absolutely, since $\sum_{i=1}^{\infty} M_i$ converges. Hence, $\sum_{i=1}^{\infty} f_i(x)$ converges absolutely for all $x \in S$.

2. Let $\varepsilon > 0$. Since $\sum M_i$ converges, then $\exists N_0 \in \mathbb{N}$ such that:

$$\sum_{i=n+1}^{\infty} M_i = \left| \sum_{i=1}^{\infty} M_i - \sum_{i=1}^{n} M_i \right| < \varepsilon, \forall n \geq N_0$$

   Take $N = N_0$. Then,

$$\begin{aligned} \left| f(x) - \sum_{i=1} n f_i(x) \right| &= \left| \sum_{i=n+1}^{\infty} f_i(x) \right| \\ &\leq \sum_{i=n+1}^{\infty} |f_i(x)| \\ &\leq \sum_{i=n+1}^{\infty} M_j \\ &< \varepsilon, \forall n \geq N, \forall x \in S \end{aligned}$$

   So, $\sum_{i=1}^{n} f_i(x) \to f(x), \forall x \in S$.

□

6.2 Properties of the convergent sequence of functions

One of the central problems related to sequences or series of functions focus on whether the limit (or derivative, or integral) can be calculated by the term-by-term operation over the series expansion. As the previous chapters have hopefully shown, derivatives and integrals are limit operations, as is the process of taking the sum of an infinite series. The multiple limit operation over an object can be problematic.

Remark. In general, interchanging limits leads to different results.

Example. For instance, consider:

$$\begin{aligned}\lim_{n\to\infty}\lim_{k\to\infty}\frac{n/k}{n/k+1} &= \lim_{n\to\infty}\frac{0}{0+1}=0\\ \lim_{k\to\infty}\lim_{n\to\infty}\frac{n/k}{n/k+1} &= \lim_{k\to\infty}1=1\end{aligned}$$

So, there are a few questions that could be made about interchanging limits applied to sequences of functions:

1. If $f_n : S \to \mathbb{R}$ is a sequence of continuous functions such that f_n converges to f pointwise or uniformly, then is f continuous?
2. If $f_n : [a,b] \to \mathbb{R}$ is a sequence of continuous differentiable functions such that f_n converges to f and $f_n' \to g$, then is f differentiable and does $g(x) = f'(x)$?
3. If $f_n : [a,b] \to \mathbb{R}$ with f_n and f continuous such that f_n converges to f, then does

$$\int_a^b f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x\ ?$$

The answer to the previous questions is yes, provided the convergence is uniform. Else, the answer is no if the convergence is pointwise. This last result can be proved by counterexamples:

Example. Finding counterexamples for each question:

1. Consider $f_n(x) = x^n$ on $[0,1]$. Then $f_n(x)$ is continuous for all n. As discussed earlier:

$$f_n(x) \to f(x) = \begin{cases} 0 \text{ if } x \in [0,1) \\ 1 \text{ if } x = 1 \end{cases}$$

   Notice $f(x)$ is not continuous, so the answer to the first of the three previous questions is negative.

2. Consider $f_n(x) = \frac{x^{n+1}}{n+1}$ on $[0,1]$. Then $f_n(x)$ converges to 0 pointwise on $[0,1]$. On the other hand,

$$f_n'(x) \to g(x) = \begin{cases} 0 \text{ if } x \in [0,1) \\ 1 \text{ if } x = 1 \end{cases}$$

   Thus, $g(x) \neq (0)' = 0$ at $x = 1$. So, the second question has a negative answer for pointwise convergence.

3. Consider

$$f_n(x) = \begin{cases} 4n^2x & \text{if } x \in \left[0, \frac{1}{2n}\right) \\ 4n - 4n^2x & \text{if } x \in \left[\frac{1}{2n}, \frac{1}{n}\right) \\ 0 & \text{if } x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

Then $f_n(x)$ converges to 0 pointwise on $[0,1]$. However,

$$\int_0^1 f_n(x)\mathrm{d}x = \frac{1}{2n}2n = 0 \not\to 0 = \int_0^1 0\mathrm{d}x$$

So, the answer to the third question is no for the case of pointwise convergence.

Next, we must show the answer to the previous questions is positive in the case of uniform convergence.

Theorem 99. If $f_n : S \to \mathbb{R}$ is continuous for all $n \in \mathbb{N}$. And $f : S \to \mathbb{R}$, such that f_n converges to f uniformly, then f is continuous.

Proof. Let $c \in S$ and $\varepsilon > 0$. Since f_n converges to f uniformly, $\exists N \in \mathbb{N}$ such that $|f_n(y) - f(y)| < \varepsilon/3, \forall n \geq N, \forall y \in S$. Since $f_N : S \to \mathbb{R}$ is continuous then $\exists \varepsilon_0 > 0$ such that $|f_N(x) - f_N(c)| < \varepsilon/3, \forall x \in (c - \delta_0, c + \delta_0)$. Choose $\delta = \delta_0$, then:

$$\begin{aligned} |f(x) - f(c)| &\leq |f(x) - f_N(x)| + |f_N(c) - f(c)| + |f_N(x) - f_N(c)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon \end{aligned}$$

□

Theorem 100. If $f_n : [a,b] \to \mathbb{R}$ is a sequence of continuous functions such that f_n converges to $f : [a,b] \to \mathbb{R}$ uniformly then

$$\int_a^b f_n(x)\mathrm{d}x \to \int_a^b f(x)\mathrm{d}x$$

Proof. Let $\varepsilon > 0$. Since $f_n \to f$ uniformly, then $\exists N \in \mathbb{N}$ such that:

$$|f_n(x) - f(x)| < \frac{\varepsilon}{b-a}, \forall n \geq N, \forall x \in [a,b]$$

Then,

$$\left| \int_a^b f_n(x)\mathrm{d}x - \int_a^b f(x)\mathrm{d}x \right| \leq \int_a^b |f_n(x) - f(x)|\mathrm{d}x < \int_a^b \frac{\varepsilon}{b-a} = \varepsilon$$

□

Remark. Notationally, this is equivalent to:

$$\lim_{n\to\infty} \int_a^b f_n(x)\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x$$

Theorem 101. If $f_n : [a,b] \to \mathbb{R}$ is a sequence of continuous differentiable functions, $f : [a,b] \to \mathbb{R}$, $g : [a,b] \to \mathbb{R}$, and:

1. f_n converges to f pointwise,
2. f_n' converges to g uniformly,

then f is continuously differentiable on $[a,b]$ and $g(x) = f'(x)$.

Proof. By the Fundamental theorem of Calculus,

$$f_n(x) - f(x) = \int_a^x f_n'(t)\mathrm{d}t, \forall n \in \mathbb{N}, \forall x \in [a,b]$$

Thus, by the previous two theorems,

$$\begin{aligned} f(x) - f(a) &= \lim_{n\to\infty} (f_n(x) - f_n(a)) \\ &= \lim_{n\to\infty} \int_a^x f_n'(t)\mathrm{d}t \\ &= \int_a^x g(t)\mathrm{d}t \end{aligned}$$

Therefore, $f(x) = f(a) + \int_a^x g(t)\mathrm{d}t$. Thus, by the Fundamental theorem of calculus, f is differentiable and $f'(x) = (\int_a^x g(t)\mathrm{d}t)' = g(x)$. □

6.3 Power series

Power series are a natural generalization of polynomials.

Definition 57 (Power series)**.** A power series about x_0 is series of the form:

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n$$

Remark. Under this setting, it is common to take $(x - x_0)^0 = 1$. Even though, 0^0 is indeterminate.

It is clear from the definition of power series that the series converges for $x = x_0$ to a_0. One question follows naturally from this observation: "for what values of x the power series converges?"

Theorem 102. Suppose $(|a_n|^{1/n})$ converges, *i.e.*:

$$R = \lim_{n\to\infty} |a_n|^{1/n}$$

and define p, the *radius of convergence*, as:

$$p = \begin{cases} \frac{1}{R} \text{ if } R > 0 \\ \infty \text{ if } R = 0 \end{cases}$$

Then:

1. If $|x - x_0| < p$, the series $\sum a_n(x - x_0)^n$ converges;
2. If $|x - x_0| > p$, the series diverges.

Proof. First, notice that:

$$\lim_{n\to\infty} |a_n(x - x_0)^n|^{1/n} = R|x - x_0|$$

So, the theorem is valid from the Root test. □

Remark. There exists power series which converge only at one point.

Suppose $\sum a_n(x - x_0)^n$ is a power series with radius of convergence p. Then, define $f : (x_0 - p, x_0 + p) \to \mathbb{R}$ such that:

$$f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

So, $f(x)$ is a limit of a sequence of functions $f_n(x)$, *i.e.*:

$$f(x) = \lim_{m\to\infty} f_m(x)$$

Where,

$$f_m(x) = \sum_{n=0}^{m} a_n(x - x_0)^n$$

Theorem 103. Let $\sum_{i=0}^{\infty} a_i(x - x_0)^i$ be a power series with radius of convergence $p \in (0, \infty]$. Then, $\forall r \in (0, p), \sum_{i=0}^{\infty} a_i(x - x_0)^i$ converges uniformly on $[x_0 - r, x_0 + r]$.

Proof. Let $r \in [0, p)$. Then, $\forall i \in \mathbb{N} \cup \{0\}, \forall x \in [x_0 - r, x_0 + r], |a_i(x - x_0)^i| \leq |a_i|r^i = M_i$. Now,

$$\lim_{i\to\infty} M_i^{1/i} = \lim_{i\to\infty} |a_i|^{1/i} r = \begin{cases} \frac{r}{p} \text{ if } p < \infty \\ 0 \text{ if } p = \infty \end{cases}$$

since $1/p = \lim_{i\to\infty} |a_i|^{1/i}$. Since $r < p$,

$$\lim_{i\to\infty} M_i^{1/i} < 1 \implies \sum_{i=1}^{\infty} M_i \text{ converges.}$$

By the Weierstrass M-test, $\sum_{i=0}^{\infty} a_i(x-x_0)^i$ converges uniformly on $[x_0-r, x_0+r]$. $\square$

Theorem 104. If a power series $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely at a point $c > 0$, then the series converges uniformly on a closed interval $[-c, c]$.

Proof. This is a straightforward implication of Weierstrass M-test. First, define $f_n : S \to \mathbb{R}$, with $f_n(x) = a_n x^n$. Then, for $x = c$, $f_n(c) = a_n c^n$, so $|a_n c^n|$ is finite, hence the sequence $(|f_n|)$ is bounded at $x = c$. We may write:

$$|f_n(c)| \leq M_n, \forall n \in \mathbb{N} \cup \{0\}$$

Since the series converges absolutely at $x = c$, then - by the comparison test - $\sum_{n=1}^{\infty} M_n$ also converges.
Furthermore, if $x \in (-c, c)$ then $|a_n x^n| \leq |a_n c^n|$. So, for $S = (-c, c)$

1. $$|f_n(x)| = |a_n x^n| < M_n, \forall x \in S$$
2. $$\sum_{n=1}^{\infty} M_n \text{ converges}$$

So, by the Weierstrass M-test:

1. $\sum_{n=0}^{\infty} f_n(x)$ converges absolutely $\forall x \in S$
2. Let $f(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} a_n x^n$, then the series converges uniformly on $S = (-c, c)$.

$\square$

It is important to notice, that if a series converges conditionally on $x = c$, then it may diverge on $x = -c$, *e.g.:*

$$\sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n}$$

converges on $x = 1$, but not on $x = -1$. More importantly, it converges uniformly on $[0, 1]$.

Lemma 2 (Abel's lemma)**.** Let (b_n) be a non-negative and monotone decreasing sequence. And consider $\sum_{n=1}^{\infty} a_n$ a series, whose partial sums are bounded, *i.e.*:

$$\left|\sum_{n=1}^{m} a_n\right| \leq A, \forall m \in N$$

Then,

$$\left|\sum_{n=1}^{m} a_n b_n\right| \leq Ab_1, \forall m \in \mathbb{N}$$

Proof. Let $s_m = \sum_{n=1}^{m} a_n$ and set $s_0 = 0$, then:

$$\begin{aligned}
\left|\sum_{i=1}^{n} a_i b_i\right| &= \left|\sum_{i=1}^{n} b_i(s_i - s_{i-1})\right| \\
&= \left|\sum_{i=1}^{n} b_i s_i - \sum_{j=0}^{n-1} b_{j+1} s_j\right| \\
&= \left|b_n s_n + \sum_{i=1}^{n-1} s_i(b_i - b_{i+1})\right| \\
&\leq \left|Ab_n + \sum_{i=1}^{n-1} A(b_i - b_{i+1})\right| \\
&= |Ab_n + A(b_1 - b_n)| = |Ab_1| = Ab_1
\end{aligned}$$

□

Theorem 105 (Abel's theorem)**.** Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series which converges on $x = c > 0$. Then, the series converges uniformly on the interval $[0, c]$.
Similarly, if it converges on $x = -c$, then it converges uniformly on $[-c, 0]$, with $c > 0$.

Proof. First, notice that:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} (a_n c^n) \left(\frac{x}{c}\right)^n$$

Let $\varepsilon > 0$, then by the Cauchy criterion, the series converges if there exists N such that:

$$\left|(a_{m+1} c^{m+1}) \left(\frac{x}{c}\right)^{m+1} + ... + (a_n c^n) \left(\frac{x}{c}\right)^n\right| < \varepsilon$$

with $n > m \geq N$. Because the series converges at $x = c$, then there exists N such that:

$$\left|a_{m+1} c^{m+1} + ... + a_n c^n\right| < \varepsilon$$

if $n > m \geq N$. First, take $\varepsilon/2$ as a bound for $\sum_{i=1}^{\infty} a_{m+i} c^{m+i}$. Second, notice $(x/c)^{m+j}$ is monotone decreasing on $x \in [0, c]$. Hence,

$$\begin{aligned}\left|(a_{m+1}c^{m+1})\left(\frac{x}{c}\right)^{m+1} + \ldots + (a_n c^n)\left(\frac{x}{c}\right)^n\right| &\leq \frac{\varepsilon}{2}\left(\frac{x}{c}\right)^{m+1} \\ &< \varepsilon\end{aligned}$$

□

Theorem 106. Consider $\sum_{i=0}^{\infty} a_i(x-x_0)^i$, a power series with radius of convergence $p \in (0, \infty]$. Then,

1. $\forall c \in (x_0 - p, x_0 + p), \sum_{i=0}^{\infty} a_i(x-x_0)^i$ is differentiable at c and:

$$\frac{\mathrm{d}}{\mathrm{d}x}\sum_{i=0}^{\infty} a_i(x-x_0)^i = \sum_{i=0}^{\infty} i a_i (x-x_0)^{i-1}$$

2. $\forall a, b$ such that $x_o - p < a < b < x_0 + p$, then:

$$\int_a^b \sum_{i=0}^{\infty} a_i(x-x_0)^i = \sum_{i=0}^{\infty}\left(\frac{(b-x_0)^{i+1}}{i+1} - \frac{(a-x_0)^{i+1}}{i+1}\right)$$

Remark. Since

$$\lim_{i\to\infty}((i+1)|a_{i+1}|)^{1/i} = \lim_{i\to\infty}\left((i+1)|a_{i+1}|^{1/(i+1)}\right)^{(i+1)/i} = \lim_{k\to\infty}|a_k|^{1/k} = p$$

the first result of the previous theorem implies $\sum a_i(x-x_0)^i$ is infinitely differentiable and:

$$k!a_k = \left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}\sum a_i(x-x_0)^i\right)\Bigg|_{x=x_0}$$

Theorem 107 (Weierstrass approximation theorem)**.** If $f \in C([a, b])$, there exists a sequence of polynomials $\{P_n\}$ such that $P_n \to f$ uniformly on $[a, b]$.

Theorem 108. Define $c_n := (\int_{-1}^{1}(1-x^2)^n \mathrm{d}x)^{-1} > 0$, and let $Q_n(x) = c_n(1-x^2)^n$. Then:

1. $\int_{-1}^{1} Q_n(x)\mathrm{d}x = 1, \forall n \in \mathbb{N}$,
2. $Q_n(x) \geq 0, \forall n \in \mathbb{N}, \forall x \in [-1, 1]$, and
3. $\forall \delta \in (0, 1), Q_n \to 0$ uniformly on $\delta \leq |x| \leq 1$.

Proof. Proving each statement:

1. $\int_{-1}^{1} Q_n(x)\mathrm{d}x = c_n \int_{-1}^{1}(1-x^2)^n \mathrm{d}x = 1$ by definition of c_n,

2. Since c_n is strictly positive, notice $(1-x^2) \geq 0, \forall x \in [-1,1]$. Any non-negative number to the n-th power will result in a non-negative number (since $n \in \mathbb{N}$). So $Q_n(x) \geq 0, \forall n \in \mathbb{N}, \forall x \in [-1,1]$,

3. First, notice:
$$(1-x^2)^n \geq 1 - nx^2, \forall n \in \mathbb{N}, \forall x \in [-1,1]$$
This can be proven by induction, beyond the scope of this particular proof. And $g(x) = (1-x^2)^n - (1-nx^2)$ satisfies $g(0) = 0$. Additionally, $g'(x) = n2x(1-(1-x^2)^{n-1}) \geq 0$ in $[0,1]$. Thus, $g(x) \geq 0$ by the Mean value theorem. So,
$$\begin{aligned}\frac{1}{c+n} &= \int_{-1}^{1}(1-x^2)^n \mathrm{d}x \\ &= 2\int_{0}^{1}(1-x^2)^n \mathrm{d}x \\ &> 2\int_{0}^{1/\sqrt{n}}(1-x^2)^n \mathrm{d}x \\ &\geq 2\int_{0}^{1/\sqrt{n}}(1-nx^2) \mathrm{d}x \\ &= 2\left(\frac{1}{\sqrt{n}} - \frac{n}{3}n^{-3/2}\right) \\ &= \frac{4}{3}\sqrt{n} && > \sqrt{n}\end{aligned}$$
Therefore $c_n < \sqrt{n}$. Let $\delta > 0$. First, notice $\lim_{n\to\infty}\sqrt{n}(1-\delta^2)^n = 0$. By the root test,
$$\begin{aligned}\lim_{n\to\infty}(\sqrt{n}(1-\delta^2)^n)^{1/n} &= \lim_{n\to\infty}(n^{1/n})^{1/2}(1-\delta^2) \\ &= 1-\delta^2 \\ &< 1\end{aligned}$$
So, $\lim_{n\to\infty}\sqrt{n}(1-\delta^2)^n = 0$. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $\sqrt{n}(1-\delta^2)^n < \varepsilon, \forall n \geq N$. Then, $\forall n \geq N, \forall \delta \leq |x| \leq 1, |c_n(1-x^2)^n| < sqrtn(1-x^2)^n \leq \sqrt{n}(1-\delta^2)^n < \varepsilon$.

$\square$

Finally, we return to proving Weierstrass approximation theorem.

Proof. Suppose $f \in C([0,1])$, with $f(0) = f(1) = 0$. f can be extended to $C(\mathbb{R})$ by setting

$f(x) = 0, \forall x \notin [0,1]$. Define:

$$\begin{aligned} P_n(x) &= \int_0^1 f(t)Q_n(t-x)\mathrm{d}t \\ &= \int_0^1 f(t)c_n(1-(t-x)^2)^n \mathrm{d}t \end{aligned}$$

Note that $P_n(x)$ is in fact a polynomial. Additionally, note that for $x \in [0,1]$,

$$\begin{aligned} P_n(x) &= \int_0^1 f(t)Q_n(t-x)\mathrm{d}t \\ &= \int_{-x}^{1-x} f(x+t)Q_n(t)\mathrm{d}t \\ &= \int_{-1}^{1} f(x+t)Q_n(t)\mathrm{d}t \end{aligned}$$

The second equality follows from change of variables, the third one comes from $f(x+t) = 0, \forall t \notin [-x, 1-x]$. Now, it is necessary to show $P_n \to f$ uniformly on $[0,1]$. Let $\varepsilon > 0$, since f is uniformly continuous on $[0,1], \exists \delta > 0$ such that $\forall |x-y| \leq \delta, |f(x) - f(y)| < \varepsilon/2$. Let $C = \sup\{f(x) : x \in [0,1]\}$, which exists by the Min/Max theorem. Choose $N \in \mathbb{N}$ such that $\sqrt{n}(1-\delta^2)^n < \varepsilon/(8C)$. Thus, $\forall n \geq N, \forall x \in [0,1]$,

$$\begin{aligned} |P_n(x) - f(x)| &= \left| \int_{-1}^{1} (f(x-t) - f(t))Q_n(t)\mathrm{d}t \right| \\ &\leq \int_{-1}^{1} |f(x-t) - f(x)|Q_n(t)\mathrm{d}t \\ &\leq \int_{|t|\leq\delta} |f(x-t) - f(x)|Q_n(t)\mathrm{d}t + \int_{\delta\leq|t|\leq 1} |f(x-t) - f(x)|Q_n(t)\mathrm{d}t \\ &\leq \frac{\varepsilon}{2} \int_{|t|\leq\delta} Q_n(t)\mathrm{d}t + \sqrt{n}(1-\delta^2)^n \int_{\delta\leq|t|\leq 1} 2C\mathrm{d}t \\ &< \frac{\varepsilon}{2} + 4C\sqrt{n}(1-\delta^2)^n \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

□

6.4 Analytic functions

Definition 58 (Analytic function)**.** A function, f, is said to be *analytic* on a non-empty, open interval (a,b) if, and only if, given $x_0 \in (a,b)$ there exists a power series centered on x_0 which converges to f near x_0. *Id est*, there exists $(a_n)_{n=0}^{\infty}$, $c, d \in (a, b)$ and $c < x_0 < d$ such that:

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \forall x \in (c, d)$$

Remark. An analytic function is said to belong to the analytic class, $\mathcal{C}^{\omega}(S)$.

Theorem 109 (Uniqueness of power series representation)**.** Consider $c, d \in \mathbb{R} \cup \{-\infty, \infty\}$, with $c < d$. Let $x_0 \in (c, d)$ and suppose $f : (c, d) \to \mathbb{R}$. If