Eidgenössische Technische Hochschule Zürich

Analysis I: One Variable

Autumn Semester 2025

Lecture Notes

Contents

References

[1] Figalli, Alessio: Skript zur Vorlesung Analysis I. ETH Zürich. 2025

[2] Von Flüe, Cyrill: Begleitskript zur Übungsstunde Analysis I. ETH Zürich. 2025

[3] Beutelspacher, Albrecht: Das ist o.B.d.A. trivial! 9. Auflage. Vieweg+Teubner Verlag

[4] Schichl, Hermann and Steinbauer, Roland: Einführung in das mathematische Arbeiten. 2. Auflage. Springer Verlag

Analysis I: One Variable

Fynn Krebser–fkrebser@student.ethz.ch

Herbstsemester 2025

0 Writing a Proof

In order to do mathematics, it is essential beeing able to write a proof. In this section we will discuss some basic rules and guidelines for writing proofs. The chapter is based on [2, 3, 4].

After having written a proof, typically it is in a state of exploration. This is the first draft of the proof. The next step is to write down a clean version of the proof. The latter of which we will now focus on.

The most important rule is probably to keep in mind normal grammar. Write in full sentences, use punctuation and capital letters. A proof is a piece of literature. It should be readable and understandable.

> Definition 0.1: **Definition**
>
> A DEFINITION is a precise description of a mathematical object.

A mathematical definition should be precise and unambiguous. It is used to abreviate long descriptions. For example, we can define prime numbers to avoid constantly repeating: "Positive integers $p > 1$ which only have 1 and p as divisors". Simply said, definitions help us talk about and understand mathematical proofs.

> Definition 0.2: **Theorem, Lemma, Proposition and Corollary**
>
> A THEOREM is a proven mathematical statement.
>
> A LEMMA is a theorem which is used to prove another theorem.
>
> A PROPOSITION is a theorem of lesser importance.
>
> A COROLLARY is a theorem which follows directly from another theorem.

Theorems, lemmas, propositions and corollaries are all statements which have been proven to be true. The difference between them is their relevance. This already implies that the names are somewhat subjective. Together with definitions, they provide the building blocks of a proof.

Two words that require extra attention are SUFFICIENT and NECESSARY. If we have $A \Rightarrow B$ then A is a sufficient for B and B is a necessary condition for A. It's best to simply memorize this.

0.1 Labelling

It is common to label certain objects, for example a line l or a matrix A. This is done to avoid repeating long descriptions however it is important to follow some rules when labelling objects. For example, one shouldn't call a function 2 or a set $9]xP$:). Instead, one should try to use meaningful labels. This can be done by labelling similar objects in a similar fashion. Similarly, one should use standards when they exist. For example, it is common to use $\mathbb{N}$ for the set of natural numbers, $\mathbb{Z}$ for the set of integers and so on.

Additionally, it can be useful to implement a hierarchy of labels. For example, one can use capital letters for sets and lower case letters for its elements.

Lastly, it is good practice to not use unnessecary labels. For example, one doesn't need to write "Every differentiable function f is continuous." Instead, one can simply write "Every differentiable function is continuous."

0.2 Proofing Techniques

A proof is a logical argument which shows that a certain statement is true. There are many possible techniques to do so, the most common ones beeing listed below.

> Definition 0.3: **Direct Proof**
>
> A DIRECT PROOF is a proof which starts from the assumptions and uses logical steps to arrive at the conclusion.

A direct proof is the most straightforward way to prove a statement.

> Example 0.4:
>
> We want to show that the sum of two even numbers is even. Let a, b be two even numbers. By definition of even numbers, there exists $k, m \in \mathbb{Z}$ such that $a = 2k$ and $b = 2m$. Thus,
>
> $$a + b = 2k + 2m = 2(k + m).$$
>
> As $k + m \in \mathbb{Z}$, we have shown that $a + b$ is even.

If one has to proof a statement for the set of all natural numbers, or similar sets, one can use induction.

> Definition 0.5: **Proof by Induction**
>
> A PROOF BY INDUCTION is a proof which consists of two steps.
>
> 1: Check it is true for $n = 1$.
>
> 2: If the equation holds for some n, then it holds for $n + 1$.

This technique is extremely powerful as can be seen in the following example.

Example 0.6: Gaussian Counting

We want to show that $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$ for all $n \in \mathbb{N}$.

Base Case: Let $n = 1$. Then, $1 = \frac{1(1+1)}{2} = 1$. Thus, the equation holds for $n = 1$.

Induction Step: Fix $n \geq 1$. We know that $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$. We want to show that $1+2+\cdots+n+(n+1) = \frac{(n+1)(n+2)}{2}$.

Notice that the first block of the LHS is exactly what we have assumed in the induction hypothesis. Thus, we can replace it.

$$\begin{aligned}\frac{n(n+1)}{2} + (n+1) &= \frac{n^2+n}{2} + n + 1 \\ &= \frac{n^2+n+2n+2}{2} \\ &= \frac{(n+1)(n+2)}{2}.\end{aligned}$$

Thus, the right hand side equals the left hand side.

Definition 0.7: Proof by Contradiction

A PROOF BY CONTRADICTION is a proof which starts by assuming the negation of the statement to be proven and then derives a contradiction.

Classic examples of a proof by contradiction are that $\sqrt{2}$ is irrational or that there are infinitely many prime numbers.

Example 0.8: $\sqrt{2}$ is irrational

We want to show that $\sqrt{2}$ is irrational. We do so by assuming the opposite, i.e. that $\sqrt{2}$ is rational. Thus, we can write $\sqrt{2} = \frac{p}{q}$ where $p, q \in \mathbb{Z}$ have no common divisors. Squaring both sides, we get $2q^2 = p^2$. This implies that p^2 is even, which in turn implies that p is even.

Thus, we can write $p = 2k$ for some $k \in \mathbb{Z}$. Plugging this back into the equation, we get $2q^2 = 4k^2$, or $q^2 = 2k^2$. This implies that q is even. However, this contradicts our initial assumption that p and q have no common divisors. $\lightning$

Thus, $\sqrt{2}$ is irrational.

Example 0.9: Infinetly many primes

We want to show that there are infinitely many prime numbers. Let us assume the opposite, i.e. that there are finitely many prime numbers.

Let $M = \{p_1, p_2, \cdots, p_n\}$ be the set of all prime numbers. Further, let $k := p_1 p_2 \cdots p_n$. Thus, every prime number divides k. However, consider $k + 1$. This number is not divisible by any prime number, as it always leaves a remainder of 1. Thus, $k+1$ is either prime itself or has a prime divisor which itself is not in M. $\lightning$

Thus, there are infinitely many prime numbers.

1 Introduction

Lec 1

We begin our journey by an exercise.

Example 1.1:

Consider the x, y plane and the parabola given by $y = x^2$. Consider the region bounded by the parabola and the x-axis and the vertical lines $x = 0$ and $x = 1$. i.E. the region

$$P = \{(x,y) \in \mathbb{R}^2 | 0 \leq x \leq 1, 0 \leq y \leq x^2\}.$$

To answer this question, we need to know what area is. For this we define three principles which we want the area to satisfy. Whatever it is, it should satisfy:

- $[a, b] \times [c, d]$ has area $(b - a)(d - c)$,
- if F is contained in G, then Area(F) $\leq$Area(G),
- if F and G are disjoint, then the Area($F \cup G$) = Area(F)+Area(G).

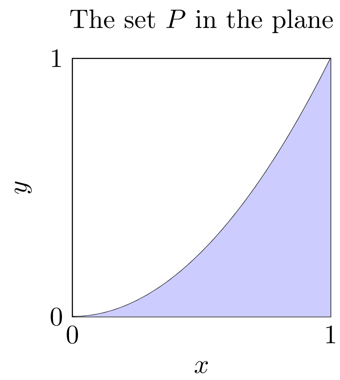


Figure 1: The set P in the plane.

Using this, we can now answer the question.

Example 1.1:

With these three principles, the area of P is $\frac{1}{3}$.

To prove this, we will need to have to use the following lemma.

Lemma 1.2:

For $n \geq 1$,

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}.$$

We want to use INDUCTION, which is a proof that is split into two steps.

Proof.

Base Case: Let $n = 1$.

$$LHS = 1^2 = 1$$
$$RHS = \frac{1}{3} + \frac{1}{2} + \frac{1}{6} = 1$$

.

Induction Step: Fix $n \geq 1$. We know that $1^2 + 2^2 + \cdots + n^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}$. We want $1^2+2^2+\cdots+n^2+(n+1)^2 = \frac{(n+1)^3}{3} + \frac{(n+1)^2}{2} + \frac{n+1}{6}$.

We notice, that the first block of the LHS is exactly what we have assumed in the induction hypothesis. Thus, we can replace it.

$$\begin{aligned}
\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} + (n+1)^2 &= \frac{(n+1)^3}{3} + \frac{(n+1)^2}{2} + \frac{n+1}{6} \\
LHS &= \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} + n^2 + 2n + 1 \\
&= \frac{n^3}{3} + \frac{3n^2}{2} + \frac{13n}{6} + 1 \\
RHS &= \frac{n^3 + 3n^2 + 3n + 1}{3} + \frac{n^2 + 2n + 1}{2} \\
&+ \frac{n+1}{6} \\
&= \frac{n^3}{3} + \frac{3n^2}{2} + \frac{13n}{6} + 1
\end{aligned}$$

Thus, the right hand side equals the left hand side. □

Using this lemma, we now want to calculate the area of our parabola.

Proof. [Proof of Theorem 1.1] To calculate the area of P we will fix a $n \geq 1$ and split the interval $[0, 1]$ into n equal parts.

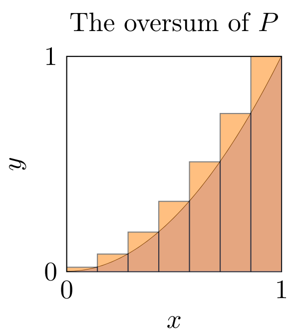


Figure 2: The oversum of P.

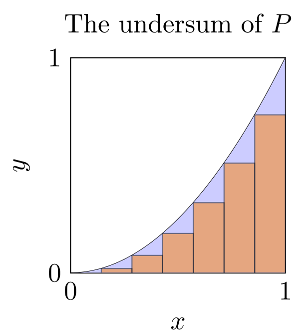


Figure 3: The undersum of P.

Let $\mathcal{A}$ =Area(P). We have

$$\mathcal{A} \leq \text{Area(figure 2)}.$$

$$\begin{aligned}
\mathcal{A} &\leq \frac{1}{n}\frac{1}{n^2} + \frac{1}{n}\frac{2^2}{n^2} + \cdots + \frac{1}{n}\frac{n^2}{n^2} \\
&= \frac{1}{n^3}\left(1^2 + 2^2 + \cdots + n^2\right) \\
&= \frac{1}{n^3}\left(\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}\right) \\
&= \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}
\end{aligned}$$

Similarly, for the second figure,

$$\begin{aligned}
\mathcal{A} &\geq \text{Area (figure 3)} \\
&= \frac{1}{n}\left(\left(\frac{0}{n}\right)^2 + \left(\frac{1}{n}\right)^2 + \cdots + \left(\frac{n-1}{n}\right)^2\right) \\
&= \frac{1}{n^3}\left(0^2 + 1^2 + \cdots + (n-1)^2\right) \\
&= \frac{1}{n^3}\left(\frac{(n-1)^3}{3} + \frac{(n-1)^2}{2} + \frac{n-1}{6}\right) \\
&= \frac{1}{n^3}\left(\frac{n^3}{3} - n^2 + \frac{n}{6}\right) \\
&= \frac{1}{3} - \frac{1}{n} + \frac{1}{6n^2}
\end{aligned}$$

Where in the last step we used the lemma with $n-1$ instead of n. Combining the both formulas, we have

$$-\frac{1}{2n} + \frac{1}{6n^2} \leq \mathcal{A} - \frac{1}{3} \leq \frac{1}{2n} + \frac{1}{6n^2}.$$

The only reasonable number between these two bounds is $\mathcal{A} = \frac{1}{3}$. □

For completeness, we give a definition of a set in naive set theory.

Definition 1.3: Set

A SET

- consists of distinct ELEMENTS,
- is uniquely characterized by its elements,
- is not an element of itself.
- If $A(x)$ is a statement about elements x in a set X, then $\{x \in X \mid A(x)\}$, denotes all elements for which $A(x)$ is true.

Finally, $\varnothing$ is the EMPTY SET.

The most useful property is probably the last one. An example of a set is $\{1, 2, 7, 10\}$, or $X = \{n \in \mathbb{N} \mid n \text{ is even}\} = \{n \in \mathbb{N} \mid \exists m \in \mathbb{N} : n = 2m\}$.

2 The Real Numbers

Lec 2

2.1 The Axioms of the Real Numbers

We want to study properties of the real numbers to formally describe them. To do this we introduce the Group.

Definition 2.1: Group

A GROUP is a set G with one operation $*$ that satisifies:

- Associativity: (a*b)*c = a*(b*c)
- Neutral element: $\exists e \in G : a * e = e * a = a \forall a$
- Inverse element: $\forall a \in G \exists a^{-1} \in G : a * a^{-1} = e$

Remark 2.2:

In general $a * b \neq b * a$. If $a * b = b * a \forall a, b \in G$ then G is a COMMUTATIVE or ABELIAN group.

Some examples of groups:

- $\mathbb{N}$ with addition $+$

  We can check this: Addition is assosciative, the neutral element is 0. However, the inverse of for example 3 is $-3 \notin \mathbb{N}$. So $\mathbb{N}$ is not a group.

- $\mathbb{Z}$ with addition $+$

  The first two points still hold. And $(-n) + n = 0$. Therefore $(\mathbb{Z}, +)$ is a commutative group as n+m = m+n.

- $\mathbb{Q}^* = \{\frac{p}{q} : p, q \in \mathbb{Z} \setminus \{0\}\}$ with multiplication $\cdot$

  This is associative, the neutral element is 1 as $\frac{p}{q} \cdot 1 = 1 \cdot \frac{p}{q} = \frac{p}{q}$. And $\frac{p}{q} \cdot \frac{q}{p} = 1$. Once again this is also commutative. Thus $(\mathbb{Q}^*, \cdot)$ is a commutative group.

Lemma 2.3:

Neutral and inverse are unique.

Proof. Let $e, e' \in G$ be two neutral elements.

$$e' \text{ (as } e \text{ is neutral) } = e * e' = e \text{ (as e' is neutral)}.$$

Given $a \in G$, assume b, c are both inverse elements of a. $a*b = b*a = e$ and $a * c = c * a = e$. Then

$$b = b * e = b * (a * c) = (b * a) * c = e * c = c.$$

□

Remark 2.4:

Since $a * a^{-1} = a^{-1} * a = e$, then a is the inverse of a^{-1}. Therefore,

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

An example of a non-commutative group would be symmetry groups. The symmetry group S_3 is the group of symmetries of an equilateral triangle. It has 6 elements: 3 rotations and 3 reflections. Interestingly, S_3 is the smallest non-commutative group.

Definition 2.5: Ring

A set R with two operations $+$ and $\cdot$[a]

- $(R, +)$ is a commutative group
- The $\cdot$ is associative, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, $\exists$ neutral element, $\cdot$ distributes over $+$ $a \cdot (b + c) = a \cdot b + a \cdot c$, $(b + c) \cdot a = b \cdot a + c \cdot a$

We denote by 0 the neutral element for $+$, and by 1 the neutral element for $\cdot$.

In general, $a \cdot b \neq b \cdot a$. If instead $a \cdot b = b \cdot a \forall a, b \in R$ then R is a COMMUTATIVE RING.

[a] + and $\cdot$ dont have to be the usual addition and multiplication

If R is a commutative ring, and every element different from 0 has an inverse for $\cdot$. Then we call it a FIELD.

Examples of rings:

$(\mathbb{Z}, +, \cdot)$ is a commutative ring. However, it is not a field as $2 \in \mathbb{Z}$ does not have an inverse for $\cdot$ in $\mathbb{Z}$.

$(\mathbb{Q}, +, \cdot)$ is a field.

Rings are generalizations for integers, and fields are generalizations for rationals.

Remember: For a group, $(a^{-1})^{-1} = a$. For a field, this means that $-(-a) = a$ and $\left(a^{-1}\right)^{-1} = a$ for all $a \neq 0$. $-$ denotes the inverse for $+$, and a^{-1} the inverse for $\cdot$.

Lemma 2.6:

Let F be a field, 0 neutral element for $+$.

- $a \cdot 0 = 0 \forall a \in F$
- $a \cdot (-b) = -(a \cdot b) = (-a) \cdot b \forall a, b \in F$
- $(-a)(-b) = a \cdot b \forall a, b \in F$

Proof.

- Since 0 is neutral, 0+0=0. By distributivity,

  $$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0.$$

  To both sides, add $-a \cdot 0$

  $$0 = a \cdot 0 + a \cdot 0 - a \cdot 0 = a \cdot 0.$$

- By the second point, $-(a \cdot b) = a \cdot (-b)$. Taking the additive inverse yields $a \cdot b = -(a \cdot (-b))$. Invoking the second point yet again but with $-b$ taking the role of b, we get $-(a \cdot (-b)) = (-a) \cdot (-b)$.

□

Remark 2.7:

Can 0=1?

If 0=1, then

$$a = a \cdot 1 = a \cdot 0 = 0 \forall a \in F.$$

So the field would only have one element. $F = \{0\}$. From now on, we will assume that $0 \neq 1$.

Something special about the real numbers is that they are ordered.

Given X, Y sets, $X \times Y = \{(x,y) : x \in X, y \in Y\}$ is the CARTESIAN PRODUCT

Given P, Q sets, $P \subseteq Q = P \subset Q$ if all elements of P are also in Q. We call P a SUBSET of Q.

$P \subsetneq Q$ if $P \subseteq Q$ but they are not equal. P is a STRICT SUBSET of Q.

$P \not\subset Q$ means P is not contained in Q.

Definition 2.8: Relation

X set, A relation is $R \subseteq X \times X$. We say that x and y are related if $(x,y) \in R$, and one writes xRy.

Symbols for relations are: $<, \leq, \sim, \cong$

Example 2.9:

Consider the set $X = \{1,2,3\}$. A possible relation is

$$R = \{(1,1),(2,2)\}.$$

A relation can be:

- REFLEXIVE: $x \sim x \forall x \in X$
- TRANSITIVE: if $x \sim y$ and $y \sim z$ implies $x \sim z$
- SYMMETRIC: if $x \sim y$ implies $y \sim x$
- ANTISYMMETRIC: if $x \sim y$ and $y \sim x$ implies $x = y$

We call an EQUIVALENCE RELATION a relation that is reflexive, transitive and symmetric.

Lec 3

If instead it is reflexive, transitive and antisymmetric, we call it a ORDER RELATION.

Example 2.10: Equivalence Relation

on $\mathbb{Z}$, $m \sim n$ if $m - n$ is divisible by 3. This is reflexive, transitive and symmetric.

Example 2.11: Order Relation

The classic is $x \leq y$ on $\mathbb{R}$. This is an order relation.

$x < y$ however is not an order, as $x < x$ is false.

If $\sim$ is an order, we use $\leq$ as a symbol.

When combining diffrent order relations, we have the following propositions:

Proposition 2.12:

Let R_1, R_2 be order relations on X. Then

1. $R_1 \cup R_2$ does not need to be an order relation.
2. $R_1 \cap R_2$ is an order relation.

Proof. We proof by counterexample and verification.

1. Let $X = \{1,2\}$ and the relation $R_1 = \{(1,1),(2,2),(1,2)\}$ and $R_2 = \{(1,1),(2,2),(2,1)\}$. Then
$$R_1 \cup R_2 = \{(1,1),(2,2),(1,2),(2,1)\}$$
is not antisymmetric. ↯
2. $R_1 \cap R_2$ is reflexive as $(x,x) \in R_1$ and $(x,x) \in R_2 \forall x \in X$.
Furthermore it is antisymmetric: if $(x,y) \in R_1 \cap R_2$ and $(y,x) \in R_1 \cap R_2$, then $(x,y) \in R_1$ and $(y,x) \in R_1$ implying $x = y$. Similarly for R_2.
Lastly, it is transitive: if $(x,y) \in R_1 \cap R_2$ and $(y,z) \in R_1 \cap R_2$, then $(x,y) \in R_1$ and $(y,z) \in R_1$ implying $(x,z) \in R_1$. Similarly for R_2. Thus $(x,z) \in R_1 \cap R_2$.

□

Definition 2.13:

$(F, \leq)$ is an ORDERED FIELD if F is a field and $\leq$ is an order relation on F such that

- Given $x, y \in F$, either $x \leq y$ or $y \leq x$
- Given $x, y, z \in F$, if $x \leq y$, then $x + z \leq y + z$
- Given $x, y \in F$, if $0 \leq x$ and $0 \leq y$, then $0 \leq x \cdot y$

Remark 2.14:

Finite fields cannot be ordered.

Proof. Exercise to the reader.

□

$(F, \leq)$ we write $y \geq x$ if $x \leq y$. Further $x < y$ if $x \leq y$ and $x \neq y$. We also say x is NON-NEGATIVE if $0 \leq x$ and x is POSITIVE if $0 < x$ and NON-POSITIVE if $x \leq 0$ and NEGATIVE if $x < 0$.

Example 2.15:

$$\mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N} \setminus \{0\} \right\}.$$

We say

$$\frac{p}{q} \leq \frac{r}{s} \text{ if } p \cdot s \leq r \cdot q.$$

Recall that $x \in F \Rightarrow -x$ is the inverse for $+$ $(x + (-x) = 0)$ and $(-x) \cdot (-y) = x \cdot y$.

Lemma 2.16:

Let $(F, \leq)$ be an ordered field. For example: (c) If $x \leq y$ and $z \leq w$, then $x + z \leq y + w$. (Same with $<$) or (f) $x^2 \geq 0 (x^2 = x \cdot x)$ or (g) $1 > 0$.

Proof. (c) $x \leq y \Rightarrow x + z \leq y + z$ and $z \leq w \Rightarrow z + y \leq w + y$. Using transitivity yields the result.

(f) Case 1: $x \geq 0$. Then $x \cdot x \geq 0$ by definition. Case 2: $x \leq 0 \Rightarrow -x \geq 0 \Rightarrow x \cdot x = (-x) \cdot (-x) \geq 0$.

(g) $1 = 1 \cdot 1 = 1^2 \geq 0$. Also $1 \neq 0$ implying $1 > 0$.

□

Lemma 2.17:

$(F, \geq), 0, 1$

- We can define the elements $2 = 1+1, 3 = 2+1, \ldots$ $-n$ is the inverse of n for $+$. Since $0 < 1 \Rightarrow 1 < 1+1 = 2 \Rightarrow 2 < 2+1 = 3 \Rightarrow \ldots$. Also $0 < 1 \Rightarrow -1 < 0, -2 \leq -1, \cdots$. Thus we have a sequence of elements

$$\ldots < -2 < -1 < 0 < 1 < 2 < \ldots \subseteq F.$$

  This is exactly the integers $\mathbb{Z}$.

- Given $p, q \in F, q \neq 0$, we define $\frac{p}{q} = p \cdot q^{-1}$. Then

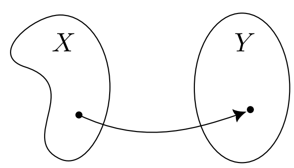
$$\mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{Z} \setminus \{0\} \right\} \subsetneq \mathbb{Z}.$$

  The inequality comes from for example $\frac{1}{2}$. Since $0 < 1 < 2 \Rightarrow \frac{1}{2} < 1$. Thus $\frac{1}{2} \notin \mathbb{Z}$.

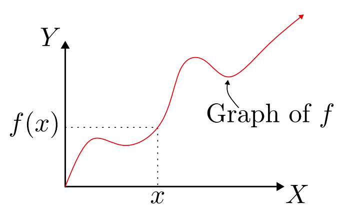
The moral of the story is that $Z \subsetneq Q \subseteq F$.

Definition 2.18: Function

Given X, Y sets, a FUNCTION f from X to Y is an assignement of an element in Y for every element of X.



Function $f : X \to Y$



Graph of a function

We call X the DOMAIN of f, Y the CODOMAIN of f. The element $y \in Y$ to which $x \in X$ is assigned is denoted by $f(x)$.

x is called the ARGUMENT of f, and $y = f(x)$ is the VALUE of f. The set $\{(x, f(x)) : x \in X\} \subseteq X \times Y$ is called the GRAPH of f.

Sometimes also the notation $f : X \to Y$ and $x \mapsto f(x)$ is used.

Remark 2.19:

Two functions $f_1 : X_1 \to Y_1, f_2 : X_2 \to Y_2$ are equal if

$$X_1 = X_2$$
$$Y_1 = Y_2$$
$$f_1(x) = f_2(x) \forall x \in X_1$$

We can define the functions sign and modulus.

Definition 2.20:

$(F, \leq)$ ordered field.

- The absolute value or the modulus is the function $|\cdot| : F \mapsto F$ defined by

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

- the SIGN is $\operatorname{sgn} : F \mapsto \{-1, 0, 1\}$ defined by

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}.$$

Lemma 2.21:

Let $(F, \leq)$ be an ordered field.

(a) $x = \operatorname{sgn}(x) \cdot |x|, |x| = |-x|, \operatorname{sgn}(-x) = -\operatorname{sgn}(x)$

(b) $|x| \geq 0$

(c) $|x \cdot y| = |x| \cdot |y|$

(e) $|x| \geq y \Leftrightarrow -y \leq x \leq y$

(g) $|x + y| \leq |x| + |y|$ (Triangle inequality)

(h) $||x| - |y|| \leq |x - y|$ (Inverse triangle inequality)

Sometimes we write xy for $x \cdot y$ when there is no ambiguity.

Proof. [Proof of (e)] Two cases: $x \geq 0$ and $x < 0$.

- If $|x| \leq y$, since in this case $x = |x| \geq 0$, $-y \leq 0 \leq x \leq y$.
  If $-y \leq x \leq y \Rightarrow |x| = x \leq y \Rightarrow |x| \leq y$.

□ Lec 4

When putting rationals on a number line, there will be gaps left. The classic example is $\sqrt{2}$, which cannot be expressed as a rational number. Let

$$X = \{q \in \mathbb{Q} : q < 0\} \cup \{q \in \mathbb{Q} : 0 < q^2 < 2\}.$$

$$Y = \{q \in \mathbb{Q} : q \geq 0 \text{ and } q^2 > 2\}.$$

Theese sets are disjoint, and on the number line, X is to the left of Y, with a gap in between.

Definition 2.22:

$(F, \leq)$. We say that it is COMPLETE if the COMPLETENESS AXIOM holds.

Let $X, Y \subseteq F$ non empty, assume that $x \leq y$ for all $x \in X, y \in Y$. Then there exists $c \in F$ such that

$$x \leq c \leq y \quad \forall x \in X, y \in Y.$$

Definition 2.23: Real numbers

The REAL NUMBERS are any complete ordered field.

This field is the only ordered field that is complete.

Definition 2.24:

Let $f : X \to Y$ be a function.

1. f is INJECTIVE if $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.
2. f is SURJECTIVE if for all $y \in Y$, there exists $x \in X$ such that $f(x) = y$.
3. f is BIJECTIVE if it is both injective and surjective. In this case, we denote $f^{-1} : Y \to X$ the inverse function of f, defined by $f^{-1} \circ f = \mathrm{id}_X$ and $f \circ f^{-1} = \mathrm{id}_Y$.

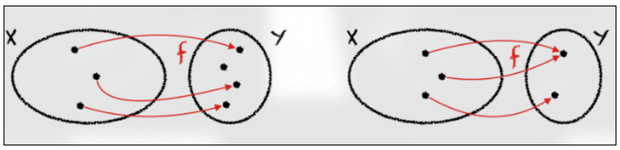


Figure 4: Injective and surjective function [1]

Definition 2.25:

$f : X \to Y$, $A \subseteq X$. Then

$$f(A) = \{y \in Y : \exists x \in A \text{ such that } y = f(x)\}$$

is the IMAGE of A through f.

If $B \subseteq Y$, then

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

is the PREIMAGE of B through f.

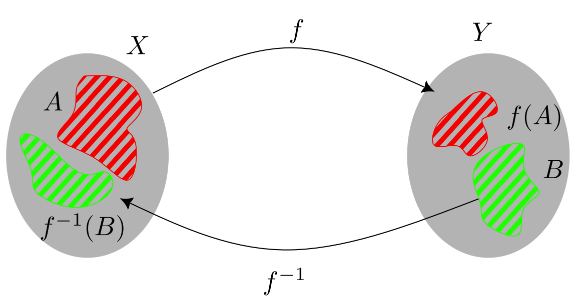


Figure 5: Image and preimage of sets

Example 2.26:

$f : \mathbb{R} \to \mathbb{R}, x \mapsto 0$.

Then $f^{-1}(\{0\}) = \mathbb{R}$, whereas $f^{-1}(\{y\}) = \varnothing$ for $y \neq 0$.

Lemma 2.27:

X, Y finite sets with n elements. $f : X \to Y$. Then f is injective iff it is surjective.

Proof. Denote the elements of X by $x_1, \dots, x_n$.

$\Rightarrow$) Assume f is injective. Then all elements $f(X) = f(x_1), \dots, f(x_n)$ are distinct. So $f(X)$ has n elements and $f(X) \subseteq Y$, Y has n elements. Therefore $f(X) = Y$ implying f is surjective.

$\Leftarrow$) We proof by contraposition. Assume f is not injective. Thus $\exists x_i \neq x_j$ such that $f(x_i) = f(x_j)$. But then f(X) $= \{f(x_1), \dots, f(x_n)\}$ has at most $n-1$ elements. And because $f(X) \subsetneq Y$ f can not be surjective. Concluding the proof. □

The lemma is not true for infinite sets.

$$f_1 : \mathbb{N} \to \mathbb{N}, n \mapsto n + 1.$$

This function is injective but not surjective ($f^{-1}(\{0\}) = \varnothing$).

$$f_2 : \mathbb{N} \to \mathbb{N}, n \mapsto \begin{cases} n = 0 & n = 0 \\ n - 1 & n \geq 1 \end{cases}.$$

f_2 is surjective but not injective ($f_2(0) = f_2(1)$).

Exercise 2.28:

Let $R_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$. Show the existance and uniqueness of the function $\sqrt{\cdot} : R_{\geq 0} \to R_{\geq 0}$ where $(\sqrt{a})^2 = a \quad \forall a \in R_{\geq 0}$.

For existance: Given $a \in R_{\geq 0}$, let $X = \{x \in R_{\geq 0} : x^2 \leq a\}$ and $Y = \{y \in R_{\geq 0} : y^2 \geq a\}$. By completeness, $\exists c \in R_{\geq 0}$ such that

$$x \leq c \leq y \quad \forall x \in X, y \in Y.$$

To proof: $c^2 = a$. (If $c^2 > a$, then $\exists \epsilon > 0$ st $(c - \epsilon)^2 > a$, so $y = c - \epsilon \in Y$ but $y < c$, contradiction. The case $c^2 < a$ is similar).

Let $a \leq b \in \mathbb{R}$. Then $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ is a CLOSED INTERVAL. a is called the LEFT ENDPOINT and b the RIGHT ENDPOINT. $b - a$ is the LENGTH of the interval.

Similarly, $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ is an OPEN INTERVAL.

The notations can be comined: $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ and $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$. These are HALF-OPEN INTERVALS.

The notation $[a, \infty) = \{x \in \mathbb{R} : x \geq a\}$ represents an UNBOUNDED CLOSED INTERVAL. Similarly, $(-\infty, b) = \{x \in \mathbb{R} : x \leq b\}$.

UNBOUNDED OPEN INTERVALS are defined as $(a, \infty) = \{x \in \mathbb{R} : x > a\}$ and $(-\infty, b) = \{x \in \mathbb{R} : x < b\}$.

Intervals which include $\pm\infty$ are called UNBOUNDED INTERVALS. Whereas the others are BOUNDED INTERVALS.

Remark 2.29:

There also exists the notation $]a, b[$ for (a, b) and $]a, b]$ for $(a, b]$.

To work with intervals, we need the following definitions for sets.

Definition 2.30:

Let P, Q be sets. Then we define

$$P \cap Q = \{x : x \in P \text{ and } x \in Q\} \quad \text{(intersection)}.$$

$$P \cup Q = \{x : x \in P \text{ or } x \in Q\} \quad \text{(union)}.$$

$$P \setminus Q = \{x : x \in P \wedge x \notin Q\} \quad \text{(relative complement)}.$$

$$P \Delta Q = (P \setminus Q) \cup (Q \setminus P) \quad \text{(symmetric difference)}.$$

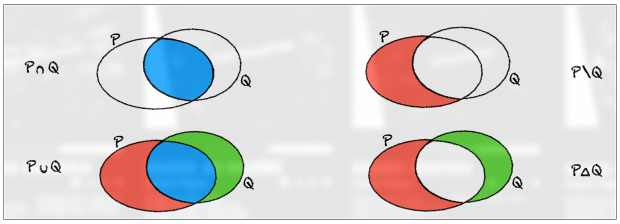


Figure 6: Set operations[1]

Given $P \subseteq X$, $P^c = X \setminus P$ is the COMPLEMENT of P in X. However this should be used with care, since the set X might not be obvious from the context.

Example 2.31: Complement Ambiguity

Given an interval $[a, b]$, then $[a, b]^c = (-\infty, a) \cup (b, \infty)$ might be the intuition. However, this only is true if we think about the reals. However, the notation might also be used in $\mathbb{Z}$ or $\mathbb{Q}$.

Definition 2.32: Union and Intersections of several sets

Given $\mathcal{A}$ a family of sets, we define

$$\bigcup_{A \in \mathcal{A}} A = \{x | \exists A \in \mathcal{A} : x \in A\}.$$

$$\bigcap_{A \in \mathcal{A}} A = \{x | x \in A \forall A \in \mathcal{A}\}.$$

If $\mathcal{A} = \{A_1, \ldots, A_n\}$ we also write

$$\bigcup_{i=1}^{n} A_i \quad \text{and} \quad \bigcap_{i=1}^{n} A_i.$$

Example 2.33:

$\mathcal{A} = \{[x, \infty) \text{ with } x \in \mathbb{R}\}$. Then

$$\bigcup_{A \in \mathcal{A}} A = \mathbb{R}$$

$$\bigcap_{A \in \mathcal{A}} A = \varnothing$$

The second one is true since for any $y \in \mathbb{R}$, we can choose $x = y + 1$. Then $y \notin [x, \infty)$.

Lec 5

Definition 2.34:

Let $x \in \mathbb{R}$. A NEIGHBORHOOD of x is any set that contains an open interval I s.t. $x \in I$.

Given $\delta > 0$, the open interval $(x - \delta, x + \delta)$ is the δ-NEIGHBORHOOD of x.

For example, $0 \in R$.

- $[-1, 1]$ is a neighborhood of 0 as it contains $0 \in (-1, 1)$.
- $[-1, 1] \cup \mathbb{Q}$ for the same reason.

For the delta neighborhood, we have

$$(x - \delta, x + \delta) = \{y \in \mathbb{R} : |x - y| < \delta\}.$$

With the notion of neighborhood, we can define further

Definition 2.35: Open and closed sets

A set $U \subseteq \mathbb{R}$ is OPEN if $\forall x \in U \, \exists I$ open interval s.t. $x \in I \subseteq U$.

A set $V \subseteq \mathbb{R}$ is CLOSED if $\mathbb{R} \setminus V$ is open.

This definition is best understood with a number line. On it, one can see that open intervals are open sets, while closed intervals are closed sets. Half-open intervals are neither open nor closed, as can be seen by looking at the third example in figure 7.

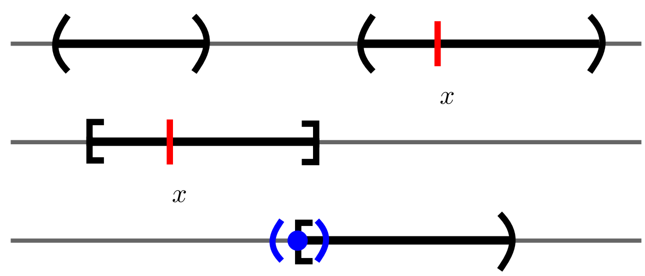


Figure 7: Open and closed sets

Theorem 2.36:

Let $U, V \subseteq \mathbb{R}$ be open. Then $U \cup V$ and $U \cap V$ are open. The same holds for arbitrary unions, but only for finite intersections.

Proof. Let U, V be open. Let $x \in U \cup V$. Wlog, assume $x \in U$. Because U is open, $\exists I$ open interval s.t. $x \in I \subseteq U \subseteq U \cup V$. Thus, $U \cup V$ is open.

Let $x \in U \cap V$. Then $x \in U$ and $x \in V$. Thus, $\exists \varepsilon_u, \varepsilon_v > 0$ s.t.

$$(x - \varepsilon_u, x + \varepsilon_u) \subseteq U, \quad (x - \varepsilon_v, x + \varepsilon_v) \subseteq V.$$

Now let $\varepsilon = \min(\varepsilon_u, \varepsilon_v)$. Then

$$(x - \varepsilon, x + \varepsilon) \subseteq (x - \varepsilon_u, x + \varepsilon_u) \cap (x - \varepsilon_v, x + \varepsilon_v) \subseteq U \cap V.$$

□

Example 2.37: Open and closed sets

$[0, 1]$ is closed in $\mathbb{R}$. Indeed, we can show that $\mathbb{R} \setminus [0, 1] = (-\infty, 0) \cup (1, \infty)$ is open. Let $x \in \mathbb{R} \setminus [0, 1]$. Wlog, assume $x \in (-\infty, 0)$. Thus, $x < 0 \Rightarrow |x| > 0$. Let $\varepsilon = \frac{|x|}{2} > 0$. By construction, is $(x - \varepsilon, x + \varepsilon) \subseteq (-\infty, 0)$. Thus, $\mathbb{R} \setminus [0, 1]$ is open, implying $[0, 1]$ is closed.

Remark 2.38:

Both $\mathbb{R}$ and $\varnothing$ are open and closed. They are the only subsets of $\mathbb{R}$ with this property.

2.2 Complex Numbers

Complex numbers can be looked at as

$$\mathbb{C} = \mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}.$$

This can be simply displayed on a plane.

We now want to define addition and multiplication on $\mathbb{C}$ to have a notion of a field. Let $z = (x, y) \in \mathbb{R}^2$, or $z = x + iy$ where x is called the REAL PART and y the IMAGINARY PART.

We first consider the map $\mathbb{R} : x \mapsto (x, 0) \in \mathbb{C}$.

Goal: Field, s.t. my operations coincide with the usual ones when considering points in $\mathbb{R}$. Also, we want $i^2 = -1$ $((0, 1) \cdot (0, 1) = (-1, 0))$.

Let $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$. We let

$$z_1 \cdot z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1).$$

Definition 2.39:

On $\mathbb{C} = \mathbb{R}^2$, define

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$
$$(x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$$

Theorem 2.40:

$\mathbb{C}$ with $+, \cdot$ as before is a field.

Proof. We take $(0, 0)$ as neutral element for $+$, and $(1, 0)$ for $\cdot$.

The addition is immediate to check. For multiplication, we have to check

$$(x_1, y_1) \cdot ((x_2, y_2) \cdot (x_3, y_3)) = ((x_1, y_1) \cdot (x_2, y_2)) \cdot (x_3, y_3),$$

Which can be done by straightforward computation.

The neutral element for multipliation is also a simple check, similar to distributivity.

Thus $\mathbb{C}$ is a commutative ring. Given $(x, y) \in \mathbb{C} \setminus \{(0, 0)\}$, consider $\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right)$. Then

$$(x, y) \cdot \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2}\right) = \left(\frac{x^2 + y^2}{x^2 + y^2}, 0\right) = (1, 0).$$

□

From now on, we write $x + iy$ for (x, y), x for $x + i \cdot 0$ and iy for $0 + iy$.

Given $z, w \in \mathbb{C}$, we write zw for $z \cdot w$, if $z \neq 0$. then z^{-1} or $\frac{1}{z}$ denotes the multiplicative inverse.

For example,

$$(i)^{-1} = \frac{1}{i} = -i.$$

Definition 2.41: Complex Conjugation

Given $z = x + iy \in \mathbb{C}$, we define $\overline{z} \in \mathbb{C}$ as $x - iy$ as the COMPLEX CONJUGATE of z.

Lemma 2.42:

The following holds for all $z, w \in \mathbb{C}$:

1. $z \cdot \overline{z} = x^2 + y^2 \geq 0$, also $z \cdot \overline{z} = 0 \Leftrightarrow z = 0$.
2. $\overline{z + w} = \overline{z} + \overline{w}$.
3. $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.

Proof. 1. Is a quick computation. For the second and third, let $z = x_1 + iy_1, w = x_2 + iy_2$. then

$$\begin{aligned}\overline{z + w} &= \overline{(x_1 + x_2) + i(y_1 + y_2)} \\ &= (x_1 + x_2) - i(y_1 + y_2) \\ &= (x_1 - iy_1) + (x_2 - iy_2) \\ &= \overline{z} + \overline{w}.\end{aligned}$$

Similarly, for the third. □

As an exercise, Given $z = x + iy$, write $x = \Re(z)$ and $y = \Im(z)$. Then

$$\Re(z) = \frac{z + \overline{z}}{2}, \quad \Im(z) = \frac{z - \overline{z}}{2i}.$$

Recall that on $\mathbb{R}$, given $x \in \mathbb{R}$, $x \mapsto |x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$. Because $\mathbb{C}$ is not an ordered field we need something different. The modulus represents the distance to 0. We want to implement something similar on $\mathbb{C}$. By the pythagorean theorem, the distance to 0 is $\sqrt{x^2 + y^2}$.

Definition 2.43:

Given $z \in \mathbb{C}$

$$|z| = \sqrt{z \cdot \overline{z}} = \sqrt{x^2 + y^2}.$$

Where $\sqrt{\cdot}$ is the square root defined before.

If $z = x \in \mathbb{R}$, then the two definitions of $|\cdot|$ coincide.

Some properties of the modulus:

$$|zw| = \sqrt{zw \cdot \overline{zw}} = \sqrt{zw \cdot \overline{z} \cdot \overline{w}} = \sqrt{z \cdot \overline{z}} \cdot \sqrt{w \cdot \overline{w}} = |z| \cdot |w|.$$

$$z^{-1} = \frac{\overline{z}}{|z|^2} \quad \text{for } z \neq 0.$$

Theorem 2.44: Triangle Inequality

For all $z, w \in \mathbb{C}$,

$$|z + w| \leq |z| + |w|.$$

Proof. We have

$$\begin{aligned}|z + w|^2 &= (x_1 + x_2)^2 + (y_1 + y_2)^2 \\ &= x_1^2 + y_1^2 + x_2^2 + y_2^2 + 2(x_1x_2 + y_1y_2) \\ &= |z|^2 + |w|^2 + 2(x_1x_2 + y_1y_2) \\ &\leq |z|^2 + |w|^2 + 2|z| \cdot |w| \\ &= (|z| + |w|)^2.\end{aligned}$$

However we still have to show $x_1x_2 + y_1y_2 \leq |z| \cdot |w|$. □ Lec 6

Lemma 2.45: Cauchy-Schwarz inequality

$z = x_1 + iy_1$, $w = x_2 + iy_2 \in \mathbb{C}$. Then

$$x_1x_2 + y_1y_2 \leq |z||w|.$$

Proof. Look at $|z|^2|w|^2 - (x_1x_2 + y_1y_2)^2$. This is equal to

$$\begin{aligned}&= (x_1^2 + y_1^2)(x_2^2 + y_2^2) - (x_1x_2 + y_1y_2)^2 \\ &= x_1^2y_2^2 + x_2^2y_1^2 - 2x_1x_2y_1y_2 \\ &= (x_1y_2 - x_2y_1)^2 \geq 0|z|^2|w|^2 \geq (x_1x_2 + y_1y_2)^2\end{aligned}$$

Taking square roots gives

$$|z||w| \geq |x_1x_2 + y_1y_2| \geq x_1x_2 + y_1y_2.$$

□

In general, $|z-w|$ represents the length of the segment between z and w.

Definition 2.46: Open and closed disks

Given $z \in \mathbb{C}$ and $r > 0$, we define

$$B(z,r) = \{w \in \mathbb{C} \mid |w-z| < r\},$$

as the OPEN DISK of radius r centered at z.

Furthermore, we define

$$\overline{B(z,r)} = \{w \in \mathbb{C} \mid |w-z| \leq r\},$$

as the CLOSED DISK of radius r centered at z.

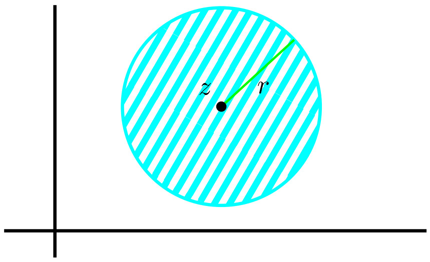


Figure 8: Disks in $\mathbb{C}$.

Notice that this definition is a generalisation of open and closed intervals, as it reduces to them by looking at $B(z,r) \cap \mathbb{R} = (z-r, z+r)$.

This definition can be used to define open and closed sets in $\mathbb{C}$.

Definition 2.47:

$U \subseteq \mathbb{C}$ is OPEN if for all $z \in U \exists r > 0$, such that $B(z,r) \subseteq U$.

A set $C \subseteq \mathbb{C}$ is CLOSED if $\mathbb{C} \setminus C$ is open.

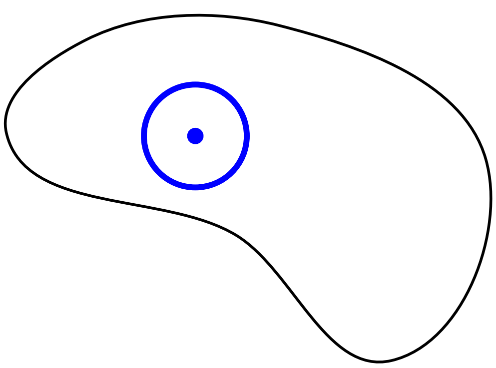


Figure 9: Open and closed sets in $\mathbb{C}$.

Again sets do not have to be either open or closed. For example,

$$A = \{(x,y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, 0 < y < 1\}.$$

2.3 Maximum and Supremum

Let us return to the real numbers for the moment.

Definition 2.48:

$X \subseteq \mathbb{R}$. We say that

- X is BOUNDED ABOVE if $\exists s \in \mathbb{R}$ such that $x \leq s \forall x \in X$. Such an s is called an UPPER BOUND of X.

  If such an s belongs to X, then s is called the MAXIMUM of X, denoted $s = \max(X)$.
- Analogous definitions hold for BOUNDED BELOW, LOWER BOUND and MINIMUM.
- A set is BOUNDED if it is both bounded above and below.

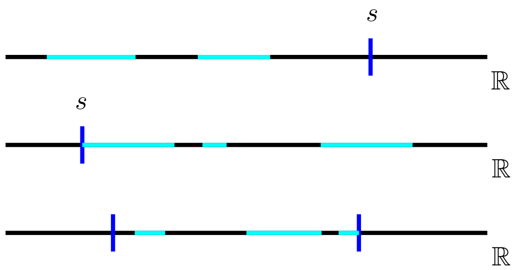


Figure 10: Bounded sets.

Remark 2.49:

The maximum, if it exists, is unique.

Proof. Indeed, if $s_1, s_2 \in X$ are maxima of X, then

$$s_1 \leq s_2 \text{ and } s_2 \leq s_1 \Rightarrow s_1 = s_2.$$

□

Definition 2.50: Supremum

Let $X \subseteq \mathbb{R}$. Define

$$A = \{a \in \mathbb{R} \mid a \geq x \forall x \in X\}.$$

This is the set containing all upper bounds of X.

If A has a minimum, then we call it the SUPREMUM of X, and denote it by $\sup(X)$.

$s = \sup(X)$ means that $x \leq s \forall x \in X$ and that $\forall t < s \exists x \in X$ such that $x > t$.

Remark 2.51:

If X has a maximum, then $\max(X) = \sup(X)$.

Example 2.52:

A few examples:

- $[a,b]$ is bounded with $\min([a,b]) = a$ and $\max([a,b]) = b$.
- (a,b) is bounded, but has no minimum or maximum. However, $\sup((a,b)) = b$.

Probably one of the most important properties of $\mathbb{R}$ is the following:

Theorem 2.53: Existence of Supremum

Let $X \subseteq \mathbb{R}$ be bounded above. Then $\sup(X)$ exists.

Proof. Let $A = \{a \in \mathbb{R} \mid a \geq x \, \forall x \in X\}$ be the set of upper bounds. Since X is bounded above, A is nonempty. By completeness, $\exists c \in \mathbb{R}$ such that

$$x \leq c \leq a \, \forall x \in X, a \in A.$$

The first inequality means that c is an upper bound of X, instead looking at the second inequality, we see that c is the least upper bound. So $c = \min(A) = \sup(X)$. □

If X is bounded below, the largest lower bound is called INFIMUM, written $\inf(X)$.

Proposition 2.54:

Given $X, Y \subseteq \mathbb{R}$, we can define the sets

$$X + Y = \{x + y \mid x \in X, y \in Y\}$$
$$XY = \{xy \mid x \in X, y \in Y\}$$

1. $$\sup(X \cup Y) = \max(\sup(X), \sup(Y)).$$
2. If $X \cap Y \neq \varnothing$, then $$\sup(X \cap Y) \leq \min(\sup(X), \sup(Y)).$$
3. $$\sup(X + Y) = \sup(X) + \sup(Y).$$
4. If $X, Y \subseteq \mathbb{R}_{\geq 0}$, then $$\sup(XY) = \sup(X) \cdot \sup(Y).$$

Proof. [Proof of 3.] Let $x_0 = \sup(X)$ and $y_0 = \sup(Y)$.

Given $z \in X + Y$, by definition $\exists x \in X \wedge y \in Y$ such that $z = x + y$. Therefore, $z = x + y \leq x_0 + y_0$. This implies that $x_0 + y_0$ is an upper bound for $X + Y$, which means that $\sup(X + Y) \leq x_0 + y_0$.

To assume equality, assume by contradiction that $x_0 + y_0 - \sup(X + Y) = \epsilon > 0$.

Since $x_0 = \sup(X), \exists \overline{x} \in X : \overline{x} > x_0 - \frac{\epsilon}{2}$. Similarly, $\exists \overline{y} \in Y : \overline{y} > y_0 - \frac{\epsilon}{2}$.

Therefore, $\sup(X + Y) \geq \overline{x} + \overline{y} > x_0 + y_0 - \epsilon = \sup(X + Y)$, which is a contradiction. ↯

□

We denote $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ for the EXTENDED REAL LINE. The rule is, that for all $x \in \mathbb{R}$, $-\infty < x < \infty$. Furthermore,

- $\infty + x = \infty$
- $-\infty + x = -\infty$
- If $x > 0$, $x \cdot \infty = \infty$, $x \cdot -\infty = -\infty$

But, $\infty - \infty$ and $0 \cdot \infty$ NO, VERBOTEN!!!

Definition 2.55:

$X \subseteq \mathbb{R}$ and X not bounded above $\Rightarrow \sup(X) = \infty$.

Lets consider $X = \varnothing$. Then $\sup(\varnothing) = -\infty$.

Lec 7

Theorem 2.56: Archimedian Principle

Given $x \in \mathbb{R} \, \exists! n \in \mathbb{Z}$ such that $n \leq x < n + 1$

Proof. Consider first $x \geq 0$. Let $E = \{n \in \mathbb{Z} \mid n \leq X\}$. Since $x \geq 0$, $0 \in E$ so $E \neq \varnothing$. Also E is bounded above. So $\exists s_0 = \sup E \in \mathbb{R}$.

Since s_0 is the least upper bound, so $s_0 \leq x$ because x is an upper bound and $\exists n_0 \in E$ such that $s_0 - 1 < n_0$, otherwise $s_0 - 1$ would be an upper bound.

The latter implies $s_0 < n_0 + 1$. Since $m \leq s_0 \, \forall m \in E$, $m < n_0 + 1 \, \forall m \in E \subseteq \mathbb{Z}$. As $m, n_0 \in \mathbb{Z}$, this implies $m \leq n_0$.

Since $n_0 \in E$, this implies that $n_0 = \max(E)$, implying $n_0 = s_0$.

Also, $n_0 = \max(E) \Rightarrow n_0 + 1 \notin E \Rightarrow n_0 + 1 > x$.

So reccalling that $s_0 \leq x$, we have

$$n_0 = s_0 \leq x < n_0 + 1.$$

If $x < 0$, apply the result to $-x$. Then $\exists m \in \mathbb{Z}$, such that

$$m \leq -x < m + 1 \Rightarrow -m - 1 < x \leq -m.$$

If $x = -m$, take $n = -m$:

$$n = -m = x < n + 1.$$

If $x < -m$, take $n = -m - 1$:

$$n < x < n + 1.$$

For uniqueness, fix an x, assume $\exists n_1, n_2 \in \mathbb{Z}$ such that

$$n_1 \leq x < n_1 + 1$$
$$n_2 \leq x < n_2 + 1.$$

Combining the inequalities diagonally, we get

$$n_1 < n_2 + 1 \Rightarrow n_1 \leq n_2$$
$$n_2 < n_1 + 1 \Rightarrow n_2 \leq n_1.$$

This however implies $n_1 = n_2$. □

Definition 2.57:

Given $x \in \mathbb{R}$, let n be given as before ($n \in \mathbb{Z}, n \leq x < n + 1$). Define

$$n = \lfloor x \rfloor,$$

as the INTEGER PART of x.

Furthermore, let

$$\{x\} = x - \lfloor x \rfloor \in [0, 1),$$

be the FRACTIONAL PART of x.

Corollary 2.58:

For every $\varepsilon > 0, \exists n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$.

Proof. If $\varepsilon \geq 1$, take $n = 1$.

If $\varepsilon < 1$, then consider $\frac{1}{\varepsilon} \geq 1$. By Archimedian principle, $\exists m \in \mathbb{Z}$ such that $m \leq \frac{1}{\varepsilon} < m + 1$. Let $n = m + 1 \in \mathbb{N}$. Then

$$\frac{1}{\varepsilon} < n \Leftrightarrow \frac{1}{n} < \varepsilon.$$

□

Corollary 2.59:

$\forall a < b \in \mathbb{R}, \exists r \in \mathbb{Q}$ such that $a < r < b$.

Proof. Let $\varepsilon = b - a$. By the corollary 2.58, $\exists m \geq 1$ such that $\frac{1}{m} < \varepsilon$. Consider $x = ma$. By the Archimedian principle, $\exists n \in \mathbb{Z}$ such that $n \leq ma < n + 1$. Then

$$\frac{n}{m} \leq a < \frac{n+1}{m} \leq a + \frac{1}{m} < a + \varepsilon = b.$$

□

Definition 2.60:

A set $X \in \mathbb{R}$ is DENSE in every open interval contains an element of X.

Proposition 2.61:

$\mathbb{Q}$ is dense in $\mathbb{R}$.

Proof. Let $\varepsilon > 0$. Choose $x \in \mathbb{R}$ and let $q \in \mathbb{N}_{\geq 0}$ such that $\frac{1}{q} < \varepsilon$. By Archimedian principle, $\exists p \in \mathbb{Z}$ such that $p \leq qx < p+1$. Dividing by q, we have

$$\frac{p}{q} \leq x < \frac{p+1}{q} < x + \frac{1}{q} < x + \varepsilon.$$

Thus $|x - \frac{p}{q}| < \varepsilon$.

□

Exercise 2.62:

For every $\varepsilon > 0 \, \exists n \geq 1$ such that

$$\frac{1}{10^n} < \varepsilon.$$

(This uses that $10^n < 10^{n+1}$)

2.3.1 Decimal Representation of Real Numbers

We have a sequence of numbers $a_0, a_1, a_2, \ldots$ where

$$a_0 \in \mathbb{Z}, a_i \in \{0, 1, 2, \ldots, 9\} \, \forall i \geq 1.$$

Start from $a_0, a_1, a_2, \ldots$ and assume $a_0 \geq 0$.

We define $x_0 = a_0$ and $y_0 = a_0 + 1$. Furthermore

$$\begin{aligned} x_1 &= a_0 + \frac{a_1}{10} & y_1 &= a_0 + \frac{a_1+1}{10} \\ x_2 &= a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} & y_2 &= x_2 + \frac{1}{10^2} \\ x_n &= \sum_{\kappa=0}^{n} a_\kappa 10^{-\kappa} & y_n &= x_n + 10^{-n}. \end{aligned}$$

From this, we have

$$x_0 < x_1 < x_2 < \ldots < x_n < y_n < \ldots < y_2 < y_1 < y_0.$$

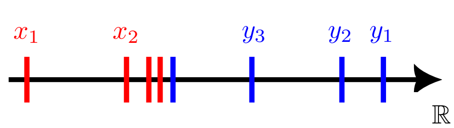


Figure 11: Decimal construction of real numbers.

Define $X = \{x_0, x_1, \cdots\}$ and $Y = \{y_0, y_1, \cdots\}$. Then we have $x \leq y \, \forall x \in X, y \in Y$. So by completeness axiom, $\exists z \in \mathbb{R}$ such that in particular $x_n \leq z \leq y_n \, \forall n \geq 0$.

Notice that c is unique. As if c and d both satisfy this, then

$$x_n \leq c, d \leq y_n \Rightarrow |c - d| \leq y_n - x_n = 10^{-n} \Rightarrow c = d.$$

□ If $a_0 < 0$, consider $-a_0$, find c and consider $-c$.

Viceversa, given $x \in \mathbb{R}$, $x \geq 0$, let

$$a_0 = \lfloor x \rfloor, a_n = \lfloor 10^n x \rfloor - 10 \lfloor 10^{n-1} x \rfloor, n \geq 1.$$

Remark 2.63:

This is not a 1:1 correspondence, as

$$0.9999999\ldots = 1.000000\ldots.$$

This only happens with infinite tails of 9s. So for example, instead of writing $0.5399\cdots$, we write 0.54.

Definition 2.64: Power Set

Given X, define $\mathcal{P}(X) = \{Y \mid Y \subseteq X\}$ as the POWER SET of X.

It is the set of all subsets of X.

Definition 2.65:

X, Y sets. We write

- $X \sim Y$ (X and Y have the same cardinality) if $\exists f : X \to Y$ bijective,
- If $\exists f : X \to Y$ injective, we write $|X| \lesssim |Y|$,
- X is COUNTABLE if $\exists f : X \to \mathbb{N}$ bijective,
- X is UNCOUNTABLE if X infinite and not countable.

Theorem 2.66:

If $X \lesssim Y$ and $Y \lesssim X$, then $X \sim Y$.

Example 2.67: Natural and Even Numbers

Given the sets

$$\mathbb{N} = \{0, 1, 2, 3, \ldots\} \text{ and } \mathbb{N}_{\text{even}} = \{0, 2, 4, 6, \ldots\},$$

we have $\mathbb{N} \sim \mathbb{N}_{\text{even}}$ as the function

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \mathbb{N}_{\text{even}} \\ n &\longmapsto f(n) = 2n. \end{aligned}$$

Example 2.68: Integers and Naturals

Given the sets $\mathbb{N}, \mathbb{Z}$, we clearly have $\mathbb{N} \lesssim \mathbb{Z}$ by the function $f : \mathbb{N} \to \mathbb{Z}, f(n) = n$. For the other direction, consider

$$\begin{aligned} f : \mathbb{Z} &\longrightarrow \mathbb{N} \\ n &\longmapsto f(n) = \begin{cases} 2n & n \geq 0 \\ -2n-1 & n < 0 \end{cases}. \end{aligned}$$

Thus there must exist a bijection $\mathbb{Z} \sim \mathbb{N}$.

Lec 8

For the set of fractions, we have the following properties:

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N} \setminus \{0\} \right\}.$$

The map

$$f : \mathbb{Q} \longrightarrow \mathbb{Z} \times \mathbb{N}$$
$$\frac{p}{q} \longmapsto f\left(\frac{p}{q}\right) = (p, q).$$

is injective, thus $\mathbb{Q} \lesssim \mathbb{Z} \times \mathbb{N} \setminus \{0\}$. Thus also $\mathbb{Q} \lesssim \mathbb{N} \times \mathbb{N}$.

We want to show that $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$. For this consider the following map:

$$\begin{aligned} f : \mathbb{N} \times \mathbb{N} &\longrightarrow \mathbb{N} \\ (0,0) \longmapsto 0 \qquad & (1,1) \longmapsto 4 \\ (1,0) \longmapsto 1 \qquad & (2,0) \longmapsto 5 \\ (0,1) \longmapsto 2 \qquad & (3,0) \longmapsto 6 \\ (0,2) \longmapsto 3 \qquad & (2,1) \longmapsto 7 \end{aligned}$$

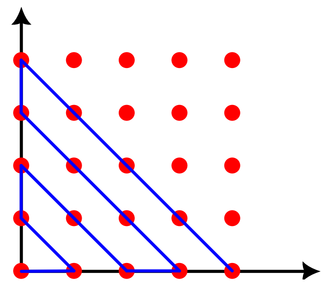


Figure 12: Mapping $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$

This map shows, that $\mathbb{N} \times \mathbb{N} \lesssim \mathbb{N}$. But clearly also $\mathbb{N} \lesssim \mathbb{N} \times \mathbb{N}$. Thus $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$, implying

$$\mathbb{Q} \lesssim \mathbb{N}.$$

When considering the set of real numbers $\mathbb{R}$, there exist more real numbers than natural numbers, i.e. $\mathbb{R} \not\sim \mathbb{N}$.

Argument 1: We proof $\nexists f : \mathbb{N} \to [0, 1)$ surjective. Assume there exists such a function. Then we can write all numbers in $[0, 1]$ as

$$\begin{aligned} r_0 &= 0, a_{00} a_{01} a_{02} a_{03} \dots \\ r_1 &= 0, a_{10} a_{11} a_{12} a_{13} \dots \\ r_2 &= 0, a_{20} a_{21} a_{22} a_{23} \dots \\ &\vdots \end{aligned}$$

Now consider the number

$$r = 0, b_0 b_1 b_2 b_3 \dots \qquad b_i = \begin{cases} 1 & a_{ii} \neq 1 \\ 2 & a_{ii} = 1 \end{cases}.$$

As this number differs from every r_i at the i-th decimal place, it follows that $r \neq r_i$ for all $i \in \mathbb{N}$. Thus $r \notin \{r_i \mid i \in \mathbb{N}\}$, contradicting the assumption that f is surjective.

2.4 Sequences of Real Numbers

Informally, a sequence is a list of numbers, indexed by $\mathbb{N}$. A more percise definition is as follows

Definition 2.69: Sequence

A SEQUENCE is a function

$$a : \mathbb{N} \longrightarrow \mathbb{R}, n \longmapsto a_n.$$

Here a_n is the n-th element of the sequence. We often write $(a_n)_{n=0}^{\infty}$ or $(a_n)_{n \in \mathbb{N}}$ or $(a_n)_{n \geq 0}$ for the sequence.

In practice, we often use $x_0, x_1, x_2, \dots$ or x_n instead of a_n.

Definition 2.70: Constant Sequences

A sequence is CONSTANT if $x_i = x_j \, \forall i, j$.

A sequence $(x_n)_{n \geq 0}$ is EVENTUALLY CONSTANT if

$$\exists N \in \mathbb{N} : x_i = x_j \, \forall i, j \geq N.$$

Example 2.71: Sequences

$(x_n)_{n \geq 0}$ is $x_0 = 0, x_n = \frac{1}{n}$ for $n \geq 1$.

Another example is $(x_n)_{n \geq 0}, x_n = (-1)^n$.

Tip 2.72:

The latter sequence, $x_n = (-1)^n$, can often be used as a counterexample.

Definition 2.73: Limit of a Sequence

Let $(x_n)_{n \geq 0}$ be a sequence. We say that it CONVERGES if

$$\exists A \in \mathbb{R} : \forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall n \geq N : |x_n - A| < \varepsilon.$$

In this case, $A = \lim_{n \to \infty} x_n$ is called the LIMIT of the sequence.

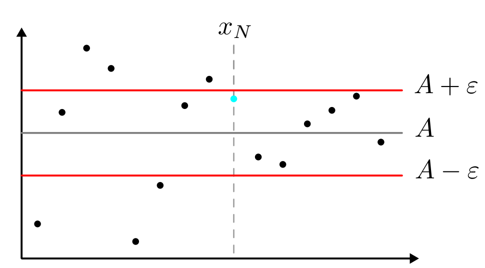


Figure 13: Definition of a Limit

Example 2.74:

$$x_n = \begin{cases} 0 & \text{if } n = 0 \\ \frac{1}{n} & \text{if } n \geq 1 \end{cases}.$$

We claim that $A = 0$ is the limit.

We fix $\varepsilon > 0$. By archimedian principle, $\exists N : 1/N < \varepsilon$. Then $\forall n \geq N$, we have

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

Proposition 2.75:
The limit, if it exists, is unique.

Proof. Assume $(x_n)_{n\geq 0}$ has two limits $A, B \in \mathbb{R}$, with $A \neq B$. Let $\varepsilon = \frac{|A-B|}{3}$. The 3 is here so the delta neighborhoods around A and B do not overlap.

By the definition of the limit, $\exists N_A \in \mathbb{N}$, such that $|x_n - A| < \varepsilon$ for all $n \geq N_A$. Similarly, $\exists N_B \in \mathbb{N}$ such that $|x_n - B| < \varepsilon$ for all $n \geq N_B$.

Take $N = \max(N_A, N_B)$. Then,

$$\begin{aligned} 3\varepsilon &= |A - B| \\ &\leq |A - x_n| + |x_n - B| \\ &< \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

This is a contradiction, thus the limit is unique. □

If we return to the example $x_n = (-1)^n$, we see that this sequence has no limit, as it does not approach any number.

We can define a subsequence which informally speaking just takes some subset of the sequence.

$$x_0, x_1, x_2, \cdots \to x_0, x_1, x_4, x_{20}, x_{79}, \cdots.$$

Definition 2.76: Subsequence
Given $(x_n)_{n\geq 0}$, a SUBSEQUENCE is a sequence of the form

$$(x_{n_k})_{k\geq 0},$$

where n_k is a strictly increasing sequence of natural numbers, that is $n_{k+1} > n_k$ for all k.

Remark 2.77:
Since $n_0 \geq 0$ and $n_{k+1} > n_k$, one can prove by induction that $n_k \geq k$.

Exercise 2.78:
Let $(x_n)_{n\geq 0}$ be a sequence such that $\lim_{n\to\infty} x_n = A$. Then for every subsequence $(x_{n_k})_{k=0}^{\infty}$, show that

$$\lim_{k\to\infty} x_{n_k} = A.$$

Lec 9

Remark 2.79:
A sequence does not need to start at $n = 1$. For example, we can denote

$$\{x_1, x_2, x_3, \cdots \leftrightarrow \{x_n\}_{n=1}^{\infty}\}.$$

This might be useful for sequences as $x_n = \frac{1}{n}, n \geq 1$.

However, there need to be infinitely many elements in the sequence.

Definition 2.80: Accumulation point
$(x_n)_{n=0}^{\infty}$. We say $A \in \mathbb{R}$ is an ACCUMULATION POINT if

$$\forall \varepsilon > 0, \forall N \in \mathbb{N}\, \exists n \geq N : |x_n - A| < \varepsilon.$$

Remark 2.81:
A limit is an accumulation point.

Proposition 2.82: Bolzano-Weierstrass
$A \in \mathbb{R}$ is an accumulation point $\Leftrightarrow \exists$ subsequence $(x_{n_k})_{k=0}^{\infty}$ such that $A = \lim_{k\to\infty} x_{n_k}$.

Proof. $\Rightarrow$: The idea of the proof is to cherry-pick the elements of the sequence that are close to A.

Assume $A \in \mathbb{R}$ is an accumulation point. First, apply the definition of accumulation point with $\varepsilon = 1$ and $N = 1$:

$$\exists n_0 \geq 1 : |x_{n_0} - A| < 1.$$

Second, apply the definition of accumulation point with $\varepsilon = \frac{1}{2}$ and $N = n_0 + 1$:

$$\exists n_1 \geq n_0 + 1 : |x_{n_1} - A| < \frac{1}{2}.$$

In general, given n_{k-1}, apply the definition of accumulation point with $\varepsilon = 2^{-k}$ and $N = n_{k-1} + 1$:

$$\exists n_k \geq n_{k-1} + 1 : |x_{n_k} - A| < 2^{-k}.$$

Now, given $\varepsilon > 0$, take $N \in \mathbb{N}$ such that $2^{-N} < \varepsilon$. Then

$$|x_{n_k} - A| < 2^{-k} \leq 2^{-N} < \varepsilon, \quad \forall k \geq N.$$

Thus $\lim_{k\to\infty} x_{n_k} = A$.

$\Leftarrow$: Fix $\varepsilon > 0$ and $N \in \mathbb{N}$. Since $A = \lim_{k\to\infty} x_{n_k} \exists N_0$ such that

$$|x_{n_k} - A| < \varepsilon, \quad \forall k \geq N_0.$$

Choose $\kappa = \max\{N_0, N\}$. Then $|x_{n_\kappa} - A| < \varepsilon$ and $n_\kappa \geq \kappa \geq N$. □

Corollary 2.83:
$A \in \mathbb{R}$ accumulation point of $(x_n)_{n=0}^{\infty}$. Then $\forall \varepsilon > 0$ infinitely many $n \in \mathbb{N}$ such that

$$x_n \in (A - \varepsilon, A + \varepsilon).$$

Proof. A accumulation point $\Rightarrow \exists (x_{n_k})_{k=0}^{\infty}$ such that $A = \lim_{k\to\infty} x_{n_k}$. Fix $\varepsilon > 0$. Then $\exists N \in \mathbb{N}$ such that $|x_{n_k} - A| < \varepsilon$ for all $k \geq N$. But this means that there are infinitely many elements of the subsequence $(x_{n_k})_{k=0}^{\infty}$ in the interval $(A - \varepsilon, A + \varepsilon)$. □

Example 2.84:
Consider the sequence $x_n = (-1)^n$. Then -1 and 1 are accumulation points.

Proposition 2.85: Limit computations
Consider two sequences $(x_n)_{n=0}^{\infty}$ and $(y_n)_{n=0}^{\infty}$ with limits A, B respectively.

(1) The sequence $(x_n + y_n)_{n=0}^{\infty}$ converges to $A + B$.

(2) The sequence $(x_n \cdot y_n)_{n=0}^{\infty}$ converges to AB.

(3) Given $\alpha \in \mathbb{R}$, the sequence $(\alpha x_n)_{n=0}^{\infty}$ converges to αA.

(4) If $x_n \neq 0 \forall n$ and $A \neq 0$, then the sequence $\left(x_n^{-1}\right)_{n=0}^{\infty}$ converges to A^{-1}.

Proof. **(1)** Fix $\varepsilon > 0$. Since $\lim x_n = A, \lim y_n = B$, $\exists N_A, N_B \in \mathbb{N}$ such that

$$|x_n - A| < \frac{\varepsilon}{2}, \quad \forall n \geq N_A$$

and

$$|y_n - B| < \frac{\varepsilon}{2}, \quad \forall n \geq N_B.$$

Therefore, if $N = \max\{N_A, N_B\}$, then

$$|(x_n + y_n) - (A+B)| \le |x_n - A| + |y_n - B| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \forall n \in \mathbb{N}.$$

In the last step we used the triangle inequality.

(4) Since $A \neq 0$, we can apply the definition of the limit with $\varepsilon = \frac{|A|}{2}$. Then $\exists N_0 \in \mathbb{N}$ such that

$$|x_n - A| < \frac{|A|}{2}, \quad \forall n \ge N_0.$$

Therefore, $|x_n| = |x_n - A + A| \ge |A| - |x_n - A|$ by triangle inequality. Then

$$|A| - |x_n - A| > |A| - \frac{|A|}{2} = \frac{|A|}{2} \forall n \ge N_0.$$

This means that all x_n are bounded below by $\frac{|A|}{2}$ for $n \ge N_0$. We do this so that we dont get too close to 0.

Therefore, for $n \ge N_0$,

$$|x_n^{-1} - A^{-1}| = \frac{|x_n - A|}{|x_n||A|} \le \frac{2}{|A|^2}|x_n - A|.$$

Where we used that $|x_n| \ge \frac{|A|}{2} \Leftrightarrow \frac{1}{|x_n|} \le \frac{2}{|A|}$.

Now, fix $\varepsilon > 0$. Choose $N_1 \in \mathbb{N}$ such that

$$|x_n - A| < \frac{|A|^2}{2}\varepsilon, \quad \forall n \ge N_1.$$

Then for $N = \max\{N_0, N_1\}$

$$|x_n^{-1} - A^{-1}| \le \frac{2}{|A|^2}|x_n - A| < \varepsilon, \quad \forall n \ge N.$$

□

Remark 2.86:

The other direction does NOT hold in general. For example, consider $a_n = (-1)^n, b_n = (-1)^{n+1}$. Then a_n, b_n do not converge, but $a_n + b_n = 0$ converges to 0.

Proposition 2.87: Limits with inequalities

$(x_n), (y_n)$ with limits A, B.

(1) If $A < B$, then $\exists N$ such that $x_n < y_n \forall n \ge N$.

(2) If $\exists N$ such that $x_n \le y_n \forall n \ge N$, then $A \le B$.

The strictness of (1) is important!

The proof to (1) is essentially the same as the proof to the uniqueness of limits.

Proof. Let $\varepsilon = \frac{B-A}{3}$ in order that $A + \varepsilon < B - \varepsilon$. For $n \ge N_A$, we have $|x_n - A| < \varepsilon$ and for $n \ge N_B$, we have $|y_n - B| < \varepsilon$. Therefore, for $N = \max\{N_A, N_B\}$,

$$x_n < A + \varepsilon < B - \varepsilon < y_n, \quad \forall n \ge N.$$

For **(2)**, assume by contradiction that $A > B$. Then we can apply (1) with x_n and y_n reversed to get $\exists N_0 \in \mathbb{N}$ such that $x_n > y_n \forall n \ge N_0$. However, this contradicts $x_n \le y_n \forall n \ge N$. □

Example 2.88:

Consider $x_n = -\frac{1}{n}$ and $y_n = \frac{1}{n}$. Then $\lim x_n = 0 = \lim y_n$. From this, we see that $x_n < y_n \nRightarrow A < B$.

Lemma 2.89: Sandwich lemma

$(x_n), (y_n), (z_n), \exists N$ such that

$$x_n \le y_n \le z_n, \quad \forall n \ge N.$$

Suppose x_n, z_n converge to the same limit $A \in \mathbb{R}$. Then

$$\lim_{n\to\infty} y_n = A.$$

Exercise 2.90:

Compute the following limits:

1. $a_n = \left(1 - \frac{1}{n^2}\right)^n$.
2. $b_n = \frac{n^n}{2^{n^2}}$.
3. $c_n = \sqrt[n]{n^2 - n + 1}$.
4. $d_n = \sum_{k=1}^{n} \frac{k}{n^2+k}$.

Lec 10

Definition 2.91: Bounded Sequences

A sequence $(x_n)_{n\ge 0}$ is BOUNDED if $\exists M > 0$ such that $|x_n| \le M \forall n \in \mathbb{N}$.

Proposition 2.92:

If (x_n) is convergent, then it is bounded.

Proof. Ket $\lim_{n\to\infty} x_n = A$. Fix $\varepsilon = 1$. Then $\exists N$ such that

$$|x_n - A| < 1, \quad \forall n \ge N.$$

Therefore, by triangle inequality,

$$|x_n| = |x_n - A + A| \le |x - A| + |A| < |A| + 1.$$

We define $M := \max\{|x_0|, |x_1|, \cdots, |x_{N-1}|, |A| + 1\}$. Then $|x_n| \le M \forall n \in \mathbb{N}$. □

In general, sequences do not need to be ordered, they can be anywhere. As an order is quite useful, this motivates the following definition.

Definition 2.93: Monotone sequences

(x_n) is

- INCREASING if $m > n \Rightarrow x_m \ge x_n$,
- STRICTLY INCREASING if $m > n \Rightarrow x_m > x_n$,
- DECREASING if $m > n \Rightarrow x_m \le x_n$,
- STRICTLY DECREASING if $m > n \Rightarrow x_m < x_n$.

Theorem 2.94:

A monotone sequence $(x_n)_{n\ge 0}$ converges iff it is bounded.

More percisely, let $X = \{x_0, x_1, \cdots, x_n\} \subseteq \mathbb{R}$. If (x_n) is increasing, then $\lim_{n\to\infty} x_n = \sup X$.

If (x_n) is instead decreasing, then $\lim_{n\to\infty} x_n = \inf X$.

Proof. $\Rightarrow$: Convergent sequences are always bounded. (Prop. 2.92)

$\Leftarrow$: Let (x_n) be increasing and bounded. (x_n) bounded implies that X is bounded, which imples that $A = \sup X$ exists.

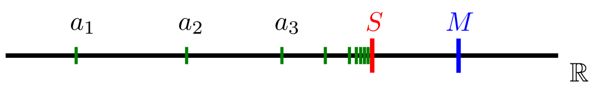


Figure 14: Bounded monotone sequences

By definition of supremum:

(i) $x \leq A \forall n \in \mathbb{N}$ because A is an upper bound of X.

(ii) $\forall \varepsilon > 0, \exists x_N \in X$ such that $x_n > A - \varepsilon$ (Since A is the smallest upper bound).

Then, given $\varepsilon > 0$, by (i) and (ii), we have for $n \geq N$

$$A - \varepsilon < x_N \leq x_n \leq A < A + \varepsilon.$$

Which implies that A is the limit of (x_n). □

Example 2.95: Herons Algorithm

Show that the following sequence can be used to approximate $\sqrt{2}$:

$$x_0 = 2, \quad x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

Solution.

For the limit x, it holds that

$$x = \frac{1}{2}\left(x + \frac{2}{x}\right) \Rightarrow x = \pm\sqrt{2}.$$

We can exclude the negative solution, since $x_n > 0 \,\forall n$.

We now show that the limit exists. We first show that (x_n) is decreasing. We can calculate

$$x_{n+1} - x_n = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) - x_n = \frac{1}{x_n} - \frac{1}{2}x_n = \frac{2 - x_n^2}{2x_n}.$$

We thus need to show that $x_n^2 \geq 2 \,\forall n$. We do this by induction. The base case $n = 0$ is clear. Assume $x_n \geq \sqrt{2}$. Then

$$x_{n+1} - \sqrt{2} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) - \sqrt{2}.$$

Using the induction hypothesis, we have

$$= \frac{x_n^2 - 2\sqrt{2}x_n + 2}{2x_n} = \frac{(x_n - \sqrt{2})^2}{2x_n} \geq 0.$$

In general, we only care about the sequence for large n. This gives rise to the following definition.

Definition 2.96: Tail of a sequence

Given $(x_n)_{n=0}^{\infty}$ and $N \in \mathbb{N}$, the TAIL of the sequence is the sequence $X_{\geq n} = \{x_n, x_{n+1}, \cdots\}$

Remark 2.97:

We can restate the definition of limit using tails of sequences:

$$\lim_{n\to\infty} x_n = A \Leftrightarrow \forall \varepsilon > 0 \,\exists N : X_{\geq N} \subseteq (A - \varepsilon, A + \varepsilon).$$

For simplicity we use the following notation:

$$\sup_{k\geq n} x_k \text{ for } \sup(X_{\geq n}) \text{ and } \inf_{k\geq n} x_k \text{ for } \inf(X_{\geq n}).$$

Let $(x_n)_{n\geq 0}$ be a bounded sequence, we define

$$s_n = \sup_{k\geq n} x_k \text{ and } i_n = \inf_{k\geq n} x_k.$$

If $m > n$, means that $X_{\geq m} \subseteq X_{\geq n}$, as we throw away more stuff. This implies that $s_m \leq s_n$ and $i_m \geq i_n$. So we have

$$i_n \leq i_m \leq s_m \leq s_n, \quad \forall m > n.$$

Thus, $(s_n)_{n\geq 0}$ is decreasing, $(i_n)_{n\geq 0}$ is increasing. since (x_n) is bounded, both (s_n) and (i_n) are bounded, this motivates to the following definition.

Definition 2.98: Superior and Inferior Limit

Let $(x_n)_{n\geq 0}$ be a bounded sequence. The numbers

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} \underbrace{\left(\sup_{k\geq n} x_k\right)}_{=s_n}$$

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty} \underbrace{\left(\inf_{k\geq n} x_k\right)}_{=i_n}$$

are called SUPERIOR LIMIT and INFERIOR LIMIT.

Also because $i_n \leq s_n$ it holds that

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

Lemma 2.99:

A bounded sequence $(x_n)_{n\geq 0}$ converges $\Leftrightarrow$

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

Proof. We use the notation of i_n and s_n from before. Denote

$$I = \liminf_{n\to\infty} x_n = \lim_{n\to\infty} i_n \text{ and } S = \limsup_{n\to\infty} x_n = \lim_{n\to\infty} s_n.$$

$\Leftarrow$: Assume $I = S$. Note that $i_n \leq x_n \leq s_n \forall n \geq 0$. Thus, the sandwich lemma yields, $I = S \Rightarrow \lim x_n = I = S$.

$\Rightarrow$: Let $\lim_{n\to\infty} x_n = A$. Then, given $\varepsilon > 0$, $\exists N \in \mathbb{N}$ such that

$$A - \varepsilon < x_n < A + \varepsilon, \quad \forall n \geq N.$$

Therefore, for $n \geq N$,

$$A - \varepsilon \leq i_n \leq s_n \leq A + \varepsilon.$$

It follows that $A - \epsilon \leq I \leq S \leq A + \varepsilon$. Because this is true for all $\varepsilon > 0$, we have $I = S = A$. □

Theorem 2.100:

$(x_n)_{n\geq 0}$ be a bounded sequence, $A = \limsup_{n\to\infty} x_n$. Then A is an accumulation point of (x_n).

Furthermore, $\forall \varepsilon > 0$,

(1) only finitely many elements x_n satisfy

$$x_n \geq A + \varepsilon.$$

(2) infinitely many elements satisfy

$$A - \varepsilon < x_n < A + \varepsilon.$$

Proof. Let $s_n = \sup_{k\geq n} x_k$. Then $A = \lim_{n\to\infty} s_n$, furthermore, (s_n) is decreasing.

Then, $\forall \varepsilon > 0 \exists N_0 \in \mathbb{N}$ such that

$$A \leq s_n < A + \varepsilon, \quad \forall n \geq N_0.$$

We want to prove, that A is an accumulation point. Fix $\varepsilon > 0$, fix $N \in \mathbb{N}$. take N_0 as before and define $N_1 = \max\{N, N_0\}$.

Since $s_{N_1} = \sup_{k\geq N_1} x_k$, $\exists n_1 \geq N_1$ such that

$$s_{N_1} - \varepsilon < x_{n_1} \leq s_{N_1}.$$

Then, $A - \varepsilon \leq s_{N_1} - \varepsilon < x_{n_1} \leq s_{N_1} < A + \varepsilon$.

This shows, that A is an accumulation point.

From the definition of N_0, we have

$$x_n \leq s_n < A + \varepsilon, \quad \forall n \geq N_0.$$

Thus only finitely many elements can be $\geq A + \varepsilon$ (at most N_0).

Finally, because A is an accumulation point, there exists a subsequence $(x_{n_k})_{k=0}^{\infty}$ such that $\lim_{k\to\infty} x_{n_k} = A$. Therefore, $\forall \varepsilon > 0$, $\exists N_2 \in \mathbb{N}$ such that

$$A - \varepsilon < x_{n_k} < A + \varepsilon, \quad \forall k \geq N_2.$$

□

Corollary 2.101:
Every bounded sequence has at least one accumulation point and therefore posesses a convergent subsequence.

Proof. $(x_n)_{n\geq 0}$ bounded $\Rightarrow$ $A = \limsup_{n\to\infty} x_n$ exists and is an accumulation point. Furthermore, every accumulation point is the limit of a subsequence.

Lec 11

□

Definition 2.102: Cauchy sequence
A sequence $(x_n)_{n\geq 0}$ is a CAUCHY SEQUENCE if

$$\forall \varepsilon > 0, \exists N : |x_n - x_m| < \varepsilon, \quad \forall n, m \geq N.$$

Lemma 2.103:
$(x_n)_{n\geq 0}$ cauchy $\Rightarrow$ (x_n) is bounded.

Proof. Fix $\varepsilon = 1$. Then $\exists N$ such that

$$|x_n - x_m| < 1, \quad \forall n, m \geq N.$$

Therefore, choosing $m = N$, we have

$$|x_n - x_N| < 1, \quad \forall n \geq N.$$

By triangle inequality, it follows that

$$|x_n| = |x_n - x_N + x_N| \leq |x_n - x_N| + |x_N| < 1 + |x_N|, .$$

Choose $M = \max\{|x_0|, |x_1|, \cdots, |x_{N-1}|, 1 + |x_N|\}$. Thus

$$|x_n| \leq M, \quad \forall n \in \mathbb{N}.$$

□

Theorem 2.104:
$(x_n)_{n\geq 0}$ is cauchy $\Leftrightarrow$ (x_n) is convergent.

Proof. $\Leftarrow$: Assume $\lim_{n\to\infty} x_n = A$. Fix $\varepsilon > 0$. Since x_n converges to A, $\exists N$, such that

$$|x_n - A| < \frac{\varepsilon}{2}, \quad \forall n \geq N.$$

Therefore, $\forall n, m \geq N$,

$$|x_n - x_m| = |x_n - A + A - x_m| \leq |x_n - A| + |x_m - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves that convergent series are cauchy.

$\Rightarrow$: If (x_n) is cauchy, we know that (x_n) is bounded, this implies that

$$\exists (x_{n_k})_{k\geq 0} : \lim_{k\to\infty} x_{n_k} = A.$$

Since (x_n) is cauchy, $\exists N_0$ such that

$$|x_n - x_m| < \frac{\varepsilon}{2}, \quad \forall n, m \geq N_0.$$

Since $lim_{k\to\infty} x_{n_k} = A$, $\exists N_1$ such that

$$|x_{n_k} - A| < \frac{\varepsilon}{2}, \quad \forall k \geq N_1.$$

Let $N = \max\{N_0, N_1\}$. Then, since $n_N \geq N$, $\forall n \geq N$,

$$|x_n - A| \leq |x_n - x_{n_N}| + |x_{n_N} - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

It is important to note that we need to have the inequality in the definition for all pairs, not just for consecutive terms.

Definition 2.105: Divergence
$(x_n)_{n\geq 0}$ DIVERGES to $+\infty$ if

$$\forall M > 0 \, \exists N : x_n > M, \quad \forall n \geq N.$$

Similarly, (x_n) DIVERGES to $-\infty$ if

$$\forall M > 0 \, \exists N : x_n < -M, \quad \forall n \geq N.$$

The sequence $0, 1, -2, 3, -4, 5, -6, \cdots$ does not diverge nor converge by our definition.

Let (x_n) be a sequence, also unbounded. We define

$$\limsup_{n\to\infty} x_n = \begin{cases} \text{finite number} & \text{if } (x_n) \text{ is bounded} \\ +\infty & \text{if } (x_n) \text{ is not bdd above} \end{cases}.$$

2.5 Complex sequences

A sequence of complex numbers is the same as before, but now each number is complex:

$$z_n = x_n + iy_n, \quad x_n, y_n \in \mathbb{R}.$$

Definition 2.106: Convergence
A complex sequence $(z_n)_{n\geq 0}$ converges if

$$(x_n)_{n\geq 0}, (y_n)_{n\geq 0} \text{ both converge.}$$

In particular, if $\lim x_n = A$ and $\lim y_n = B$, then

$$\lim_{n\to\infty} z_n = A + iB.$$

We say that (z_n) diverges if $\{|z_n|\}_{n\geq 0}$ diverges to $+\infty$.

Most ideas from real sequences carry over to complex sequences. For example, a subsequence of (z_n) is (z_{n_k}).

3 Functions of one Real Variable

3.1 Real Valued Functions

In the following we work with functions $f : D \subseteq \mathbb{R} \to \mathbb{R}$.

For two functions f_1, f_2 we define

- $(f_1 + f_2)(x) = f_1(x) + f_2(x)$,
- $(\alpha f_1)(x) = \alpha f_1(x)$,
- $(f_1 \cdot f_2)(x) = f_1(x) \cdot f_2(x)$,

We call a point $x \in D$ a ZERO of f if $f(x) = 0$. The ZERO SET of f is $\{x \in D : f(x) = 0\}$.

We say that $f_1 \leq f_2$ if $f_1(x) \leq f_2(x) \forall x \in D$. Similarly we define $f_1 < f_2$. This defines an order relation.

We call f NON-NEGATIVE if $f \geq 0$ and POSITIVE if $f > 0$.

Definition 3.1: Bounded Functions

$f : D \to \mathbb{R}$ is BOUNDED FROM ABOVE if $\exists M > 0$ such that

$$f(x) \leq M, \quad \forall x \in D.$$

Similarly, f is BOUNDED FROM BELOW if $\exists M > 0$ such that

$$f(x) \geq -M, \quad \forall x \in D.$$

Finally, f is BOUNDED if

$$|f(x)| \leq M, \quad \forall x \in D.$$

Similar to sequences, we can define monotonicity for functions.

Definition 3.2: Monotonicity

Consider $f : D \to \mathbb{R}$ and $x, y \in D$.

- f is INCREASING if $x < y \Rightarrow f(x) \leq f(y)$.
- f is STRICTLY INCREASING if $x < y \Rightarrow f(x) < f(y)$.
- f is DECREASING if $x < y \Rightarrow f(x) \geq f(y)$.
- f is STRICTLY DECREASING if $x < y \Rightarrow f(x) > f(y)$.

Example 3.3: Parabola

$f(x) = x^2$ is strictly increasing on $D = [0, 1]$. However, it is strictly decreasing on $D = [-1, 0]$. Similarly, it is neither increasing nor decreasing on $D = [-1, 1]$.

We notice that the notion of monotonicity only makes sense when the domain is specified.

Example 3.4: Floor Function

The integer part (or floor) function is increasing on the domain $\mathbb{R}$, yet not strictly increasing.

If we have a constant function, we also write $f(x) \equiv c$. This is an example of a increasing and decreasing function.

Definition 3.5: Continuity

$D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$. We call f CONTINUOUS at $x_0 \in D$ if $\forall \varepsilon > 0, \exists \delta > 0$ such that

$$\forall x \in D : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

f is called CONTINUOUS if it is continuous at every point $x_0 \in D$.

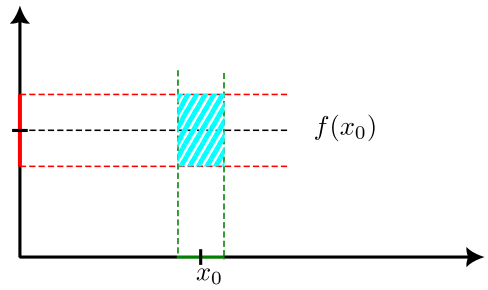


Figure 15: Epsilon-Delta definition of continuity.

Remark 3.6:

It is enough to consider ε small. More percisely, assume that $\exists \varepsilon_0$ such that $\forall \varepsilon \in (0, \varepsilon_0] \exists \delta$ such that

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Then if $\varepsilon > \varepsilon_0$, just take for δ as the one for ε_0.

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon_0 < \varepsilon.$$

Lec 12

Example 3.7: Linear Functions

$f(x) = ax + b$. Show that f is continuous for all $x_0 \in \mathbb{R}$.

Solution. Fix x_0, If $a = 0$ then $f(x)$ is constant which is continuous. Indeed, given $\varepsilon > 0$, just choose $\delta = 1$. Then

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| = |b - b| = 0 < \varepsilon.$$

If $a \neq 0$, given $\varepsilon > 0$, we want that

$$|x - x_0| < \delta \Rightarrow |ax + b - (ax_0 + b)| = |a||x - x_0| < \varepsilon.$$

Choose $\delta = \frac{\varepsilon}{|a|} > 0$. Then

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < |a||x - x_0| < |a|\delta = \varepsilon.$$

Example 3.8: Absolute Value Function

$x \to |x|$ is continuous for all $x_0 \in \mathbb{R}$.

Solution. Given $\varepsilon > 0$, we want to find $\delta > 0$ such that

$$|x - x_0| < \delta \Rightarrow ||x| - |x_0|| < \varepsilon.$$

Choose $\delta = \varepsilon$. Then, by the reverse triangle inequality,

$$|x - x_0| < \delta \Rightarrow ||x| - |x_0|| \leq |x - x_0| < \delta = \varepsilon.$$

Example 3.9: Integer part function

The function $f(x) = \lfloor x \rfloor$ is not continuous at the integers. This can be seen for example as $\lfloor 3 \rfloor = 3$ but

$$\lfloor 3 - \delta \rfloor = 2, \quad \forall \delta \in (0,1).$$

Thus,

$$|f(3) - f(3-\delta)| = |3-2| = 1 \, \forall \delta \in (0,1).$$

Exercise 3.10:

Show that $f(x) = x^2$ is continuous for all $x_0 \in \mathbb{R}$ using the $\varepsilon - \delta$ definition.

Note: $\delta = \delta(\varepsilon, x_0)$.

Definition 3.11: Restriction

Given $f : D \to \mathbb{R}$, $D' \subseteq D$. Then $f|_{D'} : D' \to \mathbb{R}$ is the RESTRICTION of f to D' where $f|_{D'}(x) = f(x)$ for all $x \in D'$.

We look at f and $f|_{D'}$ as diffrent functions.

For example, $\lfloor x \rfloor$ is not continuous on $\mathbb{R}$, but it is continuous on $D' = (0,1)$. It is also continuous on $D'' = (1,2) \cup (2,3)$.

Proposition 3.12: Sum and Products of Continuous Functions

Given $f_1, f_2 : D \to \mathbb{R}$ both beeing continuous at $x_0 \in D$. Then:

1. $f_1 + f_2$ is continuous at x_0.
2. $f_1 \cdot f_2$ is continuous at x_0.
3. af_1 is continuous at x_0 for all $a \in \mathbb{R}$.

This implies that the set of continuous functions $C^0(D)$ is a vector space.

Proof. 1. Fix $\varepsilon > 0$. Since f_1, f_2 are continuous at x_0, $\exists \delta_1, \delta_2 > 0$ such that

$$x \in D, |x - x_0| < \delta_1 \Rightarrow |f_1(x) - f_1(x_0)| < \frac{\varepsilon}{2}$$
$$x \in D, |x - x_0| < \delta_2 \Rightarrow |f_2(x) - f_2(x_0)| < \frac{\varepsilon}{2}$$

If now we choose $\delta = \min\{\delta_1, \delta_2\}$, then

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| \leq |f_1(x) - f_1(x_0)| + |f_2(x) - f_2(x_0)| < \varepsilon.$$

2. We want to estimate $|f_1(x)f_2(x) - f_1(x_0)f_2(x_0)|$. We rewrite this as

$$|f_1(x)f_2(x) - f_1(x_0)f_2(x) + f_1(x_0)f_2(x) - f_1(x_0)f_2(x_0)|.$$

Applying triangle inequality, we have

$$\begin{aligned} &|f_1(x)f_2(x) - f_1(x_0)f_2(x_0)| \\ &\leq |f_1(x) - f_1(x_0)||f_2(x)| + |f_2(x) - f_2(x_0)||f_1(x_0)| \end{aligned}$$

We apply continuity to f_2 with $\varepsilon = 1$. Thus $\exists \delta_0 > 0$ such that

$$|x - x_0| < \delta_0 \Rightarrow |f_2(x) - f_2(x_0)| < 1 \Rightarrow |f_2(x)| < 1 + |f_2(x_0)|.$$

Apply again continuity to f_1, f_2

$$\exists \delta_1 : |x - x_0| < \delta_1 \Rightarrow |f_1(x) - f_1(x_0)| < \frac{\varepsilon}{2(1 + |f_2(x_0)|)}$$
$$\exists \delta_2 : |x - x_0| < \delta_2 \Rightarrow |f_2(x) - f_2(x_0)| < \frac{\varepsilon}{2(1 + |f_1(x_0)|)}.$$

For the second inequality we use $+1$ in the denominator to avoid division by zero. Finally, choose $\delta = \min\{\delta_0, \delta_1, \delta_2\}$. Then, $\forall x \in D$ with $|x - x_0| < \delta$,

$$|f_1(x)f_2(x) - f_1(x_0)f_2(x_0)| < \frac{\varepsilon \cdot |f_2(x)|}{2(1 + |f_2(x_0)|)} + \frac{\varepsilon \cdot |f_1(x_0)|}{2(1 + |f_1(x_0)|)} < \varepsilon.$$

3. Apply the second point to $f_2 \equiv a$. □

We use the following notation:

$$\sum_{j=0}^{N} a_j = a_0 + a_1 + \cdots + a_N.$$

Remark 3.13: Polynomials are continuous

Given a POLYNOMIAL $f(x) = \sum_{j=0}^{N} a_j x^j$, where $x^j = \underbrace{x \cdot x \cdots x}_{j \text{ times}}$. Then f is a continuous function.

Given a function $f : X \to Y$ and $g : Y \to Z$, we can $g \circ f : X \to Z$, where $(g \circ f)(x) = g(f(x))$. This is called the COMPOSITION of f and g.

Proposition 3.14: Continuity of Composition

Given $f : D_1 \to D_2$ continuous at x_0, $g : D_2 \to \mathbb{R}$ continuous at $f(x_0)$. Then $g \circ f : D_1 \to \mathbb{R}$ is continuous at x_0.

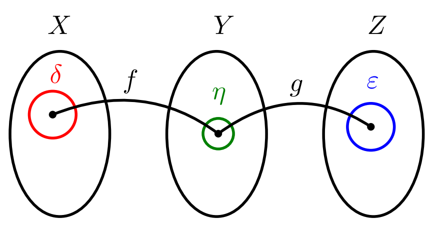


Figure 16: Composition of functions.

Proof. Fix $\varepsilon > 0$. Since g is continuous at $f(x_0)$, $\exists \eta > 0$ such that

$$y \in D_2, |y - f(x_0)| < \eta \Rightarrow |g(y) - g(f(x_0))| < \varepsilon.$$

Since f is continuous at x_0, $\exists \delta > 0$ such that

$$x \in D_1, |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \eta.$$

Combining both, we have

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \eta \Rightarrow |g(f(x)) - g(f(x_0))| < \varepsilon.$$

□

Exercise 3.15:

Show the following:

1. $f; \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{1}{x}$ is continuous,
2. $g : \mathbb{R} \to \mathbb{R}$, $g(x) \neq 0$ is continuous, then $\frac{1}{g(x)}$ is continuous,
3. $g, h : \mathbb{R} \to \mathbb{R}$ continuous, $g(x) \neq 0$, then $\frac{h(x)}{g(x)}$ is continuous.

As a notation, instead of writing $\lim_{n \to \infty} x_n = \overline{x}$, we can write $x_n \to \overline{x}$.

Theorem 3.16: Continuity = Sequential Continuity

$f : D \to \mathbb{R}$, $x \in D$. f is continuous at $\overline{x} \Leftrightarrow$

$$\forall (x_n)_{n\geq 0} \subseteq D, \text{ if } x_n \to \overline{x} \text{ then } f(x_n) \to f(\overline{x}).$$

Proof. $\Rightarrow$: Given $\varepsilon > 0\, \exists \delta > 0$ such that

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Since $x_n \to \overline{x}$, $\exists N$ such that

$$|x_n - \overline{x}| < \delta, \quad \forall n \geq N.$$

Then,

$$\forall n \geq N : |f(x_n) - f(\overline{x})| < \varepsilon.$$

So $f(x_n) \to f(\overline{x})$.

$\Leftarrow$: Assume that f is not continuous at $\overline{x}$. Then $\exists \varepsilon_0 > 0$ such that:

$$\forall \delta > 0, \exists x : |x - \overline{x}| < \delta \text{ but } |f(x) - f(\overline{x})| \geq \varepsilon_0.$$

Apply this with $\delta = 2^{-n}$ to find x_n such that

$$|x_n - \overline{x}| < 2^{-n} \text{ but } |f(x_n) - f(\overline{x})| \geq \varepsilon_0.$$

Lec 13

Then $x_n \to \overline{x}$ but $f(x_n) \not\to f(\overline{x})$. □

Remark 3.17:

If f is not continuous at $\overline{x} \exists \varepsilon > 0$ and $(x_n)_{n\geq 0}$ such that

$$x_n \to \overline{x} \text{ but } |f(x_n) - f(\overline{x})| \geq \varepsilon.$$

3.2 Continuous Functions

The following theorem is what basiclly guarantees that continuous functions can be drawn without lifting the pen.

Theorem 3.18: Intermediate Value Theorem

Given $f : [a, b] \to \mathbb{R}$ continuous. Assume $f(a) < f(b)$. Then

$$\forall c \in [f(a), f(b)]\, \exists \overline{x} \in [a, b] : f(\overline{x}) = c.$$

Proof. Define $X = \{x \in [a, b] \mid f(x) \leq c\}$. Then X is bounded since $X \subseteq [a, b]$. Also, $X \neq \varnothing$ since $a \in X$. Thus, $\exists \overline{x} = \sup(X)$.

The goal is to show that $f(\overline{x}) = c$.

1. By definition of the supremum, $\forall n \in \mathbb{N}\, \exists x_n \in [\overline{x} - 2^{-n}, \overline{x}]$ such that $f(x_n) \leq c$. This follows as the supremum is the least upper bound.

Since $x_n \to \overline{x}$, by continuity of f, $f(x_n) \to f(\overline{x})$. Since all $f(x_n) \leq c$, we have $f(\overline{x}) \leq c$.

2. Assume by contradiction that $f(\overline{x}) < c$. Define $\varepsilon = c - f(\overline{x}) > 0$. By continuity, $\exists \delta > 0$ such that

$$|x - \overline{x}| < \delta \Rightarrow |f(x) - f(\overline{x})| < \varepsilon.$$

As $|a| < b \Rightarrow -b < a < b$, we have

$$f(x) < f(\overline{x}) + \varepsilon = c.$$

Consider the two cases:

Case 1: $c = f(b)$. In that case, just take $\overline{x} = b$.

Case 2: $c < f(b)$. Then since $f(\overline{x}) \leq c$, then $\overline{x} < b$. But then $(\overline{x}, \overline{x} + \delta) \cap (\overline{x}, b) \neq \varnothing$. and there, as we have shown, $f(x) < c$. But this means

$$(\overline{x}, \overline{x} + \delta) \cap (\overline{x}, b) \subseteq X.$$

This shows, that there is a point to the right of $\overline{x}$ in X. However, this contradicts the fact that $\overline{x} = \sup(X)$. □

Remark 3.19:

The same holds, if $f(a) \geq f(b)$, with $c \in [f(b), f(a)]$.

Claim 3.20:

Given $f : [a, b] \to [a, b]$ continuous. Then f has a fixpoint i.e. $f(\overline{x}) = \overline{x}$ for some $\overline{x} \in [a, b]$.

Proof. Let $h(x) = f(x) - x$. Notice that h is continuous as it is the sum of two continuous functions. Also, zeros of h are fixpoints of f. We distinguish three cases:

1. If $f(a) = a$ then a is a fixpoint.

2. If $f(b) = b$ then b is a fixpoint.

3. If $f(a) > a$ and $f(b) < b$, then $h(a) > 0$ and $h(b) < 0$. By the intermediate value theorem, $\exists \overline{x} \in [a, b]$ such that $h(\overline{x}) = 0$. Thus, $f(\overline{x}) = \overline{x}$. □

Example 3.21:

Given $f : \mathbb{R} \to \mathbb{R}$ continuous and periodic with period 4, i.e.

$$\forall x \in \mathbb{R}, f(x + 4) = f(x).$$

Show that $\exists y \in [0, 2]$ such that $f(y) = f(y + 2)$.

Solution. Let $h(x) = f(x) - f(x + 2)$. Then h is continuous as it is the sum of two continuous functions. Looking at $h(0)$ and $h(2)$, we see

$$h(0) = f(0) - f(2)$$
$$h(2) = f(2) - f(4) = f(2) - f(0) = -h(0).$$

We have two cases: 1. If $h(0) = h(2) = 0$, then $f(0) = f(2)$ and we are done.

2. Either $h(0) > 0$ and thus $h(2) < 0$ or $h(0) < 0$ and thus $h(2) > 0$. In both cases, by the intermediate value theorem, $\exists y \in [0, 2]$ such that $h(y) = 0$. Thus, $f(y) = f(y + 2)$.

Tip 3.22:

To apply the intermediate value theorem, it is often useful to define a new function as the difference of two functions.

Definition 3.23: Inverse Function

Given: $f : X \to Y$ bijective, $\forall y \in Y \exists! x \in X : f(x) = y$. Then define $g(y) = x$. Then $g : Y \to X$ and

$$g \circ f : X \to X, g \circ f = \mathrm{id}_X, f \circ g : Y \to Y, f \circ g = \mathrm{id}_Y.$$

g is called the INVERSE of f and is denoted by f^{-1}.

Theorem 3.24: Inverse Function Theorem

Given $f : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, such that f is continuous and strictly monotone.

Then $f(I)$ is an interval and $f : I \to f(I)$ has a continuous, strictly monotone inverse $f^{-1} : f(I) \to I$.

Proof. We can assume f is strictly increasing (otherwise, consider $-f$). Let $J = f(I)$.

- f stictly increasing ($x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$) $\Rightarrow f$ is injective.

- So $f : I \to J$ is bijective by the definition of the image. Thus, $\exists g = f^{-1} : J \to I$ which is the inverse of f.

1. g is strictly increasing. Indeed,

$$x_1 < x_2 \Leftrightarrow \underbrace{f(x_1)}_{y_1} < \underbrace{f(x_2)}_{y_2} \Leftrightarrow g(y_1) < g(y_2).$$

2. For J to be an interval, $\forall y_1, y_2 \in J, y_1 < y_2$ implies $[y_1, y_2] \subseteq J$.

Fix $y_1 < y_2 \in J$. Then $\exists x_1 < x_2 : f(x_1) = y_1, f(x_2) = y_2$. Let $c \in [y_1, y_2]$. By the intermediate value theorem applied to $f : [x_1, x_2] \to \mathbb{R}$, $\exists \overline{x} \in [x_1, x_2]$ such that $f(\overline{x}) = c$. Thus, $c \in J$. So $[y_1, y_2] \subseteq J$.

This proves that J is an interval.

3. By contradiction, assume that $g = f^{-1}$ is not continuous at some point $\overline{y} \in J$. By the remark 3.17, $\exists \varepsilon > 0$ and $(y_n)_{n \geq 0} \subseteq J$ such that

$$y_n \to \overline{y} \text{ but } |g(y_n) - g(\overline{y})| \geq \varepsilon.$$

Let $x_n = g(y_n)$, $\overline{x} = g(\overline{y})$. Then $y_n = f(x_n) \to \overline{y} = f(\overline{x})$ and $|x_n - \overline{x}| \geq \varepsilon$.

Assume that there are infinitely many x_n such that $x_n \leq \overline{x} - \varepsilon$. Considering theese elements, we get a subsequence $(x_{n_k})_{k \geq 0}$ such that

$$f(x_{n_k}) \to f(\overline{x}) \text{ but } x_{n_k} \leq \overline{x} - \varepsilon.$$

Thus we have by strict monotonicity

$$f(\overline{x}) < -f(x_{n_k}) \leq f(\overline{x} - \varepsilon) < f(\overline{x}).$$

But then $f(\overline{x}) < f(\overline{x})$, which is a contradiction. □

Remark 3.25:

We know that $x \mapsto x^n$ is continuous and on $[0, \infty) \mapsto [0, \infty)$, they are strictly increasing. Thus $\exists$ inverse

$$\sqrt[n]{\cdot} : [0, \infty) \longrightarrow [0, \infty)$$
$$y \longmapsto \sqrt[n]{y}.$$

Furthermore, $x \mapsto x^{\frac{m}{n}} = \underbrace{\sqrt[n]{x} \cdots \sqrt[n]{x}}_{m \text{ times}}$ is continuous on $[0, \infty)$.

Also for $x > 0$, $x^{-\frac{m}{n}} = \frac{1}{x^{\frac{m}{n}}}$ is continuous.

To draw an inverse of a function, we can reflect the graph of the function along the line $y = x$.

Example 3.26:

Show that $f(x) = (x-5)(x-3)(x-2)(x+1)(x+2)+1$ has 5 zeros.

Solution. Since f is a polynomial, it is continuous. Also, note that for example $f(-5) < 0$, $f(-1.5) > 0$ and $f(-0.5) < 0$. Thus by intermediate value theorem, $\exists x_1 \in [-5, -1.5]$ such that $f(x_1) = 0$ and $\exists x_2 \in [-1.5, -0.5]$ such that $f(x_2) = 0$. Similarly, we can find the existance of 3 more zeros.

3.3 Continuous Functions on Compact Intervals

Lec 14

We like to show some properties of continuous functions on compact intervals.

Definition 3.27: Compact Interval

A bounded, closed interval $[a, b]$ is called COMPACT.

For example, $[a, \infty]$ is closed yet not bounded, so not compact.

Lemma 3.28:

Let $(x_n)_{n \geq 0} \subseteq [a, b]$ be a sequence. Then $\exists$ a subsequence $(x_{n_k})_{k \geq 0}$ such that $x_{n_k} \to \overline{x} \in [a, b]$.

Proof. Since $[a, b]$ is bounded, this implies that $(x_n)_{n \geq 0}$ is bounded. Thus, $\exists (x_{n_k})_{k \geq 0}$ converges to some $\overline{x} \in \mathbb{R}$, because every bounded sequence has a convergent subsequence ($\limsup$ for example). Furthermore, because $a \leq x_{n_k} \leq b$, we have $a \leq \overline{x} \leq b$. □

Tip 3.29:

When working with compact intervals, the typical strategy is to proof by contradiction and then find a convergent subsequence.

Theorem 3.30:

Given $f : [a, b] \to \mathbb{R}$ continuous. Then f is bounded.

Proof. By contradiction, assume f is not bounded. Then $\forall n \in \mathbb{N}, \exists x_n \in [a, b] : |f(x_n)| \geq n$.

By the lemma, $\exists (x_{n_k})_{k \geq 0}$ such that $x_{n_k} \to \overline{x} \in [a, b]$.

So $|f(x_{n_k})| \geq n_k \to \infty$. But since f is continuous, also $|f|$ is continuous. Then $f(x_{n_k}) \to f(\overline{x})$, which is a real number. ↯ □

Definition 3.31: Maximum & Minimum Value

Given $f : D \to \mathbb{R}$. We say that f takes its MAXIMUM VALUE at x_0 if

$$f(x) \leq f(x_0) \quad \forall x \in D.$$

We call $f(x_0)$ the MAXIMUM of f.

The same for MINIMUM VALUE and MINIMUM.

Theorem 3.32:

If $f : [a, b] \to \mathbb{R}$ is continuous, then f takes its maximum and minimum value.

Proof. We just proof it for the maximum. Since f is bounded, $f([a, b])$ is bounded. Thus, $\exists S = \sup(f([a, b]))$.

$\forall n \in \mathbb{N}, \exists y_n \in f([a, b])$ such that $S - 2^n \leq y_n \leq S$. Since y_n is in the image of f, $\exists x_n \in [a, b]$ such that $f(x_n) = y_n$.

Let $x_{n_k} \to \overline{x} \in [a, b]$ be a convergent subsequence. Since $S - 2^{-n_k} \leq f(x_{n_k}) = y_{n_k} \leq S$. Since f is continuous, we get $f(x_{n_k}) \to f(\overline{x})$. Thus,

$$S \leq f(\overline{x}) \leq S \Rightarrow f(\overline{x}) = S.$$

□

For all of the above properties, it is crucial that the interval is compact. As a conterexample, consider $f : (0, 1) \to \mathbb{R}, x \mapsto \frac{1}{x}$.

We recall the definition of a continuous function on a set $D \subseteq \mathbb{R}$:

$$\forall x_0 \in D, \forall \varepsilon > 0, \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)|.$$

A diffrent thing is uniform continuity:

Definition 3.33: Uniform Continuity

$f : D \to \mathbb{R}$ is UNIFORMLY CONTINUOUS if $\forall \varepsilon > 0, \exists \delta > 0$ such that

$$\forall x, y \in D, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

The difference to regular continuity is that δ does not depend on x_0.

Example 3.34:

Consider $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$. Then f is continuous but not uniformly continuous.

Solution. Indeed, take $\varepsilon = \frac{1}{2}$ for example. We have to show that $\forall \delta > 0, \exists x, y \in \mathbb{R}$ such that

$$|x - y| < \delta \text{ but } |f(x) - f(y)| \geq \varepsilon.$$

Now take x and $y = x + \frac{\delta}{2}$. Then, plugging in, we get $f(x) - f(y) = \delta x + \frac{\delta^2}{4}$. For $x = \frac{1}{\delta}$, we get

$$|f(x) - f(y)| = 1 + \frac{\delta^2}{4} > \frac{1}{2}.$$

Theorem 3.35: Heine-Cantor Theorem

If $f : [a, b] \to \mathbb{R}$ is continuous, then f is uniformly continuous.

Proof. If f is not uniformly continuous, then $\exists \varepsilon > 0$ such that $\forall \delta > 0, \exists x, y \in [a, b]$ such that

$$|x - y| < \delta \text{ but } |f(x) - f(y)| \geq \varepsilon.$$

For $n \in \mathbb{N}$, apply with $\delta = 2^{-n}$. Then $\exists x_n, y_n \in [a, b]$ such that

$$|x_n - y_n| < 2^{-n} \text{ but } |f(x_n) - f(y_n)| \geq \varepsilon.$$

Let $(x_{n_k})_{k \geq 0}$ be a subsequence of $(x_n)_{n \geq 0}$ with $x_{n_k} \to \overline{x}$.

Note:

$$|y_{n_k} - \overline{x}| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - \overline{x}| < 2^{-n_k} + |x_{n_k} - \overline{x}| \stackrel{k \to \infty}{\longrightarrow} 0.$$

Thus, also $y_{n_k} \to \overline{x}$. Since f is continuous, it follows that $f(x_{n_k}) \to f(\overline{x})$ and $f(y_{n_k}) \to f(\overline{x})$. Thus

$$\varepsilon \leq |f(x_{n_k}) - f(y_{n_k})| \leq |f(x_{n_k}) - f(\overline{x})| + |f(\overline{x}) - f(y_{n_k})| \to 0.$$

□

3.4 Exponential and Logarithmic Functions

We start with a lemma.

Lemma 3.36: Bernoulli inequality

$\forall a \in \mathbb{R}, a \geq -1, n \geq 1 \Rightarrow (1 + a)^n \geq 1 + na.$

Proof. We proof by induction. The base case $n = 1$ is clear. Assume the statement holds for some $n \geq 1$. Then

$$\begin{aligned}(1 + a)^{n+1} &= (1 + a)^n (1 + a) \\ &\geq (1 + na)(1 + a) \\ &= 1 + (n + 1)a + na^2 \\ &\geq 1 + (n + 1)a.\end{aligned}$$

□

The goal is to define $e^x = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n$.

Proposition 3.37:

Fix $x \in \mathbb{R}$. Then the sequence $(a_n)_{n \geq 1}$ where

$$a_n = \left(1 + \frac{x}{n}\right)^n$$

is convergent and its limit is positive.

Lemma 3.38:

Fix $x \in \mathbb{R}$, let $n_0 > -x$. Then the sequence $(a_n)_{n \geq n_0}$ is increasing.

Proof. We have

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{\left(1 + \frac{x}{n+1}\right)^{n+1}}{\left(1 + \frac{x}{n}\right)^n} \\ &= \left(1 + \frac{x}{n+1}\right)^n \cdot \frac{1 + \frac{x}{n+1}}{1 + \frac{x}{n}} \\ &= \left(\frac{n + x}{n}\right) \cdot \left(1 - \frac{x}{(n + 1)(n + x)}\right)^{n+1}\end{aligned}$$

.

Fir $n \geq n_0$, we have $\frac{x}{(n+1)(n+x)} \leq \frac{x+n}{(n+1)(n+x)} < 1$.

This allows us to apply bernoulli's inequality:

$$\begin{aligned}\frac{a_{n+1}}{a_n} &\geq \left(\frac{n + x}{n}\right) \cdot \left(1 - (n + 1)\frac{x}{(n + 1)(n + x)}\right) \\ &= \left(\frac{n + x}{n}\right) \cdot \left(1 - \frac{x}{n + x}\right) \\ &= \left(\frac{n + x}{n}\right) \cdot \left(\frac{n}{n + x}\right) = 1.\end{aligned}$$

□

Proof. [Proposition] Fix $x \in \mathbb{R}$, $n_0 > -x$. We know, that $(a_n)_{n \geq n_0}$ is increasing. We need to prove that the sequence is bounded.

1. $x \leq 0$. For $n > n_0$, $0 < 1 + \frac{x}{n} \leq 1$. This implies $0 < a_n \leq 1$. Thus, $(a_n)_{n \geq n_0}$ converges to a positive limit.

2. $x > 0$. For $n > x$, $\left(1 + \frac{x}{n}\right)^n \left(1 - \frac{x}{n}\right)^n = \left(1 - \frac{x^2}{n^2}\right)^n \leq 1$. This means

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

But the denominator, is exactly what we had in case 1. Thus, we know that $\left(1 - \frac{x}{n}\right)^n$ converges to a value in $(0, 1]$. Since this value is non-zero, we know that also $\frac{1}{\left(1 - \frac{x}{n}\right)^n}$ is bounded.

Thus, $(a_n)_{n \geq n_0}$ is bounded from above and increasing, so convergent. Lec 15

□

Definition 3.39: Exponential Function

Given $x \in \mathbb{R}$. We define the EXPONENTIAL FUNCTION as

$$\exp(x) = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n.$$

This gives $\exp : \mathbb{R} \to \mathbb{R}_{>0}$.

Definition 3.40: Euler's Number

Euler's number e is defined as $e = \exp(1) \approx 2.71828$.

Corollary 3.41:

For $x \geq -n$, we have

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

Proof. Since $(a_n)_{n \geq n_0}$ is increasing, we have

$$\exp(x) = \lim_{n \to \infty} a_n \geq a_n = \left(1 + \frac{x}{n}\right)^n.$$

□

Corollary 3.42:

If $x \geq 0$, then $\exp(x) \geq \frac{x^n}{n^n}$.

Corollary 3.43:

$\exp(x) \geq 1 + x.$

Proof. By corollary 3.41, $\exp(x) > 1 + x$ for $x > -1$. For $x \leq -1, 1 + x \leq 0 < \exp(x)$. □

Theorem 3.44: Properties of exp

For the exponential function, the following properties hold:

- $\exp(0) = 1$.
- $\exp(-x) = \frac{1}{\exp(x)}$.
- $\exp(x + y) = \exp(x) \cdot \exp(y)$.
- exp is continuous, strictly increasing and bijective from $\mathbb{R}$ to $\mathbb{R}_{>0}$.

Proof. 1. $\exp(0) = \lim_{n\to\infty} (1 + \frac{0}{n})^n = 1$.

2.

$$\exp(x) \cdot \exp(-x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n \cdot \left(1 + \frac{-x}{n}\right)^n = \lim_{n\to\infty} \left(1 - \frac{x^2}{n^2}\right)^n \leq 1.$$

By Bernoulli's inequality, we also have $1 - \frac{x^2}{n} \leq \left(1 - \frac{x^2}{n^2}\right)^n$. But as $n \to \infty$, $1 - \frac{x^2}{n} \to 1$. Thus, by sandwich lemma, the limit is 1.

3. $\exp(x + y)\exp(-x)\exp(-y) = \dots$. The proof is left as an exercise.

4. Continuity: Recall that $\exp(x) \geq 1 + x$. Therefore,

$$\exp(x) = \frac{1}{\exp(-x)} \leq \frac{1}{1 - x} \quad \text{for } x < 1.$$

Let $x \in (-\delta, \delta)$ for some $\delta > 0$. Then

i. If $x > 0$, $0 \leq \exp(x) - \exp(0) = \exp(x) - 1 \leq \frac{1}{1-x} - 1 = \frac{x}{1-x}$. This is then bounded by $\frac{\delta}{1-\delta}$.

ii. If $x < 0$, $0 \geq \exp(x) - \exp(0) = \exp(x) - 1 \geq x \geq -\delta$.

Therefore, since $\delta \leq \frac{\delta}{1-\delta}$,

$$\forall x \in (-\delta, \delta), |\exp(x) - \exp(0)| \leq \frac{\delta}{1 - \delta}.$$

Given $\varepsilon > 0$, choose $\delta = \frac{\varepsilon}{1+\varepsilon}$. Then $\frac{\delta}{1-\delta} = \varepsilon$. Thus, exp is continuous at 0.

To show, exp is continuous at arbitrary $x \in \mathbb{R}$, note that

$$\exp(x) = \exp(x - x_0)\exp(x_0).$$

For strict monotonicity, let $y > x$. This implies that

$$\exp(y) = \exp(x)\exp(y - x) > \exp(x) \cdot 1 = \exp(x).$$

The proof for bijectivity can be found in the notes. □

Definition 3.45:

We define the LOGARITHMIC FUNCTION $\log : \mathbb{R}_{>0} \to \mathbb{R}$ as the inverse of the exponential function. It satisfies the following properties:

- $\log(1) = 0$.
- $\log(xy) = \log(x) + \log(y)$.
- log is continuous, strictly increasing and bijective from $\mathbb{R}_{>0}$ to $\mathbb{R}$.

Given $a > 1$, we define

$$\log_a(x) = \frac{\log(x)}{\log(a)}.$$

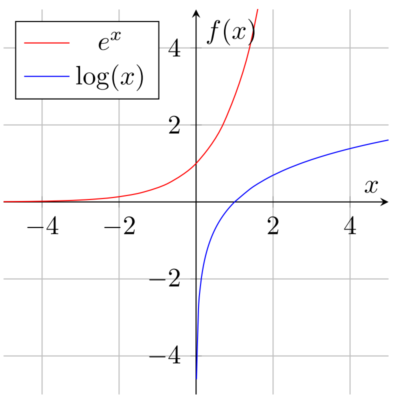


Figure 17: Exponential and Logarithmic Function

We also can define a^x

$$a^x = \exp(x \log(a)) \text{ and } x^a = \exp(a \log(x)).$$

3.5 Limits of Functions

Consider the following definition.

Definition 3.46: Accumulation Points of Sets

Given $D \subseteq \mathbb{R}, x_0 \in \mathbb{R}$. x_0 is an ACCUMULATION POINT of D if

$$\forall \delta > 0, (x_0 - \delta, x_0 + \delta) \cap D \setminus \{x_0\} \neq \varnothing.$$

Definition 3.47: Limit of a Function

Given $f : D \to \mathbb{R}$, x_0 an accumulation point of D. We say that $L \in \mathbb{R}$ is the LIMIT of f as $x \to x_0$ if

$$\forall \varepsilon > 0, \exists \delta > 0 : x \in D, |x - x_0| < \delta \Rightarrow |f(x) - L| < \varepsilon.$$

We write $L = \lim_{x\to x_0} f(x)$.

This definition is very similar to the definition of continuity, however, f does not need to be defined at x_0.

Remark 3.48:

If $\lim_{x\to x_0} f(x) = L_1$ and $\lim_{x\to x_0} g(x) = L_2$, then

- $\lim_{x\to x_0} (f(x) + g(x)) = L_1 + L_2$.
- $\lim_{x\to x_0} (f(x) \cdot g(x)) = L_1 \cdot L_2$.
- If $f \leq g$, then $L_1 \leq L_2$.
- If $f \leq h \leq g$ and $L_1 = L_2$, then $\lim_{x\to x_0} h(x) = L_1$.

Remark 3.49:

If $x_0 \in D$, then f is continuous at x_0 iff $\lim_{x\to x_0} f(x) = f(x_0)$.

Let us assume that x_0 is an accumulation point for $D \setminus \{x_0\}$. Further assume that

$$L = \lim_{x\to x_0} f|_{D\setminus\{x_0\}}(x)$$

exists.

If $x_0 \in D$, assume $f(x_0) \neq L$. In this case, x_0 is called a REMOVABLE DISCONTINUITY of f.

Instead of $\lim_{x\to x_0} f|_{D\setminus\{x_0\}}(x)$, we also write

$$\lim_{\substack{x\to x_0\\ x\neq x_0}} f(x) = L.$$

For x_0 to be a removable discontinuity, it is necessary that

$$\tilde{f} = \begin{cases} f(x) & x \neq x_0 \\ L & x = x_0 \end{cases}.$$

is continuous at x_0.

If $x_0 \notin D$, then the function

$$\tilde{f} = \begin{cases} f(x) & x \in D \\ L & x = x_0 \end{cases}.$$

Is called the CONTINUOUS EXTENSION of f at x_0.

Lemma 3.50:

$f : D \to \mathbb{R}$. Then $\lim_{x\to\overline{x}} f(x) = L$ iff

$$\forall (x_n)_{n\geq 0} \subseteq D, x_n \to \overline{x}, f(x_n) \to L.$$

Proposition 3.51:

$D, E \subseteq \mathbb{R}, f : D \to E$, assume that $L = \lim_{x\to\overline{x}} f(x)$ and $L \in E$.

Then, if $g : E \to \mathbb{R}$ with g continuous at L, we have

$$\lim_{x\to\overline{x}} g(f(x)) = g(L).$$

Proof. Let $(x_n) \subseteq D, x_n \to \overline{x}$. Then by previous lemma, $f(x_n) \to L$. Since g is continuous at L, we have

$$g(f(x_n)) \to g(L).$$

By the lemma again, this implies

$$g(f(x)) \overset{x\to\overline{x}}{\longrightarrow} g(L).$$

Lec 16

□

In the definition of the limit before, we required that $L \in \mathbb{R}$. However, we can extend this definition as follows.

Definition 3.52: Improper Limit

Let $f : D \to \mathbb{R}, x_0$ be an accumulation point of D. We say that f DIVERGES to $+\infty$ as $x \to x_0$ if

$$\forall M > 0, \exists \delta > 0 : x \in D, |x - x_0| < \delta \Rightarrow f(x) \geq M.$$

In this case, we write

$$\lim_{x\to x_0} f(x) = +\infty.$$

We say $\lim_{x\to x_0} f(x) = -\infty$ if

$$\forall M > 0, \exists \delta > 0 : x \in D, |x - x_0| < \delta \Rightarrow f(x) \leq -M.$$

Definition 3.53: One-Sided Limits

x_0 is an ACCUMULATION POINT FROM THE RIGHT if

$$(x_0, x_0 + \delta) \cap D \neq \varnothing, \quad \forall \delta > 0.$$

We say that $\lim_{x\to x_0^+} f(x) = L$ if

$$\forall \varepsilon > 0, \exists \delta > 0 : x \in D \cap (x_0, x_0 + \delta) \Rightarrow |f(x) - L| < \varepsilon.$$

Same from LEFT. In this case we write $\lim_{x\to x_0^-} f(x) = L$.

Essentially, an accumulation point of a set D is a point where we can construct a sequence in D that converges to that point.

Definition 3.54: Limits at Infinity

Let $f : D \to \mathbb{R}$. assume $D \cap (R, \infty) \neq \varnothing \forall R > 0$.

We say that $\lim_{x\to\infty} f(x) = L$ if

$$\forall \varepsilon > 0, \exists R > 0 : x \in D \cap (R, \infty) \Rightarrow |f(x) - L| < \varepsilon.$$

The limit at $-\infty$ is defined analogously.

If we define $g(x) = f(\frac{1}{x})$, then we can relate limits at infinity to limits at 0 with

$$\lim_{x\to\infty} f(x) = L \Leftrightarrow \lim_{x\to 0^+} g(x) = L.$$

Similarly for $-\infty$ and 0^-.

$$\lim_{x\to-\infty} f(x) = L \Leftrightarrow \lim_{x\to 0^-} g(x) = L.$$

Definition 3.55: One-Sided Continuity

Given $f : D \to \mathbb{R}, x_0 \in D$. Then f is CONTINUOUS FROM THE RIGHT at x_0 if

$$f(x_0) = \lim_{x\to x_0^+} f(x).$$

f is CONTINUOUS FROM THE LEFT at x_0 if

$$f(x_0) = \lim_{x\to x_0^-} f(x).$$

Example 3.56:

Let

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

This function is continuous from the right at 0, but not from the left.

If the limit from the right and left both exist and are different, then x_0 is a JUMP POINT of f.

Example 3.57:

$D = (0, \infty); f(x) = x^x$ Goal: Compute $\lim_{x\to x_0^+} f(x)$.

Solution. We perform the proof in 3 steps.

1. We compute $\lim_{y\to\infty} ye^{-y}$. Recall that $e^y \geq (1+\frac{y}{n})^n$. In particular,

$$e^y \geq (1+\frac{y}{2})^2 \Rightarrow e^{-y} \leq \frac{1}{(1+\frac{y}{2})^2}.$$

Therefore,

$$0 \leq ye^{-y} \leq \frac{y}{(1+\frac{y}{2})^2} \leq \frac{4}{y} \underset{y\to\infty}{\longrightarrow} 0.$$

Thus by sandwich lemma, $\lim_{y\to\infty} ye^{-y} = 0$.

2. Compute the limit $\lim_{x\to 0^+} x\log(x)$. By (1), given $\varepsilon > 0 \exists R > 0$: $|ye^{-y}| < \varepsilon$ for $y > R$.

Now, given $x \in (0,\delta)$, let $y = -\log(x)$. Then $y > -\log(\delta)$. Choose $\delta = e^{-R}$. Then $y > R$, which means that

$$|ye^{-y}| < \varepsilon \Rightarrow |x\log(x)| < \varepsilon.$$

Thus $\lim_{x\to 0^+} x\log(x) = 0$.

3. By definition, $x^x = e^{x\log(x)}$. Therefore,

$$\lim_{x\to 0^+} x^x = \lim_{x\to 0^+} e^{x\log(x)} = e^{\lim_{x\to 0^+} x\log(x)} = e^0 = 1.$$

Definition 3.58: Landau Big-O

Given $f, g : D \to \mathbb{R}, x_0$ an accumulation point of D. We write $f(x) = O(g(x))$ as $x \to x_0$ if $\exists M > 0, \delta > 0$ such that

$$|f(x)| \leq M|g(x)|, \forall x \in D, |x - x_0| < \delta.$$

If $g \neq 0$, this means that $\frac{f(x)}{g(x)}$ is bounded near x_0.

We write $f(x) = O(g(x))$ as $x \to +\infty$ if

$$\exists M > 0, R > 0 : |f(x)| \leq M|g(x)|, \forall x \in D \cap (R, \infty).$$

Example 3.59:

Consider the following examples.

- $x^2 = O(x)$ as $x \to 0$. This is correct as for $\delta = 1$, $|x^2| \leq |x|$ for $|x| < \delta$.
- $x = O(x^2)$ as $x \to 0$. This is incorrect as $\frac{|x|}{|x^2|} = \frac{1}{|x|}$, which is unbounded near 0.
- $x + x^2 = O(x)$ as $x \to 0$. This is correct as $|x + x^2| \leq |x| + |x|^2 \leq 2|x|$ for $|x| < 1$.
- $\frac{3x^3}{x^3+3} = O(1)$ as $x \to \infty$. This is correct as $\left|\frac{3x^3}{x^3+3}\right| \leq 3$ for all x.
- $\frac{3x^3}{x^3+3} = O(\frac{1}{x})$ as $x \to \infty$. This is incorrect as

$$\frac{\left|\frac{3x^3}{x^3+3}\right|}{\left|\frac{1}{x}\right|} = \frac{3x^4}{x^3+3} = 3x - \frac{9x}{x^3+3}.$$

This diverges as $x \to \infty$.

There also exists a related notation called little-o notation.

Definition 3.60: Landau Little-o

Given $f, g : D \to \mathbb{R}, x_0$ an accumulation point of D. We write $f(x) = o(g(x))$ as $x \to x_0$ if $\forall \varepsilon > 0, \exists \delta > 0$ such that

$$|f(x)| \leq \varepsilon|g(x)|, \forall x \in D, |x - x_0| < \delta.$$

Example 3.61:

Consider the following examples

- $x = o(x^2)$ as $x \to \infty$. This is correct as $\frac{x}{x^2} \to 0$.
- $x^2 = o(x)$ as $x \to 0$. This is correct as $\frac{x^2}{x} = |x| \to 0$.
- $\frac{3x^3}{2x^2+x^{10}} = o(1)$ as $x \to 0$. This is correct as

$$\frac{3x^3}{2x^2+x^{10}} = \frac{3x}{2+x^8} \to 0.$$

Example 3.62:

Reduce the following limit

$$\frac{x^3 - 7x^2 + 6x + 2}{x^2} = x - 7 + \frac{6}{x} + \frac{2}{x^2}.$$

Say you are interested in the limit as $x \to \infty$. It might be annoying to keep writing $\frac{6}{x} + \frac{2}{x^2}$, if we carry them through two pages of calculations. Instead, we can write

$$x - 7 + o(1) \text{ as } x \to \infty.$$

Or even if we consider for some reason only terms that grow we could write

$$x + O(1) \text{ as } x \to \infty = x + o(x) \text{ as } x \to \infty.$$

Example 3.63:

Given $f(x) = x + x^3 + 4x^4 + x^7, g(x) = x + \frac{3x^2}{1+x}$. We want to under stand the behaviour of $f \cdot g$ near 0.

Solution. Observe that $f(x) = x + o(x^2)$ and $g(x) = x + O(x^2)$. Now

$$f \cdot g = (x+o(x^2))(x+O(x^2)) = x^2 + x\cdot O(x^2) + x\cdot o(x^2) + o(x^2)\cdot O(x^2).$$

Notice that $\frac{x\cdot o(x^2)}{x\cdot x^2} = 0$ implies that $x \cdot o(x^2) = o(x^3)$. Similarly, $x \cdot O(x^2) = O(x^3)$ and thus

$$f \cdot g = x^2 + O(x^3) + o(x^3) + o(x^4) = x^2 + O(x^3).$$

Definition 3.64:

In informatics the following two definitions are also used.

$$f = \Omega(g) \text{ as } x \to x_0 \Leftrightarrow g = O(f) \text{ as } x \to x_0.$$

$$f = \omega(g) \text{ as } x \to x_0 \Leftrightarrow g = o(f) \text{ as } x \to x_0.$$

Lec 17

3.6 Sequences of Functions

We like to combine the ideas of sequences and functions.

Definition 3.65: Sequence of Functions

$D \subseteq \mathbb{R}$. $(f_n)_{n\geq 0}$ is a family of functions, $f_n : D \to \mathbb{R}$, indexed by $\mathbb{N}$.

To talk about limits of sequences of functions, we first need a notion of convergence of such sequences.

Definition 3.66: Pointwise Convergence

We say that $(f_n)_{n\geq 0}$ converges POINTWISE to $f : D \to \mathbb{R}$, if

$$\forall x \in D, f_n(x) \underset{n\to\infty}{\longrightarrow} f(x).$$

Example 3.67:

Consider $f_n(x) = x^n$ on $D = [0,1], n \geq 1$. Then if we fix $x \in [0,1)$, then $f_n(x) \to 0$ as $n \to \infty$. If we fix $x = 1$, then $f_n(1) = 1$ for all n. Therefore,

$$\lim_{n\to\infty} x^n = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1. \end{cases}.$$

So we see that the limit of continuous functions need not be continuous.

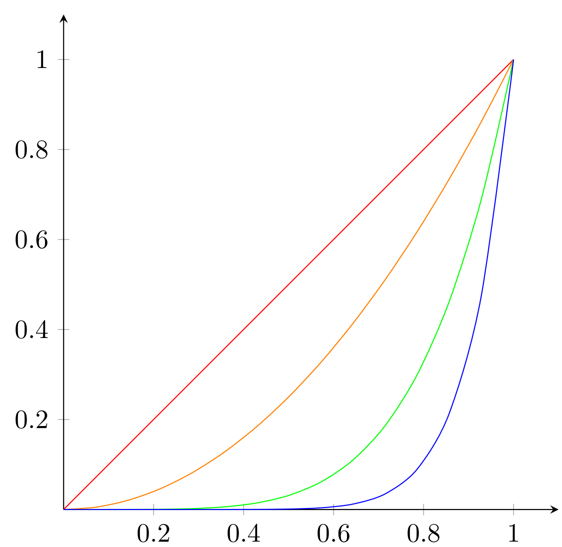


Figure 18: Family of functions

We can write pointwise convergence in terms of ε and δ.

$$\forall x \in D\, \forall \varepsilon > 0, \exists N \in \mathbb{N} : |f_n(x) - f(x)| < \varepsilon, \forall n \geq N.$$

However, note that N depends on x here. We thus define the following.

Definition 3.68: Uniform Convergence

f_n converges UNIFORMLY to $f : D \to \mathbb{R}$ if

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} : |f_n(x) - f(x)| < \varepsilon, \forall x \in D, \forall n \geq N.$$

The difference to pointwise convergence is that N does not depend on x.

Geometrically, this means that after some index N, all functions f_n lie within a band of width 2ε around f.

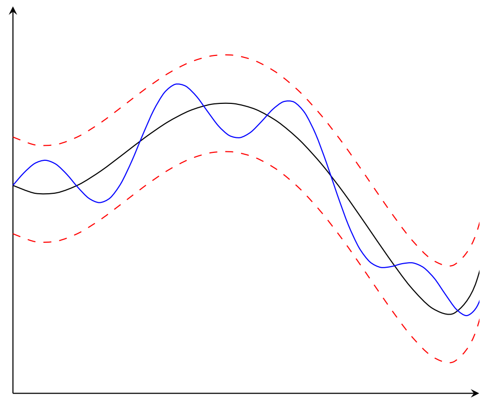


Figure 19: Uniform Convergence

Theorem 3.69:

Given $f_n : D \to \mathbb{R}$, $f : D \to \mathbb{R}$. If f_n converges uniformly to f and all f_n are continuous, then f is continuous.

Proof. Fix $\overline{x} \in D$, fix $\varepsilon > 0$. By uniform convergence, $\exists N$ such that

$$|f_N(y) - f(y)| < \frac{\varepsilon}{3}, \forall x \in D.$$

Since f_N is continuous, $\exists \delta > 0$ such that

$$x \in D, |x - \overline{x}| < \delta \Rightarrow |f_N(x) - f_N(\overline{x})| < \frac{\varepsilon}{3}.$$

Then,

$$\begin{aligned} &\forall x \in D, |x - \overline{x}| < \delta \Rightarrow \\ |f(x) - f(\overline{x})| &= |f(x) - f_N(x) + f_N(x) - f_N(\overline{x}) + f_N(\overline{x}) - f(\overline{x})| \\ &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(\overline{x})| + |f_N(\overline{x}) - f(\overline{x})| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

□

Remark 3.70:

The above theorem also works for uniform continuity.

Assume f_n converges pointwise to f. Then, we know

$$f(\overline{x}) = \lim_{n\to\infty} f_n(x).$$

Because f_n is continuous, we can write

$$f(\overline{x}) = \lim_{n\to\infty} \left(\lim_{x\to\overline{x}} f_n(x) \right).$$

On the other hand, we have

$$\lim_{x\to\overline{x}} f(x) = \lim_{x\to\overline{x}} \left(\lim_{n\to\infty} f_n(x) \right).$$

Therefore $f(\overline{x}) = \lim_{x\to\overline{x}} f(x)$ if and only if

$$\lim_{n\to\infty} \left(\lim_{x\to\overline{x}} f_n(x) \right) = \lim_{x\to\overline{x}} \left(\lim_{n\to\infty} f_n(x) \right).$$

4 Series and Power Series

4.1 Series of Real Numbers

Given a sequence $(a_n)_{n\geq 0}$, given $A \in \mathbb{R}$, we say that

$$\sum_{n=0}^{\infty} a_n = A \text{ if } s_n = \sum_{k=0}^{n} a_k \text{ converges to } A.$$

So $(s_n)_{n\geq 0}$ is a sequence and we look at its convergence. In this situation a_n is the n-th TERM of the series and A is the SUM of the series.

If the series has no limit, we say it does NOT CONVERGE. We say, that the series DIVERGES to $+\infty$ if $s_n \to +\infty$ and diverges to $-\infty$ if $s_n \to -\infty$.

For the moment beeing, we only want to consider $a_n \in \mathbb{R}$.

Proposition 4.1:
If $\sum_{n=0}^{\infty} a_n$ converges, then $a_n \to 0$ as $n \to \infty$.

Proof. Note that $s_n = s_{n-1} + a_n$. Thus $a_n = s_n - s_{n-1}$. Since $s_n \to A$, it follows that $a_n \to A - A = 0$. □

Example 4.2: Geometric Series
Given $q \in \mathbb{R}$, consider the series (We let $q^0 = 1$)

$$\sum_{n=0}^{\infty} q^n.$$

Determine for which q this series converges and find the sum in that case.

Solution. If $|q| \geq 1 \Rightarrow |q|^n \geq 1 \forall n$, thus $q^n \not\to 0$. Therefore, the series does not converge.

If $|q| < 1$, look at $s_n = \sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}$. Since $|q| < 1$, $q^{n+1} \to 0$. Thus, $s_n \to \frac{1}{1-q}$.

Example 4.3: Harmonic Series
Determine whether the series $\sum_{n=1}^{\infty} \frac{1}{n}$ converges or diverges.

Solution. For $l \in \mathbb{N}$, look at the partial sum

$$\sum_{n=1}^{2^l} \frac{1}{n} = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq \frac{1}{2}} + \cdots + \underbrace{\frac{1}{2^{l-1}+1} + \cdots + \frac{1}{2^l}}_{\geq \frac{1}{2}} \geq 1 + \frac{l}{2}.$$

Thus, the partial sums are unbounded and the series diverges.

Lemma 4.4:
The sum $\sum_{n=0}^{\infty} a_n$ is convergent if and only if

$$\sum_{n=N}^{\infty} a_n.$$

is convergent for some $N \in \mathbb{N}$.

Proof. Fix N, call s_n the partial sums for $\sum_{n=0}^{\infty} a_n$ and $\tilde{s}_n$ the partial sums for $\sum_{n=N}^{\infty} a_n$. Then,

$$s_n = \sum_{k=0}^{n} a_k = \sum_{k=0}^{N-1} a_k + \sum_{k=N}^{n} a_k = \sum_{k=0}^{N-1} a_k + \tilde{s}_n.$$

□

Proposition 4.5:
Let $a_n \geq 0 \forall n$. Then the sequence $s_n = \sum_{k=0}^{n} a_k$ is increasing. Therefore if s_n is bounded, then it converges. If s_n is not bounded, then it diverges to $+\infty$.

If $a_n \geq 0$, then s_n is bounded iff $\exists$ subsequence s.t. $(s_{n_k})_{k\geq 0}$ is bounded. Lec 18

Proposition 4.6: Majorant and Minorant Criterion
Let $0 \leq a_k \leq b_k \forall k \in \mathbb{N}$. Then,

$$0 \leq \sum_{k=0}^{\infty} a_k \leq \sum_{k=0}^{\infty} b_k.$$

So if $\sum b_k$ converges, then $\sum a_k$ converges. Similarly, if $\sum a_k$ diverges to $+\infty$, then $\sum b_k$ does too.

Proof. $a_k \leq b_k \Rightarrow \sum_{k=0}^{n} a_k \leq \sum_{k=0}^{n} b_k \forall n$. By limit computation rules, the result follows. □

In the above situation, we say that $\sum a_k$ is a MINORANT of $\sum b_k$ and that $\sum b_k$ is a MAJORANT of $\sum a_k$.

Remark 4.7:
Note that the Majorant and Minorant Criterion also works if $\exists N \in \mathbb{N}$ such that $0 \leq a_k \leq b_k \forall k \geq N$.

Example 4.8:
Consider $\sum_{k=1}^{\infty} \frac{1}{k^2}$. Observe that, for $k \geq 2$, $\frac{1}{k^2} \leq \frac{1}{k(k-1)}$. Then

$$\begin{aligned}\sum_{k=2}^{\infty} \frac{1}{k^2} \leq \sum_{k=2}^{\infty} \frac{1}{k(k-1)} &= \sum_{k=2}^{\infty} \left(\frac{1}{k-1} - \frac{1}{k}\right) \\ &= \sum_{k=2}^{\infty} \frac{1}{k-1} - \sum_{k=2}^{\infty} \frac{1}{k} \\ &= \left(1 + \frac{1}{2} + \frac{1}{3} + \dots\right) - \left(\frac{1}{2} + \frac{1}{3} + \dots\right) = 1.\end{aligned}$$

Thus, by the majorant criterion, $\sum_{k=1}^{\infty} \frac{1}{k^2}$ is bounded and therefore convergent.

Tip 4.9:
Having series with fractions, partial fraction decomposition is often useful.

Example 4.10:
Consider $\sum_{n=0}^{\infty} \frac{2n-10}{n^3-10n+100}$. For $n \geq N$, we have that $0 \leq a_n$ furthermore,

$$\begin{aligned} n^2 \frac{2n-10}{n^3 - 10n + 100} &= \frac{n^3(2 - \frac{10}{n})}{n^3(1 - \frac{10}{n^2} + \frac{100}{n^3})} \\ &\to \frac{2-0}{1-0+0} = 2. \end{aligned}$$

So for n large enough, $n^2 a_n \leq 3$. Thus, $0 \leq a_n \leq \frac{3}{n^2}$. By the majorant criterion, $\sum a_n$ converges.

In exercise 4.8, we showed that $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges. But we only could do this by doing this special trick that

$$\frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}.$$

But what if we had $\sum_{k=1}^{\infty} \frac{1}{k^{\frac{3}{2}}}$? We cannot do the same trick here.

Proposition 4.11: Cauchy Condensation Test

Let $a_k \geq 0$ and a_k be decreasing. Then

$$\sum_{k=0}^{\infty} a_k \text{ converges } \Leftrightarrow \sum_{k=0}^{\infty} 2^k a_{2^k} \text{ converges }.$$

Proof.

$$a_0 + a_1 + \underbrace{\overbrace{a_2}^{\leq a_1}}_{\geq a_2} + \underbrace{\overbrace{a_3 + a_4}^{2a_2}}_{\geq 2a_4} + \underbrace{\overbrace{a_5 + \cdots + a_8}^{\leq 4a_4}}_{\geq 4a_8} + \underbrace{\overbrace{a_9 + \cdots + a_{16}}^{8a_8}}_{\leq 8a_{16}}$$

$$\sum_{k=0}^{n} 2^k a_{2^k} \geq \sum_{j=2}^{2^{n+1}} a_j \geq \frac{1}{2} \sum_{k=1}^{n+1} 2^k a_{2^k}.$$

In the above, the first term is the part at the top, the second term is the middle part and the last term is the bottom part. If we let $n \to \infty$, the result follows. □

Example 4.12:

Consider $\sum_{k=1}^{\infty} \frac{1}{k^p}$. If $p \leq 0$, then $\frac{1}{k^p} \geq 1$ and thus the series diverges to $+\infty$. If $p > 0$, then $\sum \frac{1}{k^p}$ converges iff $\sum 2^k \frac{1}{(2^k)^p}$ converges.

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

But this is a geometric series which converges iff $|2^{1-p}| < 1$, that is $p > 1$.

When we have series that can change sign, the following definition is useful.

Definition 4.13: Absolute Convergence

$\sum_{k=0}^{\infty} a_k$ converges ABSOLUTELY if

$$\sum_{k=0}^{\infty} |a_k| \text{ converges }.$$

We say that it converges CONDITIONALLY if

$$\sum_{k=0}^{\infty} a_k \text{ converges but } \sum_{k=0}^{\infty} |a_k| = +\infty.$$

Theorem 4.14: Riemann rearrangement theorem

Let $\sum a_k$ converge conditionally, fix $A \in \mathbb{R}$. Then $\exists \varphi : \mathbb{N} \to \mathbb{N}$ bijective, such that

$$\sum_{k=0}^{n} a_{\varphi(k)} \underset{n \to \infty}{\longrightarrow} A.$$

So in essence, a conditionally convergent series can be rearranged to converge to anything we want. The proof is considered extra material.

The important takeaway from this theorem is that when dealing with series, we should be very careful when rearranging terms.

Definition 4.15: Alternating Series

Given $a_k \geq 0$, we define the ALTERNATING SERIES as

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

Proposition 4.16: Leibniz Criterion

If $a_k \geq 0$ and decreasing to 0, then the alternating series converges and

$$\sum_{k=0}^{2n+1} (-1)^k a_k \leq \sum_{k=0}^{\infty} (-1)^k a_k \leq \sum_{k=0}^{2n} (-1)^k a_k \; \forall n \in \mathbb{N}.$$

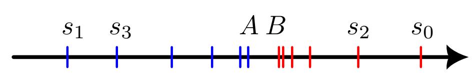


Figure 20: Leibniz Criterion

Proof. Let $s_n = \sum_{k=0}^{n} (-1)^k a_k$ be the partial sums.

For even numbers, we have

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

For odd numbers, we have

$$s_{2n+1} = s_{2n-1} + a_{2n} - a_{2n+1} \geq s_{2n-1}.$$

So the sequence $(s_{2n})_{n \geq 0}$ is decreasing and converging to some number B. Similarly, $(s_{2n+1})_{n \geq 0}$ is increasing and converging to some number A. We now show that $A = B$.

But we have $0 \leq B - A \leq s_{2n} - s_{2n-1} = a_{2n} \to 0$. Thus $A = B$ and the series converges. □

Example 4.17: Alternating Harmonic Series

Consider $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$. By the Leibniz criterion, this series converges. However, $\sum_{k=1}^{\infty} \left| \frac{(-1)^k}{k} \right| = \sum_{k=1}^{\infty} \frac{1}{k}$ diverges. Thus, the alternating harmonic series converges conditionally.

Proposition 4.18: Cauchy Criterion

A series $\sum a_k$ converges iff $\forall \varepsilon > 0 \exists N$ such that

$$\left| \sum_{k=m+1}^{n} a_k \right| < \varepsilon \quad \forall n > m \geq N.$$

Proof. $\sum_{k=0}^{\infty} a_k$ converges iff $s_n = \sum_{k=0}^{n} a_k$ converges. This however is equivalent to $(s_n)_{n \geq 0}$ being a Cauchy sequence, i.e.

$$\forall \varepsilon > 0 \exists N : |s_n - s_m| < \varepsilon \quad \forall n > m \geq N.$$

However, $s_n - s_m = \sum_{k=m+1}^{n} a_k$. Thus the result follows.

□

4.2 Absolute Convergence

We now want to study absolutely converging series a bit more.

Proposition 4.19:

If $\sum a_k$ converges absolutely, then it converges and

$$\left|\sum_{k=0}^{\infty} a_k\right| \leq \sum_{k=0}^{\infty} |a_k|.$$

Proof. Since $\sum_{k=0}^{\infty} |a_k|$ converges, $\forall \varepsilon > 0 \, \exists N$ such that for all $n \geq m \geq N$, $\sum_{k=m+1}^{n} |a_k| < \varepsilon$.

By triangle inequality,

$$\left|\sum_{k=m+1}^{n} a_k\right| \leq \sum_{k=m+1}^{n} |a_k| < \varepsilon \quad \forall n \geq m \geq N.$$

So by Cauchy, $\sum_{k=0}^{\infty} a_k$ converges.

Also $|s_n| = \left|\sum_{k=0}^{n} a_k\right| \leq \sum_{k=0}^{n} |a_k|$. Letting $n \to \infty$, the result follows.

Lec 19

□

Remark 4.20:

If $(x_n)_{n\geq 0}$ is a sequence converging to $\alpha \in \mathbb{R}$, then:

- $\forall q > \alpha, \exists N : x_n < q \, \forall n \geq N$
- $\forall r < \alpha, \exists N : x_n > r \, \forall n \geq N$

The following proposition is inspired by trying to compare a series with a geometric series.

Proposition 4.21: Cauchy Root Criterion

Given $(a_n)_{n\geq 0}$, we want to understand the absolute convergence of the series $\sum_{k=o}^{\infty} a_k$. Define

$$\alpha := \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

Then, if $\alpha < 1$, the series converges absolutely. If $\alpha > 1$, the series does NOT converge.

Proof. First, assume $\alpha < 1$. Define $q = \frac{1+\alpha}{2}$. So $q \in (\alpha, 1)$. By definition of lim sup,

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \left(\underbrace{\sup_{k\geq n} \sqrt[k]{|a_k|}}_{x_n} \right).$$

By the above remark, $\exists N$ such that

$$\sup_{k\geq n} \sqrt[k]{|a_k|} < q \quad \forall n \geq N.$$

But this means that $\sqrt[k]{|a_k|} < q$ for all $k \geq N$, which is the same as $|a_k| < q^k$ for all $k \geq N$. Since $q < 1$, $\sum_{k=0}^{\infty} q^k$ converges. By the majorant criterion, $\sum_{k=0}^{\infty} |a_k|$ converges absolutely.

Now, assume $\alpha > 1$. Recall that the lim sup is an accumulation point. Thus, $\exists$ subsequence $(\sqrt[n_k]{|a_{n_k}|})_{k\geq 0}$ such that

$$\sqrt[n_k]{|a_{n_k}|} \to \alpha > 1.$$

By the above remark with $r = 1$, $\exists K$ such that

$$\sqrt[n_k]{|a_{n_k}|} > 1 \quad \forall k \geq K.$$

But this is equivalent to $|a_{n_k}| > 1$ for all $k \geq K$. Thus, the numbers a_n do not go to 0 and thus the series cannot converge. □

Remark 4.22:

Consider the series $\sum_{k=1}^{\infty} \frac{1}{k}$. This series has $\alpha = 1$, but diverges.

The series $\sum_{k=1}^{\infty} \frac{1}{k^2}$ has $\alpha = 1$ but converges.

So for $\alpha = 1$, no conclusion can be drawn.

Tip 4.23:

The root criterion is useful when the n-th term is of the form $a_k = (b_k)^k$. Where b_k is easier to handle.

Proposition 4.24: D'Alemberts Quotient Criterion

Assume $a_0 \neq 0 \, \forall n$. Let

$$\alpha = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|.$$

Then $\alpha < 1$ implies absolute convergence, $\alpha > 1$ implies no convergence.

Proof. Assume first, $\alpha < 1$. Let $q = \frac{1+\alpha}{2} \in (\alpha, 1)$. Then, by the remark, $\exists N$ such that

$$\frac{|a_{n+1}|}{|a_n|} < q \quad \forall n \geq N.$$

Then $|a_n| = \frac{|a_n|}{|a_{n-1}|} \cdot \frac{|a_{n-1}|}{|a_{n-2}|} \cdots \frac{|a_{N+1}|}{|a_N|} |a_N| < q^{n-N} |a_N|$ for all $n \geq N$.

But all of theese terms must be $< q$. So for $n \geq N$,

$$|a_n| < q^{n-N} |a_N| = \frac{|a_N|}{q^N} q^n.$$

So the series $\sum_{n\geq N} |a_n|$ is bounded by $\frac{|a_N|}{q^N} \sum_{n\geq N} q^n$ which converges since $q < 1$. Thus, by the majorant criterion, $\sum |a_n|$ converges.

The opposite case, $\alpha > 1$ implies by the remark with $r = 1$ that

$$\frac{|a_{n+1}|}{|a_n|} > 1 \quad \forall n \geq N.$$

So $|a_{n+1}| > |a_n|$. But this implies in particular that $|a_n| > |a_{n-1}| > \cdots > |a_N| > 0$ for all $n \geq N$. Thus, the terms do not go to 0 and thus the series cannot converge. □

Again for $\alpha = 1$, no conclusion can be drawn.

Tip 4.25:

The quotient criterion is more useful than the root criterion when the term involves factorials or other operations that do not go well with n-th roots.

Example 4.26:

Consider $\sum_{k=1}^{\infty} \frac{k!}{k^k}$. We want to study its convergence.

$$\left|\frac{a_{k+1}}{a_k}\right| = \frac{(k+1)!}{(k+1)^{k+1}} \cdot \frac{k^k}{k!} = \frac{k^k}{(k+1)^k} = \left(\frac{k}{k+1}\right)^k$$

$$= \left(1 - \frac{1}{k+1}\right)^k = \left[\left(1 - \frac{1}{k+1}\right)^{k+1}\right]^{\frac{k}{k+1}}$$

$$\to e^{-1}.$$

Proposition 4.27:

Let $\sum_{k=0}^{\infty} a_k$ be absolutely convergent. Let $\varphi : \mathbb{N} \to \mathbb{N}$ be bijective. Then $\sum_{k=0}^{\infty} a_{\varphi(k)}$ is absolutely convergent and

$$\sum_{k=0}^{\infty} a_{\varphi(k)} = \sum_{k=0}^{\infty} a_k.$$

Proof. Fix $\varepsilon > 0$. Since $\sum |a_k|$ converges, $\exists N$ such that

$$\sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Let M be the smallest number s.t. $\{0, \ldots, N\} \subseteq \{\varphi(0), \ldots, \varphi(M)\}$. Equivalently, $M = \max\{\varphi^-1(0), \ldots, \varphi^-1(N)\}$. Then,

$$\sum_{k=0}^{M} a_{\varphi(k)} - \sum_{k=0}^{N} a_k = \sum_{\substack{\varphi(k)>N \\ 0\le k\le M}} a_{\varphi(k)}.$$

Therefore, for $n \geq M$,

$$\begin{aligned}\left|\sum_{l=0}^{n} a_\varphi(l) - \sum_{k=0}^{\infty} a_k\right| &= \left|\sum_{l=0}^{n} a_\varphi(l) - \sum_{k=0}^{N} a_k - \sum_{k=N+1}^{\infty} a_k\right| \\ &\leq \left|\sum_{\substack{\varphi(l)>N \\ 0\le l\le n}} a_{\varphi(l)}\right| + \left|\sum_{k=N+1}^{\infty} a_k\right| \\ &\leq \sum_{\substack{\varphi(l)>N \\ 0\le l\le n}} |a_{\varphi(l)}| + \sum_{k=N+1}^{\infty} |a_k| \\ &\leq 2 \sum_{k=N+1}^{\infty} |a_k| < \varepsilon.\end{aligned}$$

So the sum $\sum_{l=0}^{n} a_\varphi(l)$ converges to $\sum_{k=0}^{\infty} a_k$ as $n \to \infty$. Applying the same to $\sum |a_k|$ shows absolute convergence. □

An interesting question is how we can compute the product of two series.

Theorem 4.28:

Let $\sum a_k$ and $\sum b_k$ be absolutely convergent series. Fix $\alpha : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ bijection. Then

$$\sum_{k=0}^{\infty} a_k \cdot \sum_{k=0}^{\infty} b_k = \sum_{k=0}^{\infty} a_{\alpha_1(k)} b_{\alpha_2(k)}.$$

and the product series converges absolutely.

Proof. Consider first the bijection in figure 21. Then,

$$\begin{aligned}\sum_{k=0}^{n^2-1} |a_{\alpha_1(k)}||b_{\alpha_2(k)}| &= \sum_{k=0}^{n-1} |a_k| \sum_{k=0}^{n-1} |b_k| \\ &\leq \left(\sum_{k=0}^{\infty} |a_k|\right)\left(\sum_{k=0}^{\infty} |b_k|\right) \leq \infty.\end{aligned}$$

Thus, $\sum_{k=0}^{\infty} |a_{\alpha_1(k)}||b_{\alpha_2(k)}| \leq \infty$. So the series $\sum_{k=0}^{\infty} |a_{\alpha_1(k)} b_{\alpha_2(k)}|$ converges absolutely which implies that it converges.

Again, note that $\sum_{k=0}^{n^2-1} a_{\alpha_1(k)} b_{\alpha_2(k)} = \left(\sum_{k=0}^{n-1} a_k\right)\left(\sum_{k=0}^{n-1} b_k\right)$. By definition, both limits on the right converge by assumption. The limit on left we have just shown to exist. Thus, both limits are equal for this specific bijection.

Lec 20 But by proposition 4.27, the result holds for any bijection. □

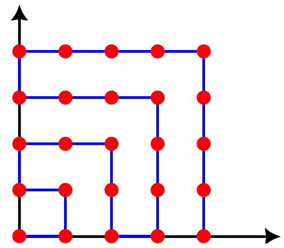


Figure 21: Specific Bijection α

This is not the best way to multiply series, so we introduce a more natural one.

Corollary 4.29: Cauchy Product Formula

If $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ are absolutely convergent series, then

$$\sum_{k=0}^{\infty} a_k \cdot \sum_{k=0}^{\infty} b_k = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k} b_k\right).$$

Proof. Plotting the bijection, it looks like this

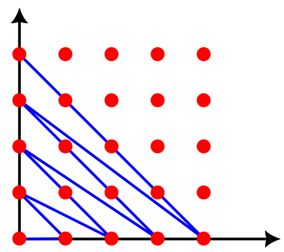


Figure 22: Diagonal Bijection

In this way, we have

$$\begin{aligned}\sum_{n=0}^{\infty} a_{\alpha_1(n)} b_{\alpha_2(n)} &= a_0 b_0 + (a_1 b_0 + a_0 b_1) \\ &\quad + (a_2 b_0 + a_1 b_1 + a_0 b_2) + \ldots \\ &= \sum_{n=0}^{\infty}\left(\sum_{k+j=n} a_k b_j\right) \\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k} b_k\right).\end{aligned}$$

□

Example 4.30:

We know that if $q < 1$, then $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$. This implies that

$$\left(\frac{1}{1-q}\right)^2 = \sum_{n=0}^{\infty} q^n \cdot \sum_{n=0}^{\infty} q^n = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} q^{n-k} q^k\right)$$

$$= \sum_{n=0}^{\infty}(n+1)q^n = \sum_{n=0}^{\infty} nq^n + \sum_{n=0}^{\infty} q^n.$$

This further implies that

$$\sum_{n=0}^{\infty} nq^n = \frac{1}{(1-q)^2} - \frac{1}{1-q} = \frac{q}{(1-q)^2}.$$

4.3 Series of Complex Numbers

Consider the following definition.

Definition 4.31:

Given $(z_n)_{n\geq 0} = (x_n + iy_n)_{n\geq 0}$, we say that the series $\sum_{k=0}^{\infty} z_k$ converges to $Z = A + iB \in \mathbb{C}$ if

$$\sum_{k=0}^{\infty} x_n = A \quad \text{and} \quad \sum_{k=0}^{\infty} y_n = B.$$

We say that the series converges absolutely if

$$\sum_{k=0}^{\infty} |z_n| = \sum_{k=0}^{\infty} \sqrt{x_n^2 + y_n^2} < \infty.$$

Oberserve that $|x_n| \leq |z_n|$ and $|y_n| \leq |z_n|$. Thus, absolute convergence of $\sum z_n$ implies absolute convergence of $\sum x_n$ and $\sum y_n$.

The converse is also true as by triangle inequality,

$$|z_n| \leq |x_n| + |y_n|.$$

4.4 Power Series

A power series is a series with powers.

Definition 4.32: Real Power Series

A POWER SERIES is a series of the form

$$\sum_{k=0}^{\infty} a_k x^k \quad a_k, x \in \mathbb{R}.$$

We define the addition of two power series as

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

The more interesting operation is multiplication. Using the cauchy product formula, we have

$$\left(\sum_{k=0}^{\infty} a_k x^k\right) \cdot \left(\sum_{k=0}^{\infty} b_k x^k\right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_{n-k} b_k\right) x^n.$$

If these sums were finite, this would be the same as polynomial multiplication. Power series serve as a generalization of polynomials.

Definition 4.33: Radius of Convergence

Let $\sum_{k=0}^{\infty} a_k x^k$ be a power series. Let

$$\rho = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \text{ and } R = \frac{1}{\rho} \in [0, \infty].$$

Where we used the convention that $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$. We call R the RADIUS OF CONVERGENCE

Theorem 4.34:

Let $\sum_{k=0}^{\infty} a_k x^k$ be a power series with $R \in (0, \infty]$. Then, the power series

- converges absolutely for $x \in (-R, R)$
- does not converge for $|x| > R$.

In particular, $f(x) = \sum_{k=0}^{\infty} a_k x^k$ is well-defined on $(-R, R)$.

Proof. Let $x \in \mathbb{R}$, let $\rho = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Look at the series $\sum_{k=0}^{\infty} a_k x^k$. Define

$$\alpha := \limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = |x| \cdot \limsup_{n\to\infty} \sqrt[n]{|a_n|} = |x|\rho.$$

By root criterion, the series converges absolutely if $\alpha < 1$ and thus if $|x| < \frac{1}{\rho} = R$.

The series does not converge if $\alpha > 1$ and thus if $|x| > \frac{1}{\rho} = R$. □

Theorem 4.35:

Let $\sum_{k=0}^{\infty} a_k x^k$ be a power series with radius of convergence $R \in (0, \infty]$. Define

$$f_n(x) = \sum_{k=0}^{n} a_k x^k.$$

Such that f_n is a polynomial of degree n. Then, $\forall r < R$, f_n converges uniformly to f on $[-r, r]$.

Proof. Since $r < R$, by the previous theorem, the series

$$\sum_{k=0}^{\infty} a_k r^k$$

converges absolutely. Thus, $\sum_{k=0}^{\infty} |a_k| r^k < \infty$. Therefore, $\forall \varepsilon > 0 \exists N$ such that

$$\sum_{k=N+1}^{\infty} |a_k| r^k < \varepsilon.$$

Therefore, $\forall x \in [-r, r]$, $|f(x) - f_n(x)|$ is nothing else but the following sum $\left|\sum_{k=n+1}^{\infty} a_k x^k\right|$ which by triangle inequality is bounded by

$$\sum_{k=n+1}^{\infty} |a_k||x|^k \leq \sum_{k=n+1}^{\infty} |a_k| r^k < \varepsilon.$$

This is the definition of uniform convergence on $[-r, r]$. □

Since f_n converges uniformly to f on $[-r, r]$ and each f_n is continuous, f is continuous on $[-r, r]$. Since this is true for all $r < R$, f is continuous on $(-R, R)$.

Remark 4.36:

If the limit $\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$ exists, then

$$R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}.$$

If $\sum a_k x^k$ and $\sum b_k x^k$ both have radius of convergence $\geq R$, then in $(-R, R)$, sum and product also converge.

Example 4.37:

Consider $\sum_{k=1}^{\infty} \frac{x^k}{k}$. In this case,

$$\rho = \limsup_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1.$$

If $|x| < 1$, the series converges absolutely. If $|x| > 1$, it does not converge.

For $|x| = 1$, no conclusion can be drawn. For example,

- $x = 1$: $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges.
- $x = -1$: $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ converges.

We now like to extend the definition of power series to complex numbers.

Definition 4.38: Complex Power Series

A COMPLEX POWER SERIES is a series of the form

$$\sum_{n=0}^{\infty} a_n z^n \quad a_n, z \in \mathbb{C}.$$

We can define radius of convergence in the same way

$$\rho = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \text{ and } R = \frac{1}{\rho} \in [0, \infty].$$

Theorem 4.39:

The series $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely if $|z| < R$ and does not converge if $|z| > R$. In particular, $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is well-defined for $z \in B(0, R)$

If we define $f_n(z) = \sum_{k=0}^{n} a_k z^k$, then f_n converges to f uniformly on $\overline{B(0, r)}$ for all $r < R$. Thus, f is continuous on $B(0, R)$.[1]

4.5 Exponential and Trigonometric Functions

We start by redefining the exponential function using power series.

Definition 4.40:

Given $x \in \mathbb{R}$, we define

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

We first observe that the radius of convergence is infinite since

$$R = \lim_{n\to\infty} \frac{\frac{1}{n!}}{\frac{1}{(n+1)!}} = \lim_{n\to\infty} (n+1) = \infty.$$

Our goal now is to show that this definition is equivalent to the previous one. Secondly, we want to extend the definition to complex numbers. After this, we will define sine and cosine using exponentials.

Lec 21

Definition 4.41: Complex Exponential Function

Given $z \in \mathbb{C}$, we define $\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$.

[1] Proof is similar to the real case. In further detail talked about in Analysis II.

Another way to see the radius of convergence is to use the root criterion. For this, we look at $\sqrt[n]{n!}$. We fix N and get

$$\sqrt[n]{n!} \geq \sqrt[n]{n \cdot (n-1) \cdots N} \geq \sqrt[n]{N^{n-N+1}} = N^{\frac{n-N+1}{n}}.$$

Now, taking the limsup as $n \to \infty$, we get

$$\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n!}} \leq \lim_{n\to\infty} \frac{1}{N^{\frac{n-N+1}{n}}} = \frac{1}{N}.$$

But N is as large as we want, so $\rho = 0$ and thus $R = \infty$.

Proposition 4.42:

$\forall x \in \mathbb{R}$,

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \exp(x).$$

Now that we have $\exp(z)$, we can define $a^z = \exp(z \ln a)$ for $a > 0$ and $z \in \mathbb{C}$.

Recall that $(z+w)^n = \underbrace{(z+w)(z+w)\cdots(z+w)}_{n \text{ times}}$. Grouping the terms, we get

$$(z+w)^n = \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k}.$$

Where we use the notation $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Theorem 4.43:

$\exp : \mathbb{C} \to \mathbb{C}$ is continuous and

$$e^{z+w} = e^z \cdot e^w, \quad |e^z| = e^{\Re(z)}.$$

Proof. Since $R = \infty$, exp is continuous on $\mathbb{C}$.

For the product,

$$\begin{aligned} e^z e^w &= \left(\sum_{n=0}^{\infty} \frac{z^n}{n!}\right) \cdot \left(\sum_{m=0}^{\infty} \frac{w^m}{m!}\right) \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{z^k}{k!} \cdot \frac{w^{n-k}}{(n-k)!}\right) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k}\right) \\ &= \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = e^{z+w}. \end{aligned}$$

For the modulus, recall that $|w| = \sqrt{w\overline{w}}$. Thus, we should look at the complex conjugate of e^z.

$$\begin{aligned} \overline{e^z} &= \overline{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{\overline{z^k}}{k!} \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{\overline{z}^k}{k!} = e^{\overline{z}}. \end{aligned}$$

But now,

$$\begin{aligned} |e^z| &= \sqrt{e^z \cdot \overline{e^z}} = \sqrt{e^z \cdot e^{\overline{z}}} = \sqrt{e^{z+\overline{z}}} \\ &= e^{\frac{z+\overline{z}}{2}} = e^{\Re(z)}. \end{aligned}$$

□

Immediately, we get the following corollary.

Corollary 4.44:
$\forall x \in \mathbb{R}$, $|e^{ix}| = e^0 = 1$.

Let us write this down explicitly.

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(ix)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(ix)^{2n+1}}{(2n+1)!} \\ &= \underbrace{\overbrace{\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}}^{\in\mathbb{R}}}_{:=\cos(x)} + i \underbrace{\overbrace{\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}}^{\in\mathbb{R}}}_{:=\sin(x)} \\ &= \cos(x) + i\sin(x). \end{aligned}$$

Definition 4.45: **Odd and Even Functions**
A function $f : \mathbb{R} \to \mathbb{R}$ is ODD if $f(-x) = -f(x)$ and is EVEN if $f(-x) = f(x)$.

Notice that

$$\cos(-x) = \sum_{n=0}^{\infty} \frac{(-1)^n(-x)^{2n}}{(2n)!} = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = \cos(x).$$

Thus, cos is even. Similarly, $\sin(x) = -\sin(-x)$, so sin is odd.

Theorem 4.46:
The following hold $\forall x, y \in \mathbb{R}$:

- $e^{ix} = \cos(x) + i\sin(x)$
- $\sin(x) = \frac{e^{ix} - e^{-ix}}{2}$
- $\cos(x) = \frac{e^{ix} + e^{-ix}}{2}$
- $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$
- $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$
- $\sin^2(x) + \cos^2(x) = 1$

Proof. The first point is just the definition, $e^{ix} = \cos(x) + i\sin(x)$. We thus also have $e^{-ix} = \cos(x) - i\sin(x)$. Adding/Subtracting these two identities, we get the second and third points.

Taking $e^{i(x+y)} = e^{ix} \cdot e^{iy}$ and expanding both sides we get

$$\begin{aligned} \cos(x+y) + i\sin(x+y) &= (\cos(x) + i\sin(x))(\cos(y) + i\sin(y)) \\ &= \cos(x)\cos(y) - \sin(x)\sin(y) \\ &\quad + i(\sin(x)\cos(y) + \cos(x)\sin(y)). \end{aligned}$$

Equating real and imaginary parts, we get the fourth and fifth points.

Finally, using $1 = |e^{ix}| = \sqrt{\cos^2(x) + \sin^2(x)}$, we get

$$\sin^2(x) + \cos^2(x) = 1.$$

□

Theorem 4.47: **Pi**
$\exists!$ number $\pi \in (0, 4)$ such that $\sin(\pi) = 0$.

For this number, it holds

$$e^{i\frac{\pi}{2}} = i, \quad e^{i\pi} = -1, \quad e^{i2\pi} = 1.$$

Proof. Look at the sequence $\left(\frac{x^n}{n!}\right)_{n\geq 0}$. This is decreasing if $x \in [0, 2]$ and all terms are non-negative. By Leibniz criterion for alternating sequences applied to $\left(\frac{x^2n}{(2n)!}\right)_{n\geq 0}$ and $\left(\frac{x^{2n+1}}{(2n+1)!}\right)_{n\geq 0}$, we get:

$$\begin{aligned} x - x^3 \leq \sin(x) \leq x - \frac{x^3}{3!} + \frac{x^5}{5!} \quad \forall x \in [0, 2] \\ -\frac{x^2}{2} \leq \cos(x) - 1 \leq -\frac{x^2}{2} + \frac{x^4}{4!} \quad \forall x \in [0, 2]. \end{aligned}$$

This is because the power series for sin and cos are

$$\begin{aligned} \sin(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \\ \cos(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots. \end{aligned}$$

Note that $\sin(0) = 0$ and $\sin(1) \geq 1 - \frac{1}{6} > \frac{1}{\sqrt{2}}$. By intermediate value theorem, $\exists p \in (0, 1)$ such that $\sin(p) = \frac{1}{\sqrt{2}}$.

This implies that $\cos^2(p) = 1 - \sin^2(p) = 1 - \frac{1}{2} \Rightarrow \cos(p) = \frac{1}{\sqrt{2}}$, as $\cos \geq 1 - \frac{1}{2} > 0$ is positive on $(0, 1)$.

So we have the number $e^{ip} = \cos(p) + i\sin(p) = \frac{1+i}{\sqrt{2}}$. Define $\pi = 4p$. Then, $e^{i\frac{\pi}{2}} = (e^{ip})^2 = i$. Furthermore, $e^{i\pi} = (e^{i\frac{\pi}{2}})^2 = i^2 = -1$ and $e^{i2\pi} = (e^{i\pi})^2 = (-1)^2 = 1$.

The second equation in particular tells us that $-1 = \cos(\pi) + i\sin(\pi)$. Implying that $\sin(\pi) = 0$ and $\cos(\pi) = -1$.

Recall that $p \in (0, 1)$, so $\pi \in (0, 4)$. Furzthermore, note that $\sin(x) \geq x - \frac{x^3}{6} > 0$ for $x \in [0, 2]$. Thus $\pi \in (2, 4)$.

Assume by contradiction $\exists s \neq \pi \in (2, 4)$ such that $\sin(s) = 0$.

Define $r = |\pi - s|$. Then $|\sin(r)| = |\sin(\pi) - \sin(s)|$. But by the addition formula,

$$|\sin(r)| = |\underbrace{\sin(\pi)}_{=0}\cos(s) - \cos(\pi)\underbrace{\sin(s)}_{=0}| = 0.$$

But $r \in (0, 2)$, and we have already shown that $\sin(x) > 0$ for $x \in (0, 2)$. This is a contradiction, so π is unique. □

From this, we can follow with the following identities.

$$\begin{aligned} \sin\left(x + \frac{\pi}{2}\right) &= \cos(x) \\ \sin(x + \pi) &= -\sin(x) \\ &\vdots \end{aligned}$$

Now given $z \in \mathbb{C} \setminus \{0\}$, define $r = |z|$ then $z = r \cdot \frac{z}{r}$. But then $|\frac{z}{r}| = 1$, so we can write $\frac{z}{r} = e^{i\theta}$ for some $\theta \in [0, 2\pi)$. This representation is called the POLAR COORDINATES. In this notation, if we have $z = re^{i\theta}$ and $w = se^{i\phi}$, then $zw = rse^{i(\theta+\phi)}$.

For the moment, we will not define the logarithm on complex numbers as e^z is not injective.

We now want to introduce two more functions. Take again e^x for $x \in \mathbb{R}$.

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \underbrace{\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}}_{:=\cosh(x)} + \underbrace{\sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}}_{:=\sinh(x)}.$$

Alternatively, we write

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \quad \sinh(x) = \frac{e^x - e^{-x}}{2}.$$

From theese identities we also get $\cosh^2(x) - \sinh^2(x) = 1$.

5 Differential Calculus

5.1 The Derivative

Lec 22

Consider $D \subseteq \mathbb{R}$, where every point $x_0 \in D$ is an accumulation point for $D \setminus \{x_0\}$. This is for example an interval.

Definition 5.1: Derivative

Given $f : D \to \mathbb{R}$, we define

$$f'(x_0) = \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(x_0 + h) - f(x_0)}{h}.$$

If $f'(x_0)$ exists, then f is DIFFERENTIABLE at x_0. We say that f is DIFFERENTIABLE on D if it is differentiable at all points $x_0 \in D$. f' is the DERIVATIVE of f.

As a convention, writing $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ means that we don't consider the case $x = x_0$. Thus

$$\begin{aligned} f'(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ \Leftrightarrow & \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) = 0 \\ &= \lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0. \end{aligned}$$

Recalling the notion of little-o notation, this is equivalent to

$$f(x) - f(x_0) - f'(x_0)(x - x_0) = o(x - x_0).$$

In particular, this formula implies that

$$\begin{aligned} \lim_{x \to x_0} f(x) &= \lim_{x \to x_0} (f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)) \\ &= f(x_0). \end{aligned}$$

Thus, if f is differentiable at x_0, then it is also continuous at x_0.

The geometric interpretation of the derivative is that $f'(x_0)$ is the slope of the tangent line to the graph of f at the point $(x_0, f(x_0))$.

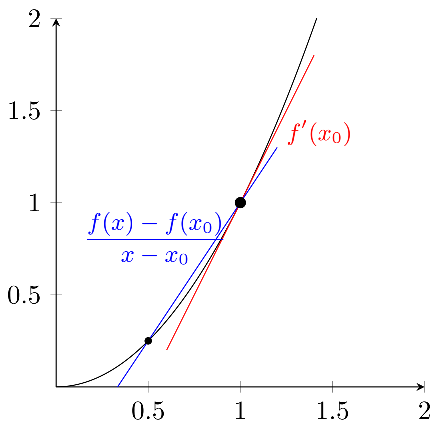


Figure 23: Graph of $f(x) = x^2$ and its tangent line at $x_0 = 1$.

Example 5.2:

Consider $f(x) = 1$. Then

$$f'(x_0) = \lim_{x \to x_0} \frac{1 - 1}{x - x_0} = 0.$$

Example 5.3:

Consider $f(x) = x$. Then

$$f'(x_0) = \lim_{x \to x_0} \frac{x - x_0}{x - x_0} = 1.$$

Example 5.4:

Consider $f(x) = e^x$. Then

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} \\ &= \lim_{h \to 0} \frac{e^x(e^h - 1)}{h} \\ &= e^x \cdot \lim_{h \to 0} \frac{e^h - 1}{h} \\ &= e^x \lim_{h \to 0} \frac{\sum_{n=0}^{\infty} \frac{h^n}{n!} - 1}{h} \\ &= e^x \lim_{h \to 0} \frac{\sum_{k=1}^{\infty} \frac{h^k}{k!}}{h} \\ &= e^x \lim_{h \to 0} \sum_{k=0}^{\infty} \frac{h^k}{(k+1)!} \end{aligned}$$

.

Define $g(x) = \sum_{n=0}^{\infty} \frac{x^n}{(n+1)!}$. Then g has radius of convergence $R = \infty$. Therefore, g is continuous on $\mathbb{R}$. So $\lim_{h \to 0} g(h) = g(0) = \frac{1}{1!} = 1$. Where we used the convention that $x^0 = 1$.

Example 5.5:

Consider $\alpha \in \mathbb{C}$, $f(x) = e^{\alpha x}$. Then

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{e^{\alpha(x+h)} - e^{\alpha x}}{h} \\ &= e^{\alpha x} \cdot \frac{e^{\alpha h} - 1}{h} \\ &= e^{\alpha x} \sum_{k=1}^{\infty} \frac{(\alpha h)^k}{hk!} \\ &= e^{\alpha x} \alpha \sum_{k=1}^{\infty} \frac{(\alpha h)^{k-1}}{k!} \\ &= e^{\alpha x} \alpha \sum_{m=0}^{\infty} \frac{(\alpha h)^m}{(m+1)!} \\ &= e^{\alpha x} \alpha g(\alpha h) \\ &= \alpha e^{\alpha x}. \end{aligned}$$

Example 5.6:

Consider $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{1}{x}$. Then

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\frac{1}{x+h} - \frac{1}{x}}{h} \\ &= \lim_{h\to 0} \frac{x-(x+h)}{(x+h)xh} \\ &= \lim_{h\to 0} \frac{-1}{(x+h)x} \\ &= -\frac{1}{x^2}. \end{aligned}$$

When we differentiate $f \to f' = f^{(1)}$. If we differentiate again, we get $f'' = f^{(2)}$. Continuing this way, we define the n-th derivative $f^{(n)}$ inductively as $f^{(n)} = (f^{(n-1)})'$. This notation is useful when we have higher order derivatives.

We say that f is N-TIMES DIFFERENTIABLE if $f^{(1)}, \dots, f^{(n)}$ exist and f is N-TIMES CONTINUOUSLY DIFFERENTIABLE if $f^{(1)}, f^{(2)}, \dots, f^{(n)}$ exist and $f^{(n)}$ is continuous. In this case, we write that $f \in C^n(D)$.

So

$$\begin{aligned} C^0(D) &= \{f : D \to \mathbb{R} \mid \text{f is continuous}\} \\ C^1(D) &= \{f : D \to \mathbb{R} \mid \text{f diff. and } f^{(1)}\text{cont.}\} \\ C^2(D) &= \{f : D \to \mathbb{R} \mid \text{f 2-x diff. and } f^{(2)} \text{ cont.}\} \\ &\vdots \\ C^\infty(D) &= \bigcap_{n=0}^{\infty} C^n(D) \\ &= \{f : D \to \mathbb{R} \mid \text{f n-x diff. } \forall n \in \mathbb{N}\} \end{aligned}$$

Example 5.7:

Consider $f(x) = \operatorname{sgn}(x)x^2$. Then as an exercise, show that

$$f'(x) = \begin{cases} 2x, & x > 0 \\ 0, & x = 0 \\ -2x, & x < 0 \end{cases}.$$

The case $x = 0$ there is

$$\lim_{h\to 0} \frac{\operatorname{sgn}(h)h^2 - 0}{h} = \lim_{h\to 0} |h| = 0.$$

Thus $f'(x) = 2|x|$. So f' is continuous, and $f \in C^1(\mathbb{R})$. But $2|x|$ is not differentiable at $x = 0$, so $f \notin C^2(\mathbb{R})$.

Proposition 5.8:

Suppose, $f, g : D \to \mathbb{R}$ are differentiable at $x_0 \in D$. Then, $f + g$ and $f \cdot g$ are differentiable at x_0 and

$$\begin{aligned} (f+g)'(x_0) &= f'(x_0) + g'(x_0) \\ (f\cdot g)'(x_0) &= f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

Proof. For the sum,

$$\begin{aligned} &\lim_{x\to x_0} \frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} \\ &= \lim_{x\to x_0} \frac{f(x) - f(x_0) + g(x) - g(x_0)}{x - x_0} \\ &= \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(x_0) + g'(x_0). \end{aligned}$$

In the case of the product,

$$\begin{aligned} &\lim_{x\to x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \lim_{x\to x_0} \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \lim_{x\to x_0} \frac{(f(x) - f(x_0))}{x - x_0} g(x) + f(x_0) \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

□

Remark 5.9:

Take $g(x) = \alpha \in \mathbb{R}$. Then

$$(\alpha f)'(x_0) = \alpha f'(x_0).$$

Proposition 5.10:

Given f, g as before but n times differentiable at x_0.

$$\begin{aligned} (f+g)^{(n)}(x_0) &= f^{(n)}(x_0) + g^{(n)}(x_0) \\ (f\cdot g)^{(n)}(x_0) &= \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x_0) g^{(n-k)}(x_0). \end{aligned}$$

Corollary 5.11:

The following hold:

- $(x^n)' = nx^{n-1}$.
- $\sin' = \cos$ and $\cos' = -\sin$.
- $\sinh' = \cosh$ and $\cosh' = \sinh$.

Proof. For the first point we use induction. For $n = 0$, $x^0 = 1$ and $1' = 0$. Assume now that $(x^n)' = nx^{n-1}$. Then,

$$\begin{aligned} (x^{n+1})' = (x^n \cdot x)' &= (x^n)' \cdot x + x^n \cdot 1 \\ &= nx^{n-1} \cdot x + x^n = (n+1)x^n. \end{aligned}$$

For the second point, $\sin(x) = \frac{e^{ix} - e^{-ix}}{2}$. Thus,

$$\begin{aligned} \sin'(x) &= \frac{(e^{ix})' - (e^{-ix})'}{2i} \\ &= \frac{ie^{ix} - (-i)e^{-ix}}{2} = \frac{ie^{ix} + ie^{-ix}}{2i} \\ &= \frac{e^{ix} + e^{-ix}}{2} = \cos(x). \end{aligned}$$

For the last point,

$$\begin{aligned} \sinh'(x) &= \frac{(e^x)' - (e^{-x})'}{2} \\ &= \frac{e^x - (-e^{-x})}{2} = \frac{e^x + e^{-x}}{2} = \cosh(x). \end{aligned}$$

□

Theorem 5.12: Chain Rule

Given $f : D \to E$ and $g : E \to \mathbb{R}$, f differentiable at $x_0 \in D$ and g differentiable at $f(x_0) \in E$. Then, $g \circ f$ is differentiable at x_0 and

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

Proof. Note that

$$g(y) = g(y_0) + g'(y_0)(y - y_0) + \varepsilon_g(y)(y - y_0),$$

where ε_g is defined tautologically as

$$\varepsilon_g(y) = \begin{cases} \frac{g(y)-g(y_0)}{y-y_0} - g(y_0), & y \neq y_0 \\ 0, & y = y_0 \end{cases}.$$

Since g is differentiable at y_0, then $\varepsilon_g(y) \to 0$ as $y \to y_0$. Thus ε_g is continuous at y_0.

Then

$$\begin{aligned} & g(f(x)) \\ & = g(f(x_0)) + g'(f(x_0))(f(x) - f(x_0)) + \varepsilon_g(f(x))(f(x) - f(x_0)) \\ & \lim_{x \to x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} \\ & = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot (g'(f(x_0)) + \varepsilon_g(f(x))) .. \end{aligned}$$

Since ε_g is continuous, $\varepsilon_g(f(x_0)) = \varepsilon_g(y_0) = 0$. So the limit becomes

$$f'(x_0) \cdot g'(f(x_0)) + 0 = g'(f(x_0)) \cdot f'(x_0).$$

□

Example 5.13:

Consider $f(x) = x^2 \sin(\frac{1}{x})$.

The derivative at $x \neq 0$ is

$$f' = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) \notin C^0.$$

But at $x = 0$, we have

$$f'(0) = \lim_{x \to 0} \frac{x^2 \sin\left(\frac{1}{x}\right) - 0}{x - 0} = \lim_{x \to 0} x \sin\left(\frac{1}{x}\right) = 0.$$

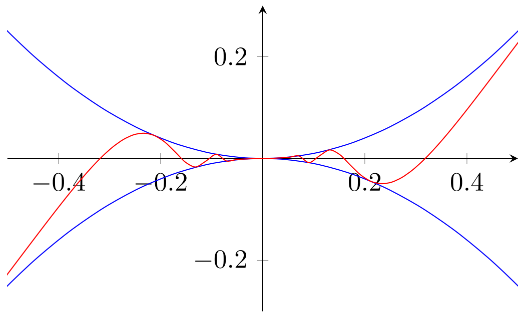


Figure 24: Graph for exercise 5.13.

Lec 23

Corollary 5.14: Quotient Rule

Given f, g differentiable at $x_0 \in D$. Then $\frac{f}{g}$, if $g(x_0) \neq 0$, is differentiable at x_0 and

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}.$$

Proof. Recall that $\frac{1}{x}$ is differentiable and

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

Call this function $\psi(x) = \frac{1}{x}$. Then,

$$\left(\frac{1}{g}\right)'(x_0) = \psi'(g(x_0)) \cdot g'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}.$$

Now we compute $\left(\frac{f}{g}\right)'(x_0)$ using the product rule:

$$\begin{aligned} \left(\frac{f}{g}\right)'(x_0) &= \left(f \cdot \frac{1}{g}\right)'(x_0) \\ &= f'(x_0) \cdot \frac{1}{g(x_0)} + f(x_0) \cdot \left(\frac{1}{g}\right)'(x_0) \\ &= \frac{f'(x_0)}{g(x_0)} - f(x_0) \cdot \frac{g'(x_0)}{(g(x_0))^2} \\ &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}. \end{aligned}$$

□

Example 5.15:

Given $f(x) = \exp(\sin\sin(x^2))$. Then we can write $f(x) = \exp(g(x))$ with $g(x) = \sin(\sin(x^2))$. Thus,

$$f'(x) = \exp(g(x)) \cdot g'(x).$$

Now, $g(x) = \sin(h(x))$ with $h(x) = \sin(x^2)$. So

$$g'(x) = \cos(h(x)) \cdot h'(x).$$

Finally, $h(x) = \sin(k(x))$ with $k(x) = x^2$. So

$$h'(x) = \cos(k(x)) \cdot k'(x) = \cos(x^2) \cdot 2x.$$

Putting everything together, we get

$$f'(x) = \exp(\sin(\sin(x^2))) \cdot \cos(\sin(x^2)) \cdot \cos(x^2) \cdot 2x.$$

Theorem 5.16: Inverse Function Theorem

Given $f : D \to E$, f continuous, bijective and f^{-1} continuous. Let f be differentiable at $\overline{x_0}$, let $\overline{y_0} = f(\overline{x_0})$ and $f'(x_0) \neq 0$ Then f^{-1} is differentiable at $\overline{y_0}$ and

$$(f^{-1})'(\overline{y_0}) = \frac{1}{f'(\overline{x_0})} = \frac{1}{f' \circ f^{-1}(\overline{y_0})}.$$

Notice that f^{-1} is automatically continuous if D is an interval.

Proof. Consider a sequence $(y_n)_{n \geq 0} \subseteq E$ such that $y_n \to \overline{y_0}$. Define $x_n = f^{-1}(y_n)$. Notice that

$$\begin{aligned} \frac{f^{-1}(y_n) - f^{-1}(\overline{y_0})}{y_n - \overline{y_0}} &= \frac{x_n - \overline{x_0}}{f(x_n) - f(\overline{x_0})} \\ &= \frac{1}{\frac{f(x_n) - f(\overline{x_0})}{x_n - \overline{x_0}}}. \end{aligned}$$

Since f^{-1} is continuous, $y_n \to \overline{y_0}$ implies $x_n \to \overline{x_0}$.

Take the limit as $n \to \infty$:

$$\lim_{n \to \infty} \frac{f^{-1}(y_n) - f^{-1}(\overline{y_0})}{y_n - \overline{y_0}} = \lim_{n \to \infty} \frac{1}{\frac{f(x_n) - f(\overline{x_0})}{x_n - \overline{x_0}}}$$

$$f^{-1}(\overline{y_0}) = \frac{1}{f'(\overline{x_0})}.$$

□

Example 5.17:
Consider $g(y) = \log(y), y > 0$. Since $g = f^{-1}$ with $f(x) = e^x$, then

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{e^{\log(y)}} = \frac{1}{y}.$$

Example 5.18:
Consider $f(x) = x^\alpha, \alpha \in \mathbb{R}$. Then

$$f(x) = \exp(\alpha \log(x)).$$

By chain rule,

$$f'(x) = \exp(\alpha \log(x)) \cdot \alpha \cdot \frac{1}{x} = \alpha x^{\alpha-1}.$$

If we know all derivatives of a function at a point, can we reconstruct the function from this information? The answer is no, as the following example shows.

Consider the following function:

$$f(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0 \\ 0, & x \leq 0 \end{cases}.$$

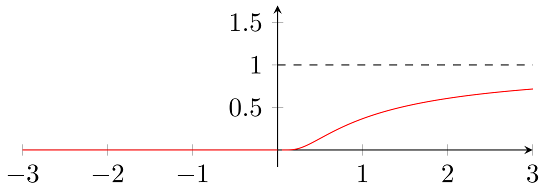


Figure 25: Graph of $f(x)$

We can compute the derivative at $x = 0$ and find that $f'(0) = 0$. In fact, one can show that $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$.

Two interesting results by Karl Weierstrass are the following:

1. There exists a function $f : \mathbb{R} \to \mathbb{R}$ that is continuous everywhere but differentiable nowhere.
2. $C^1([a,b])$ is dense in $C^0([a,b])$.

For the proof of the second result, the idea is to show that given $f \in C^0([a,b])$ and $\varepsilon > 0$, we can find $p \in \mathbb{R}[x]$ such that $f - p < \varepsilon, \forall x \in [a,b]$. Then, since polynomials are differentiable, we are done. This is known as the Weierstrass Approximation Theorem.

But this shows, that $C^1([a,b])$ is not complete! A simple example is given by $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$. Here, f_n converges uniformly to $f(x) = |x|$, which is not in $C^1([a,b])$.

Hence the inheritance property of uniformly convergent sequences of functions does not hold for differentiability.

5.2 Main Theorems of Differential Calculus

Consider the following definition.

Definition 5.19: Local Maximum/Minimum
We call $x_0 \in D$ a LOCAL MAXIMUM if $f(x_0) \geq f(x)$ for $x \in (x_0 - \delta, x_0 + \delta)$.

We call x_0 a STRICT LOCAL MAXIMUM if $f(x_0) > f(x)$ for $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$.

The same definitions hold for LOCAL MINIMUM and STRICT LOCAL MINIMUM with the inequalities reversed.

A LOCAL EXTREMUM is either a local maximum or a local minimum.

In other words x_0 is a local maximum if $\exists \delta > 0$ such that

$$x \in (x_0 - \delta, x_0 + \delta) \cap D \Rightarrow f(x) \leq f(x_0).$$

Proposition 5.20:
Given $f : D \to \mathbb{R}$, x_0 local extremum, x_0 is both a right and left accumulation point of D. Then if f is differentiable at x_0, then $f'(x_0) = 0$.

Proof. Assume x_0 is a local maximum, so $\exists \delta > 0$ such that $f(x) \leq f(x_0)$ for $x \in (x_0 - \delta, x_0 + \delta) \cap D$.

Consider first $x > x_0$. Then

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = f'_+(x_0)..$$

This is what we call the right derivative at x_0. The numerator is non-positive and the denominator is positive, so the $f'_+(x_0) \leq 0$.

Consider now $x < x_0$. Then

$$\lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} = f'_-(x_0)..$$

This is what we call the left derivative at x_0. Now both numerator and denominator are negative, so $f'_-(x_0) \geq 0$.

Since f is differentiable $f'_+(x_0) = f'_-(x_0) = f'(x_0)$. Thus,

$$0 \leq f'(x_0) \leq 0 \Rightarrow f'(x_0) = 0.$$

□

Corollary 5.21:
Given $f : I \to \mathbb{R}$, I interval, x_0 local extremum. Then

- x_0 is an endpoint of I or
- f NOT differentiable at x_0 or
- $f'(x_0) = 0$.

Theorem 5.22: Rolle's Theorem
Given $f : [a,b] \to \mathbb{R}$, continuous and differentiable on (a,b). If $f(a) = f(b)$, then $\exists \xi \in (a,b)$ such that $f'(\xi) = 0$.

Proof. Since f is continuous on $[a,b]$, $\exists x_0$ minimum point for f and $\exists x_1$ maximum point for f, where both are in $[a,b]$.

If either x_0 or $x_1 \in (a,b)$, then f' is zero at such point, so we are done.

So we only need to consider the case when both x_0 and x_1 are endpoints. If both are endpoints, then $\min f = f(x_0) = f(a) = f(b)$. But also $\max f = f(x_1) = f(a) = f(b)$. Thus $\min f = \max f$, implying that f is constant. Therefore, $f' = 0$ everywhere. □

Corollary 5.23:

Given $f : [a, b] \to \mathbb{R}$, continuous and differentiable on (a, b). If $f'(x) \neq 0 \, \forall x \in (a, b)$, then $f(a) \neq f(b)$.

Proof. If by contradiction $f(a) = f(b)$, then by Rolle's Theorem $\exists \xi \in (a, b)$ such that $f'(\xi) = 0$, contradicting the hypothesis. □

Theorem 5.24: Mean Value Theorem

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

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

The visual interpretation of the Mean Value Theorem is seen in figure 26.

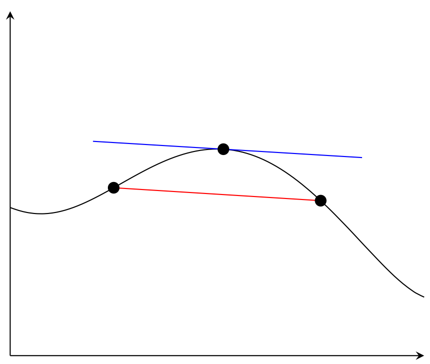


Figure 26: Graph illustrating the Mean Value Theorem.

Proof. Define $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$. Then $g(a) = f(a)$ and $g(b) = f(b) - (f(b) - f(a)) = f(a) = g(a)$.

By Rolle's Theorem, $\exists \xi \in (a, b)$ such that

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

□

Recall that uniform continuity means that $\forall \varepsilon > 0 \exists \delta > 0$ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. For example, $x \to \sqrt{x}$ is uniformly continuous. Another continuity is Lipschitz continuity, which means that $\exists L > 0$ such that $|f(x) - f(y)| \leq L|x - y|$ for all $x, y \in D$. Notice that Lipschitz continuity implies uniform continuity.

Proposition 5.25:

Given $f : [a, b] \to \mathbb{R}$, continuous and differentiable on (a, b). Then f is Lipschitz continuous $\Leftrightarrow$ f' is bounded.

Lec 24

Proof. $\Rightarrow$: Since f is Lipschitz continuous, $\exists L > 0$ such that $|f(x) - f(x_0)| \leq L|x - x_0| \, \forall x, x_0 \in [a, b]$. Dividing by $|x - x_0|$ we see

$$\left| \frac{f(x) - f(x_0)}{x - x_0} \right| \leq L.$$

Taking the limit as $x \to x_0$, we get $|f'(x_0)| \leq L$ for all $x_0 \in (a, b)$, so f' is bounded.

$\Leftarrow$: Assume now $|f'| \leq M$ on (a, b). Given $x, y \in [a, b]$, apply the Mean Value Theorem on $[x, y]$. Thus $\exists \xi \in (x, y)$ such that

$$f'(\xi) = \frac{f(y) - f(x)}{y - x} \Rightarrow |f(y) - f(x)| = |f'(\xi)||y - x| \leq M|y - x|.$$

Therefore, f is Lipschitz continuous with constant M. □

Remark 5.26:

Consider $f : [0, 2\pi] \to \mathbb{C}, f(x) = e^{ix}$. In this case, $f(0) = f(2\pi)$, but $f'(x) = ie^{ix} \neq 0$. Thus the above theorems should only be applied to real-valued functions.

Theorem 5.27: Cauchy Mean Value Theorem

Given $f, g : [a, b] \to \mathbb{R}$, continuous and differentiable on (a, b). Then $\exists \xi \in (a, b)$ such that

$$f'(\xi)[g(b) - g(a)] = g'(\xi)[f(b) - f(a)].$$

If in addition $g'(x) \neq 0$ for all $x \in (a, b)$, then

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

Proof. Define $F(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)]$. Observe that $F(a) = F(b)$. By Rolle's Theorem, $\exists \xi \in (a, b)$ such that $F'(\xi) = 0$. But then

$$0 = F'(\xi) = f'(\xi)[g(b) - g(a)] - g'(\xi)[f(b) - f(a)].$$

If $g'(x) \neq 0$ for all $x \in (a, b)$, then by corollary 5.23 $g(b) \neq g(a)$, so we get the second part of the theorem. □

Theorem 5.28: L'Hôpital

Let $f, g : (a, b) \to \mathbb{R}$ be differentiable functions such that

1. $g(x) \neq 0$ and $g'(x) \neq 0 \forall x$,
2. $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = 0$
3. The limit $\lim_{x \to a^+} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}$ exists.

Then

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = L.$$

Proof. By 2, we can extend f and g on $[a, b)$ defining $f(a) = g(a) = 0$. By 3, $\forall \varepsilon > 0 \exists \delta > 0$ such that

$$z \in [a, a + \delta] \Rightarrow \left| \frac{f'(z)}{g'(z)} - L \right| < \varepsilon].$$

Given $x \in [a, a + \delta]$, note that

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

Applying Cauchy Mean Value Theorem on $[a, x]$, we get

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi_x)}{g'(\xi_x)},$$

where $\xi_x \in (a, x)$. Then

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

□

Example 5.29:

Consider $\lim_{x\to 0} \frac{\sin x}{x}$. Both numerator and denominator go to zero as $x \to 0$. Thus we can apply L'Hôpital's rule:

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1.$$

Example 5.30:

Consider $\lim_{x\to 0} \frac{\sin x - x}{x^3}$. Both numerator and denominator go to zero as $x \to 0$. Thus we can apply L'Hôpital's rule:

$$\begin{aligned}\lim_{x\to 0} \frac{\sin x - x}{x^3} &= \lim_{x\to 0} \frac{\cos x - 1}{3x^2} \\ &= \lim_{x\to 0} \frac{-\sin x}{6x} \\ &= \lim_{x\to 0} \frac{-\cos x}{6} = -\frac{1}{6}.\end{aligned}$$

L'Hôpital's rule does not work if the limit is not of the form $\frac{0}{0}$. A simple counterexample is $\lim_{x\to 0^+} \frac{1}{x}$.

Theorem 5.31: L'Hôpital - Infinity

Let $f, g : (a, b) \to \mathbb{R}$ be differentiable functions such that

1. $g(x) \neq 0$ and $g'(x) \neq 0 \forall x$,
2. $\lim_{x\to a^+} |f(x)| = \lim_{x\to a^+} |g(x)| = +\infty$
3. The limit $\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}$ exists.

Then

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = L.$$

Theorem 5.32: L'Hôpital at Infinity

Given $R > 0, f, g : (R, +\infty) \to \mathbb{R}$ differentiable functions such that

1. $g(x) \neq 0$ and $g'(x) \neq 0 \forall x$,
2. Either $\lim_{x\to+\infty} f(x) = \lim_{x\to+\infty} g(x) = 0$ or otherwise $\lim_{x\to+\infty} |f(x)| = \lim_{x\to+\infty} |g(x)| = +\infty$
3. The limit $\lim_{x\to+\infty} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}$ exists.

Then

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = L.$$

Proof. Apply Theorem 5.28 or 5.31 to

$$x \to f\left(\frac{1}{x}\right), \quad x \to g\left(\frac{1}{x}\right).$$

□

Proposition 5.33:

Given $I \subseteq \mathbb{R}$ interval, $f : I \to \mathbb{R}$ differentiable. Then f is increasing $\Leftrightarrow f' \geq 0$.

Proof. $\Rightarrow$: If f is increasing, $f(x+h) - f(x) \geq 0$ for $h > 0$. and $f(x+h) - f(x) \leq 0$ for $h < 0$. Thus,

$$\frac{f(x+h) - f(x)}{h} \geq 0.$$

Thus, taking the limit as $h \to 0$, we get $f'(x) \geq 0$.

$\Leftarrow$: Assume by contradiction $\exists x < y \in I$ such that $f(x) > f(y)$. Then by the Mean Value Theorem $\exists \xi \in (x, y)$ such that

$$f'(\xi) = \frac{f(y) - f(x)}{y - x} < 0 \lightning$$

□

Corollary 5.34:

Given $f : I \to \mathbb{R}$ differentiable. Then f is constant $\Leftrightarrow f' = 0$.

Proof. If f is connstant, then $f' = 0$. Vice versa, if $f' = 0$, then $(-f)' = 0$ as well, so f is both increasing and decreasing, thus constant. □

Definition 5.35: Convexity

Given $f : I \to \mathbb{R}$. Then f is CONVEX if $\forall a, b \in I, \forall t \in (0, 1)$

$$f(ta + (1-t)b) \leq tf(a) + (1-t)f(b).$$

If the inequality is strict, then f is STRICTLY CONVEX. If $-f$ is (strictly) convex, then f is (STRICTLY) CONCAVE.

Visually, convexity means that the chord between two points on the graph of f lies above the graph itself.

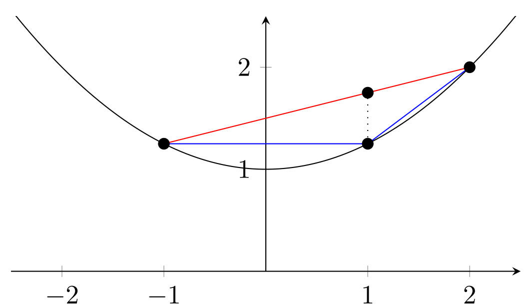


Figure 27: Graph of a convex function.

Lec 25

Another way to express this is that

$$\frac{f(b) - f(x)}{b - x} \geq \frac{f(x) - f(a)}{x - a} \forall x \in (a, b). \qquad (5.1)$$

i.e, the slope of the chord from a to x is less than the slope of the chord from x to b.

Proposition 5.36:

Assume f is differentiable, then

$$f \text{ is convex } \Leftrightarrow f' \text{ is increasing }.$$

Proof. $\Leftarrow$: Let $a < b$, fix $x \in (a, b)$. Apply the Mean Value Theorem on $[a, x]$ and $[x, b]$. Thus $\exists \xi \in (a, x), \eta \in (x, b)$ such that

$$f'(\xi) = \frac{f(x) - f(a)}{x - a}, \quad f'(\eta) = \frac{f(b) - f(x)}{b - x}.$$

But since f' is increasing, $f'(\eta) \geq f'(\xi)$, so

$$\frac{f(b)-f(x)}{b-x} \geq \frac{f(x)-f(a)}{x-a}.$$

$\Rightarrow$: Fix $a < b$, take $h > 0$ small enough so that $a+h < b-h$. Apply (5.1) on $[a, b-h]$ with $x = a+h$. Then

$$\frac{f(a+h)-f(a)}{h} \leq \frac{f(b-h)-f(a+h)}{b-h-(a+h)}.$$

Applying (5.1) on $[a+h, b]$ with $x = b-h$, we get

$$\frac{f(b-h)-f(a+h)}{b-h-(a+h)} \leq \frac{f(b)-f(b-h)}{h}.$$

Combining both inequalities, we get

$$\frac{f(a+h)-f(a)}{h} \leq \frac{f(b)-f(b-h)}{h}.$$

Letting $h \to 0$, we get $f'(a) \leq f'(b)$, hence f' is increasing. □

Recall that g is increasing $\Leftrightarrow g' \geq 0$. Thus, we can see the following corollary.

Corollary 5.37:

If f is twice differentiable, then

$$f \text{ is convex } \Leftrightarrow f' \text{ is increasing } \Leftrightarrow f'' \geq 0.$$

Remark 5.38:

The function $x \to |x|$ is convex, but it is not differentiable at $x = 0$.

Example 5.39:

We proofed that $\forall n \geq 1, a \geq -1, (1+a)^n \geq 1+na$ by means of induction. Suppose now we want to proof this $\forall p \geq 1$ but now $p \in \mathbb{R}$.

Solution. Define $f(x) = (1+x)^p - 1 - px$ for $x \in [-1, +\infty)$. The goal is to show that $f \geq 0$.

Now, note that $f'(x) = p(1+x)^{p-1} - p$. Furthermore,

$$f''(x) = p(p-1)(1+x)^{p-2} \geq 0 \forall x \in [-1, +\infty),$$

since $p \geq 1$ and $(1+x)^{p-2} \geq 0$. Thus, f is convex.

Since $f(0) = 1^p - 1 = 0$. Furthermore, $f'(0) = p - p = 0$. By convexity, f is increasing, implying that $f'(x) \geq 0$ for $x \geq 0$. Thus, f is increasing on $[0, +\infty)$, so

$$f(x) \geq f(0) = 0 \forall x \in [0, +\infty).$$

Furthermore, for $f' \leq 0$ on $[-1, 0]$, so f is decreasing on $[-1, 0]$. Thus, f is decreasing on $[-1, 0]$, so

$$f(x) \geq f(0) = 0 \forall x \in [-1, 0].$$

Combining both, we get $f(x) \geq 0$ for all $x \in [-1, +\infty)$, which was the goal.

5.3 Differentiation of the Trigonometric Functions

Recall that $\sin' = \cos$ and $\cos' = -\sin$.

This already gives us a few properties. For example, since $\cos > 0$ on $(-\frac{\pi}{2}, \frac{\pi}{2})$, sin is increasing on this interval. Thus, sin will have an inverse which we call $\arcsin : (-1, 1) \to (-\frac{\pi}{2}, \frac{\pi}{2})$. Furthermore, $\sin''$ is $-\sin$, so sin is concave on $(0, \pi)$ and convex on $(-\pi, 0)$.

For the derivative the following holds

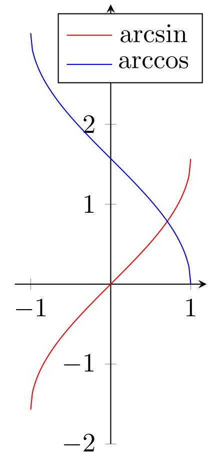


Figure 28: Graph of arcsin and arccos.

Proposition 5.40: Derivative of Arcsine

The derivative of arcsin is given by

$$\arcsin'(s) = \frac{1}{\sqrt{1-s^2}}.$$

Proof. We use the inverse function theorem:

$$\arcsin'(s) = \frac{1}{\sin'(\arcsin(s))} = \frac{1}{\cos(\arcsin(s))} = \frac{1}{\sqrt{1-s^2}}.$$

□

We can do the same for cos and its inverse $\arccos : (-1, 1) \to (0, \pi)$. Since $\cos < 0$ on $(\frac{\pi}{2}, \frac{3\pi}{2})$, arccos is decreasing.

Proposition 5.41: Derivative of Arccosine

The derivative of arccos is given by

$$\arccos'(s) = \frac{-1}{\sqrt{1-s^2}}.$$

Proof.

$$\arccos'(s) = \frac{1}{\cos'(\arccos(s))} = \frac{-1}{\sin(\arccos(s))} = \frac{-1}{\sqrt{1-s^2}}.$$

□

Next up, consider $\tan(x) = \frac{\sin x}{\cos x}$. The derivative is given by the quotient rule:

$$\begin{aligned}\tan'(x) &= \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \frac{1}{\cos^2 x}.\end{aligned}$$

The tan function is increasing on $(-\frac{\pi}{2}, \frac{\pi}{2})$, so it has an inverse $\arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$.

The derivative of arctan is

Proposition 5.42: Derivative of Arctangent

The derivative of arctan and arccot are given by

$$\arctan'(t) = \frac{1}{1+t^2} \qquad \text{arccot}'(t) = \frac{-1}{1+t^2}.$$

Proof.

$$\arctan'(t) = \frac{1}{\tan'(\arctan(t))} = \cos^2(\arctan(t))$$
$$= \frac{1}{1+\tan^2(\arctan(t))} = \frac{1}{1+t^2}.$$

Where we used the identity $1 + \tan^2 x = \frac{1}{\cos^2 x}$.

The proof for arccot is analogous. □

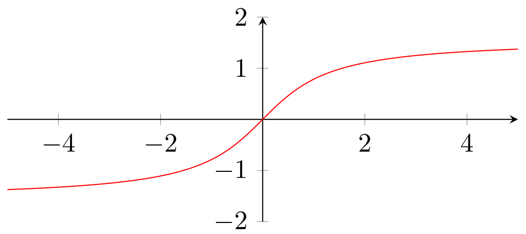


Figure 29: Graph of arctan.

Recalling the hyperbolic functions, we can introduce the inverse hyperbolic sine $\text{arcsinh} : \mathbb{R} \to \mathbb{R}$ and it holds that

Proposition 5.43: Properties of Hyperbolic Arcsine

The derivative of arcsinh is given by

$$\text{arcsinh}'(t) = \frac{1}{\sqrt{1+t^2}}.$$

Furthermore it holds that

$$\text{arcsinh} = \log\left(s + \sqrt{1+s^2}\right).$$

Proof.

$$\text{arcsinh}'(t) = \frac{1}{\sinh'(\text{arcsinh}(t))} = \frac{1}{\cosh(\text{arcsinh}(t))}$$
$$= \frac{1}{\sqrt{1+t^2}}.$$

Let now, $s = \sinh x = \frac{e^x - e^{-x}}{2}$. Then multiplying by e^x,

$$\frac{e^{2x}-1}{2} - se^x = 0.$$

Let $y = e^x$. Then

$$\frac{y^2}{2} - \frac{1}{2} - sy = 0$$
$$y = s \pm \sqrt{s^2+1} = e^x.$$

Since $e^x > 0$, we take the positive root, so

$$x = \text{arcsinh}(s) = \log\left(s + \sqrt{1+s^2}\right).$$

□

For the hyperbolic cosine, we need to be more careful as we can only invert it on $[0, +\infty)$. Thus, we define $\text{arccosh} : [1, +\infty) \to [0, +\infty)$.

Proposition 5.44: Properties of Hyperbolic Arccosine

The derivative of arccosh is given by

$$\text{arccosh}'(t) = \frac{1}{\sqrt{t^2-1}}.$$

Furthermore it holds that

$$\text{arccosh} = \log\left(s + \sqrt{s^2-1}\right).$$

Proof.

$$\text{arccosh}'(t) = \frac{1}{\sinh(\text{arccosh}(t))} = \frac{1}{\sqrt{t^2-1}}.$$

Playing the same game as before, we get

$$\text{arccosh}(s) = \log\left(s + \sqrt{s^2-1}\right) \forall s \geq 1.$$

□

For the hyperbolic tangent, we get

Proposition 5.45: Explicit form of Hyperbolic Arctangent

It holds that

$$\text{arctanh}(x) = \frac{1}{2}\log\left(\frac{1+x}{1-x}\right).$$

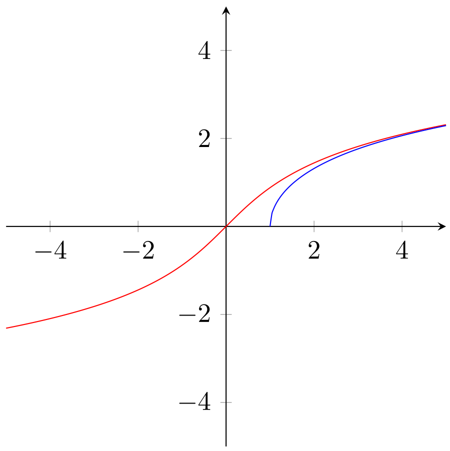


Figure 30: Graph of arcsinh and arccosh.

6 The Riemann Integral

Lec 26

6.1 Step Functions and their Integral

We start with the following definition.

Definition 6.1: Partition

A PARTITION of a set X is a collection $\mathcal{P}$ of subsets of X such that

$$X = \bigcup_{A \in \mathcal{P}} A \text{ and } A \cap B = \varnothing \forall A, B \in \mathcal{P}.$$

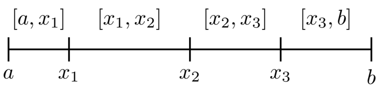


Figure 31: A partition of the interval $[a, b]$.

Definition 6.2: Decomposition

A DECOMPOSITION of an interval $[a, b]$ is a family of points $\{x_0, \dots, x_n\}$ such that

$$a = x_0 < x_1 < x_2 < \dots < x_n = b.$$

A decomposition gives a partition of $[a, b]$ into subintervals

$$[a, b] = \{a\} \cup (x_0, x_1) \cup [x_1] \cup (x_1, x_2) \cup \dots \cup (x_{n-1}, x_n) \cup \{b\}.$$

A decomposition $\{a = y_0 < y_1 < \dots < y_m = b\}$ is a REFINEMENT of $\{a = x_0 < x_1 < \dots < x_n = b\}$ if

$$\{x_0, \dots, x_n\} \subseteq \{y_0, \dots, y_m\}.$$

Definition 6.3: Step Function

A function $f : [a, b] \to \mathbb{R}$ is a STEP FUNCTION if there exists a decomposition $\{x_0, \dots, x_n\}$ of $[a, b]$ such that $f|_{(x_k, x_{k+1})}$ is constant for each $k = 0, \dots, n-1$.

In this case, we say that f is a step function with respect to the decomposition $\{x_0, \dots, x_n\}$.

Proposition 6.4:

If $f, g : [a, b] \to \mathbb{R}$ are step functions, then

$$\alpha f + \beta g \text{ is a step function } \forall \alpha, \beta \in \mathbb{R}.$$

The only critical point of this proposition is the case where the step functions are defined with respect to different decompositions.

Proof. If f is a step function for $\{x_0, \dots, x_n\}$ and g is a step function for $\{y_0, \dots, y_m\}$, consider a common refinement by taking the union of

$$\{x_0, \dots, x_n\} \cup \{y_0, \dots, y_m\} = \{z_0, \dots, z_N\}.$$

Then $f|_{(z_k, z_{k+1})}$ and $g|_{(z_k, z_{k+1})}$ are constant. Therefore, also

$$(\alpha f + \beta g)|_{(z_k, z_{k+1})}$$

is constant. Hence, it is a step function. □

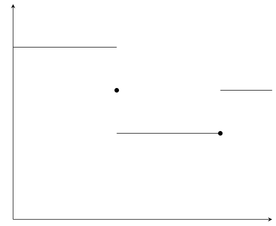


Figure 32: Graph of a step function.

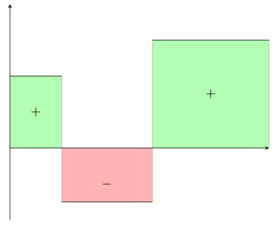


Figure 33: Integral of a step function.

Definition 6.5: Integral of a Step Function

Let $f : [a, b] \to \mathbb{R}$ be a step function with respect to the decomposition $\{a = x_0 < \dots < x_n = b\}$. Then,

$$\int_a^b f(x)\mathrm{d}x := \sum_{k=1}^n c_k (x_k - x_{k-1})$$

where $c_k = f|_{(x_{k-1}, x_k)}$ is the INTEGRAL of f from a to b.

We should check that the integral is well-defined, i.e. that it does not matter which decomposition we choose.

Remark 6.6:

The integral is independent of the decomposition.

Proposition 6.7:

Let $f, g : [a, b] \to \mathbb{R}$ be step functions. Then,

$$\int_a^b (\alpha f + \beta g)(x)\mathrm{d}x = \alpha \int_a^b f(x)\mathrm{d}x + \beta \int_a^b g(x)\mathrm{d}x.$$

Proof. By taking the union of the decompositions of f and g, we find a decomposition

$$\{a = x_0 < \dots < x_n = b\} \quad f|_{(x_{k-1}, x_k)} = c_k, \quad g|_{(x_{k-1}, x_k)} = d_k.$$

Then, $(\alpha f + \beta g)|_{(x_{k-1}, x_k)} = \alpha c_k + \beta d_k$. Then

$$\begin{aligned}\int_a^b (\alpha f + \beta g)(x)\mathrm{d}x &= \sum_{k=1}^n (\alpha c_k + \beta d_k)(x_k - x_{k-1}) \\ &= \alpha \sum_{k=1}^n c_k (x_k - x_{k-1}) + \beta \sum_{k=1}^n d_k (x_k - x_{k-1}) \\ &= \alpha \int_a^b f(x)\mathrm{d}x + \beta \int_a^b g(x)\mathrm{d}x.\end{aligned}$$

□

Proposition 6.8:

Given $f, g : [a, b] \to \mathbb{R}$ step functions. If $f \leq g$, then

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

Proof. Let $\{a = x_0 < \dots < x_n = b\}$ be a decomposition for both f and g, that is

$$f|_{(x_{k-1}, x_k)} = c_k, \quad g|_{(x_{k-1}, x_k)} = d_k.$$

Then $f \leq g \Rightarrow c_k \leq d_k$. Hence,

$$\int_a^b f(x)\mathrm{d}x = \sum_{k=1}^n c_k(x_k - x_{k-1}) \leq \sum_{k=1}^n d_k(x_k - x_{k-1}) = \int_a^b g(x)\mathrm{d}x.$$

□

Corollary 6.9:
$g \geq 0$ implies $\int_a^b g(x)\mathrm{d}x \geq 0$.

Exercise 6.10:
Show that $\forall c \in (a, b)$,

$$\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x.$$

6.2 Definition of the Riemann Integral

We start with a remark on supremum and infimum.

Remark 6.11:
Given two intervals, $A, B \subseteq \mathbb{R}$, assume that $s \leq t$ for all $s \in A, t \in B$. Then, $\sup A \leq \inf B$. Also $\sup A = \inf B$ is equivalent to $\forall \varepsilon > 0, \exists s \in A, t \in B$ such that $t - s < \varepsilon$.

Let $\mathcal{SF}$ be the set of step functions.

Definition 6.12:
Let $f : [a, b] \to \mathbb{R}$. Define the SET OF LOWER SUMS as

$$\mathcal{L}(f) \left\{ \int_a^b f(x)\mathrm{d}x \mid l, \in \mathcal{SF}, l \leq f \right\} \subseteq \mathbb{R}.$$

Similarly, define the SET OF UPPER SUMS as

$$\mathcal{U}(f) = \left\{ \int_a^b u(x)\mathrm{d}x \mid u \in \mathcal{SF}, u \geq f \right\} \subseteq \mathbb{R}.$$

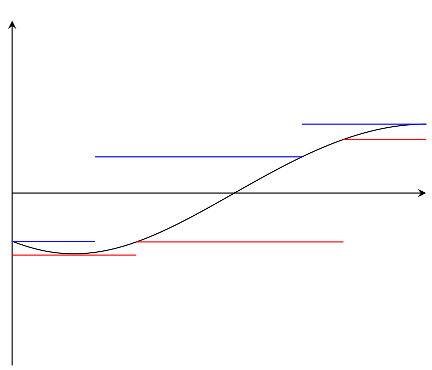


Figure 34: Lower and upper step functions for f.

If f is bounded then $\exists M > 0$ such that

$$-M \leq f(x) \leq M \forall x \in [a, b].$$

Hence $\mathcal{L}(f)$ and $\mathcal{U}(f)$ are non-empty.

If $l \leq f \leq u$, then by the previous proposition,

$$\int_a^b l(x)\mathrm{d}x \leq \int_a^b u(x)\mathrm{d}x \Rightarrow \forall s \in \mathcal{L}(f), t \in \mathcal{U}(f), s \leq t.$$

By the property of supremum and infimum, we have

$$\sup \mathcal{L}(f) \leq \inf \mathcal{U}(f).$$

Definition 6.13:
Let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is RIEMANN INTEGRABLE if

$$\sup \mathcal{L}(f) = \inf \mathcal{U}(f) = \int_a^b f(x)\mathrm{d}x.$$

Not every function is Riemann integrable. Consider for example

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

Then, for any step function $l \leq f$, we must have $l \leq 0$, But for any step function $u \geq f$, we must have $u \geq 1$. Thus,

$$\sup \mathcal{L}(f) = 0 \neq 1 = \inf \mathcal{U}(f).$$

Applying what we have seen, f is (Riemann) integrable is equivalent to

$$\forall \varepsilon > 0, \exists s \in \mathcal{L}(f), t \in \mathcal{U}(f) \text{ such that } t - s < \varepsilon.$$

This is equivalent to

$$\forall \varepsilon > 0 \exists l \leq f \leq u : \int_a^b u(x)\mathrm{d}x - \int_a^b l(x)\mathrm{d}x < \varepsilon.$$

Notice that this also means that $|\int_a^b u(x)\mathrm{d}x - \int_a^b f(x)\mathrm{d}x| < \varepsilon$ and $|\int_a^b l(x)\mathrm{d}x - \int_a^b f(x)\mathrm{d}x| < \varepsilon$.

Theorem 6.14:
Given $f, g : [a, b] \to \mathbb{R}$ integrable. Then, $\alpha f + \beta g$ is integrable and

$$\int_a^b (\alpha f + \beta g)(x)\mathrm{d}x = \alpha \int_a^b f(x)\mathrm{d}x + \beta \int_a^b g(x)\mathrm{d}x.$$

Lec 27

Proof. Since f, g are integrable, $\exists l_1 \leq f \leq u_1$ and $l_2 \leq g \leq u_2$ such that

$$\int_a^b u_1(x)\mathrm{d}x - \int_a^b l_1(x)\mathrm{d}x < \varepsilon, \quad \int_a^b u_2(x)\mathrm{d}x - \int_a^b l_2(x)\mathrm{d}x < \varepsilon.$$

and

$$\left|\int_a^b f(x)\mathrm{d}x - \int_a^b l_1(x)\mathrm{d}x\right| < \varepsilon, \quad \left|\int_a^b g(x)\mathrm{d}x - \int_a^b l_2(x)\mathrm{d}x\right| < \varepsilon$$
$$\left|\int_a^b u_1(x)\mathrm{d}x - \int_a^b f(x)\mathrm{d}x\right| < \varepsilon, \quad \left|\int_a^b u_2(x)\mathrm{d}x - \int_a^b g(x)\mathrm{d}x\right| < \varepsilon.$$

Assume $\alpha, \beta \geq 0$. Then,

$$\alpha l_1 + \beta l_2 \leq \alpha f + \beta g \leq \alpha u_1 + \beta u_2.$$

Thus, computing the integrals, we have

$$\left|\int_a^b (\alpha u_1 + \beta u_2) - (\alpha l_1 + \beta l_2)\mathrm{d}x\right| = \left|\alpha \left(\int_a^b u_1 - l_1 \mathrm{d}x\right) + \beta \left(\int_a^b u_2 - l_2 \mathrm{d}x\right)\right| < (\alpha + \beta)\varepsilon.$$

Hence, $\alpha f + \beta g$ is integrable.

Consider now

$$\left|\int_a^b (\alpha f + \beta g) - \alpha \int_a^b f - \beta \int_a^b g\right|$$
$$= \left|\int_a^b (\alpha f + \beta g) - \int_a^b (\alpha l_1 + \beta l_2) + \int_a^b (\alpha l_1 + \beta l_2) - \alpha \int_a^b l_1 - \beta \int_a^b l_2 + \alpha \int_a^b l_1 - \alpha \int_a^b f + \beta \int_a^b l_2 - \beta \int_a^b g\right|$$
$$\leq \left|\int_a^b (\alpha f + \beta g) - \int_a^b (\alpha l_1 + \beta l_2)\right| + \left|\int_a^b (\alpha l_1 + \beta l_2) - \alpha \int_a^b l_1 - \beta \int_a^b l_2\right| + \alpha \left|\int_a^b l_1 - \int_a^b f\right| + \beta \left|\int_a^b l_2 - \int_a^b g\right|$$
$$\leq \left|\int_a^b (\alpha f + \beta g) - \int_a^b (\alpha l_1 + \beta l_2)\right| + 0 + \alpha\varepsilon + \beta\varepsilon$$
$$< 2(\alpha + \beta)\varepsilon.$$

This shows the claim for $\alpha, \beta \geq 0$. If now for example for $\alpha \geq 0, \beta < 0$, we have

$$\alpha l_1 + \beta u_2 \leq \alpha f + \beta g \leq \alpha u_1 + \beta l_2.$$

The rest of the proof is analogous. □

Theorem 6.15: Given f, g integrable, then $f \leq g \Rightarrow \int f \leq \int g$.

Proof. If $l \in \mathcal{SF}$ then $l \leq f \leq g$. Thus,

$$\mathcal{L}(f) \subseteq \mathcal{L}(g) \Rightarrow \sup \mathcal{L}(f) \leq \sup \mathcal{L}(g).$$

But by definition of the integral, this is equivalent to

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

□

Definition 6.16: Positive and Negative Part

Given $f : [a, b] \to \mathbb{R}$, define

$$f^+(x) = \max\{f(x), 0\}, \quad f^-(x) = \max\{-f(x), 0\}.$$

Theese are called the POSITIVE PART and NEGATIVE PART of f.

This notion is useful because $f = f^+ - f^-$ and $|f| = f^+ + f^-$.

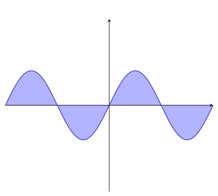


Figure 35: $f(x)$

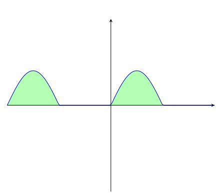


Figure 36: $f^+(x)$

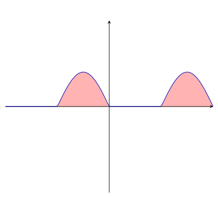


Figure 37: $f^-(x)$

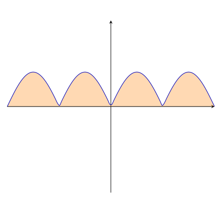


Figure 38: $|f(x)|$

Figure 39: Positive and negative parts of a function.

Exercise 6.17:

Given $z_1, z_2 \in \mathbb{R}$, show that

$$(z_1 - z_2)^+ \geq z_1^+ - z_2^+.$$

Theorem 6.18: If $f : [a, b] \to \mathbb{R}$ is integrable, then f^+, f^- and $|f|$ are integrable and

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

Proof. Fix $\varepsilon > 0$. $\exists l \leq f \leq u$ such that $\int_a^b u - l < \varepsilon$. Note that l^+ and u^+ are step functions and $l^+ \leq f^+ \leq u^+$. By the exercise, we have

$$u^+ - l^+ \leq (u - l)^+ = u(x) - l(x).$$

Thus $\int_a^b u^+ - l^+ \mathrm{d}x \leq \int_a^b u - l \mathrm{d}x < \varepsilon$. Hence, f^+ is integrable. Note that $f^- = f^+ - f$, so f^- is integrable as well.

Finally, $|f| = f^+ + f^-$ is integrable.

Now,

$$\left|\int_a^b f\right| = \left|\int_a^b f^+ - f^-\right| \leq \int_a^b f^+ + \int_a^b f^- = \int_a^b |f|.$$

□

6.3 Integrability Theorems

We now want to show that many functions are integrable so what we did is actually useful.

Theorem 6.19: Integrability of Monotone Functions

Given $f : [a, b] \to \mathbb{R}$ monotone, then f is integrable.

Proof. W.l.o.g. assume that f is increasing. Take $n \in \mathbb{N}$ large, define $x_0 = a, x_k = a + k\frac{b-a}{n}$ for $k = 1, \ldots, n$.

Let $l = f(x_k)$ on (x_k, x_{k+1}) and $u = f(x_{k+1})$ on (x_k, x_{k+1}). and $l = u = f$ at the points x_k. Then,

$$\int_a^b (u - l)\mathrm{d}x = \sum_{k=0}^{n-1} (f(x_{k+1}) - f(x_k))\frac{b-a}{n} = \frac{b-a}{n}(f(b) - f(a))$$

.

Taking n large enough, this is less than any $\varepsilon > 0$.

□

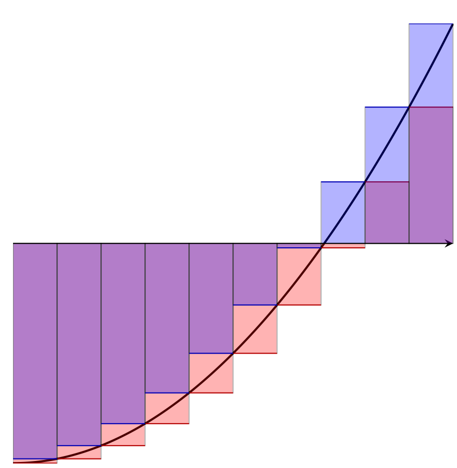


Figure 40: Monotone function and its upper and lower step functions.

Remark 6.20:

The result holds for piecewise monotone functions as well. In other words, if we can find a partition of $[a, b]$ such that f is monotone on each subinterval, then f is integrable.

Theorem 6.21: Integrability of Continuous Functions

Given $f : [a, b] \to \mathbb{R}$ continuous, then f is integrable.

Proof. Since f is continuous on $[a, b]$, implies that f is uniformly continuous. Hence $\forall \varepsilon > 0, \exists \delta > 0$ such that

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Take a partition $a = x_0 < \cdots < x_n = b$ such that

$$x_k - x_{k-1} < \delta \forall k = 1, \ldots, n.$$

Define now c_k and d_k as

$$c_k = \min_{[x_k, x_{k+1}]} f(x), \quad d_k = \max_{[x_k, x_{k+1}]} f(x).$$

So $c_k = f(z_k), z_k \in [x_k, x_{k+1}]$ and $d_k = f(y_k), y_k \in [x_k, x_{k+1}]$. Since $|y_k - z_k| \leq |x_{k+1} - x_k| < \delta$, we have

$$d_k - c_k = f(y_k) - f(z_k) < \varepsilon.$$

Define $l = c_k$ on (x_k, x_{k+1}), $u = d_k$ on (x_k, x_{k+1}) and $l = u = f$ at the points x_k. Then,

$$\begin{aligned} \int_a^b (u - l)(x)\mathrm{d}x &= \sum_{k=0}^{n-1} (d_k - c_k)(x_{k+1} - x_k) \\ &< \sum_{k=0}^{n-1} \varepsilon(x_{k+1} - x_k) \\ &= \varepsilon(b - a). \end{aligned}$$

Hence, f is integrable. □

Remark 6.22:

The result holds for piecewise continuous functions as well, i.e. if we can find a partition of $[a, b]$ such that f is continuous on each subinterval, then f is integrable.

For a function to be piecewise continuous, for each interval, the function must be continuously extendable to the endpoints, i.e. the one-sided limits at the endpoints must exist.

Lec 28

Recall that sequences of functions can converge in different ways. In particular, we have pointwise convergence if $\forall x \in [a, b]$, $\lim_{n\to\infty} f_n(x) = f(x)$ and uniform convergence if $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that

$$n \geq N \Rightarrow |f_n(x) - f(x)| < \varepsilon \quad \forall x \in [a, b].$$

The good thing about uniform convergence is that it preserves properties like continuity. We now ask ourselves what happens with integrability under theese convergences.

Example 6.23:

Consider the sequence of functions $f_n : [0, 1] \to \mathbb{R}$ defined as

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

For all $n \in \mathbb{N}$, f_n is integrable and in particular,

$$\int_0^1 f_n(x)\mathrm{d}x = 1.$$

For $x = 0$, $f_n(0) = 0$ for all $n \in \mathbb{N}$. Hence $\lim_{n\to\infty} f_n(0) = 0$. For $x > 0$, $\exists N \in \mathbb{N}$ such that $n \geq N \Rightarrow \frac{1}{n} < x$. Thus, for $n \geq N$, $f_n(x) = 0$. Hence, $\lim_{n\to\infty} f_n(x) = 0$ for all $x > 0$. So f_n tends to the zero function pointwise, but

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

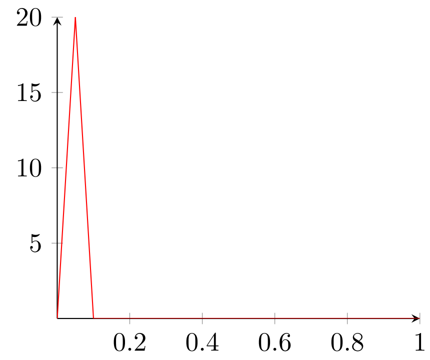


Figure 41: Function f_{10} from Example 6.23.

Theorem 6.24: Integrability under Uniform Convergence

Given $f_n : [a, b] \to \mathbb{R}$ integrable functions that converge uniformly to f. Then, f is integrable and

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

Proof. Given $\varepsilon > 0$, $\exists N$ such that

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

Since f_n is integrable, $\exists l \leq f_n \leq u$ such that $\int_a^b (u - l) < \varepsilon$. But since $|f_N - f| < \varepsilon$, we have

$$f \leq f_N + \varepsilon \leq u + \varepsilon = \hat{u}, \quad f \geq f_N - \varepsilon \geq l - \varepsilon = \hat{l}.$$

But then,

$$\int_a^b (\hat{u} - \hat{l}) = \int_a^b (u - l) + 2\varepsilon(b - a) < \varepsilon + 2\varepsilon(b - a).$$

Since ε was arbitrary, this shows that f is integrable.

Finally,

$$\left| \int f_n - \int f \right| = \left| \int (f_n - f) \right| \leq \int |f_n - f| < \varepsilon(b - a),$$

which shows that $\int f_n \to \int f$. □

7 The Derivative and the Riemann Integral

7.1 The Fundamental Theorem of Calculus

We begin the chapter with a definition.

Definition 7.1: Primitive

Given $I \subseteq \mathbb{R}$ interval, $f : I \to \mathbb{R}$. We say that $F : I \to \mathbb{R}$ is a PRIMITIVE of f if

$$F'(x) = f(x) \quad \forall x \in I.$$

Theorem 7.2: Fundamental Theorem of Calculus

Let $f : [a, b] \to \mathbb{R}$ be continuous. Then, $\forall C \in \mathbb{R}$, the function

$$F(x) = \int_a^x f(t)\mathrm{d}t + C$$

is a primitive of f. Also, all primitives are of this form.

Proof. Since f is continuous, it is integrable.

1) F is a primitive. Fix $x_0 \in [a, b]$. Since f is continuous, given $\varepsilon > 0$, $\exists \delta > 0$ such that

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Consider now first, $x > x_0$ such that $x \in [x_0, x_0 + \delta]$. Then,

$$\begin{aligned}
&\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| \\
&= \left|\frac{1}{x - x_0}\left(\int_a^x f(t)\mathrm{d}t - \int_a^{x_0} f(t)\mathrm{d}t\right) - f(x_0)\right| \\
&= \left|\frac{1}{x - x_0}\int_{x_0}^x f(t)\mathrm{d}t - f(x_0)\right| \\
&= \left|\frac{1}{x - x_0}\int_{x_0}^x f(t)\mathrm{d}t - \frac{1}{x - x_0}\int_{x_0}^x f(x_0)\mathrm{d}t\right| \\
&= \left|\frac{1}{x - x_0}\int_{x_0}^x (f(t) - f(x_0))\mathrm{d}t\right| \\
&\leq \frac{1}{x - x_0}\int_{x_0}^x |f(t) - f(x_0)|\mathrm{d}t.
\end{aligned}$$

Since $t \in [x_0, x] \subseteq [x_0, x_0 + \delta]$, we have $|f(t) - f(x_0)| < \varepsilon$. Thus,

$$\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| \leq \frac{1}{x - x_0}\int_{x_0}^x \varepsilon\mathrm{d}t = \varepsilon.$$

Now, consider $x < x_0$ such that $x \in [x_0 - \delta, x_0]$. Then,

$$\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| = \left|\frac{-1}{x - x_0}\int_x^{x_0} (f(t) - f(x_0))\mathrm{d}t\right|.$$

By the same argument as before, this is less than ε.

2) Let F be a primitive, i.e. $F' = f$. Then,

$$\left(F(x) - \int_a^x f(t)\mathrm{d}t\right)' = f(x) - f(x) = 0.$$

Where we used some primitive as found in part 1) for the integral. This implies that $F(x) - \int_a^x f(t)\mathrm{d}t = C \in \mathbb{R}$. □

Tip 7.3:

When working with integrals, it might be useful to write

$$c = \frac{1}{b - a}\int_a^b c\mathrm{d}x.$$

Corollary 7.4:

Given $F : [a, b] \to \mathbb{R}$ differentiable implies that

$$F(x) = F(a) + \int_a^x F'(t)\mathrm{d}t.$$

Proof. Clearly, F is a primitive of F'. Hence, $\exists c \in \mathbb{R}$ such that

$$F(x) = \int_a^x F'(t)\mathrm{d}t + C \quad \forall x \in [a, b].$$

Choose $x = a$ to find $C = F(a)$. □

Corollary 7.5:

Let $f : [a, b] \to \mathbb{R}$ continuous and F a primitive of f. Then,

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

Proof. Apply the previous corollary with $f = F'$ and $x = b$. □

Another notation for this is $[F(x)]_a^b$.

Example 7.6:

We know $(e^x)' = e^x$. Thus,

$$\int_a^b e^x\mathrm{d}x = e^b - e^a.$$

We know $\sin = -\cos'$. Thus,

$$\int_a^b \sin x\mathrm{d}x = -\cos b + \cos a.$$

Given $0 < a < b$,

$$\int_a^b x^\alpha\mathrm{d}x = \begin{cases} \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1} & \alpha \neq -1 \\ \log(b) - \log(a) & \alpha = -1 \end{cases}.$$

Tip 7.7:

Given $f : \mathbb{R} \to \mathbb{R}$ continuous and odd, then

$$\int_{-a}^a f(x)\mathrm{d}x = 0.$$

This might be particularly useful for the box answer section in the exam.

From this we can derive the formula for integration by parts.

Theorem 7.8: Integration by Parts

Given $f, g : [a, b] \to \mathbb{R}$ differentiable. Then,

$$\int_a^b f'(x)g(x)\mathrm{d}x = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x)\mathrm{d}x.$$

Proof. We have

$$(fg)' = f'g + g'f \Leftrightarrow f'g = (fg)' - g'f.$$

Integrating on $[a, b]$ gives

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

The first integral on the right hand side is by the Fundamental Theorem of Calculus given as

$$\int_a^b (f(x)g(x))'\mathrm{d}x = [f(x)g(x)]_a^b,$$

which concludes the proof. □

As a convention, if we use $\int_b^a f(x)\mathrm{d}x$ with $b < a$, we mean

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

Theorem 7.9: Integration by Substitution, 1

Given $I, J \subseteq \mathbb{R}$ intervals, $f : I \to J$ differentiable and $g : J \to \mathbb{R}$ continuous. Then, $\forall [a,b] \subseteq I$,

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

Proof. Let G be a primitive of g. Then $(G \circ f)' = G' \circ f \cdot f' = g(f)f'$. If we now integrat this, we see that by applying the fundamental theorem of calculus twice, we get

$$\begin{aligned}\int_a^b g(f(x))f'(x)\mathrm{d}x &= \int_a^b (G \circ f)'(x)\mathrm{d}x \\ &= G(f(b)) - G(f(a)) \\ &= \int_{f(a)}^{f(b)} g(y)\mathrm{d}y.\end{aligned}$$

□

Theorem 7.10: Integration by Substitution, 2

Let $I, J \subseteq \mathbb{R}$ intervals, $f : I \to J$ differentiable, $g : J \to \mathbb{R}$ continuous and $[a,b] \subseteq I$. Assume that $f'(x) \neq 0$ for all $x \in [a,b]$. Then,

$$\int_a^b g(f(x))\mathrm{d}x = \int_{f(a)}^{f(b)} g(y)(f^{-1})'(y)\mathrm{d}y'.$$

Proof. Since $f' \neq 0$, f is strictly monotone and hence invertible. Furthermore, by the inverse function theorem, f^{-1} is differentiable with

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

Now,

$$\begin{aligned}\int_a^b g(f(x))\mathrm{d}x &= \int_a^b g(f(x)) \cdot \frac{f'(x)}{f'(x)}\mathrm{d}x \\ &= \int_a^b \frac{g(f(x))}{f'(f^{-1}(f(x)))} f'(x)\mathrm{d}x \\ &= \int_a^b \frac{g(f(x))}{f'(x)} f'(x)\mathrm{d}x \\ &= \int_{f(a)}^{f(b)} g(y)(f^{-1})'(y)\mathrm{d}y.\end{aligned}$$

□

Remark 7.11:

We only have prooven the statement if f' is continuous. However, the general case is also true.

Lec 29

Sometimes we want to integrate functions over unbounded intervals or functions which are unbounded within a bounded interval. Then the following definition is useful.

Definition 7.12: Improper Integral

Given $f : I \to \mathbb{R}$, f is LOCALLY INTEGRABLE if

$$f|_{[a,b]} \text{ is integrable} \quad \forall [a,b] \subseteq I.$$

Let $c = \inf I, d = \sup I$. Fix $x_0 \in I$. Then

$$\int_I f(x)\mathrm{d}x = \lim_{a \to c} \int_a^{x_0} f(x)\mathrm{d}x + \lim_{b \to d} \int_{x_0}^b f(x)\mathrm{d}x.$$

If the limits exist or they diverge but the sum makes sense.

To understand the definition consider the following example.

Example 7.13:

Calculate $\int_0^\infty \frac{1}{1+x^2}\mathrm{d}x$. We have

$$\begin{aligned}\int_0^\infty \frac{1}{1+x^2}\mathrm{d}x &= \lim_{b\to\infty} \int_0^b \frac{1}{1+x^2}\mathrm{d}x \\ &= \lim_{b\to\infty} [\arctan x]_0^b \\ &= \lim_{b\to\infty} (\arctan b - \arctan 0) = \frac{\pi}{2}.\end{aligned}$$

Example 7.14:

Given $\alpha > 0$, determine

$$\begin{aligned}\int_1^\infty \frac{1}{x^\alpha}\mathrm{d}x &= \lim_{b\to\infty} \int_1^b x^{-\alpha}\mathrm{d}x \\ &= \lim_{b\to\infty} \begin{cases} \left[\frac{x^{1-\alpha}}{1-\alpha}\right]_1^b & \alpha \neq 1 \\ [\log x]_1^b & \alpha = 1 \end{cases} \\ &= \begin{cases} \frac{1}{\alpha-1} & \alpha > 1 \\ \infty & \alpha \leq 1 \end{cases}\end{aligned}$$

Lemma 7.15:

If $f \geq 0$ then,

$$\int_a^\infty f(x)\mathrm{d}x = \sup_{b \in (a,\infty)} \int_a^b f(x)\mathrm{d}x.$$

The idea is that the function $b \mapsto \int_a^b f(x)\mathrm{d}x$ is increasing. Hence the limit always exists in $[0, \infty]$ and is equal to the supremum.

Exercise 7.16:

Show that the integral

$$\int_{-\infty}^\infty e^{-x^2}\mathrm{d}x$$

is finite.

The idea here is to find $g(x) \geq e^{-x^2}$ such that it's integral is finite.

Theorem 7.17:

Let $f : [0, \infty) \to \mathbb{R}_{\geq 0}$ be decreasing. Then,

$$\sum_{k=1}^{n+1} f(k) \leq \int_0^{n+1} f(x)\mathrm{d}x \leq \sum_{k=0}^{n} f(k).$$

In particular,

$$\sum_{k=1}^{\infty} f(k) \leq \int_0^{\infty} f(x)\mathrm{d}x \leq \sum_{k=0}^{\infty} f(k).$$

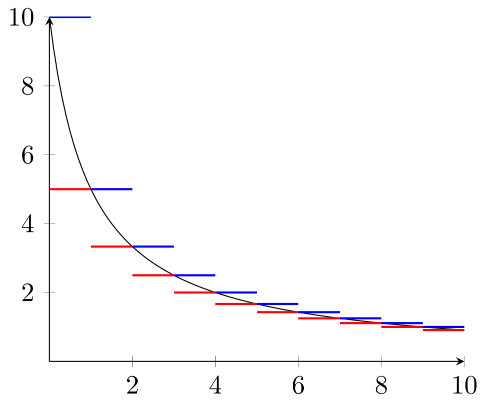


Figure 42: Upper and lower sums for a decreasing function.

Proof. Let l and u be lower and upper sums for f on $[0, n+1]$ with a partition given by the integers $0, 1, \ldots, n+1$. Since f is decreasing, we have

$$\int_0^{n+1} l(x)\mathrm{d}x \leq \int_0^{n+1} f(x)\mathrm{d}x \leq \int_0^{n+1} u(x)\mathrm{d}x.$$

Performing the integrals of the step functions we have

$$\sum_{k=1}^{n+1} f(k) \leq \int_0^{n+1} f(x)\mathrm{d}x \leq \sum_{k=0}^{n} f(k).$$

□

Example 7.18:

Consider $\sum_{k=1}^{\infty} \frac{1}{k}$. So $f(x) = \frac{1}{1+x}$. Then,

$$\int_0^{\infty} \frac{1}{1+x}\mathrm{d}x = \lim_{b\to\infty} [\log(1+x)]_0^b = \infty.$$

Hence, the series diverges.

On the other hand, consider $\sum_{k=1}^{\infty} \frac{1}{k^2}$. So $f(x) = \frac{1}{(1+x)^2}$. Then,

$$\int_0^{\infty} \frac{1}{(1+x)^2}\mathrm{d}x = \lim_{b\to\infty} \left[-\frac{1}{1+x}\right]_0^b = 1.$$

Hence, the series converges.

7.2 Integration and Differentiatiation of Power Series

We can now use the tools we have developed to integrate and differentiate power series.

Theorem 7.19:

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R > 0$. Then,

$$F(x) = \sum_{n=0}^{\infty} a_n \int_0^x t^n \mathrm{d}t = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n+1},$$

has radius of convergence R and $F' = f$.

Proof. Notice that integrating x^n gives $\frac{x^{n+1}}{n+1}$ and differentianting given $n \cdot x^{n-1}$. Thus, we know the integral and derivative of a polynomial.

Let $c_0 = 0, c_n = \frac{a_{n-1}}{n}$ for $n \geq 1$ so that

$$F(x) = \sum_{n=0}^{\infty} c_n x^n.$$

Consider $\limsup_{n\to\infty} \sqrt[n]{|c_n|}$. We have

$$\begin{aligned}
\limsup_{n\to\infty} \sqrt[n]{|c_n|} &= \limsup_{n\to\infty} \sqrt[n]{\frac{|a_{n-1}|}{n}} \\
&= \limsup_{n\to\infty} \sqrt[n]{|a_{n-1}|} \cdot \limsup_{n\to\infty} \sqrt[n]{\frac{1}{n}} \\
&= \limsup_{n\to\infty} \sqrt[n]{|a_{n-1}|} \cdot 1 = \limsup_{n\to\infty} (\sqrt[n-1]{|a_{n-1}|})^{\frac{n-1}{n}} \\
&= \limsup_{n\to\infty} \sqrt[n-1]{|a_n - 1|} = \rho = \frac{1}{R}.
\end{aligned}$$

This shows that the radius of convergence of F is also R.

We now want to show that $F' = f$. Let $[a, b] \subseteq (-R, R)$, with $x \in (a, b)$. Let

$$f_n(t) = \sum_{k=0}^{n} a_k t^k \Rightarrow \int_a^x f_n(t)\mathrm{d}t = \sum_{k=0}^{n} \frac{a_k}{k+1}(x^{k+1} - a^{k+1}).$$

Splitting the sum we have

$$\begin{aligned}
\int_a^x f_n(t)\mathrm{d}t &= \sum_{k=0}^{n} \frac{a_k}{k+1} x^{k+1} - \sum_{k=0}^{n} \frac{a_k}{k+1} a^{k+1} \\
&= F_n(x) - F_n(a),
\end{aligned}$$

since f_n converges uniformly to f on $[a, b]$. Hence, by the integrability under uniform convergence theorem,

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

Let $n \to \infty$ to get

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

So by the fundamental theorem of calculus, $F' = f$. □

The crucial part was $\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1$.

Tip 7.20:

In general its easy to pass limits through integrals, but not through derivatives.

Corollary 7.21:

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R > 0$. Then,

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1},$$

which also has radius of convergence R.

Proof. Let $c_n = (n+1)a_{n+1}$. Define $g(x) = \sum_{n=0}^{\infty} c_n x^n$. Let $\overline{R}$ be its radius of convergence. By the previous result, we have

$$G(x) = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}$$

has radius of convergence $\overline{R}$ and $G' = g$. But notice that

$$\begin{aligned} G(x) &= \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1} = \sum_{n=0}^{\infty} a_{n+1} x^{n+1} \\ &= \sum_{n=1}^{\infty} a_n x^n = f(x) - a_0. \end{aligned}$$

But then, G and f have the same radius of convergence, which implies that $\overline{R} = R$. Also $g = G' = (f - a_0)' = f'$.

□

So the moral of the story is that we can differentiate and integrate power series however we like within their radius
Lec 30 of convergence.

Corollary 7.22:
Power series are smooth (differentiable infinitely many times) inside their radius of convergence.

Example 7.23:
Show that $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{n} = \log(2)$.

Solution. Notice that

$$(\log(1+x))' = \frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty} (-1)^n x^n \quad \text{for } |x| < 1.$$

If we now integrate this on $[0, x]$ with $x \in (-1, 1)$, we get

$$\log(1+x) - \log(1) = \int_0^x \log'(1+t)\mathrm{d}t = \int_0^x \sum_{n=0}^{\infty} (-1)^n t^n \mathrm{d}t.$$

Using the previous theorem we get

$$\log(1+x) - \log(1) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k.$$

Now $\log(1) = 0$, letting $x \to 1$ the series goes to $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ and the left hand side to $\log(2)$.

7.3 Integration Methods

We define indefinite integrals. Given $f : I \to \mathbb{R}$, we write $\int f(x)\mathrm{d}x = F(x) + C$ where F is any primitive of f and $C \in \mathbb{R}$.

Recall that $\int FG'\mathrm{d}x = FG - \int F'G\mathrm{d}x + C$.

There also is a notation where we write $F' = \frac{\mathrm{d}F}{\mathrm{d}x}$. Then, we can write $\int GF'\mathrm{d}x = \int G\frac{\mathrm{d}F}{\cancel{\mathrm{d}x}}\cancel{\mathrm{d}x} = \int G\mathrm{d}F$. Notice that the last notation doesnt mean anything, its just an abuse of notation that is convenient to use. With this notation, integration by parts reads

$$\int G\mathrm{d}F = FG - \int F\mathrm{d}G + C.$$

For integration by substitution, we have

$$\int g \circ f f'\mathrm{d}x = \int (G \circ f)'\mathrm{d}x = G \circ f + C'.$$

Letting $f(x) = u$, we have that this integral can be written as

$$\int g(u)\mathrm{d}u + C.$$

If we instead are doing the second substitution method, we let

$$u = f(x) \Rightarrow x = f^{-1}(u) \Rightarrow \mathrm{d}x = (f^{-1})'(u)\mathrm{d}u.$$

And hence, can write in Leibniz notation

$$\int g(u)(f^{-1})'(u)\mathrm{d}u + C = \int g(u)\frac{\mathrm{d}x}{\mathrm{d}u}\mathrm{d}u + C.$$

Example 7.24:
Calculate $\int xe^x\mathrm{d}x$. Notice that we can write this as $\int x\mathrm{d}(e^x)$. Now, using integration by parts,

$$\begin{aligned} \int x\mathrm{d}(e^x) &= xe^x - \int x'e^x\mathrm{d}x \\ &= xe^x - \int e^x\mathrm{d}x \\ &= xe^x - e^x + C. \end{aligned}$$

Example 7.25:
Compute $\int \log(x)\mathrm{d}x$.

We can write this as $\int \log(x) \cdot 1\mathrm{d}x = \int \log(x) \cdot x'\mathrm{d}x$. Hence we can use integration by parts to get

$$\int \log(x)\mathrm{d}x = \log(x)x - \int 1 \cdot x'\mathrm{d}x = x\log(x) - x + C.$$

Tip 7.26:
When doing integration, perform a safety check to see by differentiation that you get back the original function.

Example 7.27: Substitution rule
Calculate $\int \frac{x}{1+x^2}\mathrm{d}x$. We can use substitution here. Rewrite the expression as $\frac{1}{2}\int \frac{2x}{1+x^2}\mathrm{d}x$. Now let $u = 1 + x^2$, so that $\mathrm{d}u = 2x\mathrm{d}x$. Hence

$$\begin{aligned} \int \frac{x}{1+x^2}\mathrm{d}x &= \frac{1}{2}\int \frac{1}{u}\mathrm{d}u + C \\ &= \frac{1}{2}\log|u| + C \\ &= \frac{1}{2}\log(1+x^2) + C. \end{aligned}$$

Example 7.28: Trigonometric Substitution
Let $r > 0$. Compute $\int \sqrt{r^2 - x^2}\mathrm{d}x$.

Solution. Notice that if $x = r\sin(\theta)$, then

$$r^2 - x^2 = r^2(1 - \sin^2(\theta)) = r^2\cos^2(\theta).$$

Hence we have $\mathrm{d}x = r\cos(\theta)\mathrm{d}\theta$ and thus

$$\begin{aligned} \int \sqrt{r^2 - x^2}\mathrm{d}x &= \int r\cos(\theta)\sqrt{r^2 - r^2\sin^2(\theta)}\mathrm{d}\theta \\ &= \int r\cos(\theta) \cdot r\cos(\theta)\mathrm{d}\theta \\ &= r^2\int \cos^2(\theta)\mathrm{d}\theta. \end{aligned}$$

We can now apply integration by parts to see

$$\begin{aligned}\int \cos^2 &= \cos\sin + \int \sin^2 \\ &= \cos\sin + \int (1-\cos^2) \\ &= \cos\sin + \theta - \int \cos^2 .\end{aligned}$$

We hence find $\int \cos^2 = \frac{\cos\sin + \theta}{2} + C$ and thus

$$\int \sqrt{r^2 - x^2}\mathrm{d}x = \frac{r^2}{2}(\cos\theta\sin\theta + \theta) + C.$$

Rewriting in terms of x, we have $\sin(\theta) = \frac{x}{r}$ and thus $\theta = \arcsin(\frac{x}{r})$. Also, $\cos(\theta) = \sqrt{1 - \frac{x^2}{r^2}}$.

Thus we find our result

$$\int \sqrt{r^2 - x^2}\mathrm{d}x = \frac{x\sqrt{r^2 - x^2}}{2} + \frac{r^2}{2}\arcsin\left(\frac{x}{r}\right) + C.$$

Tip 7.29:

When having an integral of the form $\int (a^2 - x^2)^{\frac{n}{2}}\mathrm{d}x$, try the substitution $x = a\sin(\theta)$.

When instead we have an integral of the form $\int (a^2 + x^2)^{\frac{n}{2}}\mathrm{d}x$, try the substitution $x = a\tan(\theta)$.

The second diffrentiation is nice as then

$$\mathrm{d}x = \frac{a}{\cos^2(\theta)}\mathrm{d}\theta \text{ and } \sqrt{a^2 + x^2} = \frac{a}{\cos(\theta)}.$$

If now however we have an extra x like in $\int (a^2 + x^2)^{\frac{n}{2}} x\mathrm{d}x$, then we can use a simpler substitution $u = a^2 + x^2$.

Example 7.30:

Compute $\int \frac{1}{(a^2+x^2)^{3/2}}\mathrm{d}x$ for $a > 0$.

Solution. Let $x = a\tan(\theta)$. Then $\mathrm{d}x = \frac{a}{\cos^2(\theta)}\mathrm{d}\theta$ and

$$\int \frac{1}{(a^2 + x^2)^{3/2}}\mathrm{d}x = \int \left(\frac{\cos(\theta)}{a}\right)^3 \cdot \frac{a}{\cos^2(\theta)}\mathrm{d}\theta.$$

This simplifies to

$$\frac{1}{a^2}\int \cos(\theta)\mathrm{d}\theta = \frac{1}{a^2}\sin(\theta) + C.$$

Writing the sine in terms of the tangent, we have

$$\int \frac{1}{(a^2 + x^2)^{3/2}}\mathrm{d}x = \frac{1}{a^2} \cdot \frac{x}{\sqrt{a^2 + x^2}} + C.$$

Example 7.31:

Compute $\int \sqrt{1 - x^2}x\mathrm{d}x$.

Solution. Let $u = 1 - x^2$. Then $\mathrm{d}u = -2x\mathrm{d}x$ and thus

$$\begin{aligned}\int \sqrt{1 - x^2}x\mathrm{d}x &= -\frac{1}{2}\int \sqrt{u}\mathrm{d}u \\ &= -\frac{1}{2}\cdot\frac{2}{3}u^{3/2} + C \\ &= -\frac{1}{3}(1 - x^2)^{3/2} + C.\end{aligned}$$

Example 7.32: Hyperbolic Substitution

Calculate $\int \sqrt{x^2 - 1}\mathrm{d}x$.

Solution. Here we can make use of the identity $\cosh^2(t) - \sinh^2(t) = 1$. Hence, we take $x = \cosh(u)$, so that $\mathrm{d}x = \sinh(u)\mathrm{d}u$ and so

$$\begin{aligned}\int \sqrt{x^2 - 1}\mathrm{d}x &= \int \sqrt{\cosh^2(u) - 1}\sinh(u)\mathrm{d}u \\ &= \int \sinh^2(u)\mathrm{d}u.\end{aligned}$$

Now, we can use integration by parts to see

$$\begin{aligned}\int \sinh^2(u)\mathrm{d}u &= \sinh(u)\cosh(u) - \int \cosh^2(u)\mathrm{d}u + C \\ &= \sinh(u)\cosh(u) - \int (1 + \sinh^2(u))\mathrm{d}u + C \\ &= \sinh(u)\cosh(u) - u - \int \sinh^2(u)\mathrm{d}u + C \\ 2\int \sinh^2(u)\mathrm{d}u &= \sinh(u)\cosh(u) - u + C.\end{aligned}$$

So we can now plug things back in to get

$$\int \sqrt{x^2 - 1}\mathrm{d}x = \frac{1}{2}(\sinh(u)\cosh(u) - u) + C.$$

Writing in terms of x we have

$$\int \sqrt{x^2 - 1}\mathrm{d}x = \frac{\sqrt{x^2 - 1}\cdot x - \operatorname{arccosh}(x)}{2} + C.$$

Another technique that is useful for sinh is to use the definition $\sinh(x) = \frac{e^x - e^{-x}}{2}$. This can then be developed using bionmials and then integrated term by term.

Another useful technique is used when we have fractions with sine and cosine. The trick is to use the HALF-ANGLE SUBSTITUTION $t = \tan(\frac{x}{2})$. Consider the following example.

Example 7.33:

Compute $\int \frac{1}{\sin(x)}\mathrm{d}x$.

Solution. Let $u = \tan(\frac{x}{2})$. Then $\sin(x) = \frac{2u}{1+u^2}$ and $\cos(x) = \frac{1-u^2}{1+u^2}$. Also, $\mathrm{d}x = \frac{2}{1+u^2}\mathrm{d}u$.

Hence,

$$\begin{aligned}\int \frac{1}{\sin(x)}\mathrm{d}x &= \int \frac{1+u^2}{2u}\cdot\frac{2}{1+u^2}\mathrm{d}u \\ &= \int \frac{1}{u}\mathrm{d}u \\ &= \log|u| + C \\ &= \log\left|\tan\left(\frac{x}{2}\right)\right| + C.\end{aligned}$$

We now want to integrate RATIONAL FUNCTIONS, i.e. functions of the form

$$\frac{P(x)}{q(x)},$$

where P and q are polynomials. The first step is to write this in a form of $g(x) + \frac{r(x)}{q(x)}$, where g is a polynomial and the degree of r is less than the degree of q.

The central idea is now that any polynomial p(x) can be written as the product of monomials and quadractic polynomials that cannot be further factored over the reals.

Now we can use PARTIAL FRACTION DECOMPOSITION to write

$$\frac{r(x)}{q(x)} = \sum_i \frac{A_i}{(x - a_i)^{m_i}} + \sum_j \frac{B_j + C_j x}{(x^2 + b_j x + c_j)^{n_j}}.$$

Example 7.34:

$\int \frac{1}{x-a}\mathrm{d}x = \log|x-a| + C.$

$\int \frac{1}{(x-a)^n}\mathrm{d}x = -\frac{1}{(n-1)(x-a)^{n-1}} + C$ for $n \geq 2$.

$\int \frac{\mathrm{d}x}{a^2+x^2} = \frac{1}{a}\arctan\left(\frac{x}{a}\right) + C.$

$\int \frac{x\mathrm{d}x}{a^2+x^2} = \frac{1}{2}\log(a^2+x^2) + C.$

$\int \frac{x\mathrm{d}x}{(a^2+x^2)^n} = \frac{(a^2+x^2)^{1-n}}{2(1-n)} + C$ for $n \neq 1$.

Example 7.35:

Integrate $\int \frac{x^4+1}{x^2(x+1)}\mathrm{d}x$.

Solution. We write

$$\frac{x^4+1}{x^3+x^2} = \frac{x(x^3+x^2)-x^3+1}{x^3+x^2} = x - \frac{x^3-1}{x^3+x^2}.$$

Again we can write

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

So we have to integrate

$$\int \frac{x^4+1}{x^2(x+1)}\mathrm{d}x = \int (x-1)\mathrm{d}x + \int \frac{x^2+1}{x^2(x+1)}\mathrm{d}x.$$

We now use partial fraction decomposition to write

$$\frac{x^2+1}{x^2(x+1)} = \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x+1}.$$

Multiplying both sides by $x^2(x+1)$ and solving the system of equations, we find $A=-1, B=1, C=2$. Hence we have to integrate

$$\begin{aligned}\int \frac{x^4+1}{x^2(x+1)}\mathrm{d}x &= \int (x-1)\mathrm{d}x + \int \left(-\frac{1}{x} + \frac{1}{x^2} + \frac{2}{x+1}\right)\mathrm{d}x \\ &= \frac{x^2}{2} - x - \log|x| - \frac{1}{x} + 2\log|x+1| + C.\end{aligned}$$

Lec 31

Example 7.36:

Integrate $I = \int \frac{1}{x(x^2+2x+2)}\mathrm{d}x$.

Solution. Since we cannot factor x^2+2x+2 further, we write

$$\frac{1}{x(x^2+2x+2)} = \frac{A}{x} + \frac{Bx+C}{x^2+2x+2}.$$

Solving this, we get $A=\frac{1}{2}, B=-\frac{1}{2}, C=-1$. Hence the integral can be written as

$$I = \frac{1}{2}\int \frac{1}{x}\mathrm{d}x - \frac{1}{2}\int \frac{x+2}{x^2+2x+2}\mathrm{d}x.$$

We can now write

$$\int \frac{x+2}{x^2+2x+2}\mathrm{d}x = \int \frac{x+2}{(x+1)^2+1}\mathrm{d}x = \int \frac{(x+1)+1}{(x+1)^2+1}\mathrm{d}x.$$

Letting $u = x+1$, we have

$$\begin{aligned}\int \frac{u+1}{u^2+1}\mathrm{d}u &= \int \frac{u}{u^2+1}\mathrm{d}u + \int \frac{1}{u^2+1}\mathrm{d}u \\ &= \frac{1}{2}\log(u^2+1) + \arctan(u) + C \\ I &= \frac{1}{2}\log|x| - \frac{1}{4}\log((x+1)^2+1) - \frac{1}{2}\arctan(x+1) + C.\end{aligned}$$

Example 7.37:

Compute $\int_0^1 \log x\mathrm{d}x = \lim_{a\to 0^+}\int_a^1 \log x\mathrm{d}x$. As seen a primitive is $x\log x - x + C$. Hence

$$\begin{aligned}\int_0^1 \log x\mathrm{d}x &= \lim_{a\to 0^+}[x\log x - x]_a^1 \\ &= \lim_{a\to 0^+}(1\cdot\log 1 - 1) - (a\log a - a) \\ &= -1 - \lim_{a\to 0^+} a\log a + \lim_{a\to 0^+} a \\ &= -1 - 0 + 0 = -1.\end{aligned}$$

Given a constant $s > 0$, we define the GAMMA FUNCTION as

$$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\mathrm{d}x.$$

We first need to check that this integral converges. Consider $\int_a^b x^{s-1}e^{-x}\mathrm{d}x$ for $0 < a < b < \infty$. Notice we can write the integral as

$$\begin{aligned}\int_a^b x^{s-1}e^{-x}\mathrm{d}x &= \int_a^b \left(\frac{1}{s}x^2\right)' e^{-x}\mathrm{d}x \\ &= \left[\frac{1}{s}x^s e^{-x}\right]_a^b + \int_a^b \frac{x^s}{s}e^{-x}\mathrm{d}x \\ &= \frac{1}{s}\left[b^s e^{-b} - a^s e^{-a}\right] + \frac{1}{s}\int_a^b x^s e^{-x}\mathrm{d}x.\end{aligned}$$

As $a \to 0$, this becomes

$$\frac{1}{s}b^s e^{-b} + \frac{1}{s}\int_0^b x^s e^{-x}\mathrm{d}x.$$

If we now let $b \to \infty$, we have that $b^s e^{-b} \to 0$ and thus

$$\Gamma(s) = \frac{1}{s}\int_0^\infty x^s e^{-x}\mathrm{d}x.$$

This further shows that $\Gamma(s+1) = s\Gamma(s)$. Observing $\Gamma(1)$, we see that

$$\Gamma(1) = \int_0^\infty e^{-x}\mathrm{d}x = [-e^{-x}]_0^\infty = 1.$$

Hence, if $s = n \in \mathbb{N}$, we have $\Gamma(n) = (n-1)!$.

7.4 Taylor Series

When we talked about derivatives, we saw that $f(x) = f(x_0) + f'(x_0)(x-x_0) + o(x-x_0)$ if f is differentiable at x_0.

Definition 7.38: **Taylor Approximation**

Given $f : I \to \mathbb{R}$ n times differentiable, $x_0 \in I$ we define the n-th TAYLOR APPROXIMATION of f at x_0 as

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Exercise 7.39:

Show that $P_n^{(k)}(x_0) = f^{(k)}(x_0)$ for all $k = 0, \ldots, n$.

Theorem 7.40: Taylor Expansion with Integral Remainder

Let $n \geq 1, f : [a, b] \to \mathbb{R}$ be n times continously differentiable, $x_0 \in [a, b]$. Then

$$f(x) = P_{n-1}(x) + \int_{x_0}^{x} f^{(n)}(t) \cdot \frac{(x-t)^{n-1}}{(n-1)!} \mathrm{d}t.$$

Proof. We show this by induction. For $n = 1$, we have $f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\mathrm{d}t$, which is just the fundamental theorem of calculus. Since $P_0(x) = f(x_0)$ and $f^{(1)}(t) = (x-t)^0/0! = 1$, the base case holds.

Assume now that the statement holds for some $n \geq 1$. Then we have

$$f(x) = P_{n-1}(x) + \int_{x_0}^{x} f^{(n)}(t) \cdot \frac{(x-t)^{n-1}}{(n-1)!} \mathrm{d}t.$$

We can now use integration by parts on the integral. Letting

$$u = -\frac{(x-t)^n}{n!} \Rightarrow \mathrm{d}u = \frac{(x-t)^{n-1}}{(n-1)!}\mathrm{d}t,$$

we find that

$$\begin{aligned} f(x) = P_{n-1}(x) &+ \left[f^{(n)}(t) \cdot -\frac{(x-t)^n}{n!} \right]_{x_0}^{x} \\ &+ \int_{x_0}^{x} f^{(n+1)}(t) \cdot \frac{(x-t)^n}{n!} \mathrm{d}t \\ &= P_{n-1}(x) + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \\ &+ \int_{x_0}^{x} f^{(n+1)}(t) \cdot \frac{(x-t)^n}{n!} \mathrm{d}t \\ &= P_n(x) + \int_{x_0}^{x} f^{(n+1)}(t) \cdot \frac{(x-t)^n}{n!} \mathrm{d}t. \end{aligned}$$

□

Recall the mean value theorem, which states that if f is continuous on $[a, b]$ and differentiable on (a, b), then there exists $\xi \in (a, b)$ such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Rearranging this gives

$$f(b) = f(a) + f'(\xi)(b - a).$$

Changing around some variable names, we can write this as

$$f(x) = f(x_0) + f'(\xi)(x - x_0).$$

Theorem 7.41: Taylor with Lagrange remainder

Let $n \geq 1, f : [a, b] \to \mathbb{R}$ n-times differentiable, $x_0 \in [a, b]$. Then $\forall x \in [a, b]\, \exists \xi_L \in (x_0, x)$ such that

$$f(x) = P_{n-1}(x) + \frac{f^{(n)}(\xi_L)}{n!}(x - x_0)^n.$$

Notice that the derivatives of P_{n-1} are taken at x_0, while the derivative in the remainder is taken at some ξ_L between x_0 and x.

Proof. Fix $x \in [a, b]$, w.l.o.g assume $x > x_0$. Define $g : (a, b) \to \mathbb{R}$ as

$$g(t) = f(t) + f'(t)(x-t) + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1}.$$

Then $g(x) = f(x)$ and $g(x_0) = P_{n-1}(x)$. Also for the derivative we have

$$g'(t) = \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1}.$$

By cauchy mean value theorem applied to $g(t)$ and $h(t) = -(x-t)^n$ $\exists \xi_L \in (x_0, x)$ such that

$$\frac{g(x) - g(x_0)}{h(x) - h(x_0)} = \frac{g'(\xi_L)}{h'(\xi_L)}.$$

This is equivalent to

$$\begin{aligned} \frac{f(x) - P_{n-1}(x)}{(x-x_0)^n} &= \frac{f^{(n)}(\xi_L)(x-\xi_L)^{n-1}/(n-1)!}{n(x-\xi_L)^{n-1}} \\ &= \frac{f^{(n)}(\xi_L)}{n!}. \end{aligned}$$

□ Lec 32

Theorem 7.42: Taylor with little-o

If $f \in C^n$, then

$$f(x) - P_n(x) = o(|x - x_0|^n).$$

Proof. For the integral remainder, we have

$$\begin{aligned} f(x) = P_{n-1}(x) &+ \int_{x_0}^{x} f^{(n)}(x_0) \frac{(x-t)^{n-1}}{(n-1)!}\mathrm{d}t \\ &+ \int_{x_0}^{x} (f^{(n)}(t) - f^{(n)}(x_0)) \frac{(x-t)^{n-1}}{(n-1)!}\mathrm{d}t \\ &= P_{n-1}(x) + \frac{f^{(n)}(x_0)}{(n-1)!} \int_{x_0}^{x} (x-t)^{n-1}\mathrm{d}t \\ &+ \int_{x_0}^{x} (f^{(n)}(t) - f^{(n)}(x_0)) \frac{(x-t)^{n-1}}{(n-1)!}\mathrm{d}t \\ &= P_{n-1}(x) + \frac{f^{(n)}(x_0)}{(n-1)!} \cdot \frac{(x-x_0)^n}{n} \\ &+ \int_{x_0}^{x} (f^{(n)}(t) - f^{(n)}(x_0)) \frac{(x-t)^{n-1}}{(n-1)!}\mathrm{d}t \\ &= P_n(x) + \int_{x_0}^{x} (f^{(n)}(t) - f^{(n)}(x_0)) \frac{(x-t)^{n-1}}{(n-1)!}\mathrm{d}t. \end{aligned}$$

The idea is now that if we fix $\varepsilon > 0$, then $\exists \delta > 0$ such that

$$|f^{(n)}(z) - f^{(n)}(x_0)| < \varepsilon \quad \text{for } |z - x_0| < \delta.$$

So if $|x - x_0| < \delta$, also $|t - x_0| < \delta$. Therefore

$$|f^{(n)}(t) - f^{(n)}(x_0)| < \varepsilon.$$

So we have

$$\begin{aligned} |f(x) - P_n(x)| &\leq \int_{x_0}^{x} |f^{(n)}(t) - f^{(n)}(x_0)| \cdot \frac{(x-t)^{n-1}}{(n-1)!}\mathrm{d}t \\ &\leq \frac{\varepsilon}{(n-1)!} \int_{x_0}^{x} (x-t)^{n-1}\mathrm{d}t \\ &= \frac{\varepsilon}{n!} \cdot |x - x_0|^n. \end{aligned}$$

This shows that

$$\frac{|f(x) - P_n(x)|}{|x - x_0|^n} \xrightarrow[x \to x_0]{} 0.$$

□

Theorem 7.43: Taylor with Big-O

If f is n times differentiable and $|f^{(n)}| \leq M$ on $[a, b]$, then

$$f(x) - P_{n-1}(x) = O((x - x_0)^n).$$

Proof. Assume $|f^(n)| \leq M$ on $[a, b]$. Then by the Lagrange form of the remainder,

$$|f(x) - P_{n-1}(x)| = \left| \frac{f^{(n)}(\xi_L)}{n!}(x - x_0)^n \right| \leq \frac{M}{n!}|x - x_0|^n.$$

□

If f is C^∞, then we can use either or the other

$$f(x) = P_n(x) + o(|x - x_0|^n) = P_n(x) + O(|x - x_0|^{n+1}).$$

Taylor expansions can be very useful to compute limits.

Example 7.44:

Compute $\lim_{x\to 0} \frac{\sin x - x}{x^3}$.

Solution. We write the Taylor expansion of $\sin x$ at 0:

$$\sin x = x - \frac{x^3}{3!} + O(x^4).$$

Hence

$$\lim_{x\to 0} \frac{\sin x - x}{x^3} = \lim_{x\to 0} \frac{x - \frac{x^3}{6} + O(x^4) - x}{x^3} = \lim_{x\to 0} \frac{-\frac{x^3}{6} + O(x^4)}{x^3} = \lim_{x\to 0} -\frac{1}{6} + O(x) = -\frac{1}{6}.$$

Suppose we have $f'(x_0) = f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0$ and $f^{(n)}(x_0) \neq 0$. Then the Taylor polynomial is

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

Suppose now $f^{(n)}(x_0) > 0$. Then if n is even, $P_n(x) \geq f(x_0)$ hence x_0 is a local minimum. If n is odd, then for x_0 is neither a local minimum nor a local maximum. Similarly, if $f^{(n)}(x_0) < 0$, then if n is even, x_0 is a local maximum, and if n is odd, x_0 is neither a local minimum nor a local maximum.

Let f now be smooth. We know that $f(x) = P_n(x) + O((x - x_0)^{n+1})$. The temptation is to let n go to infinity and conclude that $O((x-x_0)^{n+1})$ goes to 0. However, this is NOT true in general. In fact, the constant in the big-O notation may depend on n.

Definition 7.45: Analytic Functions

Given $I \subseteq \mathbb{R}$ interval, $f : I \to \mathbb{R}$ smooth. f is called ANALYTIC at $x_0 \in I$ if

1) The power series $\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$ has a positive radius of convergence.

2) $\exists \delta \in (0, R)$ such that

$$f(x) = \sum_{k=0}^{\infty} f^{(n)(x_0)} \frac{(x - x_0)^k}{k!} \quad \text{for } |x - x_0| < \delta.$$

f is ANALYTIC if it is analytic at every $x_0 \in I$.

Recall the function

$$\psi(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}.$$

This function is smooth, but not analytic at 0 since all its derivatives at 0 are 0, but $\psi(x) \neq 0$ for $x > 0$.

Theorem 7.46:

Let $f : I \to \mathbb{R} \in C^{\infty}, \exists r > 0, C_0 > 0, A > 0$ such that

$$|f^{(n)}(x)| \leq C_0 A^n n! \quad \forall x \in (x_0 - r, x_0 + r), n \in \mathbb{N}.$$

Then f is analytic at x_0.

Proof. The power series is

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad a_n = \frac{f^{(n)}(x_0)}{n!}.$$

Then

$$\sqrt[n]{|a_n|} = \sqrt[n]{\frac{|f^{(n)}(x_0)|}{n!}} \leq \sqrt[n]{C_0 A^n} = A\sqrt[n]{C_0}.$$

Hence $\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq A$. Therefore the radius of convergence R satisfies $R \geq 1/A > 0$.

Let $d < \min\{r, \frac{1}{A}\}$. Given $x \in (x_0 - d, x_0 + d) \cap I$,

$$|f(x) - P_{n-1}(x)| = \left|\frac{f^{(n)}(\xi_L)}{n!}(x - x_0)^n\right| \leq C_0 A^n |x - x_0|^n.$$

But then

$$|f(x) - P_{n-1}(x)| \leq C_0 (A\delta)^n.$$

Since $\delta < \frac{1}{A}$, we have $A\delta < 1$ and thus

$$\lim_{n\to\infty} |f(x) - P_{n-1}(x)| = 0.$$

□

8 Ordinary Differential Equations

8.1 Ordinary Differential Equations

Lec 33

In general an ORDINARY DIFFERENTIAL EQUATION (ODE) is a relation between the independent variable x and a function $u : I \to \mathbb{R}$ with its derivatives. Then for $x \in I$, $u(x)$ satisfies the relation

$$G : \mathbb{R}^{n+2} \longrightarrow \mathbb{R}$$
$$x, u(x), \ldots \longmapsto G(x, u(x), u'(x), \cdots, u^{(n)}(x)) = 0.$$

We can now define some vocabulary for ODEs.

Definition 8.1: Order

The ORDER of an ODE is the highest derivative appearing in the relation.

For example $u'' + u = 0$ is a second order ODE and $(u')^2 + u - x^3 = 0$ is a first order ODE.

Definition 8.2: Linearity

An ODE is called LINEAR if it is linear in u and its derivatives, i.e. if it can be written as

$$a_n(x)u^{(n)}(x) + \cdots + a_0(x)u(x) = f(x).$$

For example $u'' + u = 0$ is linear, while $u'' + u'u = 0$ is not linear. If we have $u'' = x^2u + x$, this is still linear as we have to treat x like a fixed parameter.

Definition 8.3: Homogeneity

A linear ODE is called HOMOGENEOUS if every term involves either u or one of its derivatives.

For example $u'' + u = 0$ is homogeneous, while $u'' + u = x$ is not homogeneous.

Example 8.4: The law of cooling

Given t as variable, $T(t)$ is the temperature of a body. Newotn's law of cooling states that

$$\dot{T}(t) = -k(T(t) - T_{\text{ext}}).$$

This is a first order linear ODE. It is not homogeneous.

Example 8.5: Harmonic oscillator

Given a mass m attached to a spring with spring constant k, the position $x(t)$ of the mass satisfies

$$m\ddot{x}(t) + kx(t) = 0.$$

Usually we write this as

$$\ddot{x}(t) + \omega^2 x(t) = 0.$$

This is a second order linear homogeneous ODE.

If there is damping, we can derive the equation

$$\ddot{x}(t) + 2\zeta\dot{x}(t) + \omega_0^2 x(t) = 0.$$

Example 8.6: Population growth

Let $P(t)$ be the population at time t. By logisitic growth, we have the differential equation

$$\dot{P}(t) = rP(t)\left(1 - \frac{P(t)}{K}\right).$$

This is a first order non-linear ODE.

Fix an interval $I \subseteq \mathbb{R}$. Consider $f, g : I \to \mathbb{R}$. Our goal is to solve

$$u'(x) + f(x)u(x) = g(x).$$

i.e. a linear first order ODE. We begin with the homogeneous case ($g \equiv 0$).

Theorem 8.7:

Let $f : I \to \mathbb{R}$ continuous. Let $F : I \to \mathbb{R}$ be a primitive of f. Then the set of all solutions $u \in C^1(I)$ of $u' + fu = 0$ is given by

$$u(x) = Ae^{-F(x)} \quad A \in \mathbb{R}.$$

Note that if we replace F by $F + c$, then $e^{-F(x)}$ is replaced by $e^{-c}e^{-F(x)}$. Hence the solutions are the same.

Proof. Let $u(x) = Ae^{-F(x)}$. Then $u'(x) = -f(x)Ae^{-F(x)} = -f(x)u(x)$. So $Ae^{-F(x)}$ solves the ODE.

Vice versa, let u be a solution of $u' + fu = 0$. Define $v(x) = u(x)e^{F(x)}$. Then

$$v'(x) = u'(x)e^{F(x)} + u(x)f(x)e^{F(x)} = e^{F(x)}(u'(x) + f(x)u(x)) = 0.$$

Hence v is constant, i.e. $v(x) = A$ for some $A \in \mathbb{R}$. Therefore

$$u(x) = Ae^{-F(x)}.$$

□

Example 8.8:

Find all solutions of

$$u' + 3u = 0.$$

Solution. A primitive of $f(x) = 3$ is $F(x) = 3x$. Hence all solutions are given by

$$u(x) = Ae^{-3x} \quad A \in \mathbb{R}.$$

Tip 8.9:

To quickly check if the solution is correct, plug it back into the ODE.

Example 8.10:

Find all solutions of

$$u' - \sin(x)u = 0.$$

Solution. A primitive of $f(x) = -\sin(x)$ is $F(x) = \cos(x)$. Hence all solutions are given by

$$u(x) = Ae^{-\cos(x)} \quad A \in \mathbb{R}.$$

Let us now consider the non-homogeneous case. The idea is to replace A in $Ae^{-F(x)}$ by $H(x)e^{-F(x)}$ for some function H to be determined.

Let $u(x) = H(x)e^{-F(x)}$. Then

$$u'(x) = H'(x)e^{-F(x)} - f(x)\underbrace{H(x)e^{-F(x)}}_{u(x)}.$$

Hence we have

$$u'(x) + f(x)u(x) = H'(x)e^{-F(x)}.$$

Impose now $H'e^{-F(x)} = g(x)$. This is equivalent to

$$H'(x) = g(x)e^{F(x)}.$$

So H is a primitive of $g(x)e^{F(x)}$. We have just shown that if F is a primitive of f and H is a primitive of $g(x)e^{F(x)}$, then $H(x)e^{-F(x)}$ is a solution.

Theorem 8.11:

Let $f, g : I \to \mathbb{R}$ continuous. Let F be a primitive of f and H a primitive of $g(x)e^{F(x)}$. Then all solutions of $u' + fu = g$ are given by

$$u(x) = H(x)e^{-F(x)} + Ae^{-F(x)} \quad A \in \mathbb{R}.$$

Proof. Let $u(x) = (H(x) + A)e^{-F(x)}$. Then, since $(H + A)' = H'$, also $H + A$ is a primitive of $g(x)e^{F(x)}$. By the previous discussion, u is a solution of $u' + fu = g$.

Vice versa, let u be a solution. Define $v(x) = u(x) - H(x)e^{-F(x)}$. Then

$$\begin{aligned} v'(x) &= u'(x) - H'(x)e^{-F(x)} + f(x)H(x)e^{-F(x)} \\ &= -f(x)u(x) + g(x) - g(x)e^F e^{-F} + fHe^{-F} \\ &= -f(u - He^{-F}) = -fv. \end{aligned}$$

Lec 34 Hence, $v' + fv = 0$. By the previous theorem $v(x) = Ae^{-F(x)}$, which concludes the theorem. □

Example 8.12:

Find the solution of $u' - 2xu = e^{x^2}$, where $u(0) = 1$.

Solution. A primitive of $f(x) = -2x$ is $F(x) = -x^2$. A primitive of $g(x)e^{F(x)} = e^{x^2}e^{-x^2} = 1$ is $H(x) = x$. Hence all solutions are given by

$$u(x) = xe^{x^2} + Ae^{x^2} \quad A \in \mathbb{R}.$$

Imposing the initial condition $u(0) = 1$ gives $A = 1$. Therefore the solution is

$$u(x) = (x + 1)e^{x^2}.$$

Definition 8.13: Autonomous 1st order ODE

An ODE of the form

$$u'(x) = f(u(x)).$$

is called an AUTONOMOUS 1ST ORDER ODE.

If $u(x) = C$ and $f(C) = 0$, then $u'(x) = f(u(x)) = 0$, so $u(x) = C$ is a solution.

If instead $f(u(x)) \neq 0$, we can divide and get

$$\frac{u'(x)}{f(u(x))} = 1 \Rightarrow \int \frac{u'(x)}{f(u(x))}\mathrm{d}x = \int 1\mathrm{d}x = x + A.$$

But by substitution, we have

$$\int \frac{\mathrm{d}u}{f(u)} = x + A.$$

Let H be a primitive of $1/f$. Then $H(u(x)) = x + A$. Hence $u(x) = H^{-1}(x + A)$.

The inverse function H^{-1} exists since H is strictly monotone as the derivative $H' = 1/f$ does not change sign.

Example 8.14:

Find all solutions of $u'(x) = ru(x)\left(1 - \frac{u(x)}{K}\right)$, where $r, K > 0$.

Solution. $u(x) = 0$ and $u(x) = K$ are constant solutions. Assume now $u(x) \in (0, K)$. Hence we get

$$\begin{aligned} \frac{u'(x)}{u(x)(1 - u(x)/K)} &= r \\ \int \frac{Ku'(x)}{u(x)(K - u(x))}\mathrm{d}x &= \int r\mathrm{d}x = rx + A \\ \int \frac{K\mathrm{d}u}{u(K-u)} &= rx + A \\ \int \frac{1}{u}\mathrm{d}u + \int \frac{1}{K-u}\mathrm{d}u &= rx + A \\ \log|u| - \log|K - u| = \log\left|\frac{u}{K-u}\right| &= rx + A \end{aligned}$$

.

We now solve for $u(x)$:

$$\begin{aligned} \frac{u(x)}{K - u(x)} &= e^{rx+A} = Be^{rx} \quad B = e^A \\ u(x) &= Be^{rx}(K - u(x)) \\ u(x) + Be^{rx}u(x) &= BKe^{rx} \\ u(x)(1 + Be^{rx}) &= BKe^{rx} \\ u(x) &= \frac{BKe^{rx}}{1 + Be^{rx}} \end{aligned}$$

.

If we impose $u(0) = u_0 \in (0, K)$, then

$$\begin{aligned} u(0) &= \frac{BK}{1 + B} = u_0 \\ BK &= u_0 + Bu_0 \\ B(K - u_0) &= u_0 \\ B &= \frac{u_0}{K - u_0} \\ u(x) &= \frac{Ku_0}{u_0 + (K - u_0)e^{-rx}} \end{aligned}$$

If $u'(x) = f(u(x))g(x)$, we get $\int \frac{\mathrm{d}u}{f(u)} = \int g(x)\mathrm{d}x$. In this case,

$$u(x) = H^{-1}(G(x) + A).$$

Theorem 8.15:

Given $I, J \subseteq \mathbb{R}$ intervals, $f : J \to \mathbb{R}$ and $g : I \to \mathbb{R}$. Assume, f, g continuous and $f(y) \neq 0$ for all $y \in J$. Let H be a primitive of $1/f$ and G a primitive of g. Then every solution of $u'(x) = f(u(x))g(x)$, $x \in I$, $u : I \to J$, is

$$u(x) = H^{-1}(G(x) + A) \quad A \in \mathbb{R}.$$

Proof. Let $u(x) = H^{-1}(G(x) + A)$. Then

$$\begin{aligned} u'(x) &= \frac{1}{H' \circ H^{-1}(G(x) + A)} = \frac{1}{H'(u(x))}G'(x) \\ &= f(u(x))g(x). \end{aligned}$$

Hence u is a solution.

Vice versa, let u be a solution, consider $H(u(x))$. Then

$$H(u(x))' = H'(u(x))u'(x) = \frac{u'(x)}{f(u(x))} = g(x) = G'(x).$$

So $(H(u(x)) - G(x))' = 0$. Hence $H(u(x)) - G(x) = A$ for some $A \in \mathbb{R}$. □

Let us now look at second order ODEs. We begin with linear, homogeneous ODEs with constant coefficients, i.e.

$$u''(x) + a_1 u'(x) + a_0 u(x) = 0.$$

Where $a_1, a_0 \in \mathbb{R}$.

To solve this, look for $e^{\alpha} x$, $\alpha \in \mathbb{C}$. Then, $(e^{\alpha x})' = \alpha e^{\alpha x}$ and $(e^{\alpha x})'' = \alpha^2 e^{\alpha x}$. So $e^{\alpha x}$ solves the ODE iff

$$\underbrace{(\alpha^2 + a_1\alpha + a_0)}_{=0} e^{\alpha x} = 0.$$

Let $p(t) = t^2 + a_1 t + a_0$ be the CHARACTERISTIC POLYNOMIAL. Then $e^{\alpha x}$ is a solution iff $p(\alpha) = 0$.

Consider the discriminant $\Delta = a_1^2 - 4a_0$. There exist three cases:

- If $\Delta > 0$, then we have two distinct real roots

$$\alpha = \frac{-a_1 + \sqrt{\Delta}}{2}, \beta = \frac{-a_1 - \sqrt{\Delta}}{2}.$$

  Where $\alpha \neq \beta$. Then $u(x) = Ae^{\alpha x} + Be^{\beta x}$, is a solution.

- If $\Delta < 0$, then we have two complex conjugate roots

$$\alpha \pm i\beta = \frac{-a_1}{2} \pm i\frac{\sqrt{|\Delta|}}{2}.$$

  Where $\alpha = -a_1/2$ and $\beta = \sqrt{|\Delta|}/2$. In this case, $e^{(\alpha+i\beta)x}$ and $e^{(\alpha-i\beta)x}$ are solutions. The first one tells us that

$$e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)),$$

  is a solution. Since real and imaginary parts of solutions do not interfere with each other, $e^{\alpha x}\cos(\beta x)$ is a solution. Since we only want real solutions, we get that the solutions are given by

$$e^{\alpha x}(A\cos(\beta x) + B\sin(\beta x)) \quad A, B \in \mathbb{R}.$$

- If $\Delta = 0$, then we have one real root

$$\alpha = -\frac{a_1}{2}.$$

  In this case, $e^{\alpha x}$ is a solution. We can guess that this isn't enough by looking at the example $u'' = 0$. Here $a_1 = a_0 = 0$ and $\alpha = 0$. The solutions are $u(x) = Ax + B$, which is not of the form $Ae^{0x} = A$.

  So we check what happens to $xe^{\alpha x}$:

$$(xe^{\alpha x})' = e^{\alpha x} + \alpha x e^{\alpha x}$$
$$(xe^{\alpha x})'' = 2\alpha e^{\alpha x} + \alpha^2 x e^{\alpha x}.$$

  Plugging this into the ODE gives

$$(2\alpha + \alpha^2 x)e^{\alpha x} + a_1(1 + \alpha x)e^{\alpha x} + a_0 x e^{\alpha x} = 0$$
$$\left(2\alpha + a_1 + x(\alpha^2 + a_1\alpha + a_0)\right) e^{\alpha x} = 0.$$

  Notice how the term in x vanishes since α is a root and $2\alpha + a_1 = 0$ by definition of α. Hence $xe^{\alpha x}$ is indeed a solution. Therefore

$$u(x) = Ae^{\alpha x} + Bxe^{\alpha x} \quad A, B \in \mathbb{R}$$

  is a solution.

Example 8.16: Damped harmonic oscillator

Find all solutions of $mu'' = -du' - ku$, where $m, d, k > 0$.

Solution. We can write this as a second order ODE as

$$u'' + \frac{d}{m}u' + \frac{k}{m}u = 0.$$

If we define $\omega = \sqrt{k/m}$ and $\zeta = d/(2\sqrt{km})$, we get

$$u'' + 2\zeta\omega u' + \omega^2 u = 0.$$

The characteristic polynomial is

$$p(t) = t^2 + 2\zeta\omega t + \omega^2.$$

The discriminant is

$$\Delta = 4(\zeta^2 - 1)\omega^2.$$

If $\Delta < 0$, i.e. $\zeta < 1$, we have weak damping and the solutions are given by

$$u(x) = e^{-\zeta\omega x}(A\cos(\gamma x) + B\sin(\gamma x)) \quad A, B \in \mathbb{R}.$$

Where $\gamma = \omega\sqrt{1 - \zeta^2}$.

If $\Delta > 0$, i.e. $\zeta > 1$, we have strong damping and the solutions are given by

$$u(x) = Ae^{-\lambda_1 x} + Be^{-\lambda_2 x} \quad A, B \in \mathbb{R}.$$

Where $\lambda_{1,2} = \omega(\zeta \pm \sqrt{\zeta^2 - 1})$.

If $\Delta = 0$, i.e. $\zeta = 1$, we have critical damping and the solutions are given by

$$u(x) = e^{-\omega x}(A + Bx) \quad A, B \in \mathbb{R}.$$

Lec 35

Theorem 8.17:

If $u : I \to \mathbb{R}, u \in C^2(I)$ is a solution of $u'' + a_1 u' + a_0 u = 0$, then $\exists A, B \in \mathbb{R}$ such that $u = Au_1 + Bu_2$, where $\Delta = a_1^2 - 4a_0$ and

- If $\Delta > 0$, then

$$u(x) = Ae^{\alpha x} + Be^{\beta x} \quad A, B \in \mathbb{R}.$$

  Where $\alpha = \frac{-a_1 + \sqrt{\Delta}}{2}$ and $\beta = \frac{-a_1 - \sqrt{\Delta}}{2}$.

- If $\Delta < 0$, then

$$u(x) = e^{\alpha x}(A\cos(\beta x) + B\sin(\beta x)) \quad A, B \in \mathbb{R}.$$

  Where $\alpha = -\frac{a_1}{2}$ and $\beta = \frac{\sqrt{|\Delta|}}{2}$.

- If $\Delta = 0$, then

$$u(x) = e^{\alpha x}(A + Bx) \quad A, B \in \mathbb{R}.$$

  Where $\alpha = -\frac{a_1}{2}$.

Proof. We have already shown that all of the above indeed are solutions. It remains to show that there are no other solutions.

Assume $\Delta > 0$ and assume $0 \in I$. The proof for the other cases of Δ is similar. Observe, that $u_1(0) = u_2(0) = 1$, $u_1'(0) = \alpha e^{\alpha 0} = \alpha > \beta = u_2'(0)$. Define $v_1(x) = \frac{\alpha u_2(x) - \beta u_1(x)}{\alpha - \beta}$ and $v_2(x) = \frac{u_1(x) - u_2(x)}{\alpha - \beta}$.

Notice that $v_1(0) = 1$, $v_1'(0) = 0$ and $v_2(0) = 0$, $v_2'(0) = 1$.

Define now $w(x) = u(x) - u(0)v_1(x)$. Note that w is still a solution since it is a linear combination of solutions. Also, $w(0) = u(0) - u(0)v_1(0) = 0$.

The goal now is to show that w is a multiple of v_2. This would then show that u is a linear combination of v_1 and v_2, which are linear combinations of u_1 and u_2.

Define $W(x) = w(x)v_2'(x) - w'(x)v_2(x)$, which is called the *Wronskian* of w and v_2.

Then

$$W' = w'v_2' + wv_2'' - w''v_2 - w'v_2' = wv_2'' - w''v_2 = wv_2'' - w''v_2.$$

Plugging in the ODE gives

$$\begin{aligned} W' &= w(-a_1v_2' - a_0v_2) + (a_1w' + a_0w)v_2 \\ &= -a_1(wv_2' - w'v_2) = -a_1W. \end{aligned}$$

This is a first order ODE for W. Its solution is

$$W(x) = Ae^{-a_1x}.$$

But $W(0) = w(0)v_2'(0) - w'(0)v_2(0) = 0$. Hence $A = 0$ and $W(x) = 0$ for all $x \in I$. Hence, $w'v_2 = v_2'w$.

Consider an interval $J \subseteq I$ where $v_2(x) \neq 0$. Then $w' = \frac{v_2'}{v_2}w$. This is a first order ODE for w. But we can rewrite it as

$$w' = (\log|v_2|)'\, w.$$

Hence, $w = A_J \exp(\log(v_2)) = A_J|v_2| = \pm A_J v_2$ on J.

In our case, v_2 is not zero on $(0,\infty)$ and $(-\infty,0)$. Hence, there exist $A_+, A_- \in \mathbb{R}$ such that

$$w(x) = \begin{cases} A_+v_2(x) & x > 0 \\ A_-v_2(x) & x < 0 \end{cases}.$$

Since w and v_2 are continuous, we have $w = A_+v_2$ on $[0,\infty)$ and $w = A_-v_2$ on $(-\infty,0]$.

It remains to show that $A_+ = A_-$. But this follows from from differentiability of w at 0:

$$w_+'(0) = \lim_{h\to0^+} \frac{w(h) - w(0)}{h} = \lim_{h\to0^+} \frac{A_+v_2(h)}{h} = A_+v_2'(0).$$

Similarly,

$$w_-'(0) = \lim_{h\to0^-} \frac{w(h) - w(0)}{h} = \lim_{h\to0^-} \frac{A_-v_2(h)}{h} = A_-v_2'(0).$$

Since $w_+'(0) = w_-'(0)$ and $v_2'(0) = 1$, we get $A_+ = A_-$. Hence

$$w(x) = Av_2(x) \quad A \in \mathbb{R}.$$

Therefore,

$$u(x) = u(0)v_1(x) + Av_2(x) = A_1u_1(x) + A_2u_2(x) \quad A_1, A_2 \in \mathbb{R}.$$

□

Corollary 8.18:

If v_1, v_2 are two solutions of $u'' + a_1u' + a_0u = 0$, and the Wronskian $W(x) = v_1(x)v_2'(x) - v_1'(x)v_2(x)$. If $W(x_0) = 0$ for some $x_0 \in I$, then v_1, v_2 are linearly dependent.

Proof. Consider $M(x) = \begin{pmatrix} v_1(x_0) & v_2(x_0) \\ v_1'(x_0) & v_2'(x_0) \end{pmatrix}$. Notice that $W(x_0) = \det(M(x_0))$. If $W(x_0) = 0$, then $\exists \begin{pmatrix} A \\ B \end{pmatrix} \in \mathbb{R}^2 \setminus \{0\}$ such that

$$0 = M(x_0)\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} Av_1(x_0) + Bv_2(x_0) \\ Av_1'(x_0) + Bv_2'(x_0) \end{pmatrix}.$$

Define $w(x) = Av_1(x) + Bv_2(x)$. Then w is a solution of the ODE with $w(x_0) = 0$ and $w'(x_0) = 0$. Since the constant function 0 is a solution, then $w = 0$ hence v_1, v_2 are linearly dependent. □

We now want to solve non-homogeneous second order ODEs of the form

$$u''(x) + a_1u'(x) + a_0u(x) = g(x).$$

We want to look for $u = H_1u_1 + H_2u_2$, where H_1, H_2 are functions.

Differentiating gives

$$u' = H_1'u_1 + H_1u_1' + H_2'u_2 + H_2u_2'.$$

Differentiating again gives

$$u'' = H_1''u_1 + 2H_1'u_1' + H_1u_1'' + H_2''u_2 + 2H_2'u_2' + H_2u_2''.$$

If we compute $u'' + a_1u' + a_0u$, and use that u_1, u_2 are solutions of the homogeneous ODE, we get

$$g = (H_1'u_1 + H_2'u_2)' + (H_1'u_1' + H_2'u_2') + a_1(H_1'u_1 + H_2'u_2).$$

The easiest way to simplify this is to impose the condition $g = H_1'u_1' + H_2'u_2'$ and $0 = H_1'u_1 + H_2'u_2$. We now want to check that we actually find a solution this way.

From the second equation, we get $H_2' = -\frac{u_1}{u_2}H_1'$. Substituting this into the first equation gives

$$H_1' = \frac{u_2g}{u_1'u_2 - u_1u_2'} = \frac{u_2g}{W}.$$

Similarly, we get

$$H_2' = -\frac{u_1g}{W}.$$

Since u_1 and u_2 are linearly independent, $W(x) \neq 0$ for all $x \in I$. Integrating both equations gives

$$H_1 = \int \frac{u_2(x)g(x)}{u_1'u_2 - u_2'u_1}\mathrm{d}x, \quad H_2 = -\int \frac{u_1(x)g(x)}{u_1'u_2 - u_2'u_1}\mathrm{d}x.$$

Theorem 8.19:

Given $u'' + a_1u' + a_0u = g$, where $a_1, a_0, g : I \to \mathbb{R}$ continuous. Let u_1, u_2 be two linearly independent solutions of the homogeneous ODE. Then all solutions of the ODE are given by

$$u = H_1u_1 + H_2u_2 + Au_1 + Bu_2 \quad A, B \in \mathbb{R}.$$

Where

$$H_1 = \int \frac{u_2(x)g(x)}{u_1'u_2 - u_2'u_1}\mathrm{d}x, H_2 = -\int \frac{u_1(x)g(x)}{u_1'u_2 - u_2'u_1}\mathrm{d}x.$$

Example 8.20:

Solve $u'' + u = 1$, $u(0) = 0, u'(0) = 1$.

Solution. Here we can guess a particular solution as $u = 1$. The homogeneous ODE has solutions $u_1(x) = \cos x$ and $u_2(x) = \sin x$. Hence,

$$u(x) = 1 + A\cos x + B\sin x.$$

Imposing the initial conditions gives $A = -1$ and $B = 1$. Therefore the solution is

$$u(x) = 1 - \cos x + \sin x.$$

The above example could've also been solved computing H_1 and H_2.

Example 8.21:

Find u solving $u'' + u = \sin(2x)$.

Solution. We can try to compute a particular solution using $v(x) = A\sin(2x)$. Then, computation gives

$$v''(x) + v(x) = -4A\sin(2x) + A\sin(2x) = -3A\sin(2x).$$

Hence, if we choose $A = -\frac{1}{3}$, we get a particular solution $v(x) = -\frac{1}{3}\sin(2x)$. The homogeneous ODE has solutions $u_1(x) = \cos x$ and $u_2(x) = \sin x$. Therefore, all solutions are given by

$$u(x) = -\frac{1}{3}\sin(2x) + A\cos x + B\sin x \quad A, B \in \mathbb{R}.$$

Example 8.22:

Solve $u'' + u = \sin(x)$.

Solution. Notice that $\sin(x)$ is already a solution of the homogeneous ODE, hence, it cannot be used as a particular solution. Instead of looking for $v(x) = A\sin(x)$, we try $v(x) = Ax\cos(x) + Bx\sin(x)$. Computing gives

$$v''(x) + v(x) = -2A\sin(x) + 2B\cos(x).$$

Hence, if we choose $A = -\frac{1}{2}$ and $B = 0$, we get a particular solution $v(x) = -\frac{1}{2}x\cos(x)$. The homogeneous ODE has solutions $u_1(x) = \cos x$ and $u_2(x) = \sin x$. Therefore, all solutions are given by

$$u(x) = -\frac{1}{2}x\cos(x) + A\cos x + B\sin x \quad A, B \in \mathbb{R}.$$

Lec 36

8.2 Existence and Uniqueness for ODEs

Let us return to first order ODEs, this time of the form

$$u'(x) = f(x, u(x)).$$

We can start with the following example

$$u'(x) = |u(x)|^\alpha, \quad \alpha > 0.$$

Let first $\alpha < 1$. Assume $u(x) > 0$ somewhere. Then we can separate variables and obtain

$$\begin{aligned}\frac{u'(x)}{u(x)^\alpha} &= 1\\ \int \frac{\mathrm{d}u}{u(x)^\alpha} &= \int 1\mathrm{d}x\\ \frac{u(x)^{1-\alpha}}{1-\alpha} &= x + A\\ u(x) &= \big((1-\alpha)(x+A)\big)^{\frac{1}{1-\alpha}}.\end{aligned}$$

If $A = 0$, then we have the solution

$$u(x) = \big((1-\alpha)x\big)^{\frac{1}{1-\alpha}}.$$

So in fact we have two solutions that satisfy the initial condition $u(0) = 0$, namely the trivial solution $u(x) = 0$ and the non-trivial solution above.

In fact, we can construct infinitely many solutions by patching together the trivial solution and the non-trivial solution. For any $x_0 > 0$, we can define

$$u(x) = \begin{cases} 0, & x \le x_0 \\ \frac{(x-x_0)^2}{4}, & x > x_0 \end{cases}..$$

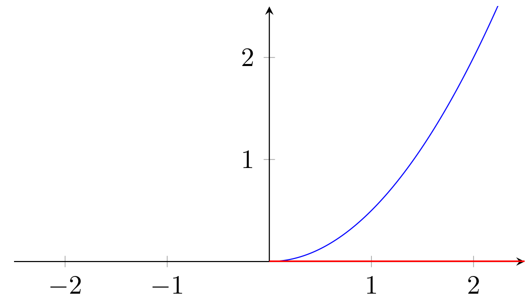


Figure 43: Solution to $u'(x) = |u(x)|^{1/2}$ with $A = 0$

Let us consider the case $\alpha > 1$, in particular $\alpha = 2$. Also let $u(0) = 1$. Then we can separate variables again and obtain

$$\begin{aligned}\frac{u'(x)}{u(x)^2} &= 1\\ \int \frac{\mathrm{d}u}{u(x)^2} &= \int 1\mathrm{d}x\\ -\frac{1}{u(x)} &= x + A\\ u(x) &= -\frac{1}{x+A}.\end{aligned}$$

From the initial condition $u(0) = 1$, we get $A = -1$. Thus $u(x) = -\frac{1}{x-1}$

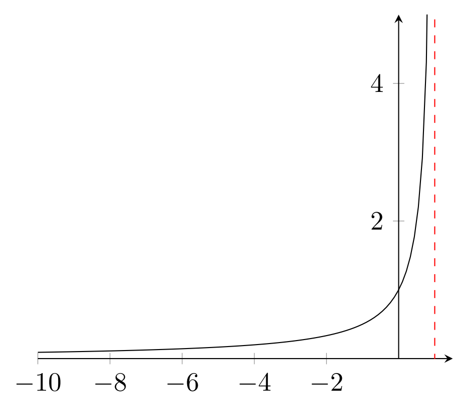


Figure 44: Solution to $u'(x) = |u(x)|^2$ with $u(0) = 1$

This is indeed the only solution that satisfies the initial condition.

Theorem 8.23: Cauchy-Lipschitz

Given $u' = f(x, u(x))$ with initial condition $u(x_0) = y_0$. If f is continuous and f is Lipschitz continuous in y, i.e. there exists $L > 0$ such that

$$|f(x, y_1) - f(x, y_2)| \le L|y_1 - y_2|.$$

Then there exists a unique solution to the ODE.

We can also require that f is only locally Lipschitz continuous in y. Then $\exists!$ solution defined $\forall x \in I \subset \mathbb{R}$, where I is an interval.

Let us now look at linear second order ODEs of the form

$$u''(x) + a_1(x)u'(x) + a_0(x)u(x) = 0.$$

In general, theese cannot be solved explicitly.

Define $U_1(x) = u(x)$ and $U_2(x) = u'(x)$. Then we have

$$\begin{cases} U_1'(x) = U_2(x) \\ U_2'(x) = -a_1(x)U_2(x) - a_0(x)U_1(x). \end{cases}.$$

We now ask for the initial conditions $u(x_0) = y_0$ and $u'(x_0) = v_0$. So the system can be written as

$$\begin{cases} U_1'(x) = U_2(x) \\ U_2'(x) = -a_1(x)U_2(x) - a_0(x)U_1(x) \end{cases}.$$

Cauchy-Lipschitz applies here as well, so there exists a unique solution if f is Lipschitz in U_1 and U_2.

Theorem 8.24:

Given $u'' + a_1(x)u' + a_0(x)u = 0$. Then the space of solutions is a 2-dimensional vector space.

Proof. Consider $u'' + a_1(x)u' + a_0(x)u = 0$ with initial conditions $u(x_0) = y_0$ and $u'(x_0) = v_0$. This is equivalent to the system

$$\begin{cases} U_1'(x) = U_2(x) \\ U_2'(x) = -a_1(x)U_2(x) - a_0(x)U_1(x) \\ U_1(x_0) = y_0, \quad U_2(x_0) = v_0 \end{cases}.$$

So by Cauchy-Lipschitz, there exists a unique solution.

Define u_1 to be the solution of the ODE with $u(x_0) = 1$ and $u'(x_0) = 0$. Define u_2 to be the solution of the ODE with $u(x_0) = 0$ and $u'(x_0) = 1$.

Then $Au_1 + Bu_2$ is a solution $\forall A, B \in \mathbb{R}$. Given any other solution u of the ODE, define $A = u(x_0)$ and $B = u'(x_0)$. Then $v = u - Au_1 - Bu_2$ solves the ODE with $v(x_0) = 0$ and $v'(x_0) = 0$.

But 0 is a solution of the ODE as well. By uniqueness of Cauchy-Lipschitz, we have $v \equiv 0$. Thus $u = Au_1 + Bu_2$. □

Proof. [Cauchy-Lipschitz] Consider the ODE $u' = f(x, u(x))$ with initial condition $u(x_0) = y_0$. This is the same as

$$u(x) = u(x_0) + \int_{x_0}^{x} u'(s)\mathrm{d}s = y_0 + \int_{x_0}^{x} f(s, u(s))\mathrm{d}s.$$

Define $u_0(x) = y_0$. Then define $u_1(x) = y_0 + \int_{x_0}^{x} f(s, u_0(s))\mathrm{d}s$. In general, define

$$u_{n+1}(x) = y_0 + \int_{x_0}^{x} f(s, u_n(s))\mathrm{d}s.$$

As a fact, u_n converges uniformly to some function u. Then we have

$$u_\infty = y_0 + \int_{x_0}^{x} f(s, u_\infty(s))\mathrm{d}s.$$

This shows existence.

Let now u_1 and u_2 be two solutions of the ODE. Consider

$$u_1(x) - u_2(x) = \int_{x_0}^{x} \big(f(s, u_1(s)) - f(s, u_2(s))\big)\mathrm{d}s.$$

Putting the modulus on both sides and using the Lipschitz condition, we get

$$\begin{aligned} |u_1(x) - u_2(x)| &\leq \int_{x_0}^{x} |f(s, u_1(s)) - f(s, u_2(s))|\mathrm{d}s \\ &\leq \int_{x_0}^{x} L|u_1(s) - u_2(s)|\mathrm{d}s. \end{aligned}$$

Consider $x \in [x_0 - \tau, x_0 + \tau]$. Then

$$|u_1(x) - u_2(x)| \leq L \max_{[x_0 - \tau, x_0 + \tau]} |u_1(s) - u_2(s)| \cdot |x - x_0|.$$

Defining $a = \max_{[x_0 - \tau, x_0 + \tau]} |u_1(s) - u_2(s)|$, we have

$$|u_1(x) - u_2(x)| \leq La\tau.$$

This further implies that $a \leq La\tau$. Choosing $\tau = \frac{1}{2L}$, we get $a \leq \frac{a}{2}$, which implies $a = 0$. Thus $u_1(x) = u_2(x)$. □

Index