Mathematical Analysis I, 2021/22 First semester

Yoh Tanimoto
Dipartimento di Matematica, Università di Roma Tor Vergata
Via della Ricerca Scientifica 1, I-00133 Roma, Italy
email: `hoyt@mat.uniroma2.it`

We use the textbook "Calculus" Vol. I by Tom M. Apostol, Wiley, but follow the chapters in a different order.

From Monday to Thursday we have lectures, and on Friday we do mostly exercises.

- Lecture notes:

  `http://www.mat.uniroma2.it/~tanimoto/teaching/2021MA1/2021MA1.pdf`

- Exercises:

  `http://www.mat.uniroma2.it/~tanimoto/teaching/2021MA1/2021MA1ex.pdf`

- Office hours: Tuesday 10:00–11:00 online, or send me a message on Teams
- Supplementary course: Basic Math

Some tips

- Writing math.
  - LaTeX. You can try it here, and you can install the full set afterwards. You need to learn some commands, but once you know it it's very powerful. All my lecture notes and slides are written in LaTeX
  - Word processor (MS Word, Apple Pages, Open Office, Libre Office (Insert → Objects → Formula)...).
- Doing quick computations.
  - Wolfram Math Alpha You can just type some formulas in and it shows the result.
  - Programming languages. Python (I used it to make the graph of the SIR model), Java, C,$\cdots$

Sep 20. Overview of the course, integers and rational numbers

Mathematical Analysis I

Summary of the course

- properties of real numbers, concept of sets.
- mathematical induction. the summation notation.
- functions. limit of functions, continuity.
- trigonometric functions ($\cos x, \sin x$), exponential function e^x, logarithmic function $\log x$.

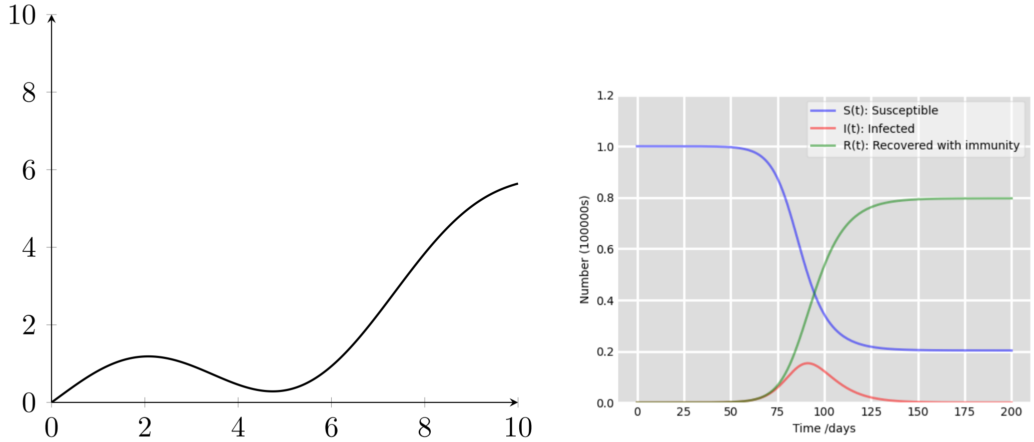


Figure 1: Left: A graph can be used to study changing quantities. Right:the SIR model.

- differential calculus and applications.
- Taylor's formula, approximation of functions
- integral of functions. the relation between integration and differentiation.
- basic differential equations.
- numerical seqeunces and series, complex numbers

What is analysis and why study it

In a real-world science, it is crucial to study **quantitative aspects** of the subject. When a quantity **changes** by time, one can study its change in a short time ($\Rightarrow$ differentiation) and then sum it up ($\Rightarrow$ integration). Another important problem is optimization: maximizing benefit or minimizing cost.

- Mechanics, the equation of motion $F(x,t) = m\frac{d^2x}{dt^2}(= ma)$
- Electrodynamics, thermodynamics, fluid mechanics (Mathematical Analysis II)
- Epidemiology (the SIR model $\frac{dS}{dt} = -\frac{\beta S(t)I(t)}{N}, \frac{dI}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t), \frac{dR}{dt} = \gamma I(t)$)

Mathematical symbols

We use **symbols** for general mathematical objects. Before using a symbol, we **declare** what kind of object it is. For example,

- x, y, z, a, b, c often for numbers (but be careful of the declaration)
- f, g, h, F, G, H often for "functions" (which we will study later)

A symbol might be "recycled", that is, can be declared to be something different (unfortunately, we have only 26×2 alphabets).

Symbols are very useful because we can express general properties of certain mathematical objects at the same time, without specifying them every time.

Integers and rational numbers

We assume that we know

- **integers**: $0, 1, 2, 3, \cdots, 100, 101, \cdots, 492837498 \cdots, -1, -2, -3, \cdots$
- **rational numbers**: $\frac{1}{2}, \frac{2}{3}, \cdots, \frac{23}{62518}, -\frac{3028746}{26543}, \cdots$ (integers are also rational numbers)
- calculations between them (sum, difference, product, division, order)

On rational numbers, we have the set of operations $+$ (summation), $\cdot$ (product): For x, y, z rational numbers (**declaration**), $x + y$ and $x \cdot y$ are again rational numbers and they satisfy

- (commutativity) $x + y = y + x, x \cdot y = y \cdot x$
- (associativity) $(x + y) + z = x + (y + z), (x \cdot y) \cdot z = x \cdot (y \cdot z)$
- (distributive law) $(x + y) \cdot z = xz + yz$
- (zero and unity) There are special distinct rational numbers, called 0 and 1, such that $x + 0 = x$ and $x \cdot 0 = 0$. And $x \cdot 1 = x$.
- (negative) There is a only one rational number, which we call $-x$, such that $x + (-x) = 0$.
- (inverse) If $x \neq 0$, there is only one rational number, which we call x^{-1}, such that $x \cdot x^{-1} = 1$.

We often simply write xy for $x \cdot y$ and $x - y$ for $x + (-y)$. xy^{-1} is also written as $\frac{x}{y}$.

Exercises Take concrete rational numbers and check these properties!

Other properties of rational numbers can be **derived** from these. Indeed, we can prove the following[1]

Theorem 1. *Let a, b, c, d be rational numbers.*

- *if $a + b = a + c$, then $b = c$.*
- $-(-a) = a$.
- $a(b - c) = ab - ac$.
- $a \cdot 0 = 0 \cdot a = 0$.
- *if $ab = ac$ and $a \neq 0$, then $b = c$.*
- *if $a \neq 0$, then $a^{-1} \neq 0$ and $(a^{-1})^{-1} = a$.*
- *if $ab = 0$, then $a = 0$ or $b = 0$.*
- *$(-a)b = -(ab)$ and $(-a)(-b) = ab$.*
- *if $b \neq 0, d \neq 0$, then $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$.*
- *if $b \neq 0, d \neq 0$, then $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$.*
- *if $a \neq 0, b \neq 0$, then $(\frac{a}{b})^{-1} = \frac{b}{a}$.*

[1] statements that can be proven are called **theorems**, and the properties that we assume are called **axioms**.

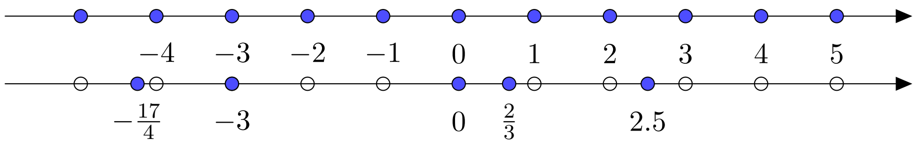


Figure 2: A geometric representation of integers and rational numbers.

.

Proof. We only prove a few of them and **leave the rest as exercises**.

Let us assume that $a + b = a + c$. Then, we take $-a$ and

$$\begin{aligned}(a + b) + (-a) &= a + (b + (-a)) \text{ (associativity)}\\ &= a + ((-a) + b) \text{ (commutativity)}\\ &= (a + (-a)) + b \text{ (associativity)}\\ &= 0 + b \text{ (definition of 0)}\\ &= b \text{ (property of 0)}\end{aligned}$$

Similarly, $(a+c)+(-a) = c$. But as $a+b = a+c$, we have $b = (a+b)+(-a) = (a+c)+(-a) = c$.

Assume that $a \neq 0$. We show that $a^{-1} \neq 0$ by contradiction. Indeed, if we had $a^{-1} = 0$, then we would have $0 \cdot a = a^{-1} \cdot a = 1$, contradiction. Therefore, $a^{-1} \neq 0$ and $1 = a \cdot a^{-1} = a^{-1} \cdot a$, hence $a = (a^{-1})^{-1}$.

If $ab = 0$ and $a \neq 0$, then we can take a^{-1} and $0 = a^{-1}0 = a^{-1}ab = 1 \cdot b = b$.

□

Integers and rational numbers can be represented on a line.

Sep 22. Concept of sets, set of numbers

Order in rational numbers

There is also an **order relation** $<$ ("x is larger than y": $y < x$) which satisfies, for x, y, z rational,

- if $0 < x, 0 < y$, then $0 < xy$ and $0 < x + y$.
- if $x < y$, then $x + z < y + z$.
- if $x \neq 0$, either $0 < x$ or $x < 0$ but not both.
- It is not true that $0 < 0$.

$x < y$ and $y > x$ have the same meaning.

We say that x is **positive** if $0 < x$ and x is **negative** if $x < 0$. If x is not positive, then either $x = 0$ or $x < 0$ and in this case we say x is **nonpositive** and write $x \leq 0$ (again, $x \geq 0$ and $0 \leq x$ mean the same thing). Similarly, if $x > 0$ or $x = 0$, we say x is **nonnegative** and write $x \geq 0$.

In addition to the "axioms", we also use the logic that, if an equality or inequality holds for some x and if $x = y$, then it also holds for y. **Example**: if $x < z$ and $x = y$, then $y < z$.

Notation: for any number x, we write $x^2 = x \cdot x$. Similarly, $x^3 = x \cdot x \cdot x$, and so on.

With the properties above, we can prove the following.

Theorem 2. *Let a, b, c, d rational numbers. Then*

- *$a < b$ if and only if $a - b < 0$.*
- *one and only one of the following holds: $a < b, b < a, a = b$.*

- *if* $a < b, b < c$ *then* $a < c$.
- *if* $a < b$ *and* $c > 0$, *then* $ac < bc$.
- *if* $a \neq 0$, *then* $a^2 (= a \cdot a) > 0$.
- $1 > 0$.
- *if* $a < b$ *and* $c < 0$, *then* $ac > bc$.
- *if* $a < b$, *then* $-a > -b$.
- *if* $ab > 0$, *then either* $a > 0, b > 0$ *or* $a < 0, b < 0$.
- *if* $a < c$ *and* $b < d$, *then* $a + b < c + d$.

Proof. We only prove a few of them and leave the rest as exercises. In general, to show "A if and only if B", it is enough to show that "if A, then B" and "if B, then A", this is because "A only if B" implies that "if not B, then not A", and by contradiction, "if A, not not B", but "not not B" means B.

If $a < b$, then by adding $-b$ to both sides, we get $a - b < 0$. Conversely, if $a - b < 0$, by adding b to both sides we get $a < b$.

If $a = b$, then $b - a = 0$ and we know that both $b - a > 0$ and $b - a < 0$ are false and hence $b > a$ and $b < a$ are false. If $a < b$, then $a - b < 0$ and $a - b = 0$ is false, and hence $a = b$ is false.

If $a < b$, then $0 < b - a$ and $0 < c \cdot (b - a) = bc - ac$, hence $ac < bc$.

If $a \neq 0$, then either $a > 0$ or $a < 0$. For the case $a > 0$, we have $a^2 = a \cdot a > 0$. For the case $a < 0$, we have $-a > 0$ and $a^2 = (-a)^2 > 0$.

□

All these "theorems" about rational numbers should be well-known to you. But it is important that we could prove them from a few axioms, which we assume to be true.

Exercises Check the remaining statements.

Naive set theory

It is very often convenient to consider sets of numbers. For example, we may consider the set $\mathbb{Q}_+$ of positive rational numbers, or the set of multiples of 2, and so on. In mathematics, a **set** is a collection of mathematical objects. The most precise treatment of sets requires a theory called axiomatic set theory, but in this lecture we think of a set simply as a collection of known objects.

We often use capital letters $A, B, C, \cdots$ for sets, but in any case we declare that a symbol is a set. For a set S, we denote by $x \in S$ the statement "x is an **element** of S". We have already seen examples of sets: let us give them special symbols.

- $\mathbb{Q}$: the set of rational numbers
- $\mathbb{Z}$: the set of integers

In general, we can consider two ways of constructing sets.

- By nomination. We can nominate all elements of a set. For example, $A = \{0, 1, 2, 3\}$ and $B = \{1, 10, 100, 1000\}$ are sets.
- By specification. We include all elements of an existing set with specific properties. For example, $A = \{x \in \mathbb{Z} : \text{ there is } y \in \mathbb{Z} \text{ such that } x = 2y\}$ (read that "A is the set of integers such that there is an integer y such that $x = 2y$") is the set of multiples of 2 (we recycled the symbol A. When we do this, we shall always declare it).

For a set constructed by nomination, the order and repetition do not matter: $\{0, 1, 2, 3\} = \{0, 3, 2, 1\} = \{0, 0, 1, 1, 1, 2, 3\}$. In other words, a set is defined by its elements.

A construction by specification appears very often. Let us introduce a more symbol.

- $\mathbb{N} = \{x \in \mathbb{Z} : x > 0\}$ is called the set of **natural numbers**.
- $\emptyset$ is the set that contains nothing and called the **empty set**. $\emptyset$ is a subset of any set: if A is a set, the statement "if $x \in \emptyset$ then $x \in A$" is satisified just because there is no such x!

Subsets

Let B be a set. We say that A is a **subset** of B if all elements of A belong to B, and denote this by $A \subset B$. It holds that $A \subset A$ for any set A.

Example 3.
- Let $A = \{1, 2, 3\}$ and $B = \{0, 1, 2, 3, 4\}$. Then $A \subset B$.
- $\mathbb{N} \subset \mathbb{Z}$.
- Let $A = \{1, 2, 3, 4, 5, 6\}$. Then $A \subset \mathbb{N}$.

It may happen that $A \subset B$ and $B \subset A$, that is, all elements of A belong to B and vice versa. This means that A and B are the same as sets, and in this case we write $A = B$.

The definition by specification $A = \{x \in B : x \text{ satisfies the property XXX...}\}$ gives always a subset, in this case of B. Note also that x in this definition has no meaning ("dummy"). One can write it equivalently $A = \{y \in B : y \text{ satisfies the property XXX...}\}$.

For $x \in A$, the set $\{x\}$ that contains only x should be distinguished from x. It is a subset of A: $\{x\} \subset A$.

Unions, intersections, complements

If A and B are sets, then we can consider the set which contains the elements of A and B, and nothing else. It is called the **union** of A and B and denoted by $A \cup B$.

Example 4.
- Let $A = \{1, 2, 3\}$ and $B = \{0, 1, 3, 4\}$. Then $A \cup B = \{0, 1, 2, 3, 4\}$.

Similarly, we can consider the set of all the elements which belong both to A and B, and nothing else. It is called the **intersection** of A and B and denoted by $A \cap B$.

Example 5.
- Let $A = \{1, 2, 3\}$ and $B = \{0, 1, 3, 4\}$. Then $A \cap B = \{1, 3\}$.

Furthermore, the **difference** of B with respect to A is all the element of A that do not belong to B and is denoted by $A \setminus B$ (note that this is different from $B \setminus A$).

Example 6.
- Let $A = \{1, 2, 3\}$ and $B = \{0, 1, 3, 4\}$. Then $A \setminus B = \{2\}$.

We can consider the union of more than two sets: $A \cup (B \cup C)$. By considering the meaning, this set contains all the elements which belong either A or $B \cup C$, which is to say all elements which belong either A or B or C. Therefore, the order does not matter and we can write $A \cup B \cup C$. Similarly, $A \cap B \cap C$ is the intersection of A, B and C.

We may consider a **family of sets** $\{A_i\}_{i \in I}$ **indexed by another set** I. For example, we can take $\mathbb{N}$ as the index set and $A_n = \{m \in \mathbb{Z} : m \text{ is a multiple of } n\}$. For a family of set, we can define the union and the intersection analogously and we denote them by

$$\bigcup_{i \in I} A_i, \qquad \bigcap_{i \in I} A_i,$$

respectively.

Sep 23. Sets and logic

Sets by specification

Let us recall that, if we fix a set A, we can define a subset of A by specification: it is the subset of elements x of A that satisfy a certain condition $\varphi(x)$:

$$\{x \in A : \varphi(x)\},$$

where $\varphi(x)$ is a condition on x. For example, $\{x \in \mathbb{Z} : x > 10\}$ is the set of integers larger than 10.

On one hand, we can consider the combined conditions: for example, the condition that $x > 10$ and the condition that $x < 15$ can be considered at the same time. The set of integers that satisfy both of the condition is

$$\{x \in \mathbb{Z} : x > 10 \text{ and } x < 15\} = \{11, 12, 13, 14\}.$$

On the other hand, we observe that this set is the intersection of two sets:

$$\{x \in \mathbb{Z} : x > 10\} = \{10, 11, 12, 13, 14, 15, 16, 17, \cdots\},$$
$$\{x \in \mathbb{Z} : x < 15\} = \{-2, -1, 0, 1, \cdots, 10, 11, 12, 13, 14, 15\}.$$

This can be generalized as follows. If $B = \{x \in A : \varphi(x)\}, C = \{x \in A : \psi(x)\}$, then $B \cap C = \{x \in A : \varphi(x) \text{ and } \psi(x)\}$.

Similarly, the union of two sets is related with "or" as follows. For example,

$$B = \{x \in \mathbb{Z} : x > 15\} = \{16, 17, 18, \cdots\},$$
$$C = \{x \in \mathbb{Z} : x < 10\} = \{\cdots - 2, -1, 0, 1, \cdots, 8, 9\}.$$

We observe that $B \cup C = \{\cdots - 2, -1, 0, 1, \cdots, 8, 9, 16, 17, 18, \cdots\}$, which is $B \cup C = \{x \in \mathbb{Z} : x > 15 \text{ or } x < 10\}$. In general, if $B = \{x \in A : \varphi(x)\}, C = \{x \in A : \psi(x)\}$, then $B \cup C = \{x \in A : \varphi(x) \text{ or } \psi(x)\}$.

Let us the consider the negation. For example,

$$B = \{x \in \mathbb{Z} : x > 15\} = \{16, 17, 18, \cdots\}.$$

Because the negation of $x < 15$ is $x \leq 15$, we have

$$\mathbb{Z} \setminus B = \{\cdots - 2, -1, 0, 1, \cdots, 14, 15\} = \{x \in \mathbb{Z} : x \leq 15\}.$$

In general, if $\neg\varphi(x)$ is the negation of $\varphi(x)$, then it holds that, for $B = \{x \in A : \varphi(x)\}$, $A \setminus B = \{x \in A : \neg\varphi(x)\}$.

Let us consider the set $\{x \in \mathbb{Q} : (x-1)(x-3) > 0\}$. To understand better this set, we need to understand the condition $(x-1)(x-3) > 0$. The left-hand side is a product of two rational numbers. The product of two rational numbers is positive if and only if one of the following cases is true.

- $x - 1 > 0$ and $x - 3 > 0$
- $x - 1 < 0$ and $x - 3 < 0$.

They are further equivalent to

- $x > 1$ and $x > 3$
- $x < 1$ and $x < 3$.

Note that $x > 1$ is true if $x > 3$. Similarly, $x < 3$ is true if $x < 1$. Therefore, these conditions are equivalent to

- $x > 3$
- $x < 1$

Altogether, we have

$$\{x \in \mathbb{Q} : (x-1)(x-3) > 0\} = \{x \in \mathbb{Q} : x > 3 \text{ or } x < 1\} = \{x \in \mathbb{Q} : x > 3\} \cup \{x \in \mathbb{Q} : x < 1\}.$$

The set of subsets, the set of pairs, graphs

We can consider also certain **sets of sets**.

Example 7.

- $\{1,2,3\}, \{2\}, \{1,4,6,7\}$ are sets. We can collect them together

$$\{\{1,2,3\}, \{0,2\}, \{1,4,6,7\}\}.$$

  This is a set of sets. It is different from the set of their elements $\{1,2,3,4,6,7\}$.

- Let $A = \{1,2,3\}$. We can collect all subsets of A:

$$\{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

- One can also consider the set of all subsets of $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, but we cannot name all the elements: they are infinite. For example, for $\mathbb{N} = \{1,2,3,4,\cdots\}$, the set of subsets of $\mathbb{N}$ is infinite.

For sets A, B, we can consider **ordered pairs** of elements in A and B.

Example 8.

- Let $A = \{1,2,3\}, B = \{3,4\}$. Then the sef $A \times B$ of the ordered pairs of A, B is

$$A \times B := \{(1,3), (2,3), (3,3), (1,4), (2,4), (3,4)\}.$$

- If we take $\mathbb{N}$, then $\mathbb{N} \times \mathbb{N}$ is the set of all ordered pairs of natural numbers. $\mathbb{N} \times \mathbb{N} = \{(1,1), (1,2), (2,1), (1,3), (2,2), (3,1), \cdots\}$.

Ordered pairs can be described using **graphs**. If $A, B \subset \mathbb{Z}$ have finitely many points, say m, n respectively, then there are $m \cdot n$ ordered pairs. We take the horizontal axis for A and the vertical axis for B.

To obtain the graph of $A \times B$, we should mark the point (x, y) if and only if $x \in A$ and $y \in B$. For any subset X of $A \times B$, we should mark the point (x, y) if and only if $(x, y) \in X$. See Figure 3.

The graph can be understood in terms of ordered pairs. Let $A = \{1,2,3,4,5,6\}$, and $B = \{(x, y) \in A \times A : y = 2x\}$. Let us give all elements of B and draw its graph. We check all $6 \times 6 = 36$ elements. See Figure 4, it is $\{(1,2), (2,4), (3,6)\}$. Notice that it is on a **straight line**!

Sep 27. Real numbers

Are rational numbers all we need?

It is true that, in the real world, we can measure quantities to a certain accuracy, so we get numbers in a decimal representation:

- $c = 299792458[\mathrm{m} \cdot \mathrm{s}^{-1}]$ (the speed of light)

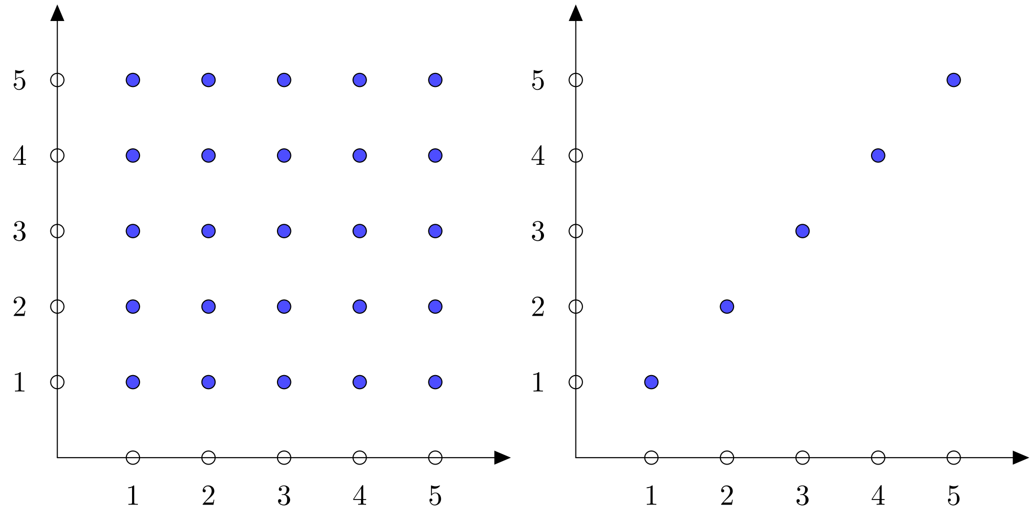


Figure 3: Left: the set of all ordered pairs $\{1,2,3,4,5\} \times \{1,2,3,4,5\}$. Right: a subset $\{(1,1),(2,2),(3,3),(4,4),(5,5)\} \subset \{1,2,3,4,5\} \times \{1,2,3,4,5\}$.

.

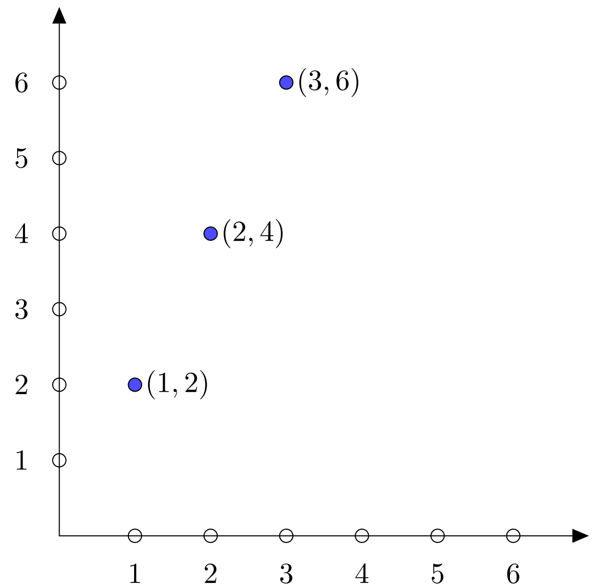
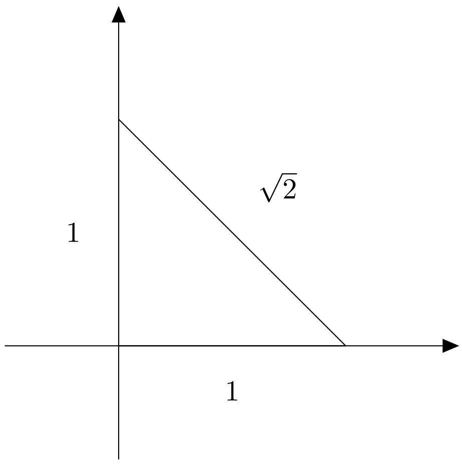


Figure 4: The set of all ordered pairs $(x, y) \in \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$ with $y = 2x$.



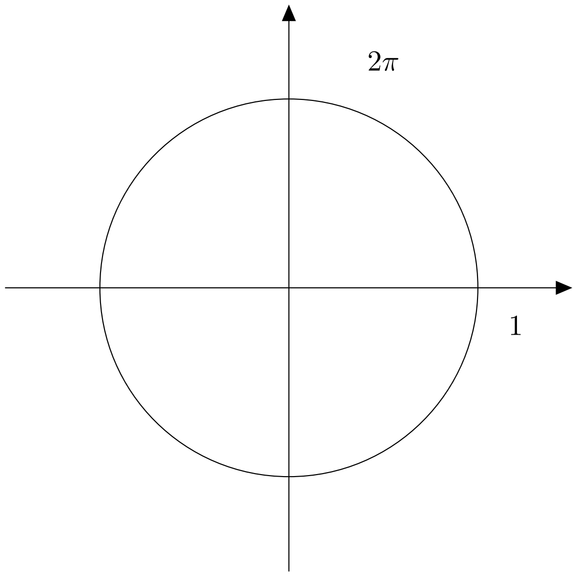


Figure 5: Left: the right triangle with equal sides 1. By the Pytagoras' theorem, the longest side is $\sqrt{1^2+1^2} = \sqrt{2}$. Right: the unit circle with radius 1 (diameter 2). The length of the circle (the circumference) is 2π.

- $G = 0.0000000000667430(15)[\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2}]$ (the gravitational constant), where (15) means these digits might be incorrect.
- Any other measured quantity in the real world.

And any experiment has a certain accuracy, so it makes sense only to write a certain number of digits, so rational numbers seem to suffice.

But for certain cases, we know that we should consider **irrational numbers**. For example,

- $\sqrt{2} = 1.41421356\cdots$, the number x such that $x^2 = 2$.
- $\pi = 3.1415926535\cdots$, the circumference of the circle with diameter 1.
- $e = 2.718281828\cdots$, Napier's number (we will define it in the lecture).
- Any decimal number which is not repeating.

For the next theorem, we need a **proof by contradiction**: by assuming the converse of the conclusion, we derive a contradiction, then we can conclude that the converse of the conclusion is false, that is, the conclusion is correct.

Recall that an integer p is **even** if it is a multiple of 2 (there is another integer r such that $p = 2r$), and p is **odd** if it is not even.

Theorem 9. *$\sqrt{2}$ is not a rational number.*

Proof. We prove this by contradiction, that is, we assume that $\sqrt{2}$ is a rational number. So there are integers p, q such that $\sqrt{2} = \frac{p}{q}$. We may assume that this is already reduced (that is, not a fraction like $\frac{4}{8}$ but like $\frac{1}{2}$. It's the form which you cannot simplify further).

As $\sqrt{2}\cdot\sqrt{2} = 2$, we have $\frac{p}{q}\cdot\frac{p}{q} = \frac{p^2}{q^2} = 2$, hence $p^2 = 2q^2$. As $\frac{p}{q}$ is reduced, there are two cases.

- if p is odd, then the equality $p^2 = 2q^2$ is even = odd, contradiction.
- if p is even, then q is odd and we can write is as $p = 2r$, with another integer r, and $p^2 = 4r^2 = 2q^2$, and $2r^2 = q^2$. This is even = odd, contradiction.

So, in all cases we arrived a contradiction from the assumption that $\sqrt{2}$ is rational. This means that $\sqrt{2}$ is irrational. □

Exercise. Prove that $2\sqrt{2}$ is irrational.

It has been proven that π and e are irrational, but they are more difficult. Instead, it can be easily proven that any nonrepeating decimal number cannot be rational. This means *there are many irrational numbers.*

In other words, the set of rational numbers have "many spaces between them". We should fill them in with irrational numbers, so that the set of real numbers is a "continuum".

The axioms of the real numbers

Here we start the study of Mathematical Analysis, based on the set of **real numbers**. Our approach is synthetic, in the sense that we take the axioms for real numbers for granted, and develop the theory on them. It is also possible to "costruct" real numbers from rational numbers, and rational numbers from integers, integers from natural numbers, and so on, but at some point we have to assume certain axioms for simpler objects. If you are interested, look at "Dedekint's cut" (for real numbers), or "Peano's axioms" (for natural numbers).

We assume that, the set $\mathbb{R}$ of real numbers is equipped with operations $+$ (summation), $\cdot$ (product) and for x, y, z real numbers, $x + y$ and $x \cdot y$ are again real numbers and they satisfy (*just the same properties for rational numbers* $\mathbb{Q}$)

- (commutativity) $x + y = y + x, x \cdot y = y \cdot x$
- (associativity) $(x + y) + z = x + (y + z), (x \cdot y) \cdot z = x \cdot (y \cdot z)$
- (distributive law) $(x + y) \cdot z = xz + yz$
- (zero and unity) There are special distinct rational numbers, called 0 and 1, such that $x + 0 = x$ and $x \cdot 0 = 0$. And $x \cdot 1 = x$.
- (negative) There is only one real number, which we call $-x$, such that $x + (-x) = 0$.
- (inverse) If $x \neq 0$, there is only one real number, which we call x^{-1}, such that $x \cdot x^{-1} = 1$.

There is also an **order relation** $<$ which satisfies, for x, y, z real,

- if $0 < x, 0 < y$, then $0 < xy$ and $0 < x + y$.
- if $x < y$, then $x + z < y + z$.
- if $x \neq 0$, either $0 < x$ or $x < 0$ but not both.
- $0 \nless 0$

We can prove Theorems for real numbers corresponding to Theorems 1, 2. Therefore, the real numbers have the same properties as the rational numbers, concerning the sum, product and order.

We say that $S \subset \mathbb{R}$ is **bounded above** if there is $x \in \mathbb{R}$ such that for any $y \in S$ it holds that $y \leq x$, and we write $S \leq x$. S is said to be **bounded below** if there is $x \in \mathbb{R}$ such that for any $y \in S$ it holds that $y \geq x$, and we write $S \geq x$.

If S is both bounded above and below, we say that S is **bounded**.

If S is bounded above, then any $x \in \mathbb{R}$ such that $S \leq x$ is called an **upper bound** of S. Similarly, if $x \leq S$, then x is said to be a **lower bound** of S.

If S has a least upper bound, that is there is x such that $S \leq x$ and $x \leq y$ for any upper bound y of S, then x is called the **supremum** of S and we denote it by $x = \sup S$. Similarly, if S has a largest lower bound x, then it is called the **infimum** of S and we denote it by $x = \inf S$.

$\mathbb{R}$ includes $\mathbb{Z}$ and $\mathbb{Q}$: $1 \in \mathbb{R}$, hence $2 = 1 + 1, 3 = 1 + 1 + 1, \cdots$ and $-1, -2, \cdots \in \mathbb{R}$. Also, if $p, q \in \mathbb{Z}, \frac{p}{q} \in \mathbb{R}$.

What distinguishes $\mathbb{R}$ from $\mathbb{Q}$ is the following.

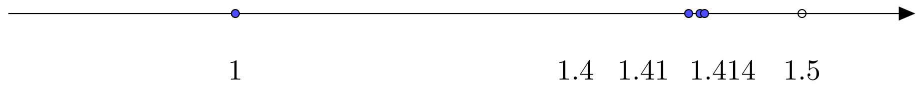


Figure 6: The set S approximating $\sqrt{2}$, which is bounded by 1.5.

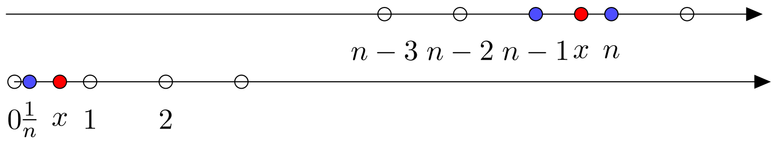


Figure 7: Any $x \in \mathbb{R}$ falls between $n-1$ and n (including equality) for some $n \in \mathbb{N}$. For any $x > 0$, there is $n \in \mathbb{N}$ such that $\frac{1}{n} < x$.

- (the least upper bound axiom, or the completeness axiom) every nonempty subset S of $\mathbb{R}$ which is bounded above has a supremum: there is $B \in \mathbb{R}$ such that $B = \sup S$.

This should imply that $\sqrt{2} = 1.41421356\cdots$ belongs to $\mathbb{R}$! Indeed, let us take, by chopping the digits of $\sqrt{2}$, $S = \{1, 1.4, 1.41, 1.414, 1.4142, \cdots\}$. S is bounded above, indeed, $1.5 > 1, 1.4, 1.41, 1.414, \cdots$. On the other hand, if x has a decimal representation, e.g. 1.415, then there is a smaller number $x' = 1.4149$. So, $\sup S$ should be exactly $\sqrt{2}$. We will see this more precisely later.

(A lemma is a theorem (a consequence of axioms) used to prove a more important theorem)

Lemma 10. *If $S \subset \mathbb{R}$ is bounded above and $B = \sup S$, then for any $\epsilon > 0$, there is $x \in S$ such that $B - \epsilon < x$.*

Proof. By contradiction, assume that there is $\epsilon > 0$ such that $B - \epsilon \geq x$ for all $x \in S$. Then B is not the *least* upper bound, because $B - \epsilon$ is an upper bound of S and $B - \epsilon < B$. □

Theorem 11 (Archimedean property)**.** *The set $\mathbb{N} = \{1, 2, 3, \cdots\}$ is not bounded above.*

Proof. By contradiction, assume that $\mathbb{N}$ were bounded above. Then by the completeness axiom, there is $x = \sup \mathbb{N}$. By the lemma above, for $\epsilon = \frac{1}{2}$, there is $n \in \mathbb{N}$ such that $x - \frac{1}{2} < n$. But then $x < n + \frac{1}{2} < n + 1 \in \mathbb{N}$, and this contradicts the assumption that x were the upper bound of $\mathbb{N}$. This implies that $\mathbb{N}$ is not bounded above. □

(A corollary is a theorem which follows easily from a more complicated theorem)

Corollary 12. *For any $x \in \mathbb{R}$, there is $n \in \mathbb{N}$ such that $x < n$. For any $y, z \in \mathbb{R}$ and $z > y$, there is $n \in \mathbb{N}$ such that $\frac{1}{n} < z - y$.*

Proof. By the theorem above, x is not an upper bound of $\mathbb{N}$, so there is n such that $x < n$. By applying this to $\frac{1}{z-y}$, there is n such that $\frac{1}{z-y} < n$, which implies that $\frac{1}{n} < z - y$. □

Therefore, we can represent the set of real numbers by a straight line, and any point $x \in \mathbb{R}$ is on the line and it falls between an integer n and another $n-1$ (possibly $x = n$). Conversely, any point on the line gives an element in $\mathbb{R}$.

Any real number $\mathbb{R}$ has a decimal representation (next lecture).

Note that $\mathbb{Q}$ does not have the completeness property!

Proposition 13. *Let $A = \{x \in \mathbb{Q} : x^2 < 2\} \subset \mathbb{R}$. Then A is bounded above, and $s^2 = 2$, where $s = \sup A$.*

Proof. A is bounded above, indeed, if $x^2 < 2$, then $x^2 < 4 = 2^2$, and hence $x < 2$.

Let $s = \sup A \in \mathbb{R}$. Then $s^2 = 2$. We prove this by contradiction.

- if $s^2 < 2$, then we take $\epsilon > 0$ such that $0 < \epsilon < \frac{2-s^2}{s}$ (or $s\epsilon < 2 - s^2$) and $\epsilon < s$. Then $(s + \frac{\epsilon}{4})^2 = s^2 + s\frac{\epsilon}{2} + \frac{\epsilon^2}{16} < s^2 + \frac{s\epsilon}{2} + \frac{s\epsilon}{2} < s^2 + s\epsilon < 2$, therefore, s is not an upper bound of A (because $s + \frac{\epsilon}{4} \in A$), contradiction.
- if $s^2 > 2$, then we take $\epsilon > 0$ such that $0 < \epsilon < \frac{s^2-2}{s}$ (or $s\epsilon < s^2 - 2$) and $\epsilon < s$. Then $(s - \frac{\epsilon}{4})^2 = s^2 - s\frac{\epsilon}{2} + \frac{\epsilon^2}{16} > s^2 - s\epsilon > 2$, therefore, s is not the least upper bound of A (because $s - \frac{\epsilon}{4} \in A$ is another upper bound, smaller than s), contradiction.

But we know that there is no rational number $s \in \mathbb{Q}$ such that $s^2 = 2$. Hence $s = \sup A \notin \mathbb{Q}$. □

This also says that $s = \sqrt{2}$ belongs to $\mathbb{R}$.

Sep 29. Some sets in real numbers.

Intervals

In the set of real numbers, we can consider **intervals**: let $a, b \in \mathbb{R}$ and $a < b$. We introduce

- $(a, b) = \{x \in \mathbb{R} : a < x, x < b\}$ (an **open** interval)
- $(a, b] = \{x \in \mathbb{R} : a < x, x \leq b\}$
- $[a, b) = \{x \in \mathbb{R} : a \leq x, x < b\}$
- $[a, b] = \{x \in \mathbb{R} : a \leq x, x \leq b\}$ (a **closed** interval)
- $(a, \infty) = \{x \in \mathbb{R} : a < x\}$
- $[a, \infty) = \{x \in \mathbb{R} : a \leq x\}$
- $(-\infty, b) = \{x \in \mathbb{R} : x < b\}$
- $(-\infty, b] = \{x \in \mathbb{R} : x \leq b\}$

Remember that, a, b are given numbers, and x is a "dummy" number. You can write them in a different way, without using x:

- (a, b) is the set of all numbers larger than a and smaller than b
- $[a, b]$ is the set of all numbers larger than or equal to a and smaller than or equal to b

Example 14. Consider $(0, 1)$.

- $0.1, 0.2, 0.5, 0.999 \in (0, 1)$.
- $0, 1, 2, 3, 10, -1, -2 \notin (0, 1)$.
- $\sup(0, 1) = 1$.
- $\inf(0, 1) = 0$.

Consider $[0, 1]$.

- $0, 0.1, 0.2, 0.5, 0.999, 1 \in [0, 1]$.
- $2, 3, 10, -1, -2 \notin [0, 1]$.
- $\sup[0, 1] = 1$.
- $\inf[0, 1] = 0$.

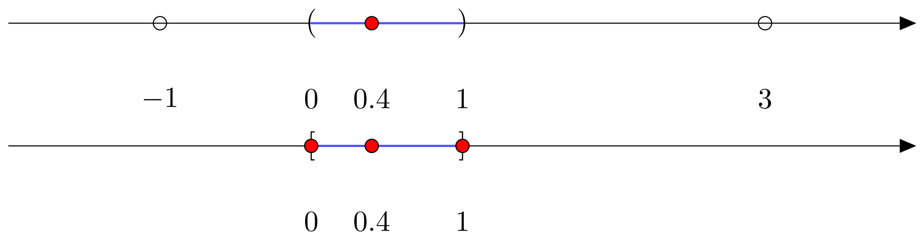


Figure 8: Open and closed intervals $(0,1)$ and $[0,1]$. The open interval does not include the edges $0,1$, while the closed interval $[0,1]$ does.

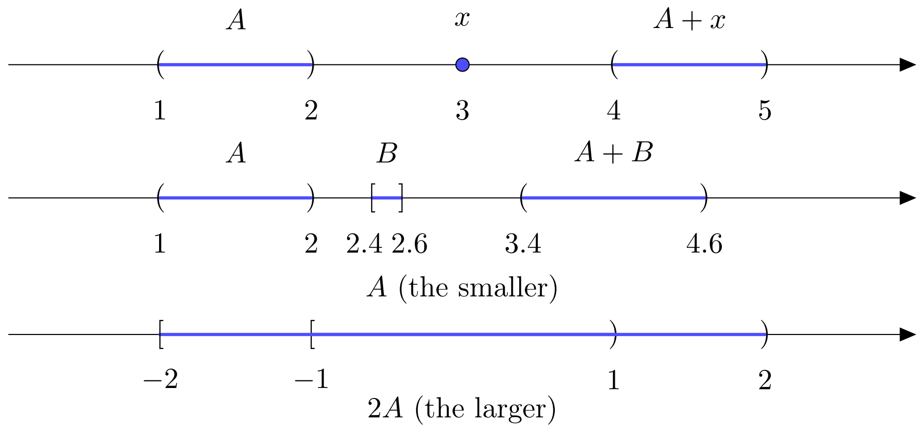


Figure 9: Intervals and their operations. Top: $(1,2)+3=(4,5)$. Middle: $(1,2)+[2.4,2.6]=(3.4,4.6)$. Bottom: $2[-1,1)=[-2,2)$.

Operations on sets

Let A, B be subsets of $\mathbb{R}$ and $a \in \mathbb{R}$. We denote various subsets $\mathbb{R}$ as follows.

- $A + a = \{x \in \mathbb{R} : x = y + a \text{ for some } y \in A\} = \{y + a : y \in A\}$
- $A - a = \{x \in \mathbb{R} : x = y - a \text{ for some } y \in A\} = \{y - a : y \in A\}$
- $aA = \{x \in \mathbb{R} : x = ay \text{ for some } y \in A\} = \{ay : y \in A\}$
- $A + B = \{x \in \mathbb{R} : x = y + z \text{ for some } y \in A, z \in B\} = \{y + z : y \in A, z \in B\}$
- $A - B = \{x \in \mathbb{R} : x = y - z \text{ for some } y \in A, z \in B\} = \{y - z : y \in A, z \in B\}$
- $AB = \{x \in \mathbb{R} : x = yz \text{ for some } y \in A, z \in B\} = \{yz : y \in A, z \in B\}$

We write $a < x < b$ as a shorthand notation for $a < x$ and $x < b$.

Example 15.

- Consider $A = (0,1)$, $x = 2$. Then $A + x = (2,3)$, because if $0 < y < 1, 2 < y + 2 < 3$. Note that the boundary $2,3$ is not included.
- Consider $A = [1,2]$, $B = (2.4, 2.6)$. Then $A + B = (3.4, 4.6)$. Note that the boundary $2,3$ is not included, because there is no $x \in A, y \in B$ such that $x + y = 3.4$ or 4.6.
- Consider $A = [-1,1), a = 2$. Then $2A = [-2,2)$.

Some properties of upper and lower bounds

Note that $\sup A, \inf A$ are only defined for nonempty sets (otherwise the definition is meaningless).

Lemma 16. *If $x, y \in \mathbb{R}$ and $x - \epsilon < y$ for any $\epsilon > 0$, then $x \le y$.*

Proof. By contradiction. If $x > y$, then by Archimedean property, we have n such that $\frac{1}{n} < x - y$, in other words, $x - \frac{1}{n} > y$, which contradicts the assumption that $x - \epsilon < y$ for arbitrary $\epsilon > 0$. □

Theorem 17. *Let $A, B \subset \mathbb{R}$ and define $C = A + B$.*

- *if A, B are bounded above, then $A + B$ is bounded above and $\sup A + \sup B = \sup C$.*
- *if A, B are bounded below, then $A + B$ is bounded below and $\inf A + \inf B = \inf C$.*

Proof. We prove only the first one, because the second one is analogous.

By the completeness axiom, A and B have the supremum $\sup A, \sup B$. As $\sup A$ and $\sup B$ are upper bounds of A and B respectively, for any element $z \in C$ we have $x \in A, y \in B$ such that $z = x + y$ and $x \le \sup A$, $y \le \sup B$ hence $z = x + y \le \sup A + \sup B$. In particular, $\sup A + \sup B$ is an upper bound of C, hence $\sup C \le \sup A + \sup B$.

Conversely, we know from Lemma 10 that, for any $\epsilon > 0$, there is $x \in A$ (and $y \in B$) such that $\sup A - \frac{\epsilon}{2} < x$ (and $\sup B - \frac{\epsilon}{2} < y$). Therefore, $\sup A + \sup B - \frac{\epsilon}{2} - \frac{\epsilon}{2} = \sup A + \sup B - \epsilon < x + y \le \sup C$ for arbitrary $\epsilon > 0$, hence by Lemma 16, $\sup A + \sup B \le \sup C$. Altogether, hence $\sup C = \sup A + \sup B$. □

Remember that $\sup A$ is the least (smallest) upper bound and $\inf B$ is the greatest (largest) lower bound.

Theorem 18. *Let $A, B \subset \mathbb{R}$. If for any $x \in A$ and $y \in B$ it holds that $x < y$, then $\sup A \le \inf B$.*

Proof. Any $y \in B$ is an upper bound of A, hence $\sup A \le y$. This means that $\sup A$ is a lower bound of B, hence $\sup A \le \inf B$. □

The square roots of real numbers

Theorem 19. *For any $a \in \mathbb{R}, a > 0$, there is $s \in \mathbb{R}, s > 0$ such that $s^2 = a$.*

Proof. Let $A = \{x \in \mathbb{R} : x^2 < a\} \subset \mathbb{R}$. Then A is bounded above: Indeed, as $x^2 < a$, there are two cases:

- if $a > 1$, then $x^2 < a^2$ and hence $x < a$.
- if $a \le 1$, then $x^2 < 1$ and hence $x < 1$.

In either case, A is bounded.

Note that A is not empty, because $0 \in A$. Let $s = \sup A \in \mathbb{R}$. $s \ne 0$ because we can take n large enough by the Archimedean property that $\frac{1}{n} < a$, therefore, $(\frac{1}{n})^2 < a$ (because $\frac{1}{n} < 1$), therefore, $\frac{1}{n} \le s$.

We prove $s^2 = a$ by contradiction.

- if $s^2 < a$, then we take $\epsilon > 0$ such that $0 < \epsilon < \frac{a - s^2}{s}$ (or $s\epsilon < a - s^2$) and $\epsilon < s$. Then $(s + \frac{\epsilon}{4})^2 = s^2 + s\frac{\epsilon}{2} + \frac{\epsilon^2}{16} < s^2 + \frac{s\epsilon}{2} + \frac{s\epsilon}{2} < s^2 + s\epsilon < a$, therefore, s is not an upper bound of A (because $s + \frac{\epsilon}{4} \in A$), contradiction.
- if $s^2 > a$, then we take $\epsilon > 0$ such that $0 < \epsilon < \frac{s^2 - a}{s}$ (or $s\epsilon < s^2 - a$) and $\epsilon < s$. Then $(s - \frac{\epsilon}{4})^2 = s^2 - s\frac{\epsilon}{2} + \frac{\epsilon^2}{16} > s^2 - s\epsilon > a$, therefore, s is not the least upper bound of A (because $s - \frac{\epsilon}{4} \in A$ is another upper bound, smaller than s), contradiction.

□

We denote it by $s = \sqrt{a}$.

For any $n \in \mathbb{N}$, we can define the n-th root of any positive number a and we denote it by $a^{\frac{1}{n}}$. The existence can be proved similarly.

Decimal representation of real numbers

We denote $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Any (positive) real number $x \in \mathbb{R}$ can be written in the form $x = a_0 + \frac{a_1}{10} + \frac{a_2}{100} + \cdots$, where a_0 is an integer and $a_1, a_2, \cdots$ are integers between 0 and 9 (negative numbers can be most commonly written as $-\sqrt{2} = -1.41421\cdots$, although an analogous representation can apply to negative numbers).

Examples:

- $\frac{1}{3} = 0.33333\cdots$
- $\sqrt{2} = 1.41421\cdots$
- $\pi = 3.14159\cdots$

Indeed, let $x \in \mathbb{R}$ be a real number and $x > 0$. By the Archimedean property, there is a natural number $n \in \mathbb{N}_0$ such that $n - 1 \leq x < n$ (this is possible, because any subset of $\mathbb{N}$ has the minimal element, which we prove below). We take $a_0 = n - 1$.

Note that $0 < x - a_0 < 1$. Therefore, $0 < 10(x - a_0) < 10$. Take $a_1 \in \mathbb{N}_0$ the largest natural number such that $a_1 \leq 10(x - a_0)$. As it is the largest, we have again $0 < 10(x - a_0) - a_1 < 1$.

We can repeat this procedure and obtain a_n, and it always hold that $x - a_0.a_1 \cdots a_n < 0.\underbrace{0\cdots 01}_{n\text{-digits}}$.

Let $A = \{a_0, a_0.a_1, a_0.a_1a_2, a_0.a_1a_2a_3, \cdots\}$. This A is bounded (by $a_0 + 1$), hence it has the supremum s. Note that x is un upper bound of A, hence $\sup A \leq x$. On the other hand, if for any $\epsilon = 0.\underbrace{0\cdots 01}_{n\text{-digits}}$, we have $x - \epsilon < a_0.a_1 \cdots a_n \in A$, therefore, $x \leq \sup A$. Altogether, $x = \sup A = s$.

Proposition 20. *A real number that is has nonrepeating decimal representation is irrational.*

Proof. We prove that any (positive) rational number has a repeating decimal representation. Then the claim follows by contradiction.

Let $x = a_0.a_1a_2 \cdots = \frac{p}{q}$, $p, q \in \mathbb{N}$. We can write $p = nq + r$, where $n, r \in \mathbb{N}$ and $0 \leq r < q$ (division with remainder). We set $a_0 = n$. Then we write $10r = n_1q + r_1$ again, and wet $a_1 = n_1$. In this way, we obtain the decimal representation of $\frac{p}{q}$, but there are only finitely many possible values $0, 1, \cdots q - 1$ of r_1 because we are doing the division with remainder by q. This means that the numbers repeat after at largest q digits. □

The converse of this (any irrational number has a nonrepeating decimal representation) will be proven later.

Sep 30. Natural numbers and induction.

Mathematical induction

The set $\mathbb{N}$ of natural numbers can be caracterized by the Peano axioms:

- $1 \in \mathbb{N}$
- For every $n \in \mathbb{N}$, $n + 1 \in \mathbb{N}$

- For every $n \in \mathbb{N}$, $n + 1 \neq 1$
- Let $S \subset \mathbb{N}$. If $1 \in S$ and $n + 1 \in S$ for any $n \in S$, then $S = \mathbb{N}$.

In other words, $\mathbb{N}$ consists of 1 and all other numbers obtained by adding 1 repeatedly to 1, and that is all. This is the precise definition of $\mathbb{N}$.

With this characterization, we obtain the mathematical induction. Let $\varphi(n)$ be a set of propositions depending on $n \in \mathbb{N}$. If $\varphi(1)$ is true, and if we can prove $\varphi(n+1)$ from $\varphi(n)$, then $\varphi(n)$ is true for all natural numbers. Indeed, let $S = \{n \in \mathbb{N} : \varphi(n) \text{ is true }\}$. S is a subset of $\mathbb{N}$, $1 \in S$ and if $n \in S$, then $n + 1 \in S$. From the Peano axioms, we have $S = \mathbb{N}$. In other words, $\varphi(n)$ holds for all $n \in \mathbb{N}$.

Example 21. $n^2 \geq 2n - 1$ for all n.

Indeed, we apply mathematical induction to $\varphi(n) =$ "$n^2 \geq 2n - 1$". With $n = 1$, we have $1 \geq 2 \cdot 1 - 1 = 1$.

If we assume that this holds for n, then $(n+1)^2 = n^2 + 2n + 1 \geq 2n - 1 + 2n + 1 = 4n = 2n + 2n \geq 2n + 1 = 2(n+1) - 1$, therefore, we proved $\varphi(n+1)$ from $\varphi(n)$. We can now conclude that $\varphi(n)$ is true for all $n \in \mathbb{N}$.

Exercise: prove that $n > 0$ for all $n \in \mathbb{N}$.

The well-ordering principle

First we need the following.

Lemma 22. *Let $m, n \in \mathbb{N}$, $m > n$. Then $m - n \in \mathbb{N}$.*

Proof. This is proved by a double induction. Let $\varphi(m, n) =$ "if $m > n$, then $m - n \in \mathbb{N}$". Let us first set $n = 1$, $m = 1$. In this case, $m > n$ is not true, so we do not have to prove anything.

Assume that $\varphi(m, 1)$ is true, that is, if $m > 1$, then $m - 1 \in \mathbb{N}$. To prove $\varphi(m+1, 1)$, assume that $m + 1 > 1$, but $m + 1 - 1 = m \in \mathbb{N}$. By induction, $\varphi(m, 1)$ is true for all $m \in \mathbb{N}$.

Assume that $\varphi(m, n)$ is true for all $m \in \mathbb{N}$. Assume that $m > n + 1$. Then, $m - 1 > n > 0$, and $m - 1 \in \mathbb{N}$ by $\varphi(m, 1)$. Then by $\varphi(m - 1, n)$, $m - (n + 1) = m - 1 - n \in \mathbb{N}$. that is, we proved $\varphi(m, n + 1)$. □

We combine a proof by contradiction and mathematical induction.

Theorem 23. *For any nonempty subset $S \in \mathbb{N}$, there is the smallest element in S. That is, there is $n \in S$ such that $n \leq m$ for all $m \in S$.*

Proof. Let us call **A** the assumption that S is not empty.

Let us assume the contrary, that S does not admit the smallest element (call this assumption **B**). It means that, for any $n \in S$, there is $m \in S$ such that $m < n$.

Let $T = \{n \in \mathbb{N} : m > n, \text{ for all } m \in S\}$. We show that $T = \mathbb{N}$ by induction.

- First, $1 \in T$. To prove this, assume that $1 \notin T$ (call this $\mathbf{C}_1$). Then, there must be $m \in S$ such that $m \leq 1$. This means $1 \in S$. But 1 is always the smallest element of any subset of $\mathbb{N}$, contradicting **B**. Therefore, $\mathbf{C}_1$ is false and we obtain $1 \in T$.
- Next, let $n \in T$ and we prove that $n + 1 \in T$. Assume that $n + 1 \notin T$ (call this $\mathbf{C}_n$). Then, there is $m \in S$ such that $m \leq n + 1$, but since $n \in T$, it must hold that $n < m$. This means that $m = n + 1$ by the previous lemma, and any $\ell \leq n$ does not belong to S. Therefore, $m = n + 1$ would be the smallest element of S, contradicting **B**. Therefore, $\mathbf{C}_n$ is false and we obtain $n + 1 \in T$.

Then by induction (the Peano axioms) we have $T = \mathbb{N}$. This implies that for any $m \in S$ it it holds for $m < n$ for all $n \in T = \mathbb{N}$. But there is no such m (larger than any natural number), hence $S = \emptyset$. This contradicts the assumption **A** of the theorem. Therefore, the assumption **B** made in the proof is wrong. That is, S admits the smallest element. □

Alternatively, this can be proved as follows, but using the axiom of the least upper bound (the proof above uses only the Peano axioms). As $\mathbb{N}$ is bounded below, S is bounded below as well. Let $a = \inf S$. We show that $a \in S$. If not, for $\epsilon = \frac{1}{2}$, there is $n \in S$ such that $n < a + \frac{1}{2}$. But then $0 < n - a < \frac{1}{2}$, which is impossible by the previous lemma.

Corollary 24. *Let $x \in \mathbb{R}, x > 0$. There is $n \in \mathbb{N}$ such that $n - 1 \le x < n$.*

Proof. By the Archimedean principle, there is n such that $x < n$. Therefore, the set $\{m \in \mathbb{N} : x < m\}$ is nonempty, and by the well-ordering principle, it has the smallest element n. As this is the smallest element, $n - 1 \le x$. □

We have used this property before to find the decimal representation of x.

The summation and product notations

Assume that we have a sequence of numbers, that is a family $\{a_n\}_{n\in S}$ of real numbers indexed by $S \subset \mathbb{N}$. This means that we have numbers $a_1, a_2, a_3, \cdots$. Sometimes we start the index from 0, and have $a_0, a_1, a_2, \cdots$.

Example 25.

- $a_1 = 1, a_2 = 2, a_3 = 3, \cdots$
- $a_1 = 1, a_2 = 4, a_3 = 9, \cdots$
- $a_1 = 4, a_2 = 2534, a_3 = \frac{3}{361}$ (a *finite* sequence stops at some $n \in \mathbb{N}$)

When we have a (finite) sequence, we can sum all these numbers up: $a_1 + \cdots + a_n$. We denote this by the following symbol.

$$\sum_{k=1}^{n} a_k = a_1 + \cdots + a_n$$

In this symbol, k is a dummy index and plays no specific role. We have

$$\sum_{k=1}^{n} a_k = \sum_{k=0}^{n-1} a_{k+1} = a_1 + \cdots + a_n.$$

On the other hand, the number on the top (n in this example) is where the sequence stops. Similarly, we can define $\sum_{k=m}^{n} a_k = a_m + a_{m+1} + \cdots + a_n$ for $n \ge m$.

More precisely, this is a recursive definition: We define $\sum_{k=1}^{1} = a_1$ and $\sum_{k=1}^{n+1} a_k = \sum_{k=1}^{n} a_k + a_{n+1}$ Similarly to mathematical induction, we define in this way $\sum_{k=1}^{n} a_k$ for all natural numbers $n \in \mathbb{N}$.

Example 26.

- $a_1 = 1, a_2 = 2, a_3 = 3$. $\sum_{k=1}^{3} a_k = 1 + 2 + 3 = 6$.
- $a_1 = 1, a_2 = 4, a_3 = 9, a_4 = 16$. $\sum_{k=1}^{4} a_k = 1 + 4 + 9 + 16 = 30$.

Let us also introduce a symbol for product.

$$\prod_{k=1}^{n} a_k = a_0 \cdot a_1 \cdot \cdots \cdot a_n$$

Example 27.

- $a_1 = 1, a_2 = 2, a_3 = 3$. $\prod_{k=1}^{3} a_k = 1 \cdot 2 \cdot 3 = 6$.

- $a_1 = 1, a_2 = 4, a_3 = 9, a_4 = 16$. $\prod_{k=1}^{4} a_k = 1 \cdot 4 \cdot 9 \cdot 16 = 576$.

In particular, we denote

- For $a \in \mathbb{R}$, $a^n = \prod_{k=1}^{n} a$. For example, $a^1 = a, a^2 = a \cdot a, a^3 = a \cdot a \cdot a$. By convention, for $a \neq 0$, we set $a^0 = 1$.
- $n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot \cdots \cdot n$. By convention, we set $0! = 1$.

  For example, $2! = 2, 3! = 6, 4! = 24, \cdots$.
- For $n, k \in \mathbb{N}, n \geq k$, we define $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. For example, $\binom{4}{2} = \frac{4!}{2!2!} = 6$.

Some useful formulas

The summation formulas

Proposition 28. *We have the following.*

- $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.
- $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.
- *For* $a \neq 1$, $\sum_{k=1}^{n} a^k = \frac{a(1-a^n)}{1-a}$.

Proof. We prove them by induction.

- $\sum_{k=1}^{1} k = 1 = \frac{1 \cdot 2}{2} = 1$ is correct. Assume the formula $\sum_{k=1}^{n} k = \frac{k(k+1)}{2}$ for n, then

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

  Then by induction the formula holds for all $n \in \mathbb{N}$.

- $\sum_{k=1}^{1} k^2 = 1^2 = \frac{1 \cdot 2 \cdot 3}{6} = 1$ is correct. Assume the formula $\sum_{k=1}^{n} k = \frac{n(n+1)(2n+1)}{6}$ for n, then

$$\begin{aligned}
\sum_{k=1}^{n+1} k^2 &= \sum_{k=1}^{n} k^2 + (n+1)^2 \\
&= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\
&= \left(\frac{n(2n+1)}{6} + (n+1)\right)(n+1) \\
&= \frac{(2n^2 + n + 6n + 6)(n+1)}{6} \\
&= \frac{(2n+3)(n+2)(n+1)}{6} \\
&= \frac{(2(n+1)+1)((n+1)+1)(n+1)}{6}.
\end{aligned}$$

  Then by induction the formula holds for all $n \in \mathbb{N}$.

- $\sum_{k=1}^{1} a^k = a = \frac{a(1-a)}{1-a}$ is correct. Assume the formula $\sum_{k=1}^{n} a^k = \frac{a(1-a^n)}{1-a}$ for n, then

$$\begin{aligned}\sum_{k=1}^{n+1} a^k &= \left(\sum_{k=1}^{n} a^k\right) + a^{n+1} \\ &= \frac{a(1-a^n)}{1-a} + a^{n+1} \\ &= \frac{a - a^{n+1} + a^{n+1} - a^{n+2}}{1-a} \\ &= \frac{a(1-a^{n+1})}{1-a}\end{aligned}$$

Then by induction the formula holds for all $n \in \mathbb{N}$.

□

The binominal theorem

Lemma 29. $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$ *for* $n \geq k$.

Proof. We prove this by induction, but in a slightly different form: we prove that the formula is correct for $n = k$, and prove that it holds for $n+1$ assuming the formula for n. In this way, we prove the formula for $n \geq k$.

If $n = k$, we have $\binom{k+1}{k} = \frac{(k+1)!}{k!(k+1-k)!} = k+1 = \frac{k!}{(k-1)!} + 1 = \binom{k}{k-1} + \binom{k}{k}$.

Assuming the formula for n, we have

$$\begin{aligned}\binom{n+2}{k} &= \frac{(n+2)!}{k!(n+2-k)!} \\ &= \frac{n+2}{n+2-k}\binom{n+1}{k} \\ &= \frac{n+2}{n+2-k}\left(\binom{n}{k-1} + \binom{n}{k}\right) \\ &= \frac{n+2}{n+2-k}\left(\frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!}\right) \\ &= \frac{(n+1)!}{(k-1)!(n+2-k)!} + \frac{n!}{(k-1)!(n+2-k)!} + \frac{n+2}{n+2-k} \cdot \frac{n!}{k!(n-k)!} \\ &= \frac{(n+1)!}{(k-1)!(n+2-k)!} + \frac{n!}{k!(n+1-k)!} \cdot \frac{k + (n+2)(n+1-k)}{n+2-k} \\ &= \frac{(n+1)!}{(k-1)!(n+2-k)!} + \frac{n!}{k!(n+1-k)!} \cdot \frac{k + (n+1)(n+1-k) + n + 1 - k}{n+2-k} \\ &= \frac{(n+1)!}{(k-1)!(n+2-k)!} + \frac{n!}{k!(n+1-k)!} \cdot \frac{(n+1)(n+1-k) + n + 1}{n+2-k} \\ &= \frac{(n+1)!}{(k-1)!(n+2-k)!} + \frac{n!}{k!(n+1-k)!} \cdot \frac{(n+1)(n+2-k)}{n+2-k} \\ &= \frac{(n+1)!}{(k-1)!(n+2-k)!} + \frac{(n+1)!}{k!(n+1-k)!} = \binom{n+1}{k-1} + \binom{n+1}{k}.\end{aligned}$$

□

Theorem 30. *For any* $a, b \in \mathbb{R}, n \in \mathbb{N}$, $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$, *where in this theorem we mean* $0^0 = 1$.

Proof. By induction. For $n = 0$, this holds in the sense of $1 = 1$.

Assume that this holds for n. Then,

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n \cdot (a+b) \\
&= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&= \sum_{k=1}^{n+1}\binom{n}{k-1}a^k b^{n+1-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&= \sum_{k=1}^{n}\left(\binom{n}{k-1}+\binom{n}{k}\right)a^k b^{n+1-k} + a^{n+1}b^0 + a^0 b^{n+1} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^k b^{n+1-k}
\end{aligned}$$

□

For example, we have

- $(x+y)^2 = x^2 + 2xy + y^2$
- $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$
- $(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$

and so on.

Oct. 4. Functions, domains and ranges.

Functions

By a function we usually mean "a map" which assigns to a number x another number $f(x)$, or an assignment

$$x \longmapsto f(x).$$

There are many "real-world" examples of functions: When a quantity changes with time, you can use x as time (or often you denote it by t) and the quantity by $f(x)$. Or we can plot a set of data that depend on a parameter (more concretely: you take a path on a mountain and set x as the horizontal distance from the house and $f(x)$ as the height at the point x).

More precisely, we can consider it as follows: for each number x there is another number $f(x)$, and nothing else. We can express this situation using ordered pairs.

Let us assume that we know the correspondence $x \longmapsto f(x)$, defined on a subset ("domain") S. Then we can draw the graph, namely, the subset $\{(x,y) \in S \times \mathbb{R} : y = f(x)\}$, or in other words, we collect all points (x,y) where $y = f(x)$.

More generally we can define a **function** to be a subset f of $\mathbb{R} \times \mathbb{R}$ such that for each $x \in f$ there is one and only one y. Also in this case we denote the relation by $y = f(x)$. In this sense, the graph and the function are the same thing.

Let us introduce some terminology.

- $\{x \in \mathbb{R} :$ there is some $(x,y) \in f\}$ is called the **domain** of f.

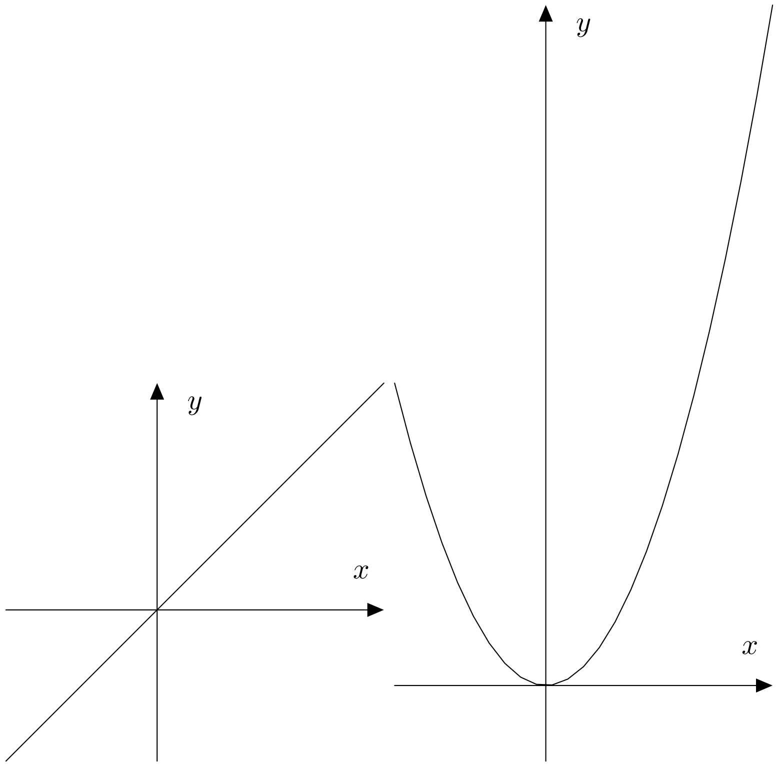


Figure 10: Left: the graph of $y = x$. Right: the graph of $y = x^2$.

- $\{y \in \mathbb{R} :$ there is some $(x, y) \in f\}$ is called the **range** of f.

Example 31.
- $f(x) = x$. Namely, $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x\}$. The domain is $\mathbb{R}$, the range is $\mathbb{R}$.
- $f(x) = x^2$. Namely, $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x^2\}$. The domain is $\mathbb{R}$, the range is $[0, \infty)$.
- $f(x) = x^5 - 2x^3 + 1$. $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x^5 - 2x^3 + 1\}$. The domain is $\mathbb{R}$, the range is $\mathbb{R}$.
- $f(x) = \sqrt{x}$ for $x \geq 0$. Namely, $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \geq 0, y = \sqrt{x}\}$. The domain is $[0, \infty)$, the range is $[0, \infty)$.
- $f(x) = \sqrt{1 - x}$ for $1 - x \geq 0$, or $x \leq 1$. Namely, $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \leq 1, y = \sqrt{1 - x}\}$. The domain is $(-\infty, 1]$, the range is $[0, \infty)$.

The set $\{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}$ is *not* a function. Indeed, for each $x \in (-1, 1)$, there are two numbers $y = \sqrt{1 - x^2}, -\sqrt{1 - x^2}$ that satisfy the equation $x^2 + y^2 = 1$.

Let us introduce the **absolute value** of $x \in \mathbb{R}$:

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

where := means we define the left-hand side by the right-hand side. This is also a function with the domain $\mathbb{R}$ and the range $[0, \infty)$.

We define the **sign** of $x \in \mathbb{R}$:

$$\operatorname{sign} x := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

We define the factorial of $x \in \mathbb{N}_0$: $f(n) = n!$. The domain is $\mathbb{N}_0$.

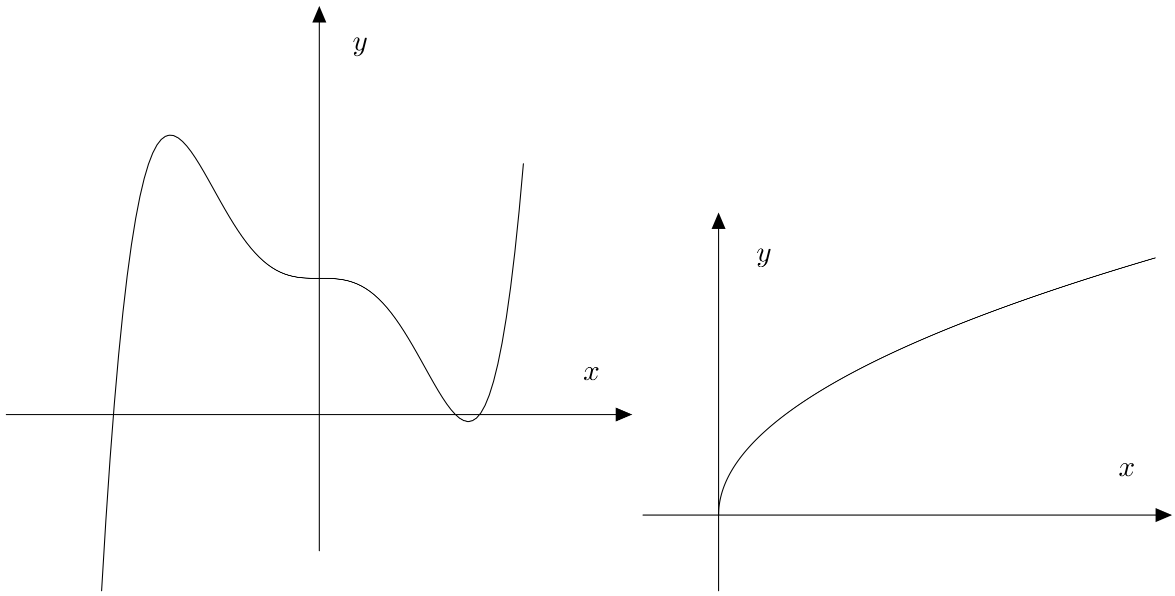


Figure 11: Left: the graph of $y = x^5 - 2x^3 + 1$. Right: the graph of $y = \sqrt{x}$.

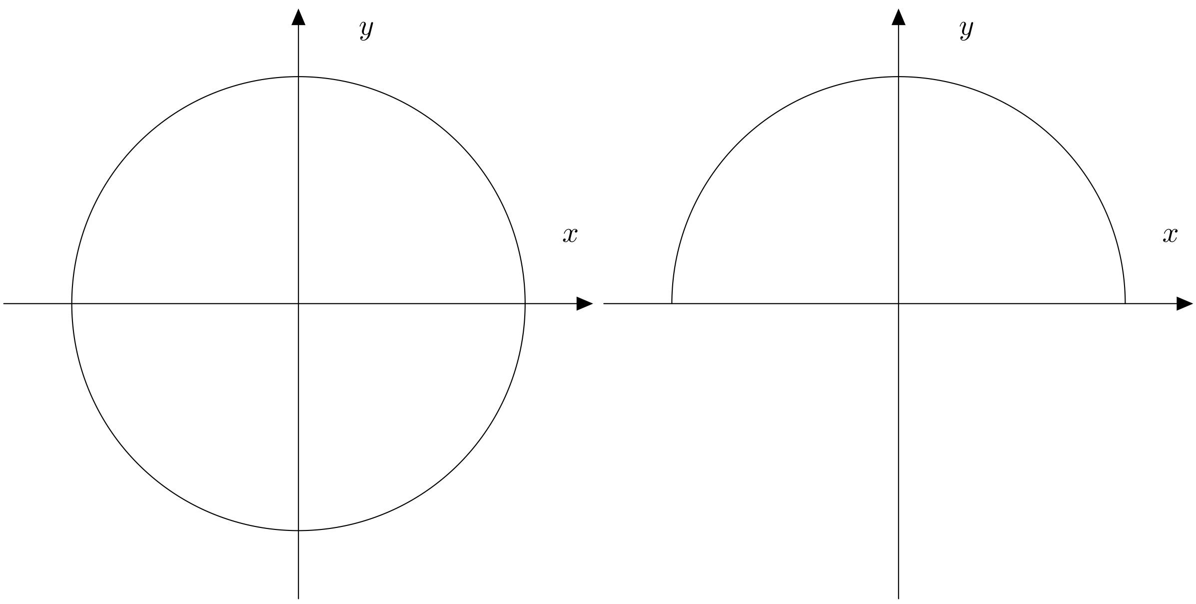


Figure 12: Left: the graph of $x^2 + y^2 = 1$, not a function of x. Right: the graph of $y = \sqrt{1 - x^2}$.

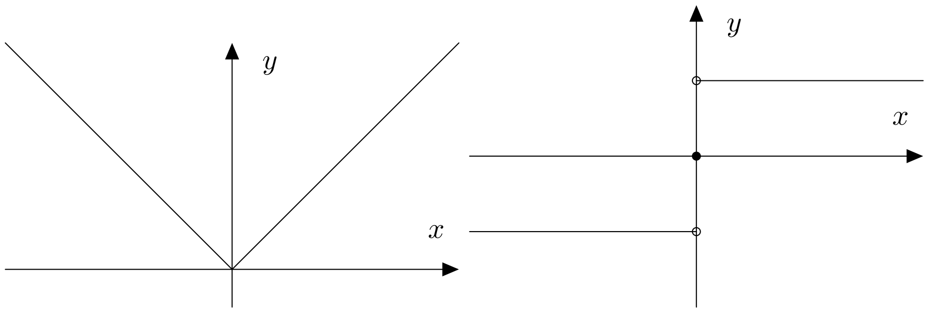


Figure 13: Left: the graph of $y = |x|$. Right: the graph of $y = \operatorname{sign} x$, with a “jump” at $x = 0$.

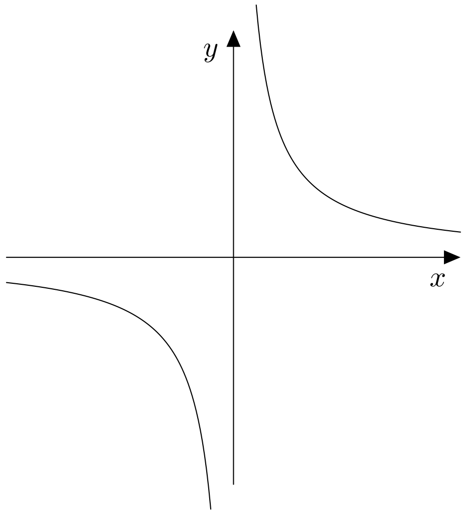


Figure 14: The graphs of $y = \frac{1}{x}$ on $(-\infty, 0) \cup (0, \infty)$.

Operations on functions

When we have two or more functions, we can produce more functions. Let $f(x)$ be a function with domain S and $g(x)$ a function with domain T.

- Sum. We can define the sum $h(x) = f(x) + g(x)$, defined on $S \cap T$.

  Example: with $f(x) = x, g(x) = x^2, h(x) = x + x^2$.

- Product. We can define the product $h(x) = f(x) \cdot g(x)$, defined on $S \cap T$.

  Example: with $f(x) = x, g(x) = x^2, h(x) = x^3$.

- Division. We can define the division $h(x) = \frac{f(x)}{g(x)}$ defined on $S \cap \{x \in T : g(x) \neq 0\}$.

  Example: with $f(x) = x + 1, g(x) = (x + 2)(x - 1), h(x) = \frac{x+1}{(x+2)(x-1)}$, defined on $\mathbb{R} \setminus \{1, -2\} = (-\infty, -2) \cup (-2, 1) \cup (1, \infty)$.

- Composition. We can define the composed function $h(x) = f(g(x))$, defined on $\{x \in T : g(x) \in S\}$.

  Example: with $f(x) = \sqrt{x}, g(x) = x + 1, h_1(x) = \sqrt{x + 1}$, defined on $\{x \in \mathbb{R} : x + 1 \geq 0\}$.

  Note that this is different from the composition in the reversed order: $h_2(x) = g(f(x)) = \sqrt{x} + 1$, defined on $[0, \infty)$.

We say that a function $f(x)$ is **injective** if for any pair $x_1 \neq x_2$ in the domain, it holds that $f(x_1) \neq f(x_2)$. Similarly, we say that a function $f(x)$ is **surjective** if the range is $\mathbb{R}$. A function which is both injective and surjective is said to be **bijective**.

For example, $f(x) = x$ is injective and surjective (hence bijective), but $f(x) = x^2$ is neither injective nor surjective. But if we consider $f(x) = x^2$ with the restricted domain $[0, \infty)$, it is injective: for positive numbers $x_1 \neq x_2$, $x_1^2 \neq x_2^2$.

For an injective function $f(x)$, we can define the **inverse function** f^{-1}: the domain of f^{-1} is the range R of f, and it assigns to $f(x)$ the number x: it is characterized by $f^{-1}(f(x)) = x$. Its graph (its formal definition) is given by $\{(x, y) \in \mathbb{R} \times \mathbb{R} : x \in R, x = f(y)\}$. The range of f^{-1} is the domain of f.

For example, consider $f(x) = x^2$ on the domain $[0, \infty)$. The range of f is $[0, \infty)$, hence the domain of f^{-1} is $[0, \infty)$. For any $x \in [0, \infty)$, we should have $f^{-1}(f(x)) = f^{-1}(x^2) = x$, therefore, $f^{-1}(x) = \sqrt{x}$.

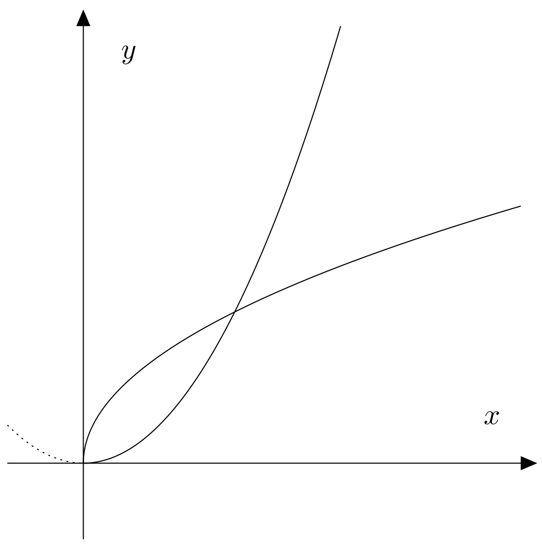


Figure 15: The graphs of $y = \sqrt{x}$ and $y = x^2$ on $[0, \infty)$.

The triangle inequality

Lemma 32. *Let $x, a \in \mathbb{R}, a \geq 0$. Then $|x| \leq a$ if and only if $-a \leq x \leq a$.*

Proof. Assume that $x \geq 0$.

- If $|x| = x \leq a$, then $-a < 0 \leq x \leq a$.
- If $-a \leq x \leq a$, then $|x| = x \leq a$.

Instead, if we assume that $x < 0$, then

- If $|x| = -x \leq a$, then $-a \leq x < 0 \leq a$.
- If $-a \leq x \leq a$, then $|x| = -x \leq a$.

□

Theorem 33. *For any $x, y \in \mathbb{R}$, it holds that $|x + y| \leq |x| + |y|$.*

Proof. We have $-|x| \leq x \leq |x|, -|y| \leq y \leq |y|$ by Lemma, therefore, $-|x| - |y| \leq x + y \leq |x| + |y|$, and again by Lemma this implies that $|x + y| \leq |x| + |y|$. □

Corollary 34. *For any $x_1, x_2, \cdots, x_n \in \mathbb{R}$, it holds that $|\sum_{k=1}^n a_k| \leq \sum_{k=1}^n |a_k|$.*

Proof. By induction. For $n = 1$, $\left|\sum_{k=1}^1 a_1\right| = |a_1| = \sum_{k=1}^1 |a_k|$ is obvious. Assuming the inequality for n, we have

$$\begin{aligned}
\left|\sum_{k=1}^{n+1} a_k\right| &= \left|\sum_{k=1}^{n} a_k + a_{n+1}\right| \\
&\leq \left|\sum_{k=1}^{n} a_k\right| + |a_{n+1}| && \text{by Theorem} \\
&\leq \sum_{k=1}^{n} |a_k| + |a_{n+1}| && \text{by induction hypothesis} \\
&= \sum_{k=1}^{n+1} |a_k|
\end{aligned}$$

which concludes the induction. □

Oct 06. Sequence and convergence of sequence.

Convergence of sequences

We saw sequences of real numbers $a_1, a_2, \cdots$. A sequence can be infinite, that is, it continues infinitely. For example,

- $a_1 = 1, a_2 = 2$ and in general, $a_n = n$.
- $a_1 = 1, a_2 = 4$ and in general, $a_n = n^2$.

A sequence can be considered as a function with the domain $\mathbb{N}$.

Among sequences, we have seen the following:

- $a_1 = 1, a_2 = \frac{1}{2}$ and $a_n = \frac{1}{n}$.
- $a_1 = \frac{1}{2}, a_2 = \frac{3}{4}$ and $a_n = 1 - \frac{1}{2^n}$.

Intuitively, the first of them gets closer and closer to 0, while the second one gets closer and closer to 1. But what does it mean that it gets closer to a number?

We make precise the notion that a sequence get "arbitrarily" close to a number as follows.

Definition 35. Let $\{a_n\}$ be a sequence of real numbers. If there is $L \in \mathbb{R}$ such that for each $\epsilon > 0$ there is N_ϵ such that for $n \geq N_\epsilon$ it holds that $|a_n - L| < \epsilon$, we say that $\{a_n\}$ **converges** to L.

We write this situation as $\lim_{n\to\infty} a_n = L$, or simply $a_n \to L$.

Example 36. Let us see some convergent sequences.

- $a_1 = 1, a_2 = \frac{1}{2}$ and $a_n = \frac{1}{n}$. We expect that this sequence converges to 0. Indeed, for any $\epsilon > 0$, there is N_ϵ such that $\frac{1}{N_\epsilon} < \epsilon$ (the Archimedean property). Furthermore, if $n > N_\epsilon$, then $|\frac{1}{n} - 0| = \frac{1}{n} < \frac{1}{N_\epsilon} < \epsilon$, therefore, with $L = 0$, we have that $\{a_n\}$ converges to 0.
- $a_1 = \frac{1}{2}, a_2 = \frac{3}{4}$ and $a_n = 1 - \frac{1}{2^n}$. We expect that this sequence converges to 1. Indeed, for any $\epsilon > 0$, there is N_ϵ such that $\frac{1}{N_\epsilon} < \epsilon$ and note that $\frac{1}{2^{N_\epsilon}} < \frac{1}{N_\epsilon}$. Furthermore, if $n > N_\epsilon$, then $\frac{1}{2^n} < \frac{1}{N_\epsilon}$ and hence $|1 - \frac{1}{2^n} - 1| = \frac{1}{2^n} < \frac{1}{N_\epsilon} < \epsilon$, therefore, with $L = 1$, we have that $\{a_n\}$ converges to 1.
- The sequence $a_n = \frac{1}{\sqrt{n}}$ converges to 0. Indeed, for each ϵ, there is N_ϵ such that $\frac{1}{N_\epsilon} < \epsilon$, and hence if $n > N_\epsilon^2$, then $\frac{1}{\sqrt{n}} < \frac{1}{\sqrt{N_\epsilon^2}} = \frac{1}{N_\epsilon} < \epsilon$.

Note that

- If $\{a_n\}$ converges to L, then it does not converge to any other number. Indeed, if $x \neq L$, then take N such that $|a_n - L| < \frac{1}{2}|L - x|$ for $n > N$. Then by the triangle inequality $|L - x| < |a_n - x| + |a_n - L|$, and hence $|a_n - x| > |x - L| - |a_n - L| > \frac{1}{2}|L - x| \neq 0$. Therefore, $\{a_n\}$ does not converge to x.
- The sequence $a_1 = 1, a_2 = 0, a_3 = 1, \cdots, a_n = \frac{1}{2}(1 - (-1)^n)$ does not converge to any number.
- The sequence $a_1 = 1, a_2, \cdots, a_n = n$ does not converge to any number.
- In general, if for any x there is an $N_x \in \mathbb{N}$ such that for $n > N_x$ it holds that $|a_n| > x$, then we say that $\{a_n\}$ **diverges**.
- The sequence $a_n = 2^n$ diverges.

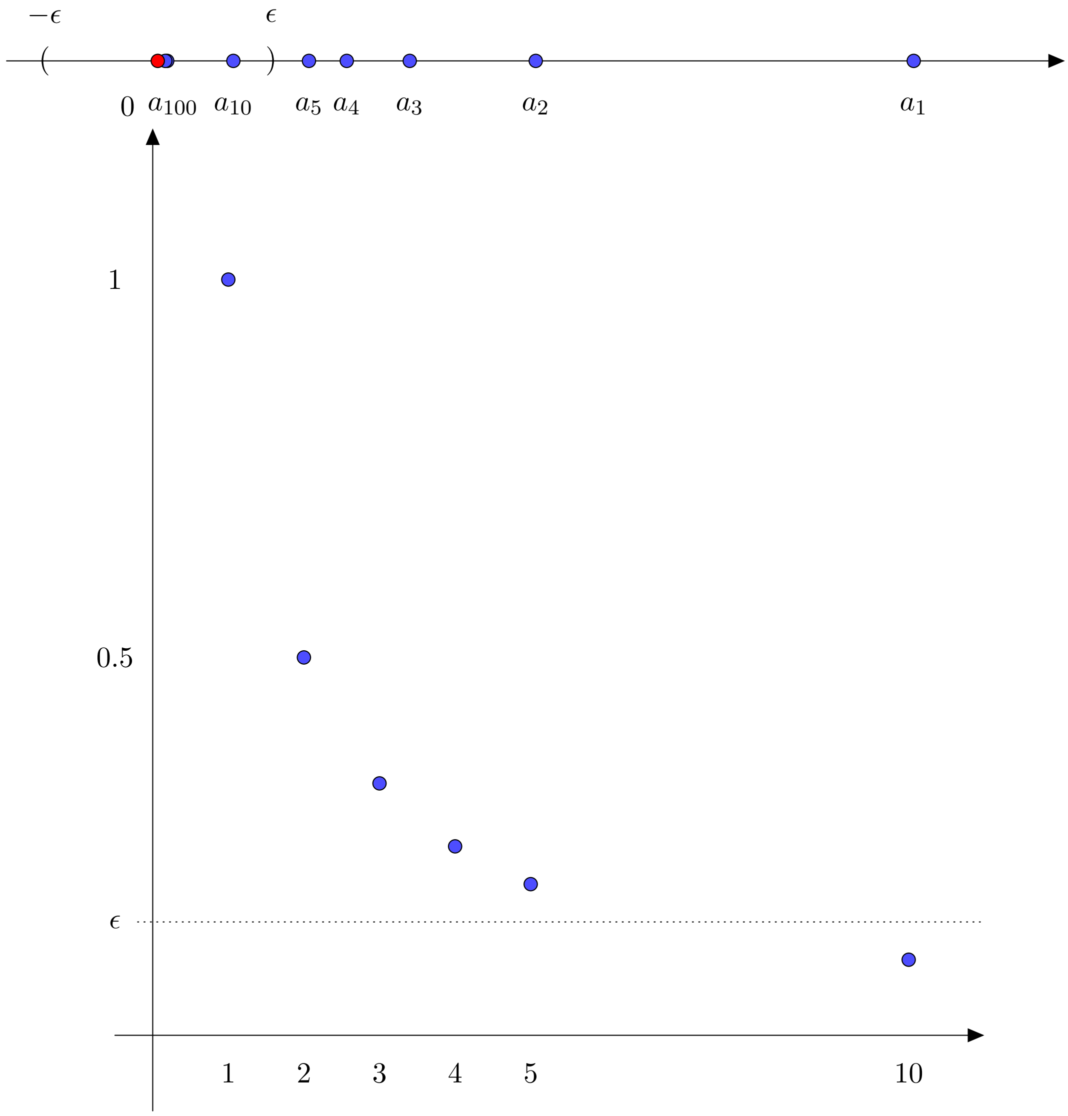


Figure 16: Up: the sequence $a_n = \frac{1}{n}$ plotted on the line. Bottom: the sequene $a_n = \frac{1}{n}$ as a function on $\mathbb{N}$.

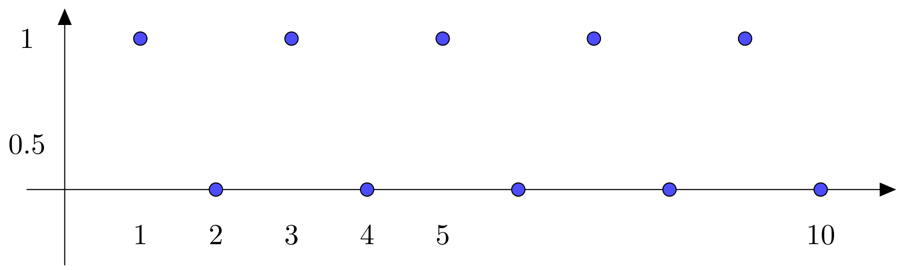


Figure 17: The sequence $a_n = \frac{1}{2}(1 - (-1)^n)$ as a function on $\mathbb{N}$.

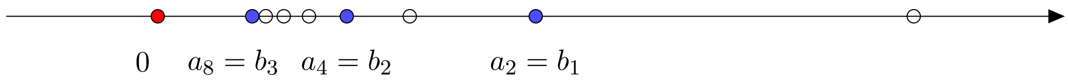


Figure 18: The subsequence a_{2^n} of the sequence $a_n = \frac{1}{n}$.

Some properties of convergent sequences

Given a sequence $\{a_n\}$, one can take a **subsequence** of it. That is, we take an increasing sequence of natural numbers $m_1 < m_2 < m_3 < \cdots$ and define a new sequence $b_n = a_{m_n}$.

Example 37. Given $a_n = \frac{1}{n}$ and $m_n = 2^n$, the subsequence is $a_{2^n} = \frac{1}{2^n}$.

If $\{a_n\}$ is convergent to L, then any subsequence $\{a_{m_n}\}$ is convergent to L. Indeed, as $m_1 < m_2 < m_3 \cdots$, we have $n \le m_n$ and hence, for any $\epsilon > 0$, we take N such that $|a_n - L| < \epsilon$ for $n > N$, hence all $n > N$, $|a_{m_n} - L| < \epsilon$.

We say that $\{a_n\}$ is **nondecreasing** (respectively **nonincreasing**) if $a_n \le a_{n+1}$ (respectively $a_n \ge a_{n+1}$) holds for all $n \in \mathbb{N}$. A sequence $\{a_n\}$ is said to be **bounded above** (respectively **bounded below**) if there is $M \in \mathbb{R}$ such that $a_n \le M$ (respectively $a_n \ge M$) for all $n \in \mathbb{N}$).

Lemma 38. *Let $\{a_n\}$ be a nondecreasing sequence and bounded above. Then a_n converges to a certain real number $L \in \mathbb{R}$.*

Proof. Let $A = \{a_n : n \in \mathbb{N}\}$. As $\{a_n\}$ is bounded above, A is bounded above. We put $L = \sup A$. By Lemma 10, for each $\epsilon > 0$, there is $N \in \mathbb{N}$ such that $L - \epsilon < a_N$. As a_n is nondecreasing, we have $L - \epsilon < a_n$ for all $n > N$. On the other hand, we have $a_n \le L$ because $L = \sup A$. Altogether, $|a_n - L| < \epsilon$ for such n. As n was arbitrary, a_n converges to L. □

Note that $|ab| = |a||b|$.

Theorem 39. *The following hold.*

- *If $a_n \to L$, then there is $\tilde{L}$ such that $|a_n| < \tilde{L}$ for all n.*
- *If $a_n \to L, b_n \to M$, then $a_n + b_n \to L + M, a_n \cdot b_n \to LM$. If $M \neq 0$, then $b_n \neq 0$ for sufficiently large n and $\frac{a_n}{b_n} \to \frac{L}{M}$.*
- *If $a_n > 0$ diverges, then $\frac{1}{a_n}$ converges to 0.*

Proof.
- Assume that $a_n \to L$. Given, say 1, there is N such that $|a_n - L| < 1$ for $n > N$, hence $|a_n| < L+1$ for $n > N$. Then, we can take a number $\tilde{L}$ such that $|a_1|, \cdots, |a_{N-1}| < \tilde{L}$ and $L + 1 < \tilde{L}$.
- Let $\epsilon > 0$ be arbitrary. There are $N_1, N_2 \in \mathbb{N}$ such that for $n > N_1$ (respectively $n > N_2$) it holds that $|a_n - L| < \frac{\epsilon}{2}$ (respectively $|b_n - M| < \frac{\epsilon}{2}$). Let N be the largest of N_1, N_2. Then we have
$$|a_n + b_n - L - M| \le |a_n - L| + |b_n - M| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$
hence $a_n + b_n$ converges to $L + M$.
As for the product, given $\epsilon > 0$, we take N such that $|a_n - L| < \frac{\epsilon}{2(|M|+1)}, |b_n - M| < \frac{\epsilon}{2(|L|+1)}$ and $|b_n| < |M| + 1$ for $n > N$ (this can be done as in the case of sum). Then
$$|a_n b_n - LM| = |a_n b_n - b_n L + b_n L - LM| \le |(a_n - L)b_n)| + |(b_n - M)L|$$
$$\le |a_n - L||b_n| + |b_n - M||L| \le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

which shows the desired convergence.

We prove $\frac{1}{b_n} \to \frac{1}{M}$. If $b_n \to M$ and $M \neq 0$, then $|b_n - M| < \frac{|M|}{2}$ for sufficiently large n, and hence $|b_n| > \frac{|M|}{2}$, in particular $b_n \neq 0$. We can now show that $\frac{1}{b_n} \to \frac{1}{M}$. Indeed, by taking N such that $|b_n - M| < \frac{\epsilon M^2}{2}$

$$\left|\frac{1}{b_n} - \frac{1}{M}\right| = \frac{|M - b_n|}{|M||b_n|} < \frac{\frac{\epsilon M^2}{2}}{\frac{M^2}{2}} = \epsilon,$$

which shows $\frac{1}{b_n} \to \frac{1}{M}$. Now $\frac{a_n}{b_n} \to \frac{L}{M}$ follows from this and the product with a_n.

- For any $\epsilon > 0$, there is N such that for $n > N$ it holds that $|a_n| > \frac{1}{\epsilon}$, that is $\frac{1}{a_n} < \epsilon$, hence $\frac{1}{a_n}$ converges to 0.

□

We denote $a^{-n} = \frac{1}{a^n}$.

Proposition 40. *The following hold.*

- *Let $a > 1$. Then a^n diverges.*
- *Let $0 < a < 1$. Then a^n converges to 0.*
- *Let $0 < a < 1$. Then $b_n = \sum_{k=1}^n a^k$ converges to $\frac{a}{1-a}$.*

Proof.

- If $a > 1$, we can write $a = 1 + y$ where $y > 0$. By the binomial theorem, we have

$$a^n = (1+y)^n = \sum_{k=0}^{n} \binom{n}{k} 1^k y^{n-k} > 1 + ny,$$

by only taking the terms $k = n, n-1$. Now it is clear that for any x there is large enough n such that $1 + ny > x$, therefore, $x < 1 + ny < a^n$, that is, a^n diverges.

- If $0 < a < 1$, then $\frac{1}{a} > 1$ and $(\frac{1}{a})^n$ diverges. Therefore, $a^n = (\frac{1}{a})^{-n}$ converges to 0.
- We know that $b_n = \sum_{k=1}^n a^k = \frac{a(1-a^n)}{1-a}$, and $a^n \to 0$.

□

Oct. 07. Continuity of functions.

Decimal representation of real numbers

Now that we have defined convergence of sequences, we can make sense of all decimal representations as real numbers.

Theorem 41. *Let $a_n \in \mathbb{N}_0$ and $0 \leq a_n \leq 9$. Then $b_n = \sum_{k=0}^n a_k 10^{-k}$ converges to a real number.*

Proof. Let $b_n = \sum_{k=0}^n a_k 10^{-k}$. This is nondecreasing and bounded above by $a_0 + 1$. By Lemma 38, this converges to a real number. □

When the sequence converges, it converges to only one number. In this way, we can say that a decimal representation $a_0.a_1a_2a_3\cdots$ defines a real number.

Now we can prove that any repeating decimal representation gives a rational number. For example consider $0.123123123\cdots$. This can be written as

$$0.1 + 0.02 + 0.003 + 0.0001 + 0.00002 + 0.000003 + \cdots = \sum_{k=0}^{n} a_k 10^{-k},$$

where $a_1 = 1, a_2 = 2, a_3 = 3, a_4 = 1, a_5 = 2, a_6 = 3, \cdots$. It is easy to see that this is equal to $0.123 + 0.000123 + \cdots = \sum_{k=1}^{n}(100a_{3k+1} + 10a_{3k+2} + a_{3k+3})1000^{-k}$. We know that this sum converges and compute

$$\begin{aligned}\sum_{k=1}^{n}(100a_{3k+1} + 10a_{3k+2} + a_{3k+3})1000^{-k} &= 123\sum_{k=1}^{n}1000^{-k} \\ \to 123\frac{1000^{-1}}{1-1000^{-1}} &= \frac{123}{999}.\end{aligned}$$

Proposition 42. *Any real number given by a repeating decimal representation is rational.*

Proof. Indeed, let us take a repeating sequence $0 \le a_n \le 9$ as above. That is, there is $m \in \mathbb{N}$ such that $a_{n+m} = a_m$.

Then, for $j, \ell \in \mathbb{N}$,

$$\begin{aligned}\sum_{k=0}^{j\ell} a_k &= a_0 + \sum_{j=1}^{\ell} 10^{-jm}\sum_{k=1}^{m} a_k 10^{m-k} \\ &= a_0 + (\sum_{k=1}^{m} a_k 10^{m-k})\frac{10^{-m}(1-10^{-j\ell})}{1-10^{-m}} \\ &\to a_0 + (\sum_{k=1}^{m} a_k 10^{m-k})\frac{10^{-m}}{1-10^{-m}} = a_0 + (\sum_{k=1}^{m} a_k 10^{m-k})\frac{1}{10^m - 1}\end{aligned}$$

as $\ell \to \infty$. The last expression is evidently rational. □

Theorem 43. *For any real number a there is a sequence a_n of rational numbers such that $a_n \to a$.*

Proof. Take the decimal representation of a, truncate it to the n-th digit, and call it a_n. Then $\{a_n\}$ are rational and $a_n \to a$. □

Continuity of functions

Let us go back to studying functions. Among functions, we saw the sign function

$$\operatorname{sign} x := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

and its graph (Figure 13) has a "jump" at $x = 0$.

Intuitively, the "jump" means that, the value at $x = 0$ is 0, but if one approaches to 0 from the right, the value of the function remains 1, while it is -1 from the left.

Let us make this precise.

Definition 44. Let f be a function defined on S (the domain), and let $a \in \mathbb{R}$ such that there is a sequence $x_n \in S, x_n \neq a$ such that $x_n \to a$. We write

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

if for any $\epsilon > 0$ there is $\delta > 0$ such that $|f(x) - L| < \epsilon$ for any $x \neq a, |x - a| < \delta$.

Example 45. Let $f(x) = \begin{cases} 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$.

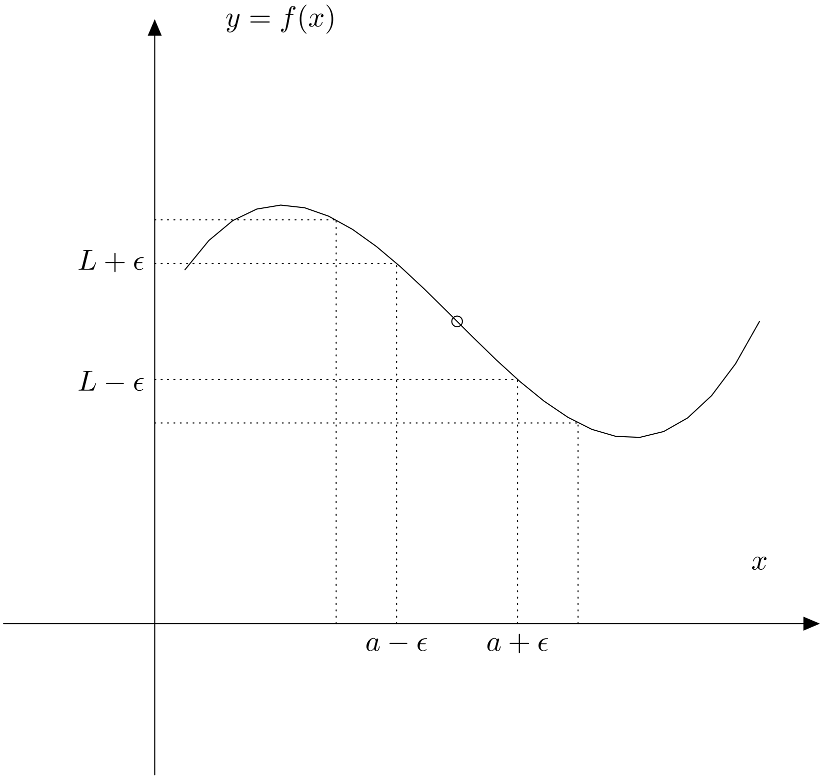


Figure 19: The limit $\lim_{x \to a} f(x)$.

- Consider $a = 2$. Then, for any ϵ, we can take $\delta = \frac{1}{2}$ and $|f(x) - 1| = |1 - 1| = 0$ for any $x \in (2 - \delta, 2 + \delta) = (\frac{3}{2}, \frac{5}{2})$. Therefore, $\lim_{x \to 2} f(x) = 1$. A similar situation holds for any $x \neq 0$.
- Consider $a = 0$. Then, for any $x \neq 0$, $f(x) = 1$, hence again we have $\lim_{x \to 0} f(x) = 1$, although $f(0) = 0$ by definition.
- For the function $\operatorname{sign} x$ (Figure 13), there is no limit $\lim_{x \to 0} f(x)$ at $x = 0$.

The limit makes precise the concept of "approaching a point". The absence of "jump" can also be formalized using limit.

Definition 46. Let f be a function defined on S (the domain), and let $a \in S$ (this time a is in the domain) such that there is a sequence $x_n \in S, x_n \neq a$ such that $x_n \to a$. We say that f is continuous at a if $\lim_{x \to a} f(x) = f(a)$. We say that f is continuous on S if it is continuous at each point in S.

Now we can understand the "jumps" in terms of limit and continuity.

Example 47.
- The function $\operatorname{sign} x$ is not continuous at $x = 0$, because it does not have $\lim_{x \to 0} \operatorname{sign} x$.
- The function $f(x) = \begin{cases} 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ is not continuous at $x = 0$, because $\lim_{x \to 0} f(x) = 1 \neq 0 = f(0)$.
- The function $f(x) = c$ is continuous. Indeed, let us fix $a \in \mathbb{R}$. For any ϵ, $|f(x) - c| = |c - c| = 0 < \epsilon$, hence $\lim_{x \to a} f(x) = c = f(a)$.

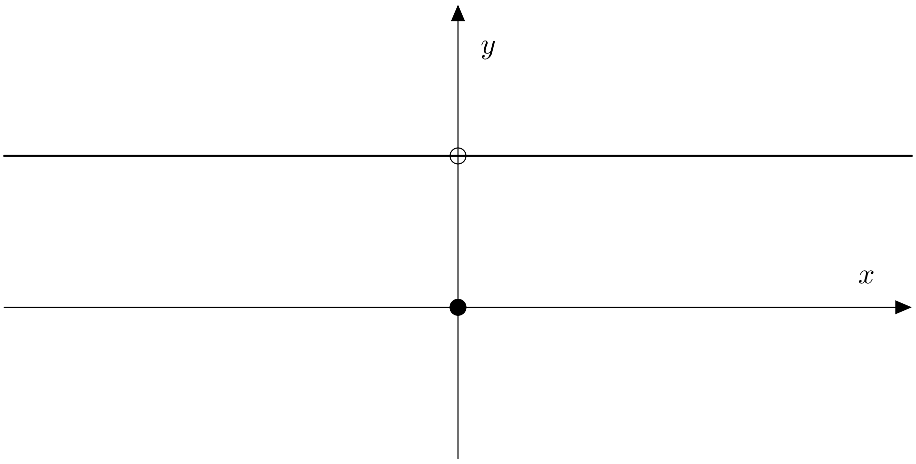


Figure 20: The graph of $y = \begin{cases} 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$.

- The function $f(x) = x$ is continuous. Indeed, let us fix $a \in \mathbb{R}$. Then, for each $\epsilon > 0$, we take $\delta = \epsilon$ and for $|h| < \delta = \epsilon$ it holds that $|f(a+h) - a| = |a + h - a| = |h| < \delta = \epsilon$, therefore, $\lim_{x\to a} f(x) = a = f(a)$.

Theorem 48. *Let f, g be functions defined on S, and let a such that there is $\{x_n\} \subset S$, $x_n \to a$. Assume that $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Then*

- *There is $\delta > 0$ such that if $|x - a| < \delta, x \neq a$ then $|g(x)| \leq |M| + 1$.*
- *$\lim_{x\to a}(f(x) + g(x)) = L + M$ and $\lim_{x\to a}(f(x)g(x)) = LM$.*
- *Assume that $M \neq 0$, then there is $\delta > 0$ such that, if $|x - a| < \delta$, then $|g(x)| > \frac{|M|}{2}$ for x such that $|x - a| < \delta, x \neq a$.*
- *If $L \neq 0$, then $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{L}{M}$.*

Furthermore, if $a \in S$ and if f, g are continuous at a, then $f + g, fg$ are continuous at a. If $g(a) \neq 0$, then $\frac{f}{g}$ is continuous at a.

Proof. The proof is similar to that of Theorem 39.

- Let $\delta > 0$ such that $|g(x) - M| < 1$ for x such that $|x - a| < \delta, x \neq a$. Then $|g(x)| < |M| + 1$.
- For a given $\epsilon > 0$, let $\delta > 0$ such that $|f(x) - L| < \frac{\epsilon}{2}, |g(x) - M| < \frac{\epsilon}{2}$ for $|x - a| < \delta, x \neq a$. Then $|f(x) + g(x) - L - M| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, which shows the desired limit.

  For the product, for a given $\epsilon > 0$, let $\delta > 0$ such that $|f(x) - L| < \frac{\epsilon}{2(|M|+1)}, |g(x) - M| < \frac{\epsilon}{2(|L|+1)}$ and $|g(x)| < |M| + 1$ for $|x - a| < \delta, x \neq a$. Then $|f(x)g(x) - LM| = |f(x) - L||g(x)| + |g(x) - M||L| < \frac{\epsilon(|M|+1)}{2(|M|+1)} + \frac{\epsilon|L|}{2(|L|+1)} < \epsilon$, which shows the desired limit.
- Let $\delta > 0$ such that $|g(x) - M| < \frac{|M|}{2}$ for x such that $|x - a| < \delta, x \neq a$. Then, by the triangle inequality, $|g(x)| \geq |M| - |g(x) - M| \leq \frac{|M|}{2}$.
- We show that $\lim_{x\to a} \frac{1}{g(x)} = \frac{1}{M}$. Then the general case follows from this and the limit of product. Assume $M \neq 0$, and let $\epsilon > 0$. Then there is $\delta > 0$ such that $|g(x) - M| < \frac{|M|}{2}$ for

$x \neq a, |x-a| < \delta$ and hence $|g(x)| > \frac{|M|}{2}$, in particular $g(x) \neq 0$. Now, there is $\tilde{\delta} > 0, \tilde{\delta} < \delta$ such that for $x \neq a, |x-a| < \tilde{\delta}$ it holds that $|g(x) - M| < \frac{\epsilon M^2}{2}$. Then

$$\left|\frac{1}{g(x)} - \frac{1}{M}\right| = \frac{|M - g(x)|}{|M||g(x)|} < \frac{\frac{\epsilon M^2}{2}}{\frac{M^2}{2}} = \epsilon,$$

which shows the desired limit.

If f, g are continuous, then $\lim_{x\to a} f(x) = f(a), \lim_{x\to a} g(x) = g(a)$, hence $\lim_{x\to a}(f(x)+g(x)) = f(a) + g(a), \lim_{x\to a} f(x)g(x) = f(a)g(a), \lim_{x\to a} \frac{f(x)}{g(x)} = \frac{f(a)}{g(a)}$. □

From this, we know that

- If $f(x) = a_0 + a_1x^1 + \cdots a_nx^n$ (a polynomial), then f is continuous. $f(x) = x^2, f(x) = x^5 + 34x^3 - 454$...
- If $f(x) = \frac{P(x)}{Q(x)}$ and $P(x), Q(x)$ are polynomial, then f is continuous at x if $Q(x) \neq 0$. $f(x) = \frac{x-2}{x^2}$ is continuous on $x \neq 0$ (actually defined on $\{x \in \mathbb{R} : x \neq 0\}$, $f(x) = \frac{x^3}{x^2-1} = \frac{x^3}{(x-1)(x+1)}$ is continuous on $x \neq -1, 1$.

Oct. 11. Properties of continuous functions.

Sequences and continuity of functions

We can use sequences to study functions, especially regarding continuity. Let f be a function defined on a certain domain S and $\{x_n\}$ a sequence in S. Then we can construct a new sequence by $\{f(x_n)\}$.

Theorem 49. *Let f be a function defined on S. f is continuous at $a \in S$, that is, $\lim_{x\to a} f(x) = f(a)$ if and only if it holds that $f(x_n) \to f(a)$ for all sequences $\{x_n\}$ in S such that $x_n \to a, x_n \neq a$.*

Proof. Assume that $\lim_{x\to a} f(x) = f(a)$. Then, for each $\epsilon > 0$, there is $\delta > 0$ such that if $|x - a| < \delta$, then it holds that $|f(x) - f(a)| < \epsilon$. Let us take any sequence $\{x_n\}$ such that $x_n \to a$. This means that, for δ above, there is N such that $|x_n - a| < \delta$ for $n > N$. Then by the observation above, we have $|f(x_n) - f(a)| < \epsilon$. This shows that, for $n > N$, we have $|f(x_n) - f(a)| < \epsilon$. Therefore, for the given ϵ we found N such that $|f(x_n) - f(a)| < \epsilon$ for $n > N$. This means that $f(x_n) \to f(a)$.

Conversely, assume that $f(x_n) \to f(a)$ for all sequences $\{x_n\}$ such that $x_n \to a, x_n \neq a$. To do a proof by contradiction, let us assue that there is $\epsilon > 0$ for which for all δ there is $x \in S, x \neq a$ such that $|x - a| < \delta$ but $|f(x) - f(a)| > \epsilon$. Let us take $\delta_n = \frac{1}{n}$. For each δ_n there is $x_n \in S$ such that $|x_n - a| < \frac{1}{n}, x \neq a$ but $|f(x_n) - f(a)| > \epsilon$. Then, it is clear that $x_n \to a$, but $f(x_n)$ is not converging to $f(a)$, which contradicts the assumption. Therefore, it must hold that $\lim_{x\to a} f(x) = f(a)$.

□

Lemma 50. *Let $x_n \to x$ and $x_n \leq a$. Then $x \leq a$.*

Proof. Assume the contrary, that is, $x > a$. Then there is N such that $|x_n - x| < \frac{x-a}{2}$, and $x_n - a = x_n - x + x - a > |x - a| - \frac{|x-a|}{2} = \frac{|x-a|}{2}$, which contradicts $x_n \leq a$. Therefore, $x \leq a$. □

It also holds that, if $x_n \to x$ and $x_n \geq a$, then $x \geq a$.

We can show that, if $A \subset [a, b]$, then $\sup A \in [a, b]$: indeed, we can take a sequence $\{x_n\} \subset A$ such that $x_n \to \sup A$ by Lemma 10.

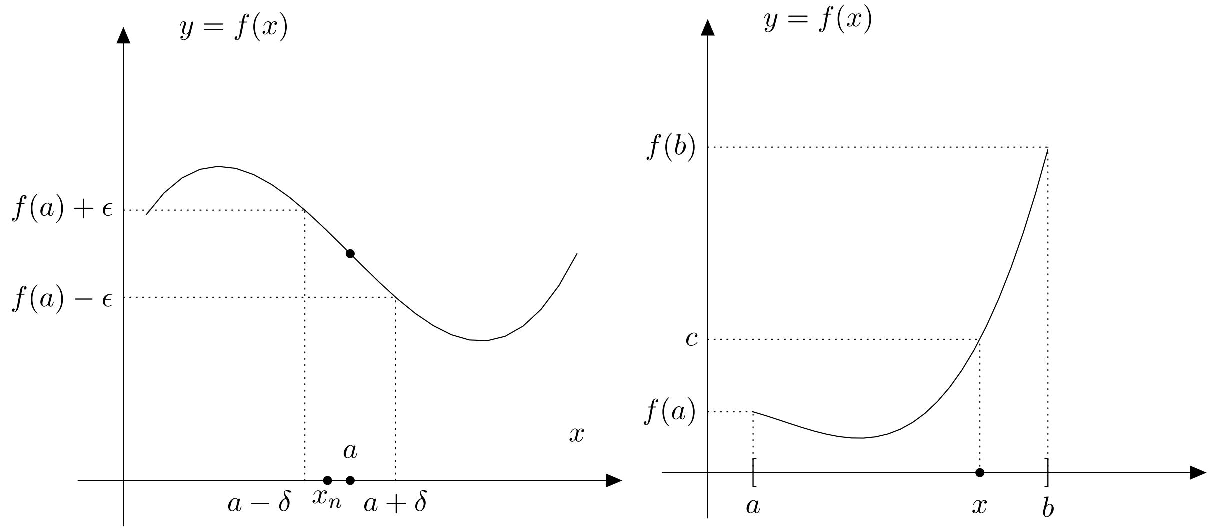


Figure 21: Left: continuity of f and a sequence $x_n \to a$. Right: the intermediate value c is taken at x.

Theorem 51 (the intermediate value theorem)**.** *Let f be a continuous function on a closed interval $[a, b]$. Assume that $f(a) < f(b)$. Then, for any value $c \in (f(a), f(b))$, there is $x \in (a, b)$ such that $c = f(x)$.*

Proof. Let $c \in (f(a), f(b))$, and we define $A = \{x \in [a, b] : f(x) < c\}$. A is bounded above, because it is contained in $[a, b]$, therefore, we can take $x = \sup A$. By Lemma 10, for each n, there is $x_n \in A$ such that $x - \frac{1}{n} < x_n$, hence $x_n \to x$. Since f is continuous, we have $f(x) = \lim_{n\to\infty} f(x_n)$. On the other hand, $x_n \in A$, hence $f(x_n) < c$ and hence $f(x) \leq c$ by Lemma 50.

$x \neq b$ because $f(b) > c \geq f(x)$. Therefore, we can take a sequence $x_n > x, x_n \to x$ in the interval $[x, b]$, and then $f(x_n) \geq c$ because $x_n \notin A$. By continuity of f, we have $f(x) = \lim_n f(x_n) \geq c$. Altogether, we have $f(x) = c$. □

Composition and inverse functions

Let f, g be two functions, f defined on S and g defined on the image (range) of f: $f(S) = \{y \in \mathbb{R} : \text{ there is } x \in S, y = f(x)\}$. Recall that we can compose two functions: for $x \in S$, $g(f(x))$ gives a number, hence the correspondence $x \to g(f(x))$ is a function on S. We denote this **composed function** by $g \circ f$.

Theorem 52. *In the situation above, if f and g are continuous, then $g \circ f$ is continuous as well.*

Proof. We take $a \in S$. Given $\epsilon > 0$, there is $\delta_1 > 0$ such that $|g(y) - g(f(a))| < \epsilon$ if $|y - f(a)| < \delta_1$, by continuity of g. For this δ_1, there is $\delta_2 > 0$ such that $|f(x) - f(a)| < \delta_1$ if $|x - a| < \delta_2$. Altogether, we have $|g(f(x)) - g(f(a))| < \epsilon$ if $|x - a| < \delta_2$, hence we have proved the continuity of $g \circ f$. □

Definition 53. Let f be a function on S. We say that f is **monotonically increasing (nondecreasing, decreasing, nonincreasing, respectively)** if for each $x_1, x_2 \in S, x_1 < x_2$ it holds that $f(x_1) < f(x_2)$ $(f(x_1) \leq f(x_2), f(x_1) > f(x_2), f(x_1) \geq f(x_2)$, respectively).

Example 54. (Non)examples of monotonic functions.

- $f(x) = x$ is monotonically increasing.

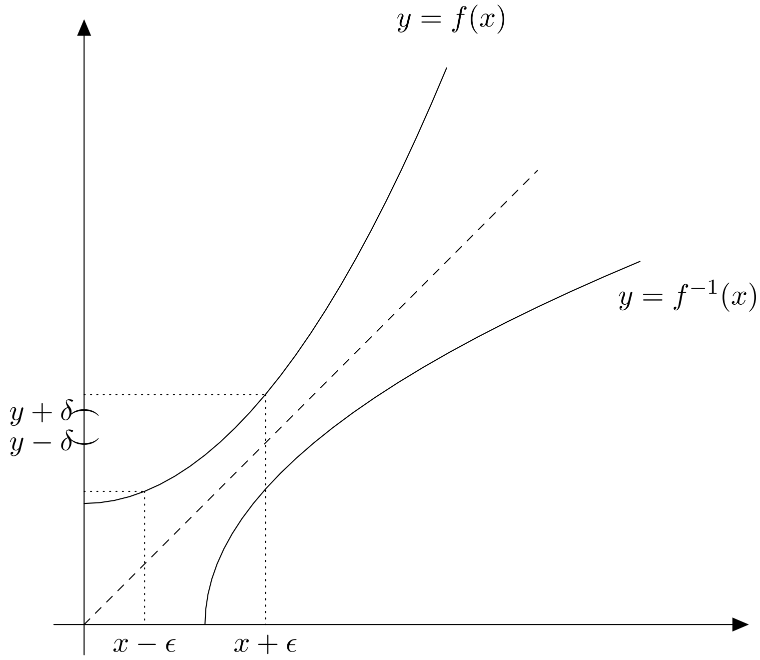


Figure 22: The continuity of the inverse function. For a given ϵ, we can take δ.

- $f(x) = x^2$ is not monotonically increasing on $\mathbb{R}$, but it is so on $\mathbb{R}_+$.
- $f(x) = \operatorname{sign} x$ is monotonically nondecreasing.

If a function f is monotonically increasing (or decreasing), it is injective: for $x_1 \neq x_2$, it holds that $f(x_1) \neq f(x_2)$. Therefore, we can consider its inverse function.

Theorem 55. *Let f be a monotonically increasing continuous function on an interval $[a, b]$. Then the inverse function f^{-1} defined on $[f(a), f(b)]$ is monotonically increasing and continuous.*

Proof. Note that the domain of f^{-1} is $[f(a), f(b)]$ by the intermediate value theorem (continuity of f is needed here).

Let us first show that f^{-1} is monotonically increasing. For each $y_1 < y_2, y_1, y_2 \in [f(a), f(b)]$, there are $x_1, x_2 \in S$ such that $y_1 = f(x_1)$ and $y_2 = f(x_2)$ by the intermediate value theorem and we have $x_1 < x_2$ by monotonicity of f. This means that $f^{-1}(y_1) = x_1 < x_2 = f^{-1}(y_2)$, that is, f^{-1} is monotonically increasing.

Let $x_0 \in (a, b)$. For a given $\epsilon > 0$, we take δ as the smaller of $f(x_0 + \epsilon)$ and $f(x_0 - \epsilon)$ (if $x_0 \pm \epsilon$ are not in S, replace them by a or b). Then for any $y \in (f(x_0) - \delta, f(x_0) + \delta)$, we have $f^{-1}(y) \in S \cap (x_0 - \epsilon, x_0 + \epsilon)$ by monotonicity of f. This is the continuity of f^{-1}.

If $x_0 = a$ or b, then we only have to consider one side. □

Roots and power functions

Let us consider $f(x) = x^n$ defined on $\mathbb{R}_+ \cup \{0\}$. This is monotonically increasing (because, if $x_1 < x_2$, then $x_2^n = (x_1 + (x_2 - x_1))^n > x^n$ by the binomial theorem). Therefore, we can define the inverse function $f^{-1}(x)$ and denote it by $x^{\frac{1}{n}}$. This shows that, for any $x \in \mathbb{R}_+ \cup \{0\}$, there is one and only one y such that $y^n = x$. The function $g(x) = x^{\frac{1}{n}}$ is monotonically increasing and continuous by Theorem 55.

Let $p, q \in \mathbb{N}$, $x \geq 0$. Note that we have $(x^p)^q = x^{pq} = (x^q)^p$. Then it is easy to see that $(x^p)^{\frac{1}{q}} = (x^{\frac{1}{q}})^p$: if $y = (x^p)^{\frac{1}{q}}$, then $y^q = x^p = ((x^{\frac{1}{q}})^q)^p = (x^{\frac{1}{q}})^{pq}$ and hence $y = (x^{\frac{1}{q}})^p$.

Furthermore, let $m \in \mathbb{N}$. Then for $y = (x^{mp})^{\frac{1}{mq}}$ we have $(y^q)^m = y^{mq} = x^{mp} = (x^p)^m$, and hence $y^q = x^p$ and $y = (x^p)^{\frac{1}{q}}$.

Therefore, we can write $y = x^{\frac{p}{q}}$ and no confusion arises.

Oct. 13. Exponential functions.

For $a > 0$ and $p, q \in \mathbb{N}$, we have defined $a^{\frac{p}{q}}$. Then the natural question arises whether a^x can be defined for real numbers x.

For a fixed $a > 0$, we can consider $f(x) = a^x$ as a function defined on the set of rational numbers $\mathbb{Q}$.

Lemma 56. *We have the following.*

- *For $p, q, r, s \in \mathbb{N}$, we have $a^{\frac{p}{q}} a^{\frac{r}{s}} = a^{\frac{p}{q}+\frac{r}{s}}$.*
- *For $p, q, r, s \in \mathbb{N}$, we have $(a^{\frac{p}{q}})^{\frac{r}{s}} = a^{\frac{pr}{qs}}$.*
- *If $a > 1$, then $f(x) = a^x$ is monotonically increasing (as a function on $\mathbb{Q}$).*
- *If $0 < a < 1$, then $f(x) = a^x$ is monotonically decreasing.*

Proof.

- Recall that we have $a^{\frac{p}{q}} = a^{\frac{ps}{qs}}$ and $a^{\frac{r}{s}} = a^{\frac{qr}{qs}}$, and hence

$$a^{\frac{p}{q}} a^{\frac{r}{s}} = a^{\frac{ps}{qs}} a^{\frac{qr}{qs}} = (a^{\frac{1}{qs}})^{ps} (a^{\frac{1}{qs}})^{qr} = (a^{\frac{1}{qs}})^{ps+qr} = a^{\frac{ps+qr}{qs}} = a^{\frac{p}{q}+\frac{r}{s}}.$$

- We will prove this as an exercise.
- Let us take $a > 1$. First, for any $q \in \mathbb{N}$, $a^{\frac{1}{q}} > 1$, indeed, if $a^{\frac{1}{q}} \le 1$, we would have $a = (a^{\frac{1}{q}})^q \le 1$, contradiction.

  If $x_1, x_2 \in \mathbb{Q}$ and $x_1 < x_2$, we may assume that $x_1 = \frac{p}{q}, x_2 = \frac{r}{q}$ and $p < r$. Then

$$a^{x_1} = a^{\frac{p}{q}} = (a^{\frac{1}{q}})^p < (a^{\frac{1}{q}})^r = a^{\frac{r}{q}} = a^{x_2}.$$

- The case $0 < a < 1$ is similar.

□

We would like to define a^x by $\lim_{n\to\infty} a^{x_n}$, where $x_n \in \mathbb{Q}$ and $x_n \to x \in \mathbb{R}$. For this purpose, we need some properties of sequences.

Lemma 57. *If $a_n \le b_n$ and $a_n \to L, b_n \to M$, then $L \le M$.*

Proof. Consider $b_n - a_n \ge 0$. By Lemma 50, $b_n - a_n \to M - L \ge 0$, hence $M \ge L$. □

We write $a_n \to \infty$ if for any $x \in \mathbb{R}$ there is N such that $a_n > x$ for $n > N$.

Theorem 58. *Let $a_n \le b_n \le c_n$ be three sequences. If $a_n \to L$ and $c_n \to L$, then $b_n \to L$. Similarly, if $a_n \to \infty$, then also $b_n \to \infty$.*

Proof. For a given $\epsilon > 0$, we take N such that for $n > N$ it holds that $|a_n - L| < \epsilon$ and $|c_n - L| < \epsilon$. For a fixed $n > N$, this means that $L - \epsilon < a_n \le b_n \le c_n < L + \epsilon$, and hence $|b_n - L| < \epsilon$. This means that $b_n \to L$.

If $a_n \to \infty$, then for a given x there is N such that $x < a_n \le b_n$, hence $b_n \to \infty$. □

For a statement like "there is N such that for $n > N$..." we say simply that "for sufficiently large n...".

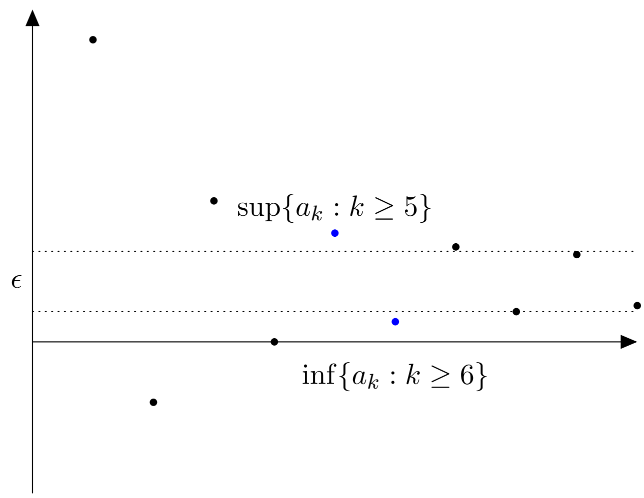


Figure 23: A Cauchy sequence.

Proposition 59. *We have the following.*

- *For $a > 1, p \in \mathbb{N}$, we have $\frac{a^n}{n^p}$ diverges.*
- *It holds that $n^{\frac{1}{n}} \to 1$.*
- *For $a > 1$, we have $a^{\frac{1}{n}} \to 1$.*

Proof.
- Let us consider first $p = 1$. Then, writing $a = 1 + y$ with $y > 0$, we have, for $n \geq 2$,

$$a^n = (1+y)^n = \sum_{k=0}^{n} \binom{n}{k} 1^k y^{n-k} > 1 + \frac{n(n-1)}{2} y^2,$$

and hence $\frac{a^n}{n} > \frac{(n-1)y^2}{2}$. As $\frac{(n-1)y^2}{2} \to \infty$, so does it hold $\frac{a^n}{n} \to \infty$.

For a general $p \in \mathbb{N}$, we take $a^{\frac{1}{p}}$, then $1 < a^{\frac{1}{p}}$ and $\frac{a^{\frac{n}{p}}}{n} \to \infty$, hence $\frac{a^n}{n^p} = \left(\frac{a^{\frac{n}{p}}}{n}\right)^p \to \infty$.

- Let $\epsilon > 0$. We need prove that $n^{\frac{1}{n}} < 1 + \epsilon$ for sufficiently large n. Equivalently, $n < (1+\epsilon)^n$. This follows from the previous point that $\frac{(1+\epsilon)^n}{n} \to \infty$, in particular, $\frac{(1+\epsilon)^n}{n} > 1$ for sufficiently large n.
- $1 < a^{\frac{1}{n}} < n^{\frac{1}{n}}$ for $a < n$, therefore the claim follows from Theorem 58.

□

Definition 60. A sequence a_n is said to be a **Cauchy sequence** if for each $\epsilon > 0$ there is $N \in \mathbb{N}$ such that $|a_m - a_n| < \epsilon$ for $m, n > N$.

Differently from the convergence to a number L, this says that two elements in the sequence are close to each other for large enough m, n.

Lemma 61. *A sequence a_n is convergent if and only if it is a Cauchy sequence.*

Proof. If $a_n \to L$, then for any $\epsilon > 0$ we can take N such that $|a_n - L| < \frac{\epsilon}{2}$ for $n > N$, therefore, if $n, m > N$, then $|a_m - L| < \frac{\epsilon}{2}$ as well and hence $|a_m - a_n| \leq |a_m - L| + |L - a_n| < \epsilon$.

Conversely, if a_n is Cauchy, then it is bounded. Indeed, we take N such that $|a_m - a_{N+1}| < 1$, then this means that $|a_m| < |a_{N+1}| + 1$. Then we can take the largest number among $|a_1|, \cdots, |a_N|, |a_{N+1}| + 1$ as a bound. Next, we consider the sequence

$$b_n = \inf\{a_k : k \geq n\}.$$

This is well-defined because $\{a_k : k \geq n\}$ is bounded. Furthermore, this sequence is increasing because $\{a_k : k \geq n+1\} \supset \{a_k : \kappa \geq n\}$. Therefore, b_n converges to some number L. Similarly, with $c_n = \sup\{a_k : k \geq n\}$, this is bounded and decreasing, hence converges to M.

Note that $b_n \leq a_n \leq c_n$, therefore, $L \leq M$. Actually, we have $L = M$. Indeed, for given $\epsilon > 0$, we can find sufficiently large ℓ, m, n such that $|c_n - M| < \frac{\epsilon}{5}, |a_\ell - c_n| < \frac{\epsilon}{5}, |b_n - L| < \frac{\epsilon}{5}, |a_m - b_n| < \frac{\epsilon}{5}$ and $|a_\ell - a_m| < \frac{\epsilon}{5}$. This implies that $|M - L| < \epsilon$ for arbitrary $\epsilon > 0$, hence it must hold $M = L$.

Now, as $b_n, c_n \to L = M$ and $b_n \leq a_n \leq c_n$, we have $a_n \to L$ by Theorem 58. □

Finally, we can define a^x for all real number x.

Proposition 62. *Let $a > 0, x_n \in \mathbb{Q}, x_n \to x$. Then a^{x_n} converges. If $y_n \in \mathbb{Q}, y_n \to x$, then $\lim_{n\to\infty} a^{x_n} = \lim_{n\to\infty} a^{y_n}$.*

Proof. Note that $\{x_n\}$ is bounded, hence $\{a^{x_n}\}$ is bounded as well, say by M, because the exponential function on $\mathbb{Q}$ is monotonic. We show that a^{x_n} is convergent. To see this, it is enought to see that a^{x_n} is Cauchy by Lemma 61.

For a given $\epsilon > 0$, we take $\delta > 0$ such that $|a^z - 1| < \frac{\epsilon}{M}$ for $0 < z < \delta$. For sufficiently large m, n, we have $|x_m - x_n| < \delta$ and in that case,

$$|a^{x_m} - a^{x_n}| = |a^{x_m}||1 - a^{x_n - x_m}| \leq M|1 - a^{x_n - x_m}| < M\frac{\epsilon}{M} = \epsilon.$$

This means that $\{a^{x_n}\}$ is a Cauchy sequence, and hence it converges to a certain real number, which we call a^x.

If $\{y_n\}$ is another sequence converging to x, then we can consider a further new sequence $x_1, y_1, x_2, y_2, \cdots$, and this converges to some number. But the subsequence $\{x_n\}$ converges to a^x, and hence the whole sequence and hence $\{y_n\}$ must converge to a^x as well. □

As we said in the proof, for an arbitrary real number $x \in \mathbb{R}$, we define the **exponential function** by

$$a^x := \lim_{n\to\infty} a^{x_n}, \text{ where } x_n \in \mathbb{Q}, x_n \to x.$$

The exponential functions appear in various natural phenomena. It happens typically when we consider a collection of objects that increase or decrease independently (such as colonies of bacteria, radioactive nuclei, and so on). See Figure 24.

Oct. 14. Logarithm.

Some properties of exponential functions

Proposition 63. *We have the following.*

- *For $a > 1$, $f(x) = a^x$ is monotonically increasing and continuous.*
- *$a^x a^y = a^{x+y}$.*
- *$(a^x)^y = a^{xy}$.*

Proof.
- Let $x < y$. Then we take sequences $x_n \to x, y_n \to y$, where $x_n, y_n \in \mathbb{Q}$. Then for sufficiently large n we have $x_n < z_1 < z_2 < y_n$ where $z_1, z_2 \in \mathbb{Q}$, and therefore, $a^x \leq a^{z_1} < a^{z_2} \leq a^y$.

  As for continuity, let us take $x, x_n \in \mathbb{R}$ and $x_n \to x$. Then there is $y_n \in \mathbb{Q}$ such that $|a^{x_n} - a^{y_n}| < \frac{1}{n}$ and $|x_n - y_n| < \frac{1}{n}$. Then $y_n \to x$ as well, hence $a^{y_n} \to a^x$, while $a^{y_n} - \frac{1}{n} < a^{x_n} < a^{y_n} + \frac{1}{n}$, therefore, $a^{x_n} \to a^x$.
- Take sequences $x_n \to x, y_n \to y, x_n, y_n \in \mathbb{Q}$. We have $a^{x_n} a^{y_n} = a^{x_n + y_n}$, and $x_n + y_n \to x + y$, therefore, $a^x a^y = a^{x+y}$.

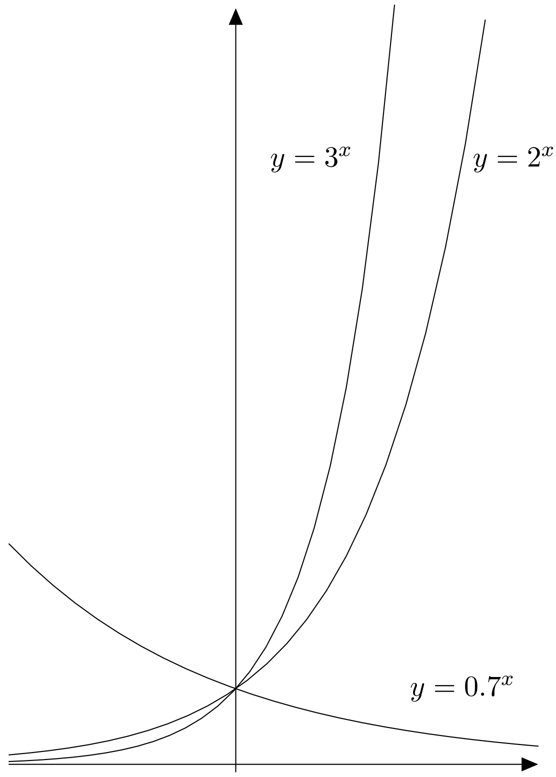


Figure 24: The exponential functions.

- Take sequences $x_n \to x, y_n \to y$, $x_n, y_n \in \mathbb{Q}$. For fixed m, we have $(a^{x_n})^{y_m} \to (a^x)^{y_m}$ and this is equal to $a^{x_n y_m} \to a^{x y_m}$. Now we take the limit $m \to \infty$ and obtain $(a^x)^y = a^{xy}$ by continuity.

□

Napier's number

Let us introduce Napier's number. We take

$$e_n = \left(1 + \frac{1}{n}\right)^n, \qquad E_n = \left(1 + \frac{1}{n}\right)^{n+1} = \left(1 + \frac{1}{n}\right) e_n$$

Lemma 64. *For $x \geq -1$, we have $(1+x)^n \geq 1 + nx$ for all n.*

Proof. By induction. With $n = 0, 1$, we have $(1+x)^0 = 1 = 1$ and $1 + x = 1 + x$. Assuming that this holds for n, we expand

$$\begin{aligned}(1+x)^{n+2} &= (1+x)^n(1+x)^2 \geq (1+nx)(1+x)^2 \\ &= 1 + nx + 2x + x^2 + 2nx^2 + nx^3 = 1 + (n+2)x + x^2(1 + 2n + nx) \geq 1 + (n+2)x\end{aligned}$$

because $1 + 2n + nx \geq 0$. This completes the induction for even and odd numbers.

□

Theorem 65. *e_n and E_n converge to the same number*[2] *e.*

Proof. The proof of this theorem requires several steps.

- We have $1 < e_n < E_n$. Indeed, $1 < 1 + \frac{1}{n}$, and this follows easily.

[2]This proof is take from L. Chierchia "Corso di analisi. Prima parte." McGrow Hill.

- e_n is monotinically increasing, that is, $e_n < e_{n+1}$. Indeed,

$$\begin{aligned}\frac{e_n}{e_{n-1}} &= \frac{(1+\frac{1}{n})^n}{(1+\frac{1}{n-1})^{n-1}} = \frac{(1+\frac{1}{n})^n}{(\frac{n}{n-1})^{n-1}} = (1+\tfrac{1}{n})^n \cdot (\tfrac{n-1}{n})^{n-1} \\ &= \frac{(1+\frac{1}{n})^n \cdot (\frac{n-1}{n})^n}{\frac{n-1}{n}} = \frac{(1+\frac{1}{n})^n \cdot (1-\frac{1}{n})^n}{\frac{n-1}{n}} \\ &= \frac{(1-\frac{1}{n^2})^n}{\frac{n-1}{n}} \geq \frac{1-\frac{1}{n}}{1-\frac{1}{n}} = 1.\end{aligned}$$

  Similarly E_n is monotonically decreasing. Indeed.

$$\begin{aligned}\frac{E_n}{E_{n-1}} &= \frac{(1+\frac{1}{n})^{n+1}}{(1+\frac{1}{n-1})^n} = \frac{1+\frac{1}{n}}{(\frac{n}{n-1})^n(\frac{n}{n+1})^n} = \frac{1+\frac{1}{n}}{(\frac{n^2}{n^2-1})^n} \\ &= \frac{1+\frac{1}{n}}{(1+\frac{1}{n^2-1})^n} \leq \frac{1+\frac{1}{n}}{1+\frac{n}{n^2-1}} < \frac{1+\frac{1}{n}}{1+\frac{1}{n}} = 1.\end{aligned}$$

- Now we have that $\{e_n\}$ and $\{E_n\}$ are convergent. Note also that $E_n - e_n = e_n(1+\frac{1}{n}-1) = e_n \cdot \frac{1}{n} \to 0$, because e_n is bounded, say by M, and $\frac{1}{n} \to 0$, therefore, $E_n - e_n \leq \frac{M}{n} \to 0$.

□

We call this limit e, the **Napier's number** (sometimes **Euler's number**). The function e^x plays a special role in analysis, as we will see below.

Logarithm

Let $a > 0, a \neq 1$. We have defined the exponential function $f(x) = a^x$, and we have seen that it is continuous, monotonically increasing if $a > 1$. If $0 < a < 1$, it is monotonically decreasing.

Let $a > 1$. We know that a^n diverges, and hence $a^{-n} \to 0$. By the intermediate value theorem, we see that the range of a^x is $\mathbb{R}_+$. Now we can define the inverse function (everything is analogous for $0 < a < 1$).

Definition 66. The **logarithm base** a **of** x $\log_a x$ is the inverse function $f(y) = a^y$: $\log_a : \mathbb{R}_+ \longmapsto \mathbb{R}$ and it holds that

$$\log_a a^x = x = a^{\log_a x}.$$

We denote $\log x = \log_e x = \ln x$.

Example 67. $\log_2 8 = 3, \log_9 3 = \frac{1}{2}$.

We say that $\lim_{x\to\infty} f(x) = \infty$ if for each $Y > 0$ there is $X > 0$ such that if $x > X$ then $f(x) > Y$. Similarly, we define $\lim_{x\to\pm\infty} f(x) = \pm\infty$.

Proposition 68. *Let $a, b > 0, a \neq 1 \neq b$, $x, y > 0, t \in \mathbb{R}$. Then*

(i) $\log_a a = 1, \log_a 1 = 0$.

(ii) $\log_a(xy) = \log_a x + \log_a y$.

(iii) $\log_a(x^t) = t \log_a x$.

(iv) $\log_{a^{-1}} x = -\log_a x$.

(v) $\log_a x = \log_a b \cdot \log_b x$.

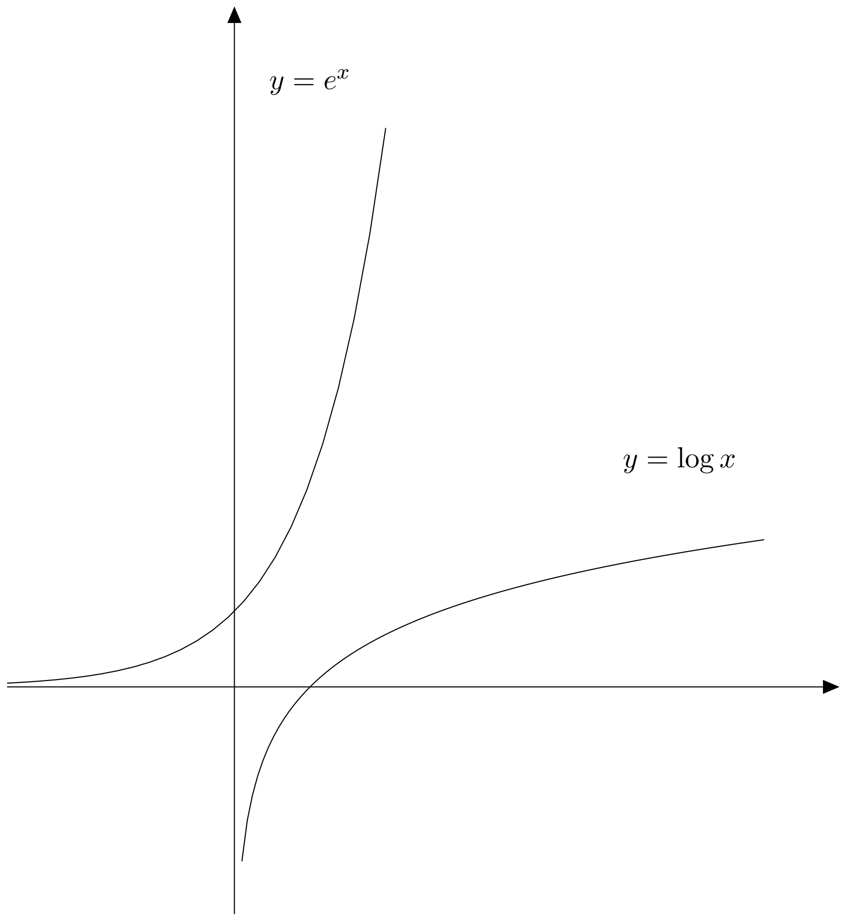


Figure 25: The logarithm and the exponential.

(vi) Let $a > 1$. Then $f(x) = \log_a x$ is monotonically increasing and continuous. $\log_a x > 0$ if and only if $x > 1$.

(vii) Let $a > 1, \alpha > 0$. Then $\lim_{x\to+\infty} \frac{x^\alpha}{\log_a x} = +\infty$.

Proof. (i) $a^1 = a, a^0 = 1$.

(ii) $a^{\log_a x + \log_a y} = a^{\log_a x} a^{\log_a y} = xy$.

(iii) $a^{t\log_a x} = (a^{\log_a x})^t = x^t$.

(iv) $(1/a)^{-\log_a x} = 1/a^{-\log_a x} = 1/(a^{\log_a x})^{-1} = 1/x^{-1} = x$.

(v) $a^{\log_a b \cdot \log_b x} = (a^{\log_a b})^{\log_b x} = b^{\log_b x} = x$.

(vi) This follows from Theorem 55.

(vii) First we show that $\lim_{n\to\infty} \frac{(a^{n-1})^\alpha}{\log_a(a^n)} = \infty$. This is straightforward because $\frac{(a^{n-1})^\alpha}{\log_a(a^n)} = \frac{a^{(n-1)\alpha}}{n} \to \infty$. To show the given limit, we take for $y > 0$ $n \in \mathbb{N}$ such that $n - 1 \le y < n$. Then $\frac{(a^y)^\alpha}{\log_a a^y} > \frac{(a^{n-1})^\alpha}{\log_a(a^n)}$, and hence the left-hand side grows as y grows. That is, $\lim_{y\to\infty} \frac{(a^y)^\alpha}{\log_a a^y} = \infty$. Finally, recall that $x = a^y$ is monotonic, and x grows infinitely as y grows. That is, given $Z > 0$, there is $Y > 0$ such that $\frac{(a^y)^\alpha}{\log_a a^y} > Z$ for $y > Y$, which implies that $\frac{x^\alpha}{\log_a x} > Z$ for $x > a^Y$. This means that $\lim_{x\to\infty} \frac{x^\alpha}{\log_a x} = \infty$.

□

Logarithm is extremely useful in natural science. When we have a data which grows exponentially, we can take the log of the value and plot it to a plane, then they lie on a straight line. The exponent can be read from the slope of the line (this is called the logarithmic scale). In that case, the logarithm base 10 is often used.

When $y = x^p$, then we can consider $z = \log y, w = \log x$, hence $e^z = y, e^w = y$. We have $e^z = y = x^p = (e^w)^p = e^{wp}$. By taking log of both side, we obtain $z = pw$. That is, by the log-log plot, a power relation $y = x^p$ is translated into a linear relation $z = pw$.

Oct. 18. Notable limits, hyperbolic functions.

Some notable limits

Proposition 69. *Let $a \in \mathbb{R}$. The function $f(x) = x^a$ defined on $\mathbb{R}_+$ satisfies $x^a y^a = (xy)^a$ and is continuous.*

Proof. Note that these properties hold if $f(x) = x^q$, where q is rational.

Let $x, y > 0$. For a rational q we have $(xy)^q = x^q y^q$ and hence by taking $q_n \to a$ we have $(xy)^a = x^a y^a$. As for continity, assume that $x \ne y$, then take $a < q \in \mathbb{Q}$. We have $|f(y) - f(x)| = x^a|y^a x^{-a} - 1| = x^a \left|\left(\frac{y}{x}\right)^a - 1\right| < x^a \left|\left(\frac{y}{x}\right)^q - 1\right|$ and $\lim_{y\to x} \left|\left(\frac{y}{x}\right)^q - 1\right| = 1$ by the continuity of the rational case. Therefore, by squeezing we have $\lim_{y\to x} f(y) = f(x)$. □

Let $L \in \mathbb{R}$, and f is a function defined on (a, ∞). If for each $\epsilon > 0$, there is X such that $|f(x) - L| < \epsilon$ for $x > X$, then we write that $\lim_{x\to\infty} f(x) = L$.

Example 70. $\lim_{x\to\infty} \frac{1}{x} = 0$. $\lim_{x\to\infty} \frac{x}{x-1} = 1$.

Let $f(x)$ defined on (a, b) and $L \in \mathbb{R}$. If for each $\epsilon > 0$ there is δ such that $|f(x) - L| < \epsilon$ for $x \in (a, a + \delta)$, we denote it by $\lim_{x\to a^+} f(x)$, and we call it the **right limit** of f at a. Similarly, we write $\lim_{x\to b^-} f(x)$ for the **left limit**.

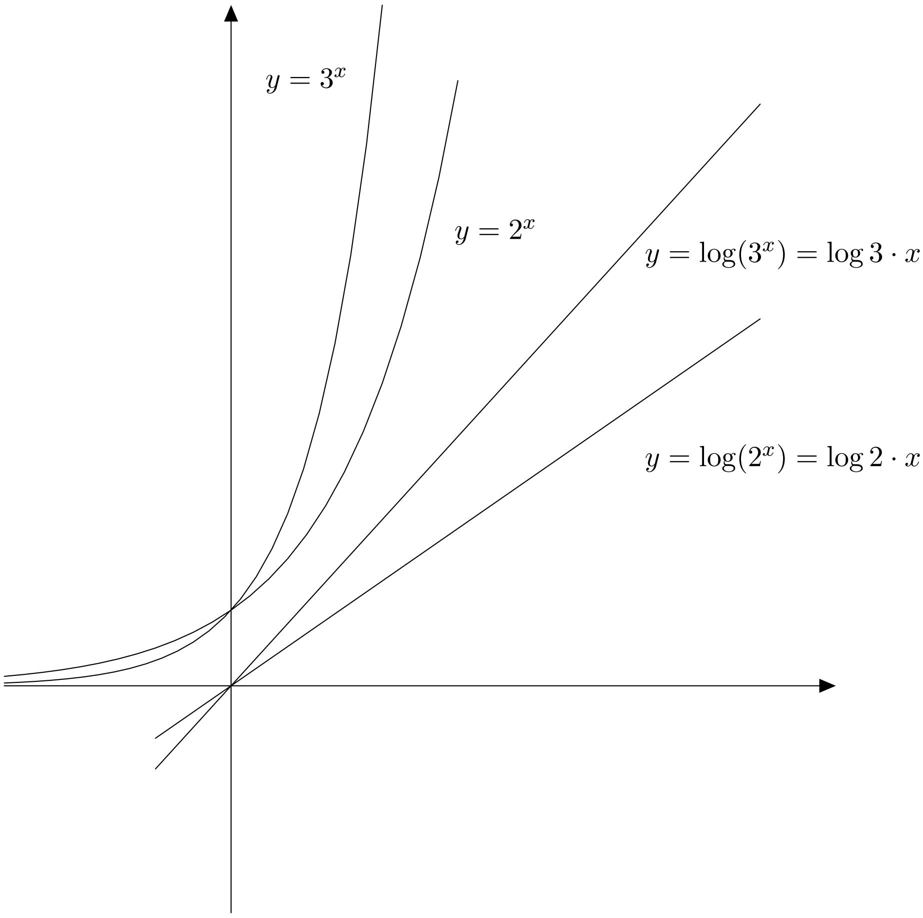


Figure 26: The exponential functions composed with the logarithm.

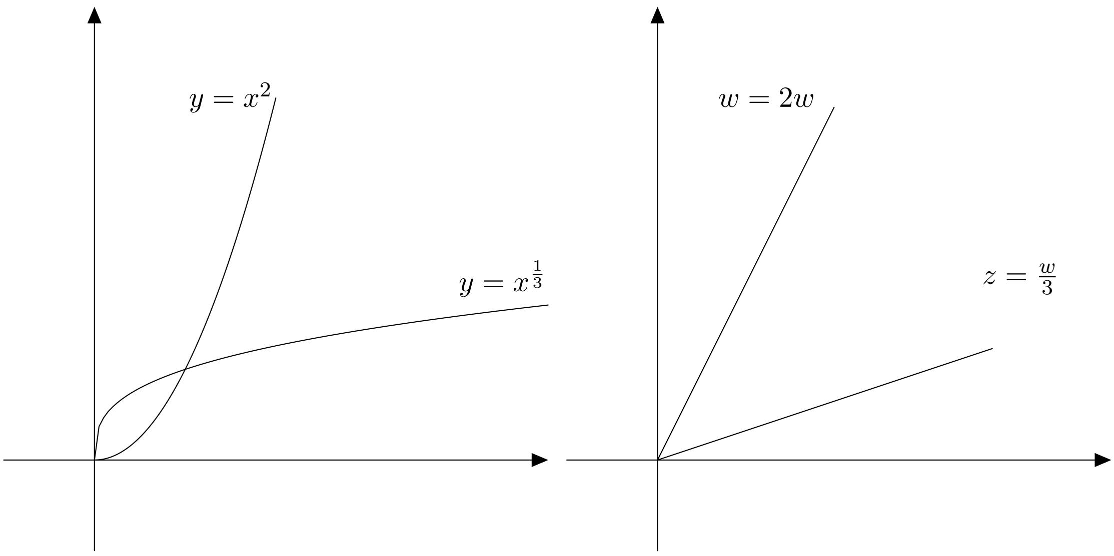


Figure 27: The log-log plot of the relation $y = x^p$.

Example 71. Let $f(x) = \operatorname{sign} x$. $\lim_{x\to 0^+} f(x) = 1$, $\lim_{x\to 0^-} f(x) = -1$.

If $f(x)$ is defined on $(b,a)\cup(a,c)$, $\lim_{x\to a} f(x) = L$ exists if and only if both the left and right limits exist and $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x)$ and it is L. We leave the proof to the reader.

Lemma 72. *Let f be a function on S, $\lim_{y\to y_0} f(y) = L$. Assume that g is a function on T, continuous at x_0 and $g(x_0) = y_0$, $g(x) \neq y_0$ if $x \neq 0$, $|x - x_0 < \epsilon$ for some ϵ. Then $\lim_{x\to x_0} f(g(x)) = \lim_{y\to y_0} f(y) = L$. Similarly, if $\lim_{y\to\infty} f(y) = L$ and $\lim_{x\to\infty} g(x) = \infty$, then $\lim_{x\to\infty} f(g(x)) = L$.*

Proof. The first statement can be proven similarly to the continuity of the composed function $f(g(x))$.

As for the second point, for a given ϵ we take Y such that $|f(y) - L| < \epsilon$ for $y > Y$. Then, there is X such that $g(x) > Y$ for $x > X$. Altogether, $|f(g(x)) - L| < \epsilon$ if $x > X$. □

We call this the change of variables, in the sense that we can calculate $\lim_{y\to y_0} f(y)$ by calculating $\lim_{x\to x_0} f(g(x))$ and vice versa.

For $x \in \mathbb{R}$, we denote by $[x]$ the largest integer n such that $n \le x$, and call it the **integer part of** x. For example, $[\sqrt{2}] = 1, [\pi] = 3$, and so on.

In this Proposition, $n \in \mathbb{N}$ and $x \in \mathbb{R}$.

Proposition 73. *We have the following.*

(i) $\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1}$.

(ii) $\lim_{n\to\infty} \left(1 + \frac{1}{n^2}\right)^n = 1$.

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

(iv) $\lim_{x\to 0} (1 + x)^{\frac{1}{x}} = e$.

(v) $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$.

(vi) $\lim_{x\to\infty} \left(1 + \frac{t}{x}\right)^x = e^t$.

(vii) $\lim_{x\to 0} \frac{e^x - 1}{x} = 1$.

Proof. (i) Note that

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

Note that $\frac{1}{x}$ is continuous at $x = 1, e$, and hence $\left(1 + \frac{1}{n-1}\right)^{-1} \to 1$ and $\frac{1}{\left(1+\frac{1}{n-1}\right)^{n-1}} \to \frac{1}{e}$. Altogether, $\left(1 - \frac{1}{n}\right)^n = \frac{1}{e} = e^{-1}$.

(ii) $\lim_{n\to\infty} \left(1 + \frac{1}{n^2}\right)^n = \lim_{n\to\infty} \left(\left(1 + \frac{1}{n^2}\right)^{n^2}\right)^{\frac{1}{n}}$. As $\lim_{n\to\infty} \left(1 + \frac{1}{n^2}\right)^{n^2} = e$, this sequence is bounded by, say M. Then $1 < \left(1 + \frac{1}{n^2}\right)^n < M^{\frac{1}{n}}$ but $M^{\frac{1}{n}} \to 1$, then by squeezing we have $\left(1 + \frac{1}{n^2}\right)^n \to 1$.

(iii) Note that, if $a_n \to a$, then $b_n = a_{n+1} \to a$ as well. Furthermore, if $a < b < c$ and if $|a - x| < \epsilon, |c - x| < \epsilon$, then by the triangle inequality we have $-\epsilon < a - x < \epsilon$, hence $a - \epsilon < x < a + \epsilon$. Similarly, $c - \epsilon < x < c + \epsilon$, and therefore, $b - \epsilon < x < b + \epsilon$ and hence $|b - x| < \epsilon$.

We know that $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty}\left(1+\frac{1}{n+1}\right)^{n+1} = e$. Let $n = [x]$, then $n \le x < n+1$ and

$$\frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{1+\frac{1}{n+1}} < \left(1+\frac{1}{x}\right)^x < \left(1+\frac{1}{n}\right)^{n+1} = \left(1+\frac{1}{n}\right)^n \cdot \left(1+\frac{1}{n}\right).$$

Note that the left-hand side and the right-hand side tend to e, because $1+\frac{1}{n+1} \to 1, 1+\frac{1}{n} \to 1$. This means that, for a given ϵ, $\left|\frac{(1+\frac{1}{n+1})^{n+1}}{1+\frac{1}{n+1}} - e\right| < \epsilon, \left|\left(1+\frac{1}{n}\right)^n \cdot \left(1+\frac{1}{n}\right) - e\right| < \epsilon$ for sufficiently large n. This implies that $\left|\left(1+\frac{1}{x}\right)^x - e\right| < \epsilon$.
Altogether, this says that, if x is sufficiently large, then we apply this argument with $n = [x]$, and obtain that $\left|\left(1+\frac{1}{x}\right)^x - e\right| < \epsilon$. This is $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$.

(iv) By the previous point and a change of variables $\frac{1}{x}$, note that $\frac{1}{x} > 0$, $\lim_{x\to 0^+}(1+x)^{\frac{1}{x}} = e$. We have $\lim_{x\to 0-}(1+x)^{\frac{1}{x}} = e$ as well. So we have checked both the right and left limits.

(v) As $\log y$ is continuous at $y = e$,

$$\lim_{x\to 0}\frac{\log(1+x)}{x} = \lim_{x\to 0}\log(1+x)^{\frac{1}{x}} = \log\lim_{x\to 0}(1+x)^{\frac{1}{x}} = \log e = 1,$$

where we used $\lim_{x\to 0}(1+x)^{\frac{1}{x}} = e$.

(vi) Note that $\lim_{x\to\infty}\left(1+\frac{t}{x}\right)^x = \lim_{x\to\infty}\left(\left(1+\frac{t}{x}\right)^{\frac{x}{t}}\right)^t = e^t$, where we used the continuity of $g(a) = a^t$.

(vii) With $y = e^x - 1$, we have $\log(y+1) = x$ and $\lim_{x\to 0}\frac{e^x-1}{x} = \lim_{y\to 0}\frac{y}{\log(1+y)} = 1$. □

Definition 74.

- $\sinh x = \frac{e^x - e^{-x}}{2}$
- $\cosh x = \frac{e^x + e^{-x}}{2}$
- $\tanh x = \frac{\sinh x}{\cosh x}$

Proposition 75. *(i)* $\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y$.

(ii) $\sinh(x+y) = \cosh x \sinh y + \sinh x \cosh y$.

(iii) $(\cosh x)^2 - (\sinh x)^2 = 1$.

Proof. (i) $\cosh x \cosh y + \sinh x \sinh y = \frac{1}{4}(e^x + e^{-x})(e^y + e^{-y}) + \frac{1}{4}(e^x - e^{-x})(e^y - e^{-y}) = \frac{1}{4}(2e^{x+y} + 2e^{-x-y}) = \cosh(x+y)$.

(ii) analogous.

(iii) analogous.

□

Proposition 76. *(i)* $\sinh^{-1}(x) = \log(x + \sqrt{x^2+1})$.

(ii) $\cosh^{-1}(x) = \log(x + \sqrt{x^2-1})$ *for* $x > 1$.

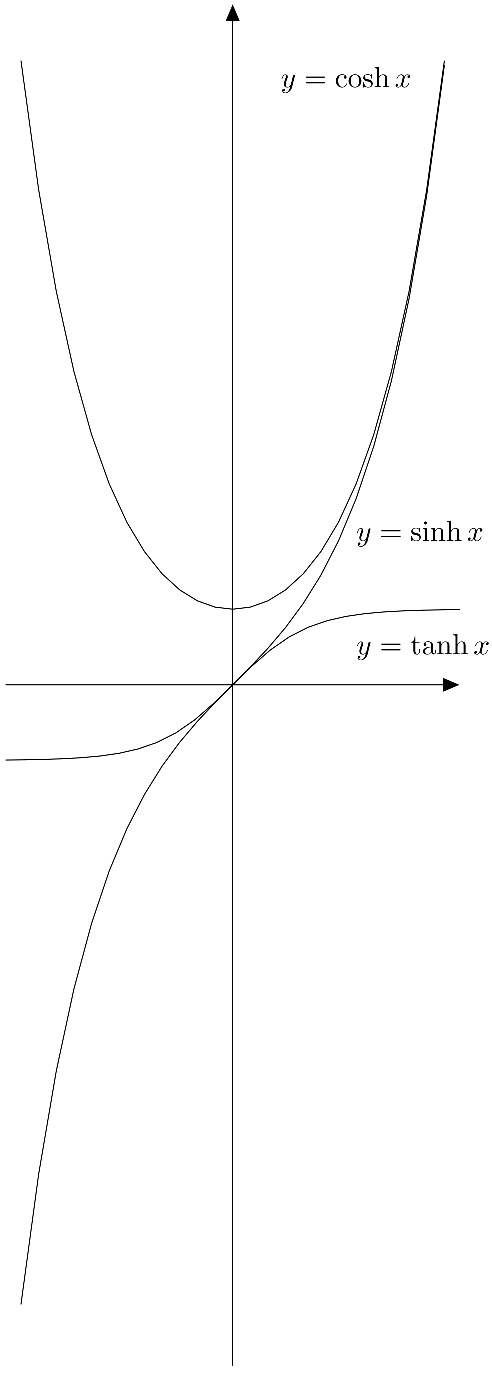


Figure 28: The hyperbolic functions.

Proof. (i)

$$\sinh(\log(x+\sqrt{x^2+y})) = \frac{1}{2}\left((x+\sqrt{x^2+1}) - \frac{1}{x+\sqrt{x^2+1}}\right) = \frac{1}{2}\frac{(x+\sqrt{x^2+1})^2-1}{x+\sqrt{x^2+1}}$$
$$= \frac{1}{2}\frac{x^2+2x\sqrt{x^2+1}+x^2+1-1}{x+\sqrt{x^2+1}} = x.$$

(ii) analogous.

□

Definition 77. (i) $\operatorname{Arcsinh} x = \sinh^{-1}(x) = \log(x+\sqrt{x^2+1})$.

(ii) $\operatorname{Arccosh} x = \cosh^{-1}(x) = \log(x+\sqrt{x^2-1})$ for $x > 0$.

Oct. 20. Review of trigonometric functions and some limits.

Trigonometric functions

The functions $\sin\theta$ and $\cos\theta$ are usually defined as the length of the horizontal and vertical sides of the right triangle obtained from a point p on the unit circle (the circle centered at $(0,0)$ with radius 1) such that the x-axis and the segment from the point of origin to p makes an angle of degree θ. However, to make this definition precise, we would first need to define the **angle**, that is the **length of the arc** on the unit circle, then consider the right triangle...

That is possible, but we would have to wait until we define integral before define trigonometric functions (or define the trigonometric function by something called power series). In this lecture, we prefer practicality, therefore,

- We assume that there are functions called $\sin\theta, \cos\theta$.
- We use the figures and the elementary geometry to derive their elementary properties.
- Then we study their analytic aspects: limit, derivative, integral, Taylor expansion, and so on.

Now, to obtain $\cos\theta$ and $\sin\theta$, we draw the unit circle, and take the point p on the unit circle such that the x-axis and the segment from the point of origin to p makes an angle of degree θ going **anticlockwise**, $0 \le \theta \le 90$ (degrees). Then $\cos\theta$ is defined to be the x-coordinate of the point p, and $\sin\theta$ is defined to be the y-coordinate of p.

We can make a right triangle by drawing the vertical line from this point. If $0 \le \theta \le 90$ (degrees), then $\cos\theta$ is the length of the horizontal side of the triangle, while $\sin\theta$ defined to be the length of the vertical side. When $\theta \ge 90$ (degrees), then $\cos\theta$ becomes negative.

There are various ways to represent the angle. Often we use the **degrees**, which devide the circle into 360 degrees. Another is called the **radian**, which defines the angle by the lenght of the arc on the unit circle. In radian, we have 360 (degrees) $= 2\pi$ (radian), 180 (degrees) $= \pi$ (radian), 90 (degrees) $= \frac{\pi}{2}$ (radian), 45 (degrees) $= \frac{\pi}{4}$ (radian) and so on. In this lecture, from this point **we use radian**, unless otherwise specified.

Some important values:

- $\sin 0 = 0, \cos 0 = 1$.
- $\sin\frac{\pi}{6} = \frac{1}{2}, \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$.
- $\sin\frac{\pi}{4} = \frac{1}{\sqrt{2}}, \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$.
- $\sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}, \cos\frac{\pi}{3} = \frac{1}{2}$.

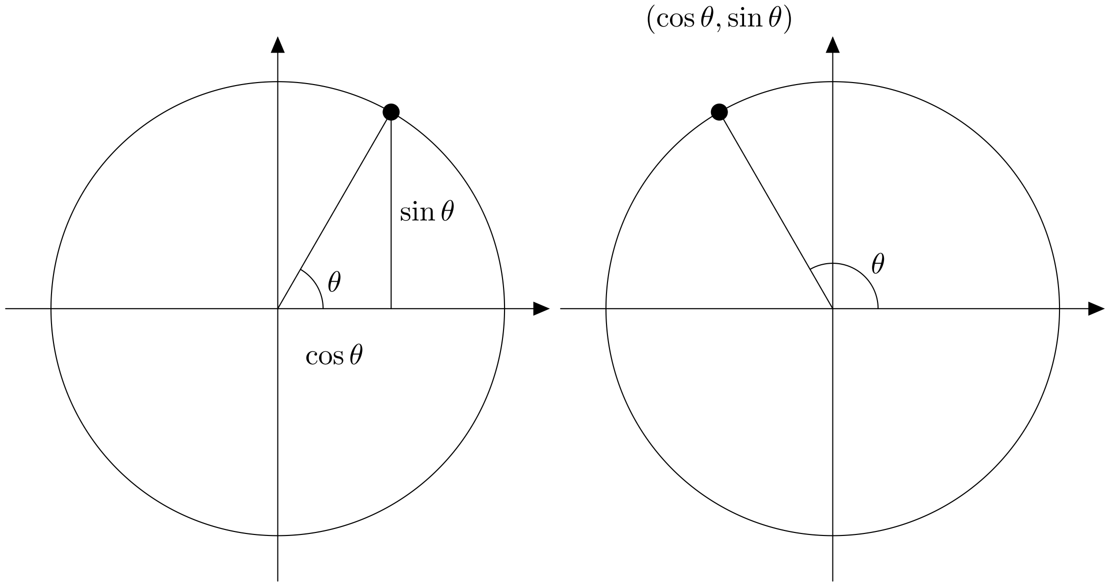


Figure 29: The trigonometric functions and their values for general angle θ.

- $\sin\frac{\pi}{2} = 1, \cos\frac{\pi}{2} = 0$.

We can extend $\cos\theta$ and $\sin\theta$ to all real numbers, considering that for $\theta > 2\pi$ we go around the circle more than once, and for $\theta < 0$ we go around the circle clockwise. With this understanding, we have

- $\cos(\theta + 2\pi) = \cos\theta$
- $\sin(\theta + 2\pi) = \sin\theta$.
- $\cos(-\theta) = \cos\theta$
- $\sin(-\theta) = -\sin\theta$.

In this way, we can consider cos and sin as **functions** on $\mathbb{R}$. They are continuous, because if we change slightly the degree, the point p moves only slightly (we do not prove this, as we introduce these functions only by geometry, without defining the arg length).

They are related by the formulas $\cos(\theta + \frac{\pi}{2}) = -\sin\theta$ and $\sin(\theta + \frac{\pi}{2}) = \cos\theta$ (see Figure 30).

We introduce also $\tan\theta = \frac{\sin\theta}{\cos\theta}$.

Some formulas

We often write $\cos^2\theta = (\cos\theta)^2, \sin^2\theta = (\sin\theta)^2, \cos^3\theta = (\cos\theta)^3, \sin^3\theta = (\sin\theta)^3$, etc.

- $\cos^2\theta + \sin^2\theta = 1$. This is because of the Pytagorean theorem: $\cos\theta$ and $\sin\theta$ are the length of the horizontal and vertical sides of the right triangle, while the length of the longest side is 1.
- $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$. See Figure 32
- $\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$.

From these formulas, we can derive various useful formulas.

- $\cos 2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta$. Indeed, $\cos 2\theta = \cos\theta\cos\theta - \sin\theta\sin\theta = \cos^2\theta - \sin^2\theta$ and use $\cos^2\theta + \sin^2\theta = 1$.

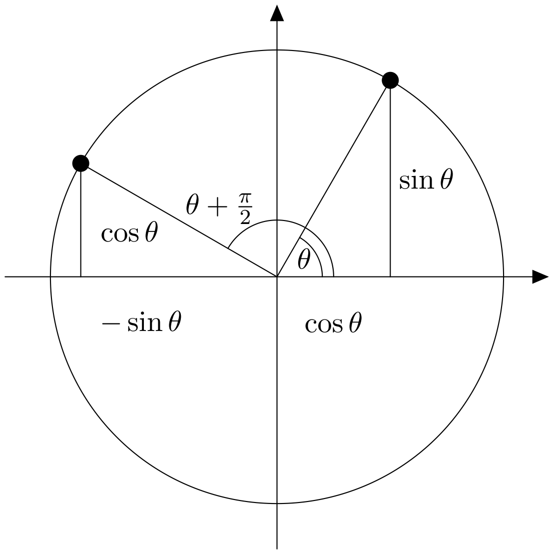


Figure 30: A relation $\cos(\theta + \frac{\pi}{2}) = -\sin\theta$ and $\sin(\theta + \frac{\pi}{2}) = \cos\theta$.

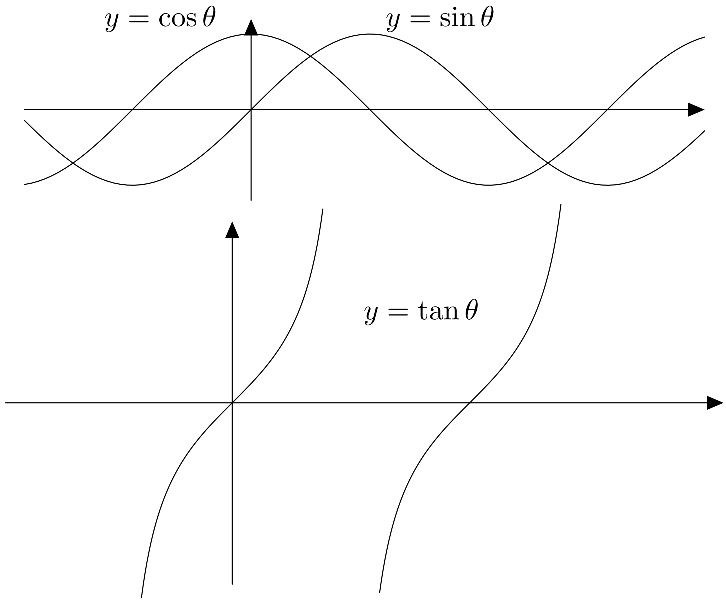


Figure 31: The graphs of $\cos\theta, \sin\theta$ and $\tan\theta$.

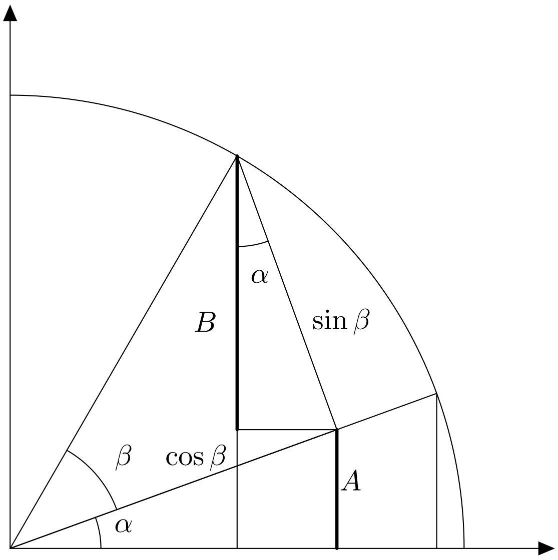


Figure 32: The formula $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$. $A = \cos\beta\sin\alpha, B = \sin\beta\cos\alpha$ and $A + B = \sin(\alpha+\beta)$.

- $\sin 2\theta = 2\sin\theta\cos\theta$. Indeed, $\sin 2\theta = \sin\theta\cos\theta + \cos\theta\sin\theta$.
- $\sin\alpha\cos\beta = \frac{1}{2}(\sin(\alpha+\beta) + \sin(\alpha-\beta))$. Indeed,

$$\begin{aligned}
&\frac{1}{2}(\sin(\alpha+\beta)+\sin(\alpha-\beta)) \\
&= \frac{1}{2}((\sin\alpha\cos\beta+\cos\alpha\sin\beta)+(\sin\alpha\cos(-\beta)+\cos\alpha\sin(-\beta))) \\
&= \frac{1}{2}(\sin\alpha\cos\beta+\cos\alpha\sin\beta+\sin\alpha\cos\beta-\cos\alpha\sin\beta) \\
&= \sin\alpha\cos\beta.
\end{aligned}$$

- $\cos\alpha\sin\beta = \frac{1}{2}(\sin(\alpha+\beta) - \sin(\alpha-\beta))$.
- $\cos\alpha\cos\beta = \frac{1}{2}(\cos(\alpha+\beta) + \cos(\alpha-\beta))$.
- $\sin\alpha\sin\beta = \frac{1}{2}(-\cos(\alpha+\beta) + \cos(\alpha-\beta))$.

For example, we can compute $\cos\frac{\pi}{8}$. Indeed, $2\cos^2\frac{\pi}{8} - 1 = \cos(\frac{\pi}{8}\cdot 2) = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$, and hence $\cos\frac{\pi}{8} = \sqrt{\frac{\frac{1}{\sqrt{2}}+1}{2}}$.

Some limit

By comparing the areas of the triangles of the sector, we see $\frac{1}{2}\cos\theta\sin\theta < \frac{\theta}{2} < \frac{1}{2}\frac{\sin\theta}{\cos\theta}$ (see Figure 33), and hence $\cos\theta < \frac{\sin\theta}{\theta} < \frac{1}{\cos\theta}$. As we assumed that sin and cos are continuous, and $\cos 0 = 1$, we obtain $\lim_{\theta\to 0}\frac{\sin\theta}{\theta} = \lim_{\theta\to 0}\cos\theta = \lim_{\theta\to 0}\frac{1}{\cos\theta} = 1$ by squeezing.

Oct. 21. Open and closed sets, Bolzano-Weierstrass theorem.

Definition 78. Let $O \subset \mathbb{R}$. We say that O is **open** if for any $p \in O$ there is $\epsilon > 0$ such that $(p-\epsilon, p+\epsilon) \subset O$ (this ϵ depends on p). Let $F \subset \mathbb{R}$. We say that F is **closed** if for any convergent sequence $\{a_n\} \subset F, a_n \to a$, it holds that $a \in F$.

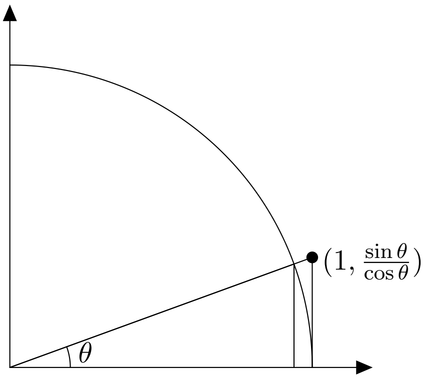


Figure 33: By comparing the areas of the triangles and the sector, we see $\frac{1}{2}\cos\theta\sin\theta < \frac{\theta}{2} < \frac{1}{2}\frac{\sin\theta}{\cos\theta}$.

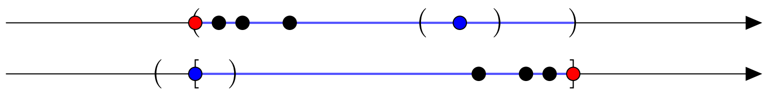


Figure 34: Open and closed intervals. An open set include a small "neighborhood" of any point in it, but a sequence in it might converge to a point outside. A closed subset contains the limit of any sequence in it, but a point might "touch" other points outside.

Example 79.

- Consider the open interval $A = (0, 1)$. This is open, because for any point $p \le \frac{1}{2}$ we can take $\epsilon = \frac{p}{2}$ and $(\frac{p}{2}, \frac{3p}{2}) \subset (0, 1)$. If $p > \frac{1}{2}$, we can take $\epsilon = \frac{1-p}{2}$. On the other hand, $(0, 1)$ is not closed. Indeed, the sequence $a_n = \frac{1}{n}$ belongs to $A = (0, 1)$, but the limit 0 does not belong to A.
- Consider the closed interval $B = [0, 1]$. This is closed. Indeed, for any convergent sequence $\{a_n\} \subset B, a_n \to a$, it holds that $0 \le a_n \le 1$ and hence $0 \le a \le 1$. On the other hand, for $p = 0$, for any ϵ, $(-\epsilon, \epsilon) \not\subset B$, therefore, B is not open.

Therefore, the terminology "open" and "closed" for intervals are consistent with those for general sets we have just introduced.

For any set $A \subset \mathbb{R}$, we denote its complement by $A^c = \mathbb{R} \setminus A$.

Lemma 80. *$O \subset \mathbb{R}$ is open if and only if O^c is closed.*

Proof. Let O be open and assume that O^c is not closed. That is, there is a sequence $\{a_n\} \subset O^c$ that converges to a, but $a \in O^c$. Therefore, it must holds $a \in O$. But we can take $\epsilon > 0$ such that $(a - \epsilon, a + \epsilon) \subset O$, and if $a_n \to a$, it would have to hold that $a_n \in (a - \epsilon, a + \epsilon) \subset O$, which contradicts the assumption that $\{a_n\} \subset O^c$. Therefore, O^c is closed.

Conversely, let O^c be closed, and assume that O is not open. As O is not open, there is $a \in O$ such that for any $\frac{1}{n} > 0$ there is a_n such that $|a_n - a| < \frac{1}{n}$, but $a_n \notin O$. Hence $a_n \in O^c$. But with this condition $a_n \to a$, which contradicts the assumption that O^c is closed. Therefore, O must be open. □

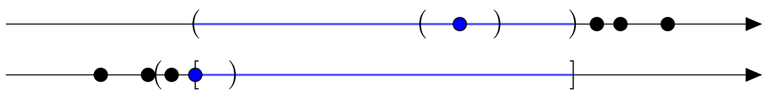


Figure 35: Any point in an open set is "protected" from outside. On the other hand, if a set is not open, there is a point which is not "protected".

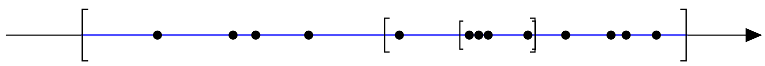


Figure 36: Nested invertals. As the sequence $\{a_n\}$ contains infinitely many points, one of two intervals must contain infinitely many of them.

It is not difficult to prove that any union (even if infinite!) of open sets is again open. Similarly, any intersection of closed sets is again closed.

Let us recall that a sequence $\{a_n\}$ is called Cauchy if for any given $\epsilon > 0$ there is N such that for $m, n > N$ it holds that $|a_m - a_n| < \epsilon$.

Furthermore, we said that b_n is a subsequene of a_n if there is a growing sequence $N_n \in \mathbb{N}$ such that $b_n = a_{N_n}$, that is, b_n is obtained by skipping some elements in a_n. Recall that we consider *infinite* sequences, that is, the sequence does not stop at any a_n, but continues infinitely.

Theorem 81 (Bolzano-Weierstrass)**.** *Let $\{a_n\}$ be a bounded sequence. Then there is a convergent subsequence of $\{a_n\}$.*

Proof. As $\{a_n\}$ is bounded, we can find M sufficiently large such that $a_n \in [-M, M]$. As the sequence $\{a_n\}$ infinitely many elements, one of the intervals $[-M, 0], (0, M]$ must contain infinitely many of them. Therefore, we can take a subsequence $b_n = a_{m_n}$ such that b_n are contained one of them. To fix the idea, assume that $b_n \in (0, M]$ (the other case is just analogous).

As $(0, M] = (0, \frac{M}{2}] \cup (\frac{M}{2}, M]$, one of them must contain infinitely many elements of b_n. Therefore, we can take a subsequence $c_n = b_{k_n}$ such that c_n are contained one of them.

By continuing this procedure, for each n we obtain a subsequence that is contained in an interval of length $\frac{M}{2^{n-1}}$, and the later one is a subsequence of the former. Let us take a subsequence $a_1, b_2, c_3, \cdots$ of the original sequence. Then, for $n, m > N$, any two elements are contained in an interval of length $\frac{M}{2^{N-1}}$. Therefore, this subsequence is Cauchy. Then it is a convergent sequence by Lemma 61. □

It is important that a_n is bounded. Indeed, if not, it is obviously impossible in general to extract a convergent sequence: consider $a_n = n$, which is not bounded and not convergent to any point. In addition, the possibility to extract a convergent subsequence does not mean that the original sequence is convergent, or there is only one convergent subsequence.

Theorem 82. *Let f be a continuous function defined on a bounded closed interval F. Then f is bounded, that is, there is $M > 0$ such that $|f(x)| < M$ for $x \in F$.*

Proof. Let us suppose the contrary, that for any $n > 0$ there is $x_n \in F$ such that $|f(x_n)| \geq n$. As $\{x_n\}$ is a sequence in a bounded set F, we can take a convergent subsequence $\{y_n\}$ of $\{x_n\}$. As F is closed, $y_n \to y$ and $y \in F$. By assumption f is continuous, therefore, it must hold that $\lim_{n\to\infty} f(y_n) = f(y)$. But this is impossible because $|f(y_n)| \geq n$ by our choice. Therefore, f is bounded. □

Example 83.

- Consider the function $f(x) = \frac{1}{x}$ defined on $\mathbb{R} \setminus \{0\}$. This is not bounded, but when we restrict it to an interval $[\frac{1}{n}, n]$, it is bounded by n.
- Consider the function $f(x) = \begin{cases} \frac{1}{x} & \text{if } x \in [-1, 1], x \neq 0 \\ 0 & x = 0 \end{cases}$. This is defined on a closed interval $[-1, 1]$, but not continuous. Therefore, the previous theroem does not apply. Indeed, it is not bounded.

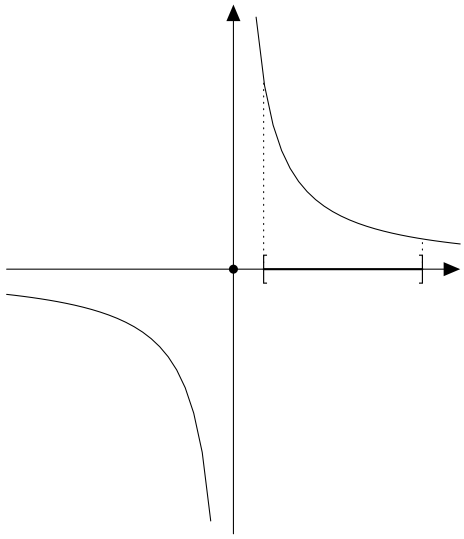


Figure 37: A continuous function on a bounded closed interval is bounded. If either of these conditions are violated, then function can be unbounded.

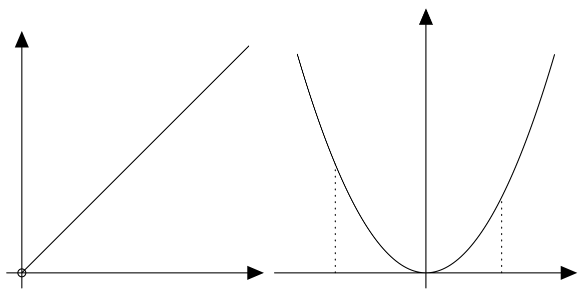


Figure 38: Left: $y = x$ on $x > 0$. There are no minimum or maximum. Right $y = x^2$ on $\mathbb{R}$. The minimum is 0 at $x = 0$, but there is not maximum. When restricted to $[a, b]$, either a^2 or b^2 is the maximum.

Oct. 25. Maximum and minimum of functions, the Weierstrass theorem, uniform continuity.

The maximum and minimum of functions

Definition 84. Let f be a function defined on S.

- We say that f takes its **maximum** at x_0 if $f(x_0) \geq f(x)$ for all $x \in S$.
- We say that f takes its **minimum** at x_0 if $f(x_0) \leq f(x)$ for all $x \in S$

Example 85. Note that a function does not necessarily admit maximum or minimum. If it has, they may depend on the domain.

- $f(x) = x$, defined on $x > 0$, has no maximum or minimum. Indeed, for any $x > 0$, $f(\frac{x}{2}) = \frac{x}{2} < x$ and $f(2x) = 2x > x$.
- $f(x) = x^2$, defined on $x \in \mathbb{R}$, has no maximum but the minimum is at $x = 0$ with $f(0) = 0$. If it is restricted to the interval $[a, b]$, then the maximum is the larger one of a^2, b^2.

Theorem 86 (Weierstrass)**.** *Let $F \subset \mathbb{R}$ be a bounded closed set (or interval), and f be a continuous function on F. Then f admits both a maximum and a minimum in F.*

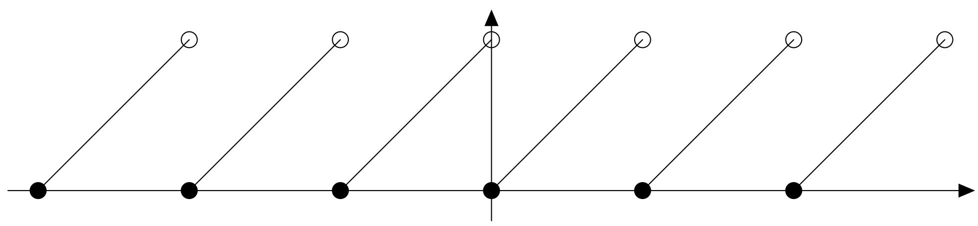


Figure 39: The graph of the function $y = x - [x]$, the decimal part of x. This is bounded, but has no maximum. The minimum is 0 at $x \in \mathbb{Z}$.

Proof. By Theorem 82, f is bounded, say $-M < f(x) < M$. Then the image $A = \{f(x) : x \in F\}$ is a bounded set in $\mathbb{R}$, therefore, it admits $\sup A$ and $\inf A$. Let us prove that f admits a maximum (the case for minimum is analogous). For each n there is $x_n \in F$ such that $\sup A - \frac{1}{n} < f(x_n)$.

As F is bounded, x_n admits a convergent subsequence y_n, $y_n \to y$ and $y \in F$ because F is closed. Now, as f is continuous, we have $f(y) = \lim_{n\to\infty} f(y_n)$. As y_n is a subsequence, it holds that $\sup A - \frac{1}{n} < f(y_n) \leq \sup A$. This implies that $f(y) = \sup A$. That is, f attains a maximum at y. □

Example 87. (and non example)

- $f(x) = x^2$ is continuous, hence on any closed and bounded F f admits a maximum and a minimum. But not on the whole real line $\mathbb{R}$, which is not bounded.
- $f(x) = x - [x]$ is not continuous, and indeed it does not admit a maximum on $[0, 1]$, although $[0, 1]$ is close and bounded.

Often it is said that a closed and bounded set $F \subset \mathbb{R}$ is **compact**. We have seen that in any sequence $\{a_n\}$ in a compact set admits a convergent subsequence (the Bolzano-Weierstrass theorem), and the limit is in F. Conversely, if a set A has a property that any sequence in it has a convergent subsequence with the limit in A, then it is compact (bounded and closed): indeed, A must be bounded because otherwise we could take an unbounded sequence. Furthermore, A must be closed, because if $a_n \in A$ is a convergent sequence, we can take a convergent subsequence with the limit a in A, but there is only one limit for a_n, hence $a_n \to a \in A$, that is, A is closed.

Let us see another strong property of continuous functions defined on bounded and closed sets.

Definition 88. Let $S \subset \mathbb{R}$, $f : S \to \mathbb{R}$. f is said to be **uniformely continuous on** S if, for any $\epsilon > 0$, there is $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $x, y \in S, |x - y| < \delta$.

Note the difference with the continuity: a function f is continuous if for each $x \in S$ and for each ϵ there is δ such that $|f(y) - f(x)| < \epsilon$ if $|y - x| < \delta$. In other words, the number δ may change from point x to others.

On the other hand, uniform continuity asserts that for each $\epsilon > 0$ there is δ that applies to all $x, y \in S$, hence *uniformly* in S.

Example 89. (functions that are not uniformly continuous)

- $f(x) = \frac{1}{x}$ is continuous on $\{x \in \mathbb{R} : x > 0\}$. However, it is not uniformly continuous. Indeed, for $\epsilon = 1$ for any $\delta > 0$, we can take N such that $\frac{1}{N} < \delta$ and $N > 2$. Then $x = \frac{1}{N}, y = \frac{2}{N}$, hence $f(y) - f(x) = \frac{N}{2} > 1 = \epsilon$ but $x - y = \frac{1}{N} < \delta$.
- $f(x) = \sin\frac{1}{x}$ is continuous on $\{x \in \mathbb{R} : x > 0\}$ but not uniformly continuous. Indeed, for $\epsilon = \frac{1}{2}$ for any $\delta > 0$, we can take N such that $\frac{2}{\pi N} < \delta$ and N odd. Then $x = \frac{2}{\pi N}, y = \frac{1}{\pi N}$, hence $|f(\frac{1}{\pi N}) - f(\frac{2}{\pi N})| = |\sin(\pi N) - \sin(\frac{\pi N}{2})| = 1 > \epsilon$ but $x - y = \frac{1}{\pi N} < \delta$.

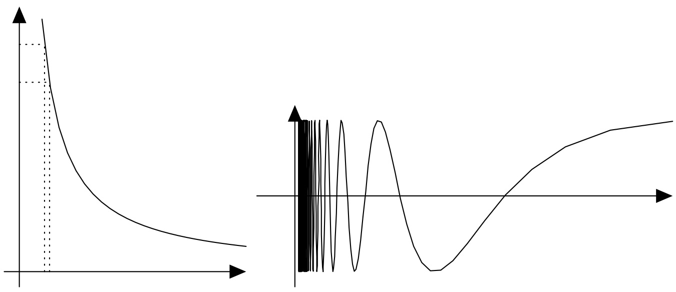


Figure 40: Functions continuous but not uniformly continuous.

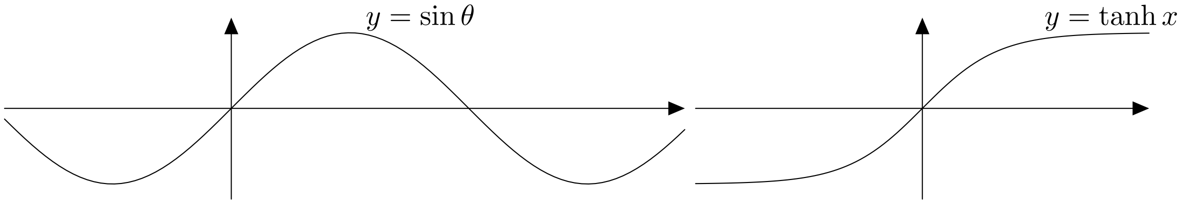


Figure 41: Functions defined on $\mathbb{R}$ but uniformly continuous.

Note that the function $f(x) = |x|$ is continuous. Indeed, if $x > 0$, then $f(x) = x$ and this is continuous at x. Similarly, f is continuous at $x < 0$. Finally, if $x = 0$, for any $\epsilon > 0$, we take $\delta = \epsilon$. Then if $|y - x| = |y - 0| < \delta$, then $|y| - |0| = |y - 0| < \delta = \epsilon$.

Theorem 90 (Heine-Cantor)**.** *Let F bounded and closed, $f : F \to \mathbb{R}$ a continuous function. Then f is uniformly continuous.*

Proof. To prove this by contradiction, assume that there is $\epsilon > 0$ such that for any $\delta > 0$ there are $x, y \in F, |x - y| < \delta$ but $|f(x) - f(y)| > \epsilon$. In particular, for $\delta = \frac{1}{n} > 0$ there are $x_n, y_n \in F$ such that $|x_n - y_n| < \frac{1}{n}$ but $|f(x_n) - f(y_n)| > \epsilon$. Let x_{N_n} be a convergent subsequence of x_n (which exists by Theorem 81) to $\tilde{x} \in F$. Let us extract a subsequence $\{y_{N_n}\}$ of $\{y_n\}$. As $|\tilde{x} - y_{N_n}| \leq |\tilde{x} - x_{N_n}| + |x_{N_n} - y_{N_n}| \to 0$, also $\{y_{N_n}\}$ must be convergent to $\tilde{x} \in F$.

Then $\lim_{n\to\infty} |f(x_{N_n}) - f(y_{N_n})| = |f(\tilde{x}) - f(\tilde{x})| = 0$, as f is continuous (note that the absolute value is continuous). But this contradicts the assumption that $|f(x_{N_n}) - f(y_{N_n})| > \epsilon$.

Therefore, for all ϵ there exists δ such that for all $x, y \in F, |x-y| < \delta$ vale $|f(x)-f(y)| < \epsilon$. □

Until now, we have studied continuity of functions. A function f is continuous at point x if for each $\epsilon > 0$ there is δ such that $|f(y) - f(x)| < \epsilon$ for y such that $|y - x| < \delta$. This tells us that "the graph is connected", but does not tell us how fast the function f changes.

We would like to know such information. For example, if f represents the motion of a car (in one direction), then how can we determine the **speed** of the car? Or if f represents the height of the mountain in a path and x represents the distance from the starting point, what is the **slope** of the mountain?

In the case of the speed, if the car has travelled 100km in two hours, then the *average speed* is 50km/h per hour. But it might be that the car travelled with the constant speed 50km/h, or it travelled with 40km/h in the first one hour and then 60km/h in the second one hour. Is it possible to determine the speed at a moment? In the case of a mountain, what is the slope at a point?

They should be approximated by secant lines.

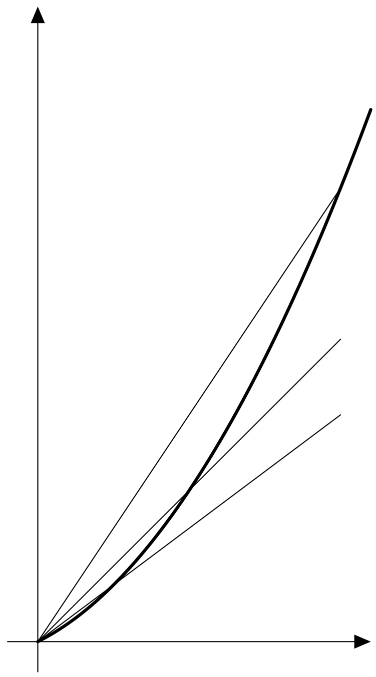


Figure 42: The slope at a point as the limit of the slopes of secant lines.

Oct. 27. Derivative. First examples.

Derivative

As we discussed, we can define the average speed of a car, or the average slope of a curve in an interval. By taking the limit of the interval that tends to 0, we should obtain the speed or the slope at one point.

Definition 91. Let $I \subset \mathbb{R}$ an open interval, f a function defined on I.

- Let $x_0 \in I$ and h small such that $x_0 + h \in I$.

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

  is called the average rate of change of f between x_0 and $x_0 + h$.

- the function f is said to be **differentiable at** x_0 if the following limit exists:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

  If this limit exists, it is called the **derivative of** f **at** x_0 and it is denoted by $f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$, $Df(x_0)$ or $\frac{df}{dx}(x_0)$.

The derivative at the point x_0 is defined to be the limit of average rates of change. In this sense, the derivative represents the rate of change af the point x_0. If $f(t)$ represents the position of a car at time t, then $f'(t)$ is the speed of the car at time t.

Derivatives of elementary functions.

- Let $f(x) = c$ for $x \in \mathbb{R}$ (constant). For any $x \in \mathbb{R}$, $\frac{f(x+h)-f(x)}{h} = \frac{c-c}{h} = 0$, therefore, $f'(x) = 0$.

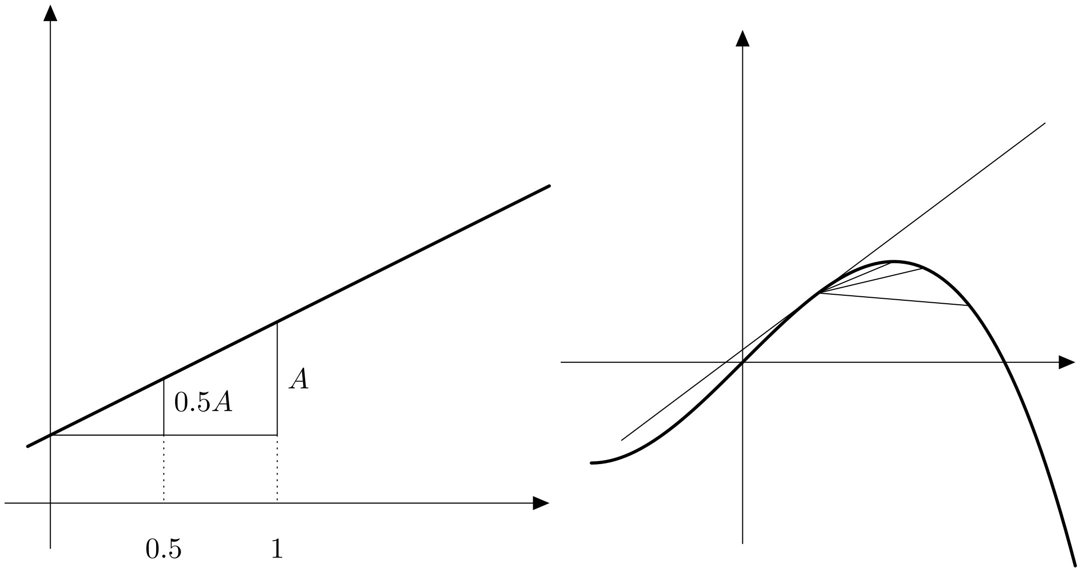


Figure 43: The slope of the straight line at a point as the limit of the slopes of secant lines.

- Let $A \in \mathbb{R}$ and $f(x) = Ax$ for $x \in \mathbb{R}$ (a straight line). For any $x \in \mathbb{R}$, $\frac{f(x+h)-f(x)}{h} = \frac{A(x+h)-Ax}{h} = \frac{Ah}{h} = A$, $f'(x) = A$.
- Let $A \in \mathbb{R}$ and $f(x) = Ax^2$ for $x \in \mathbb{R}$ (parabola). For any $x \in \mathbb{R}$, $\frac{f(x+h)-f(x)}{h} = \frac{A(x+h)^2-Ax^2}{h} = \frac{A(2xh+h^2)}{h} = A(2x+h)$, therefore, $f'(x) = \lim_{h\to 0} A(2x+h) = 2xA$.
- Let $n \in \mathbb{N}$ and $f(x) = Ax^n$ for $x \in \mathbb{R}$. It holds that $(x+h)^n = \sum_{k=0}^{n} \binom{n}{k} x^k h^{n-k} = x^n + nx^{n-1}h + \frac{n(n-1)}{2}x^{n-2}h^2 + \dots$. For any $x \in \mathbb{R}$,

$$\frac{f(x+h)-f(x)}{h} = \frac{A(x+h)^n - Ax^n}{h} = \frac{A(x^n + nx^{n-1}h + \frac{n(n-1)}{2}x^{n-2}h^2 + \dots h^n - x^n)}{h}$$
$$= Anx^{n-1} + A \cdot \frac{n(n-1)}{2}x^{n-2}h + \cdots h^{n-1},$$

therefore, $f'(x) = \lim_{h\to 0} A(nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}h + \cdots h^{n-1}) = Anx^{n-1}$.
- Let $f(x) = \frac{1}{x}$ for $x \in \mathbb{R}, x \neq 0$. For any $x \in \mathbb{R}, x \neq 0$,

$$\frac{f(x+h)-f(x)}{h} = \frac{\frac{1}{x+h} - \frac{1}{x}}{h} = \frac{x-(x+h)}{hx(x+h)} = -\frac{1}{x(x+h)}$$

therefore, $f'(x) = \lim_{h\to 0} -\frac{1}{x(x+h)} = -\frac{1}{x^2}$.
- Let $f(x) = \log x$, $x > 0$. Then

$$\frac{\log(x+h) - \log x}{h} = \log\left(1 + \frac{h}{x}\right)^{\frac{1}{h}} = \frac{1}{x}\log\left(1+\frac{h}{x}\right)^{\frac{x}{h}},$$

therefore, $f'(x) = \lim_{h\to 0} \frac{1}{x}\log(1+\frac{h}{x})^{\frac{x}{h}} = \lim_{y\to 0} \frac{1}{x}\log(1+y)^{\frac{1}{y}} = \frac{1}{x}$ (this is one of the notable limits we have learned)
- Let $f(x) = e^x$, $x \in \mathbb{R}$. Then

$$\frac{e^{x+h} - e^x}{h} = e^x \frac{e^h - 1}{h},$$

therefore, $f'(x) = \lim_{h\to 0} e^x \frac{e^h-1}{h} = e^x$ (this is one of the notable limits).

- $f(x) = \sin x, x \in \mathbb{R}$. Recall the formula $\cos\alpha \sin\beta = \frac{1}{2}(\sin(\alpha+\beta) - \sin(\alpha-\beta))$. Then, with $\alpha = x + \frac{h}{2}$, $\beta = \frac{h}{2}$, we have $\sin(x+h) - \sin x = 2\cos(x + \frac{h}{2})\sin\frac{h}{2}$, therefore,

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} \\ &= \lim_{h\to 0} \frac{2\cos(x+\frac{h}{2})\sin\frac{h}{2}}{h} \\ &= \lim_{h\to 0} \cos\left(x + \frac{h}{2}\right) \lim_{h\to 0} \frac{\sin\frac{h}{2}}{\frac{h}{2}} \\ &= \cos x \cdot 1 = \cos x \end{aligned}$$

(by the continuity of $\cos x$ and one of the notable limits $\lim_{h\to 0} \frac{\sin h}{h} = 1$ and the change of variable $\frac{h}{2}$ replacing h.

- $f(x) = \cos x, x \in \mathbb{R}$. Recall the formula $-\sin\alpha\sin\beta = \frac{1}{2}(\cos(\alpha+\beta) - \cos(\alpha-\beta))$ Then, with $x + \frac{h}{2}$, $\beta = \frac{h}{2}$, we have $\cos(x+h) - \cos x = -2\sin(x+\frac{h}{2})\sin\frac{h}{2}$

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\cos(x+h) - \cos x}{h} \\ &= \lim_{h\to 0} \frac{-2\sin(x+\frac{h}{2})\sin\frac{h}{2}}{h} \\ &= -\lim_{h\to 0} \sin\left(x + \frac{h}{2}\right) \lim_{h\to 0} \frac{\sin\frac{h}{2}}{\frac{h}{2}} \\ &= -\sin x \cdot 1 = -\sin x \end{aligned}$$

(by the continuity of $\sin x$ and $\lim_{h\to 0} \frac{\sin h}{h} = 1$ and the change of variables).

Lemma 92. *If $f(x)$ is differentiable at x_0, then f is continuous at x_0.*

Proof. We compute the limit:

$$\lim_{x\to x_0} f(x) - f(x_0) = \lim_{h\to 0} f(x_0+h) - f(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} \cdot h = f'(x_0) \cdot 0 = 0.$$

That is, $\lim_{x\to x_0} f(x) = f(x_0)$.

□

Definition 93. Let $f : [x_0 - \delta, x_0] \to \mathbb{R}$ where $\delta > 0$. If the following limit $\lim_{h\to 0-} \frac{f(x_0+h)-f(x_0)}{h}$ exists (from the left), f is said to be left-differentiable at x_0, and this limit is denoted by $D_- f(x_0)$, the left derivative. Similarly, we define the right derivative.

Example 94. Let $f(x) = |x|, x_0 = 0$. $D_- f(0) = \lim_{h\to 0-} \frac{|0+h|-0}{h} = \lim_{h\to 0-} \frac{-h}{h} = -1$, while $D_+ f(0) = \lim_{h\to 0+} \frac{h}{h} = 1$.

Definition 95. Let f be defined on an open interval I. If f is differentiable at each point x of I, then $x \mapsto f'(x)$ defines a new function I. This is called the **derivative of** $f(x)$.

Example 96.
- The derivative of $f(x) = C$ (constant) is $f'(x) = 0$.
- The derivative of $f(x) = x$ is $f'(x) = 1$.
- The derivative of $f(x) = x^2$ is $f'(x) = 2x$.
- The derivative of $f(x) = \sin x$ is $f'(x) = \cos x$.

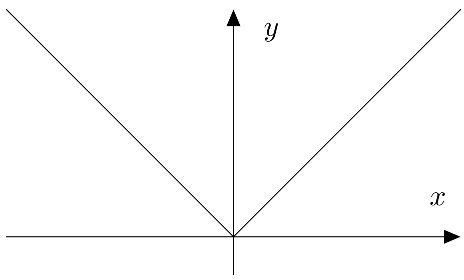


Figure 44: The graph of $y = |x|$, which has left and right derivatives, but they do not coincide.

Oct. 28. More examples of derivatives.

For a function f defined on an open interval I and $x \in I$, we have defined the derivative $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$, and we say that f is differentiable at x if this limit exists. Sometimes we denote this as $f'(x) = (Df)(x)$.

This is equivalent to write $Df(x) = f'(x) = \lim_{w\to x} \frac{f(w)-f(x)}{w-x}$.

Let f, g be functions. We write this $x \mapsto f(x)$. We denote by $f+g$ the function that maps $x \mapsto f(x)+g(x)$. Similarly, $f \cdot g$ is the function $x \mapsto f(x)g(x)$, $\frac{f}{g}$ is the function $x \mapsto \frac{f(x)}{g(x)}$, and the composition is $f \circ g$ that is given by $x \mapsto f(g(x))$.

Theorem 97. *Let f, g be functions on open intervals. The following hold if f, g are differentiable at x (or f at $g(x)$ for the chain rule):*

- *For $a, b \in \mathbb{R}$, $D(af+bg)(x) = aDf(x) + bDg(x)$ (linearity).*
- *$D(fg)(x) = Df(x)g(x) + f(x)Dg(x)$ (Leibniz rule).*
- *If $g(x) \neq 0$, then $D(\frac{f}{g})(x) = \frac{Df(x)g(x)-f(x)Dg(x)}{g(x)^2}$.*
- *$D(f \circ g)(x) = Dg(x)Df(g(x))$ (the chain rule).*
- *If $Df(x) \neq 0$ and f is monotonically increasing or decreasing and continuous in $(x-\epsilon, x+\epsilon)$ for some $\epsilon > 0$. Then f^{-1} is differentiable at $y = f(x)$ and $D(f^{-1}(y)) = \frac{1}{Df(x)}$.*

Proof.

- This is straightforward from the algebra of limits:

$$\begin{aligned}
&\lim_{h\to 0} \frac{af(x+h)+bg(x+h)-af(x)-bg(x)}{h} \\
&= \lim_{h\to 0} a\frac{f(x+h)-f(x)}{h} + \lim_{h\to 0} b\frac{g(x+h)-g(x)}{h} \\
&= a\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} + b\lim_{h\to 0} \frac{g(x+h)-g(x)}{h} \\
&= aDf(x) + bDg(x).
\end{aligned}$$

- Note that $f(x+h)g(x+h) - f(x)g(x) = f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) -$

$f(x)g(x)$, and g is continuous at x because it is differentiable there:

$$\begin{aligned}
&\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h)}{h} + \lim_{h\to 0} \frac{f(x)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \lim_{h\to 0} g(x+h) + f(x) \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} \\
&= Df(x)g(x) + f(x)Dg(x).
\end{aligned}$$

- As $g(x) \neq 0$, we have $\lim_{h\to 0} \frac{1}{g(x+h)} = \frac{1}{g(x)}$ and

$$\begin{aligned}
&\lim_{h\to 0} \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x+h)}{g(x+h)g(x)h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x) + f(x)g(x) - f(x)g(x+h)}{g(x+h)g(x)h} \\
&= \lim_{h\to 0} \frac{(f(x+h) - f(x))g(x) - f(x)(g(x+h) - g(x))}{g(x+h)g(x)h} \\
&= \frac{Df(x)g(x) - f(x)Dg(x)}{g(x)^2}.
\end{aligned}$$

- Note first that the difference $u(k) = \frac{f(g(x)+k) - f(g(x))}{k} - Df(g(x))$ tends to 0 as $k \to 0$. Let us also set $u(0) = 0$, then u is continuous around 0. We can write this as $f(g(x)+k) - f(g(x)) = k(Df(g(x)) + u(k))$, and this holds also for $k = 0$.

  We compute

$$\begin{aligned}
&\lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h} \\
&= \lim_{h\to 0} \frac{f(g(x) + (g(x+h) - g(x))) - f(g(x))}{h} \\
&= \lim_{h\to 0} \frac{(g(x+h) - g(x))(Df(g(x)) + u(g(x+h) - g(x)))}{h} \\
&= \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} \cdot Df(g(x)) + \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} \cdot u(g(x+h) - g(x)) \\
&= Dg(x)Df(g(x)),
\end{aligned}$$

  because $g(x+h)$ tends to $g(x)$, $u(k)$ is continuous and $u(0) = 0$.

- Let us assume that f is monotonically increasing and continuous in $(x - \epsilon, x + \epsilon)$. Then, with $y = f(x)$,

$$\begin{aligned}
\lim_{h\to 0} \frac{f^{-1}(y+h) - f^{-1}(y)}{h} &= \lim_{z\to y} \frac{f^{-1}(z) - f^{-1}(y)}{z - y} \\
&= \lim_{w\to x} \frac{f^{-1}(f(w)) - f^{-1}(f(x))}{f(w) - f(x)} \\
&= \lim_{w\to x} \frac{w - x}{f(w) - f(x)} = \frac{1}{Df(x)},
\end{aligned}$$

where in the second equality we used the change of variables $z = f(w)$. The case where f is monotonically decreasing is analogous.

□

Example 98. • Let $f(x) = x^4 + 3x^2 - 34$. Then $Df(x) = 4x^3 + 6x$.

- Let $f(x) = \frac{x^2+1}{x-2}$. Then, for $x \neq 2$, $Df(x) = \frac{2x(x-2)-(x^2+1)\cdot 1}{(x-2)^2} = \frac{x^2-4x-1}{(x-2)^2}$.
- Let $f(x) = \sin x, g(x) = x^2$. By linearity, $D(\sin x + x^2) = \cos x + 2x$.

  By Leibniz rule, $D(x^2 \sin x) = 2x \sin x + x^2 \cos x$.

  Let us take the composition $\sin(x^2) = f(g(x))$. By the chain rule, $D(\sin(x^2)) = D(x^2) \cdot (D\sin)(x^2) = 2x \cdot \cos(x^2)$. For $(\sin x)^2 = g(f(x))$, $D((\sin x)^2) = D(\sin x) \cdot 2(\sin x) = 2 \sin x \cos x$.
- By the chain rule, $D(\exp(-x)) = D(-x) \cdot (D\exp)(-x) = -\exp(-x)$. By linearity, $D\sinh x = D(\frac{1}{2}(e^x - e^{-x})) = \frac{1}{2}(e^x + e^{-x}) = \cosh x$. Analogously, $D\cosh x = \sinh x$.
- For $a > 0$, it holds that $a^x = (e^{\log a})^x = e^{\log a \cdot x}$. Indeed, by the chain rule,

  $$D(a^x) = D(\exp(\log a \cdot x)) = D(\log a \cdot x) \cdot (D\exp)(\log a \cdot x) = \log a \cdot \exp(\log a \cdot x) = \log a \cdot a^x.$$
- Let $a > 0$ and $f(x) = x^a$ for $x > 0$. $f(x) = \exp(\log x \cdot a)$, and by the chain rule,

  $$Df(x) = D(\log x \cdot a)D(\exp)(\log x \cdot a) = \frac{a}{x} \cdot \exp(\log x \cdot a) = \frac{a}{x} \cdot x^a = ax^{a-1}.$$

  For $a < 0$, we consider $f(x) = x^a = \frac{1}{x^a}$ and we obtain the same formula $f'(x) = ax^{a-1}$. For $a = 0$, because $x^a = 1$, we have $D(x^0) = D(1) = 0$.
- $D\tan x = D(\frac{\sin x}{\cos x}) = \frac{\cos x \cdot \cos x - \sin x \cdot (-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x}$.
- $f(y) = \arctan y$. That is, $f(y) = g^{-1}(y)$, where $g(x) = \tan x$ restricted to $[-\frac{\pi}{2}, \frac{\pi}{2}]$. By the formula for the inverse function, we have $Df(y) = \frac{1}{Dg(x)} = \cos^2 x$, where $y = g(x) = \tan x$. Therefore, $y^2 = \frac{\sin^2 x}{\cos^2 x} = \frac{1-\cos^2 x}{\cos^2 x}$, and $\cos^2 x = \frac{1}{1+y^2}$. By substituting this in the previous result, $D\arctan y = Df(y) = \frac{1}{1+y^2}$.
- $f(x) = \tanh x$. $f'(x) = \frac{1}{\cosh^2 x}$.
- $f(x) = \arcsin x$ (the inverse function of $\sin x$ restricted to $[-\frac{\pi}{2}, \frac{\pi}{2}]$). $f'(x) = \frac{1}{\sqrt{1-x^2}}$.

Nov. 3. Meaning of derivative, some applications.

Tangent line

We defined derivative as the limit of average slope of a graph, and expected that it should represent the slope at one point. If we have the slope at one point, then we should be able to draw the tangent line to the graph at that point.

Recall that the **slope** of a segment (x_0, y_0)–(x_1, y_1) is defined by $\frac{y_1-y_0}{x_1-x_0}$. The graph of $y = Ax+B$ has the slope A. Therefore, if the graph of the function $y = f(x)$ passes the point (x_0, y_0) and the derivative is $f'(x_0)$, the tangent line should be

$$y = f'(x_0)(x - x_0) + y_0 = f'(x_0)x + y_0 - f'(x_0)x_0.$$

Indeed, this is of the form $y = Ax + B$ with $A = f'(x_0)$ and $B = y_0 - f'(x_0)x_0$, and passes the point (x_0, y_0).

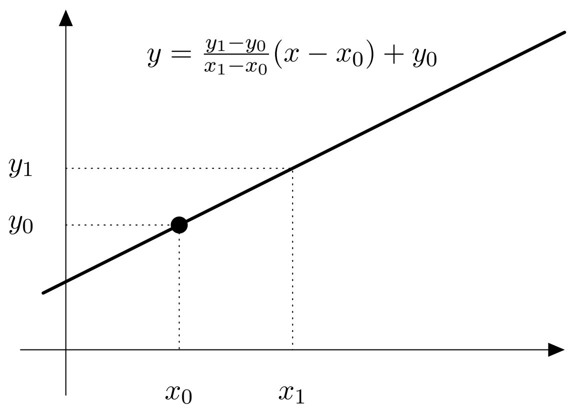


Figure 45: The slope of the straight line is $\frac{y_1-y_0}{x_1-x_0}$.

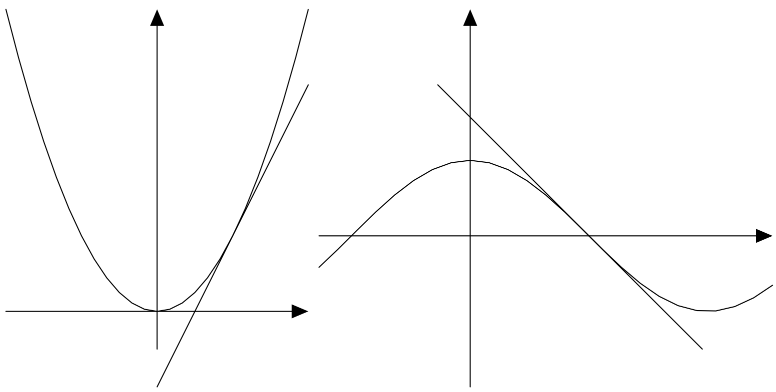


Figure 46: The tangent lines to the graphs of $x^2, \cos x$. Their equations are $y = 2(x-1)+1, y = -(x-\frac{\pi}{2})$, respectively.

When the slope is positive, the line goes upwards (when one goes to the right), while the line goes downwards when the slope is negative. When the slope is 0, it is a holizontal line. The vertical line is represented by the equation $x = a$, and this is not of the form $y = Ax + B$.

If we draw these lines, they are almost always indeed tangent, but in some cases they closs the graph.

Extrema and stationary points

Definition 99. Let $x \in \mathbb{R}$. For $\epsilon > 0$, we call the interval $(x-\epsilon, x+\epsilon)$ the **ϵ-neighbourhood of** x.

Let f be defined on an interval I. We say that f takes a **local minimum** or **relative minimum** (**local maximum** or **relative maximum** , respectively) at $x \in I$ if there is an $\epsilon > 0$ of x such that x is the minimum (maximum, respectively) of f in $(x-\epsilon, x+\epsilon) \cap I$.

If x is the minimum (maximum) of f on I we may say that x is the **global** or **absolute maximum** (**minimum**), to distinguish them from local (relative) minimum (maximum).

Example 100. Let $f(x) = x^3 - x$. When we consider this as a function on $\mathbb{R}$, there is no global maximum or mininum, but there are local maximum and minimum at $x = \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$, respectively (we will see why they are these points later). If we restrict the funcion to $[-2, 2]$, then $-2, 2$ are the global minimum and the global maximum, respectively.

Theorem 101. *Let f be defined on an open interval I and assume that f takes a local minimum (or a local maximum) at the point $c \in I$. If f is differentiable at c, then $f'(c) = 0$.*

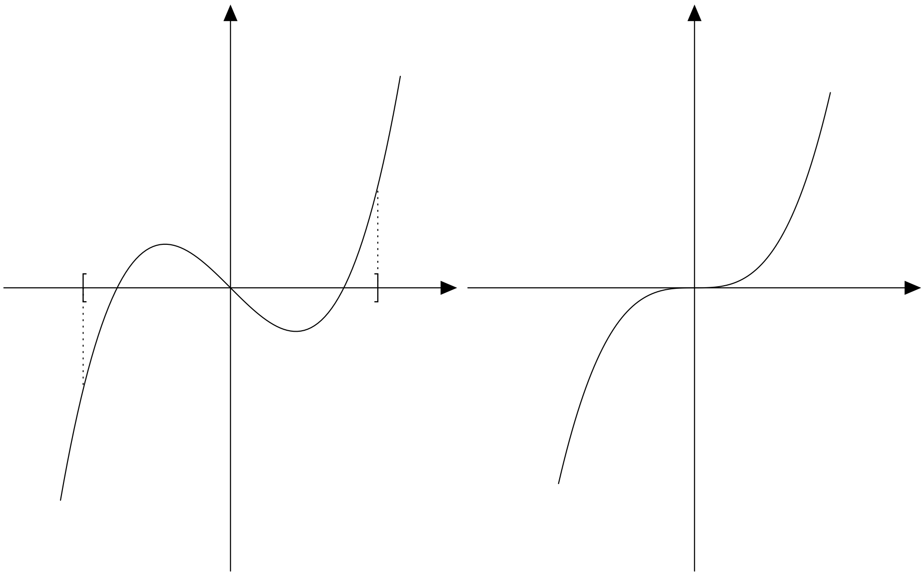


Figure 47: Left:The graph of $y = x^3 - x$. The local maximum and minimum are $x = \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$, respectively. When restricted to a closed bounded interval, it has global maximum and minimum. Right: The graph of $y = f(x) = x^3$. $f'(x) = 3x^2$, hence $x = 0$ is a stationary point.

Proof. Let c be a local maximum (the case for minimum is analogous). Then $f(x) \leq f(c)$ for all $x \in (c - \epsilon, c + \epsilon)$. As $f(x)$ is differentiable at $x = c$, both of its left and right derivatives must coincide.

On the other hand, $\lim_{h \to 0^+} \frac{f(c+h)-f(c)}{h} \leq 0$, and $\lim_{h \to 0^-} \frac{f(c+h)-f(c)}{h} \geq 0$, therefore, $f'(c) = 0$. □

Definition 102. A local minimum or a local maximum of a function f is called an **extremum**. A point x where $f'(x) = 0$ holds is called a **stationary point**.

Any extremum of a differentiable function is a stationary point by Theorem 101, but a stationary point is not necessarily an extremum.

Example 103.

- $y = f(x) = x^3$. Then $f'(x) = 3x^2$, hence $x = 0$ is a stationary point, but as $f(x)$ is monotonically increasing, it is not an extremum.
- $y = f(x) = x^3 - x$. Then $f'(x) = 3x^2 - 1$, hence $x = \pm\frac{1}{\sqrt{3}}$ are stationary points. They are local maximum and minimum, respectively.
- $y = |x|$. This function has the minimum at $x = 0$, but the function does not have derivative there. In particular, it does not hold $f'(0) = 0$ there ($f'(0)$ has no meaning there).

Concrete situation of composed function

Imagine that we have a balloon and a gas is pumped into it at a rate of $50\text{cm}^3/\text{s}$. If the pressure remains constant, how fast is the radius of the balloon increasing when the radius is 5cm?

- The volume $V(t)$ of the balloon at time t (second): $V(t) = 50t\text{cm}^3$. This implies $\frac{dV}{dt} = 50\text{cm}^3/\text{s}$.
- The radius $r(t)$ of the sphere with volume $V(t)$: $\frac{4\pi r(t)^3}{3} = V(t)$, By differentiating both sides by t, $4\pi\frac{dr}{dt}(t)r(t)^2 = \frac{dV}{dt}$.
- By solving this with $r(t_0) = 5$, $\frac{dr}{dt}(t_0) = \frac{50}{4\pi 5^2} = \frac{1}{2\pi}$.

Some shape can be represented by an equation, and the equation may define a function **implicitly**. For example, we know that the circle centered at $(0,0)$ with radius r is given by

$$x^2 + y^2 = r^2.$$

As we saw before, if we consider only the part $y \geq 0$, it defines the function $y = \sqrt{r^2 - x^2}$.

It is not always possible to find an **explicit** expression for y of a given equation. Yet, an equation may define a function in an abstract way. Let us write it $y(x)$.

With the explicit expression, $y(x) = \sqrt{r^2 - x^2} = (r^2 - x^2)^{\frac{1}{2}}$, therefore,

$$y'(x) = \frac{1}{2}\frac{-2x}{\sqrt{r^2 - x^2}} = -\frac{x}{\sqrt{r^2 - x^2}}.$$

It holds that $y(x)y'(x) = -x$.

This last relation can be also derived as follows: by taking the derivative of $x^2 + y(x)^2 = r^2$, we obtain $2x + 2y(x)y'(x) = 0$, hence $y(x)y'(x) = -x$.

If we know some concrete values of y, x (even if we do not know the general formula), then we can compute $y'(x)$ at that point.

The inverse trigonometric functions

$\sin x, \cos x, \tan x$ are injective on certain domains, and hence have the inverse functions. The standard choices are the following.

- $\sin x$: consider the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The range is $[-1, 1]$. The inverse function is denoted by $\arcsin x$, defined on $[-1, 1]$.
- $\cos x$: consider the interval $[0, \pi]$. The range is $[-1, 1]$. The inverse function is denoted by $\arccos x$, defined on $[-1, 1]$.
- $\tan x$: consider the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The range is $\mathbb{R}$. The inverse function is denoted by $\arctan x$, defined on $\mathbb{R}$.

Let us compute the derivative of $\arcsin y$ by putting $y = \sin x$. Then $D(\sin x) = \cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2}$. By the general formula, $D(\arcsin y) = \frac{1}{D(\sin x)} = \frac{1}{\cos x} = \frac{1}{\sqrt{1-y^2}}$.

Nov. 3 (14:00). More applications of derivative.

Theorem 104 (Rolle). *Let f be continuous in $[a, b]$ and differentiable in (a, b). If $f(a) = f(b)$, then there is $x_0 \in (a, b)$ such that $f'(x_0) = 0$.*

Proof. If f is constant, then $f'(x) = 0$ for all $x \in (a, b)$.

If f is not constant, then by Theorem 86 of Weierstrass, f has a minimum and a maximum. As f is not constant, one of them must be different from $f(a) = f(b)$. Therefore, we take x_0 that is either minimum or maximum, and $a \neq x_0 \neq b$. Let us take an open interval containing x_0. Now by Theorem 101, $f'(x_0) = 0$. □

Proposition 105 (Lagrange's mean value theorem)**.** *Let f be continuous in $[a, b]$ and differentiable in (a, b). Then there is $x_0 \in (a, b)$ such that $\frac{f(b)-f(a)}{b-a} = f'(x_0)$.*

Proof. Let $g(x) = f(x) - \frac{(f(b)-f(a))x}{b-a}$, which is continuous in $[a, b]$ and differentiable in (a, b). Then $g(a) = \frac{f(a)b-f(b)a}{b-a} = g(b)$, and by Theorem 104 there is x_0 such that $g'(x_0) = 0$. This implies $f'(x_0) - \frac{f(b)-f(a)}{b-a} = 0$. □

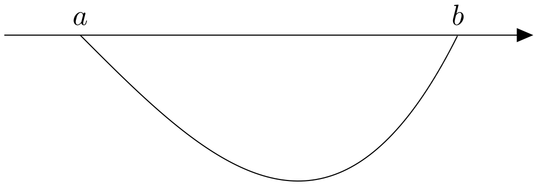


Figure 48: A non constant function, continuous in $[a, b]$ and differentiable in (a, b), must have a stationary point.

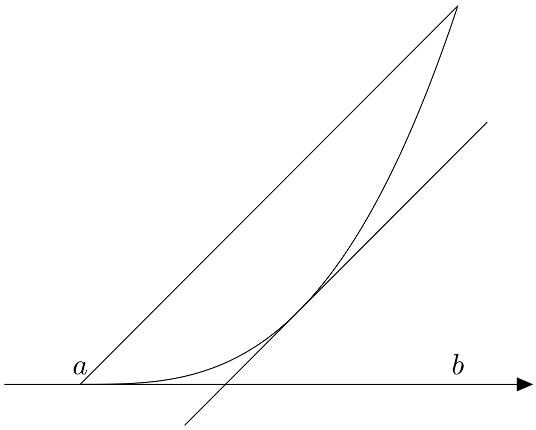


Figure 49: A function continuous in $[a, b]$ and differentiable in (a, b), must have a point where the deriative is equal to the mean slope.

Corollary 106. *Let f be continuous in $[a, b]$ and differentiable in (a, b). If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant.*

Proof. Let $x < y \in [a, b]$. By Theorem 105, there is $x_0 \in (x, y)$ such that $\frac{f(y)-f(x)}{y-x} = f'(x_0) = 0$, therefore, $f(x) = f(y)$. □

Corollary 107. *Let f be continuous in $[a, b]$ e derivabile in (a, b).*

- *If $f'(x) \geq 0$ (> 0, respectively) for all $x \in (a, b)$, then f is monotonically non decreasing (increasing, respectively).*
- *If $f'(x) \leq 0$ (< 0, respectively) for all $x \in (a, b)$, then f is monotonically non increasing (decreasing, respectively).*

Proof. Let $x < y \in (a, b)$. By Theorem 105, there is $x_0 \in (x, y)$ such that $\frac{f(y)-f(x)}{(y-x)} = f'(x_0)$. If $f'(x_0) \geq 0 (> 0)$, then $f(y) - f(x) \geq 0 (> 0)$, that is f is monotonically non decreasing (increasing, respectively).

The case $f'(x) \leq 0 (< 0)$ is analogous. □

Example 108.
- $f(x) = x^2$. $f'(x) = 2x$, hence f is decreasing if $x < 0$, $x = 0$ is the only one stationary point, and is increasing if $x > 0$.
- $f(x) = \sin x$. $f'(x) = \cos x$, hence f is increasing if $x \in (-\frac{\pi}{2} + 2\pi n, \frac{\pi}{2} + 2\pi n)$ for $n \in \mathbb{Z}$, $x = \frac{\pi}{2} + 2\pi n, -\frac{\pi}{2} + 2\pi n$ are stationary points, and f is decreasing if $x \in (\frac{\pi}{2} + 2\pi n, \frac{3\pi}{2} + 2\pi n)$.

Theorem 109. *Let f be continuous in $[a, b]$ and differentiable in (a, b). Let $c \in (a, b)$.*

- *If $f'(x) > 0$ for $x \in (a, c)$ and $f'(x) < 0$ for $x \in (c, b)$, then f has a maximum at c.*
- *If $f'(x) < 0$ for $x \in (a, c)$ and $f'(x) > 0$ for $x \in (c, b)$, then f has a minimum at c.*

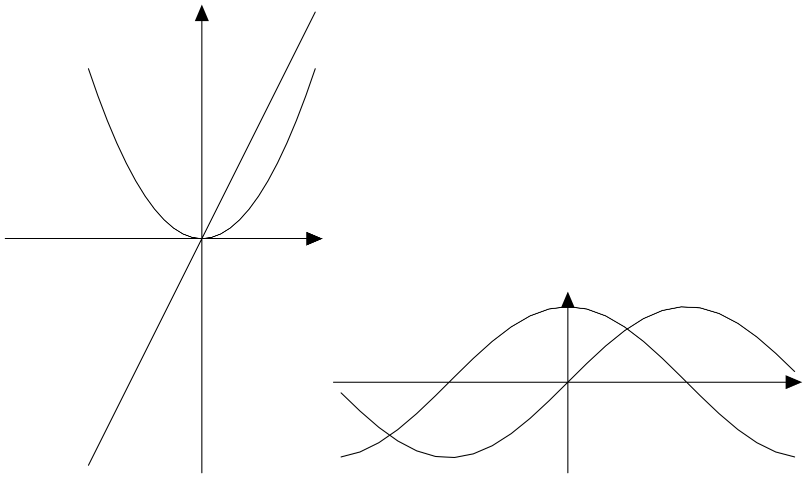


Figure 50: A function and its derivative. When the derivative is positive (negative) in an interval, the function is increasing (decreasing).

Proof. If $f'(x) > 0$ for $x \in (a, c)$, then it is increasing there and continuous at c, therfore, for any $x \in (a, b)$ it holds that $f(c) \geq f(x)$. On the other hand, as $f'(x) < 0$ for $x \in (c, b)$, and $f(c) \geq f(x)$ for $x \in (c, b)$.

The second case is analogous. $\square$

Example 110.

- $f(x) = x^3 - x$. $f'(x) = 3x^2 - 1$, and $f'(x) > 0$ if and only if $x < -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} < x$, and $f'(x) < 0$ if and only if $-\frac{1}{\sqrt{3}} < x < \frac{1}{\sqrt{3}}$. Therefore, x takes a local maximum at $x = -\frac{1}{\sqrt{3}}$ and a local minimum $x = \frac{1}{\sqrt{3}}$. As f is differentiable in $\mathbb{R}$, there is no other local maximum or minimum.
- $f(x) = \cosh x$. $f'(x) = \sinh x$, and $f'(x) > 0$ if and only if $x > 0$, and $f'(x) < 0$ if and only if $x <$. Therefore, x takes a minimum at $x = 0$ and no other minimum or maximum.
- $f(x) = \sinh x$. $f'(x) = \cosh x$ and $\cosh x > 0$, and hence $f(x)$ is monotonically increasing.

Note that, even if $f'(x) > 0$ at one point, it does not mean that f is monotonically increasing in a neighbourhood of x. Indeed, a counterexample is given by

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) + \frac{x}{2} & \text{for } x \neq 0 \\ \frac{1}{2} & \text{for } x = 0 \end{cases}.$$

As we have seen, this function without the part $\frac{x}{2}$ is differentiable, and it has the derivative 0 at $x = 0$. Therefore, with $\frac{x}{2}$, it is still differentiable and $f'(0) = \frac{1}{2} > 0$.

Yet, f is not monotonically increasing in any interval $(-\epsilon, \epsilon)$. To see this, note that

$$f'(x) = \begin{cases} 2x \sin\left(\frac{1}{x}\right) - \frac{x^2}{x^2} \cos\left(\frac{1}{x}\right) + \frac{1}{2} = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) + \frac{1}{2} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}.$$

and for any $\epsilon > 0$, there is $x < \epsilon$ such that $f'(x) < 0$: for example, one can take $x = \frac{1}{2\pi n}$ for sufficiently large n. Then the term $2x \sin \frac{1}{x} = 0$, while $-\cos \frac{1}{x} = -1$, and then $f'(x) = -\frac{1}{2}$.

Note that the derivative $f'(x)$ is discontinuous in this case.

Proposition 111 (Cauchy's mean value theorem)**.** *Let $a < b$, f, g be continuous in $[a, b]$ and differentiable in (a, b). Then there is $x_0 \in (a, b)$ such that $f'(x_0)(g(b) - g(a)) = g'(x_0)(f(b) - f(a))$.*

Proof. Let $h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a))$, then $h(x)$ is continuous in $[a, b]$ and differentiable on in (a, b). $h(a) = f(a)g(b) - f(b)g(a) = h(b)$. By Rolle's theorem 104, there is $x_0 \in (a, b)$ such that $0 = h'(x_0) = f'(x_0)(g(b) - g(a)) - g(x_0)(f(b) - f(a))$. $\square$

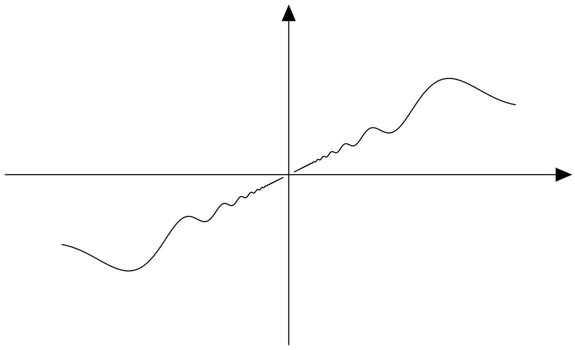


Figure 51: A function f such that $f'(0) > 0$ but is not monotonically increasing in any interval containing $x = 0$.

Nov. 4. Higher derivatives, convexity and concavity, asymptotes.

Higher derivatives

As we saw before, if f is defined on an open interval and is differentiable on each point of I, then f' defines a new function on I, the (first) derivative. It may happen that f' is again differentiable on each point of I, and it defines a further new function f'', the **second derivative**. If f'' is again differentiable, one can also define the third derivative, and so on. We denote the n-th derivative by $f^{(n)}$, or $D^n f$, $\frac{d^n f}{dx^n}$.

Example 112.
- If $f(x) = x^4$, then $f'(x) = 4x^3$, $f''(x) = 12x^2$, $f^{(3)}(x) = 24x$, and so on. This f is infinitely many times differentiable.
- If $f(x) = \sin x$, then $f'(x) = \cos x, f''(x) = -\sin x, f^{(3)}(x) = -\cos x, f^{(4)}(x) = \sin x$ and so on. Again this is infinitely many times differentiable.
- Let $f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) + \frac{x}{2} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$, then we have

$$f'(x) = \begin{cases} 2x \sin\left(\frac{1}{x}\right) - \frac{x^2}{x^2}\cos\left(\frac{1}{x}\right) + \frac{1}{2} & \\ = 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) + \frac{1}{2} & \text{for } x \neq 0 \\ \frac{1}{2} & \text{for } x = 0 \end{cases}$$

  and this derivative is not continuous. In particular, f is only once differentiable.

The second derivative is useful to study whether the stationary point (or a **critical point**) is a maximum or a minimum, and also to study the shape of the graph.

Lemma 113. *Suppose that f is differentiable in an open interval I and at x_0 it is twice differentiable.*

- *If x_0 is a stationary point and $f''(x_0) > 0$ ($f''(x_0) < 0$, respectively), then f takes a local minimum (a local maximum, respectively) at x_0 .*
- *If x_0 is a local minimum (a local maximum, respectively), then $f''(x_0) \geq 0$ ($f''(x_0) \leq 0$, respectively).*

Proof.
- Let $f''(x_0) > 0$. Then there is $\epsilon > 0$ such that $\frac{f'(x_0+h)-f'(x_0)}{h} = \frac{f'(x_0+h)}{h} > 0$ for $|h| < \epsilon$. This means that $f'(x_0 + h) > 0$ for $h > 0$ and $f'(x_0 + h) < 0$ for $h < 0$, and hence f is monotonically decreasing in $(x_0 - \epsilon, x_0)$ and increasing in $(x_0, x_0 + \epsilon)$, that is, f takes a minimum at x_0.

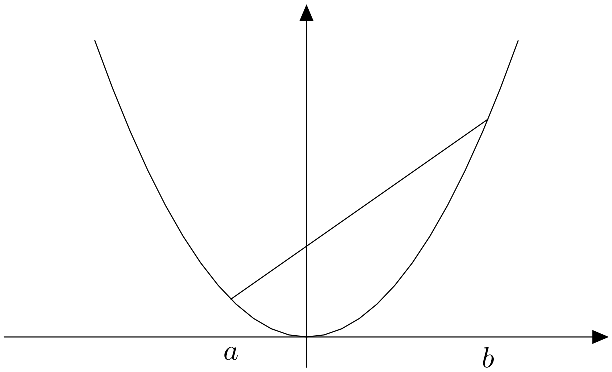


Figure 52: A convex function. The graph is below the segment between any pair of points $(a, f(a)), (b, f(b))$.

- If x_0 a local minimum and suppose that $f''(x_0) < 0$, then x_0 would be a local maximum and it would contradict the previous point.
- Other cases are analogous.

□

Example 114.
- Let $f(x) = x^2$. We have $f'(x) = 2x$ and $x = 0$ is a stationary point. As $f''(x) = 2$, x is a minimum.
- Let $f'(x) = x^3 - 3x$. We have $f'(x) = 3x^2 - 3$ and $x = 1, -1$ are stationary points. As $f''(x) = 6x$, f takes a maximum at $x = -1$ and a minimum at $x = 1$.

Convexity and concavity

Note that, for $a, b \in \mathbb{R}$ and $t \in [0, 1]$. Then $ta + (1 - t)b$ is a point between a, b. Indeed, if $a < b$, then $a = ta + (1 - t)a < ta + (1 - t)b < tb + (1 - t)b = b$ (the case $b < a$ is analogous).

Definition 115. Let f be defined on an interval I. We say that f is convex (concave, respectively) if for any $a, b \in I$ and $t \in [0, 1]$ it holds that

$$f(ta + (1 - t)b) \le tf(a) + (1 - t)f(b) \text{ (respectively } f(ta + (1 - t)b) \ge tf(a) + (1 - t)f(b)).$$

Note that $(ta + (1 - t)b, tf(a) + (1 - t)f(b))$ defines a segment between $(a, f(a))$ and $(b, f(b))$. Indeed, the slope from the point $(a, f(a))$ to such a point is $\frac{(1-t)(f(b)-f(a))}{(1-t)(b-a)} = \frac{f(b)-f(a)}{b-a}$, which does not depend on t.

Theorem 116. *Assume that f is continuous on $[a, b]$, differentiable on (a, b). If f' is monotonically nondecreasing (nonincreasing, respectively), then f is convex (concave, respectively). In particular, if $f''(x) > 0$ ($f''(x) < 0$, respectively) for $x \in (a, b)$, then f is convex (concave, respectively).*

Proof. Let $x < y$ in $[a, b]$ and $t \in (0, 1)$. Let $z = tx + (1 - t)y$. We have to prove that $f(z) \le tf(x) + (1 - t)f(y)$, or equivalently, $t(f(z) - f(x)) \le (1 - t)(f(y) - f(z))$.

By Theorem 105, there are points c, d such that $x < c < z$ and $z < d < y$ such that $f(z) - f(x) = f'(c)(z - x)$ and $f(y) - f(z) = f'(d)(y - z)$. As f' is nondecreasing, $f'(c) \le f'(d)$ and hence, using $t(z - x) = (1 - t)(y - z)$,

$$t(f(z) - f(x)) = tf'(c)(z - x) \le f'(d)t(z - x) = f'(d)(1 - t)(y - z) = (1 - t)(f(y) - f(z)).$$

□

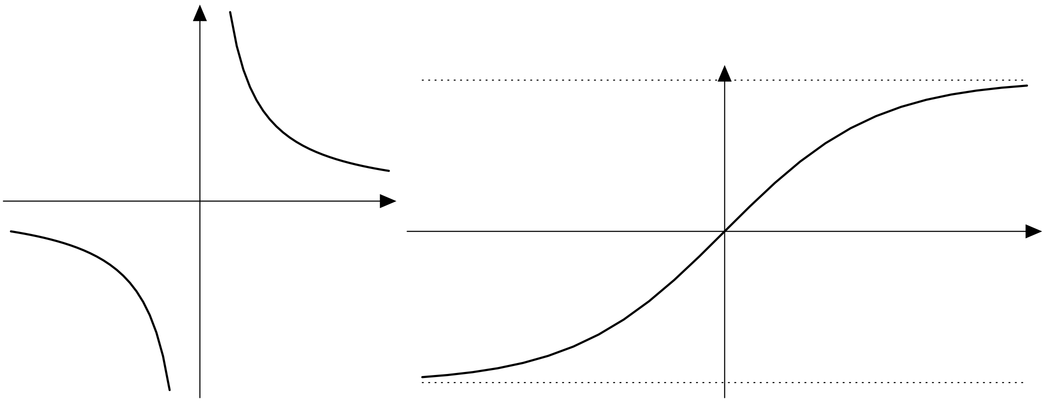


Figure 53: The asymptotes for $\frac{1}{x}$ and $\tanh x$.

Under certain conditions, it can also be shown that f is convex, then $f'' > 0$. We omit the proof.

Example 117.
- Let $f(x) = x^2$. As $f''(x) = 2$, f is convex.
- Let $f'(x) = x^3$. As $f''(x) = 6x$, f is concave on $(-\infty, 0)$ and convex on $(0, \infty)$.

Asymptotes

The graph of some function may approach a straight line. A more precise concept of this is asymptotes.

Definition 118.
- Let f be defeind on (a, ∞). If $\lim_{a\to\infty} f(x) = L$, then we say that $y = L$ is a **horizontal asymptote** (analogous for $-\infty$).
- Let f be defined on (a, b). If $\lim_{x\to a^+} |f(x)| \to \infty$, then $x = a$ is called a **vertical asymptote**. (analogous for b).
- Let f be defined on (a, ∞). If there is $A, B \in \mathbb{R}$ such that $\lim_{x\to\infty} \frac{f(x)}{x} = A$ and $\lim_{x\to\infty} f(x) - Ax = B$, then we say that $y = Ax + B$ is an **oblique asymptote** (analogous for $-\infty$).

Example 119.
- Let $f(x) = \tanh x$. We know that $\lim_{x\to\infty} \tanh x = 1, \lim_{x\to-\infty} \tanh x = -1$, hence $y = 1, -1$ are the horizontal asymptotes of $\tanh x$.
- Let $f(x) = \frac{1}{x}$ on $(-\infty, 0) \cup (0, \infty)$. We know that $\lim_{x\to 0^+} \frac{1}{x} = \infty$, $\lim_{x\to 0^-} \frac{1}{x} = -\infty$, and hence $x = 0$ is a vertical asymptote of $\frac{1}{x}$. $y = 0$ is a horizontal asymptote of $\frac{1}{x}$ because $\lim_{x\to\pm\infty} \frac{1}{x} = 0$.
- Let $f(x) = x \tanh x$. Then, we see that $\lim_{x\to\infty} \frac{x \tanh x}{x} = \lim_{x\to\infty} \tanh x = 1$ and

$$\begin{aligned} \lim_{x\to\infty} x \tanh x - x &= \lim_{x\to\infty} x \left(\frac{e^x - e^{-x}}{e^x + e^{-x}} - 1 \right) \\ &= \lim_{x\to\infty} \frac{-2xe^{-x}}{e^x + e^{-x}} = 0, \end{aligned}$$

hence $y = x$ is an oblique asymptote. Similarly, $y = -x$ an oblique asymptote for $x \to -\infty$.

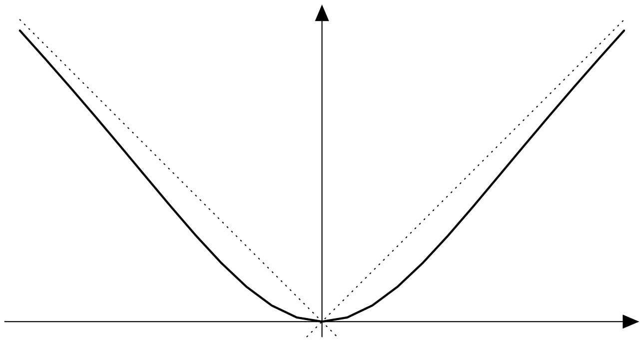


Figure 54: The oblique asymptotes for $x \tanh x$.

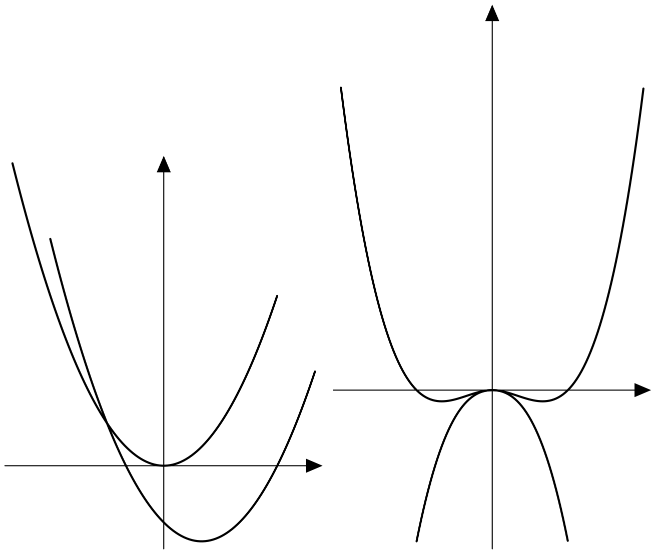


Figure 55: Left:the graphs of x^2 and $(x - \frac{1}{2})^2 - 1$. Right:the graphs of $x^3 - x^2$ and $-x^3 - x^2$.

Nov. 10. Curve sketching

Symmetry of functions

Recall that a function is a subset in $\mathbb{R} \times \mathbb{R}$ in the sense that it collects all the points $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = f(x)\}$. This is the graph itself.

We can consider certain operations on a function.

- **Translation**. If $g(x) = f(x-a)+b$ for some function f, g, then the graph of g is obtained by translating the graph of f by (a, b). Indeed, if (x, y) is on the graph of f, then $(x+a, y+b)$ is on the graph of g.
- **Reflection**. If $g(x) = f(-x)$ for some function f, g, then the graph of g is obtained by reflecting the graph of f with respect to $x = 0$. Indeed, if (x, y) is on the graph of f, then $(-x, y)$ is on the graph of g.
- If $g(x) = f(-(x - 2a))$ for some function f, g, then the graph of g is obtained by reflecting the graph of f with respect to $x = a$.
- **Scaling**. If $g(x) = bf(x/a)$ for some function f, g and $a, b > 0$, then the graph of g is obtained by scaling the graph of f by a in the x-direction and b in the y-direction. Indeed, if (x, y) is on the graph of f, then (ax, by) is on the graph of g.

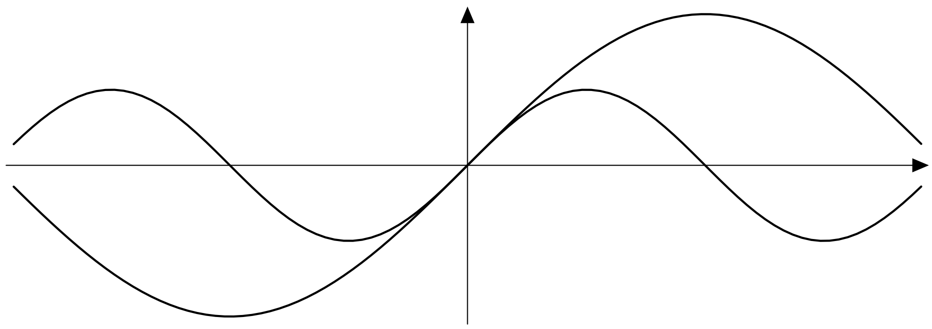


Figure 56: The graphs of $\sin x$ and $2\sin(x/2)$.

A graph or a function may have a **symmetry**. A function f is said to have a symmetry if it is invariant under certain operations.

- **Translation symmetry**. If $f(x) = f(x-a)$, then the graph of f remains invariant under the translation $(a, 0)$.
- **Reflection**. If $f(x) = f(-x)$, then the graph of f is invariant under the reflection respect to $x = 0$ and f is said to be even.
- $f(x) = -f(-x)$, f is said to be **odd**.
- $f(x) = f(-(x-2a))$ has reflection symmetry with respect to $x = a$.

Example 120.

- The graph of $\sin x$ is invariant under 2π translation and under the reflection with respect to $\frac{\pi}{2}$, because $\sin(x+2\pi) = \sin(x)$ and $\sin(-(x-\pi)) = -\sin(x-\pi) = \sin(x)$. On the other hand, $\sin(-x) = -\sin x$, hence $\sin x$ is an odd function.
- If $f(x) = (x-\frac{1}{2})^2 - 1$ is invariant under the reflection with respect to $x = \frac{1}{2}$ because $((-(x-1)) - \frac{1}{2})^2 - 1 = (-x+\frac{1}{2})^2 - 1 = (x-\frac{1}{2})^2 - 1$.

Curve sketching

The graph of a function f can be **qualitatively** drawn as follows.

(0) Determine the (natural) domain A of definition of f.

(0.5) Check if f has a symmetry or a period.

(1) Study the sign of f: where $f(x) > 0, = 0, < 0$ hold.

(2) Determine the asymptotes.

(3) Study the sign of f' and find stationary points (where $f'(x) = 0$).

(4) Study the stationary points and find local minima and maxima (either by the second derivative or the first).

Example 121.

- $f(x) = e^{-(2x-1)^2}$.

  (0) $f(x)$ is defined for all $x \in \mathbb{R} = A$ in a natural way.

  (0.5) $f(x+\frac{1}{2}) = f(-x+\frac{1}{2})$, that is $f(x)$ is even with respect to $x = \frac{1}{2}$.

  (1) $e^{-(2x-1)^2} > 0$ for all $x \in \mathbb{R}$.

  (2) Consider $x \to \pm\infty$. $\lim_{x\to\pm\infty} f(x) = 0$. The asymptote is $y = 0$.

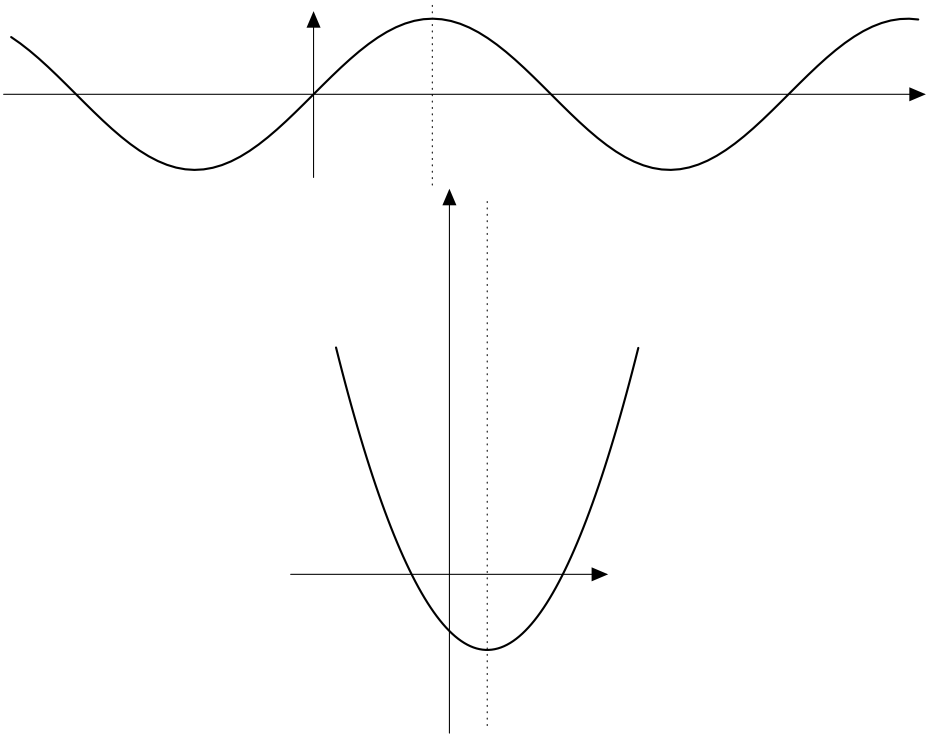


Figure 57: Above:The graph of $\sin x$ is invariant under 2π translation and under the reflection with respect to $\frac{\pi}{2}$. Below:The graph of $(x - \frac{1}{2})^2 + 1$ is invariant under the reflection with respect to $x = \frac{1}{2}$.

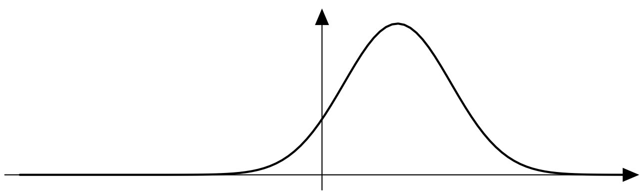


Figure 58: The graph of $f(x) = e^{-(2x-1)^2}$.

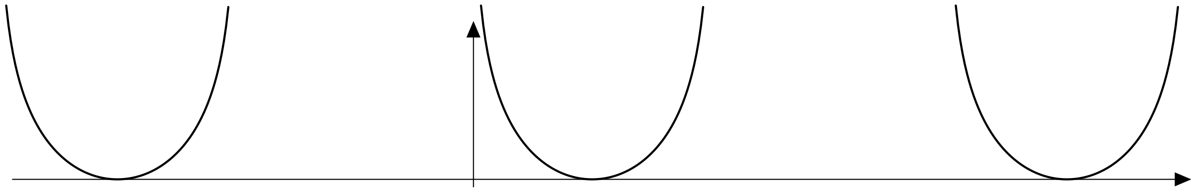


Figure 59: The graph of $f(x) = \log(\frac{1}{\sin x})$.

(3) $f'(x) = -4(2x-1)e^{-(2x-1)^2}$. $f'(x) = 0 \Leftrightarrow 2x-1 = 0 \Leftrightarrow x = \frac{1}{2}$. $f(\frac{1}{2}) = 1$.

(4) $f''(x) = (16(2x-1)^2 - 8)e^{-(2x-1)^2} = (64x^2 - 64x + 8)e^{-(2x-1)^2}$.

x		$\frac{1}{2}$	
$f'(x)$	$+$	0	$-$
$f''(x)$		$-$	
$f(x)$	$\nearrow$	1	$\searrow$

- $f(x) = \log(\frac{1}{\sin x})$.

(0) $\log y$ is defined for $y > 0$, hence $\frac{1}{\sin x} > 0$, that is $\sin x > 0 \Leftrightarrow x \in (2n\pi, (2n+1)\pi)$ for $n \in \mathbb{Z}$.

(0.5) $\sin(x + 2\pi) = \sin x$, hence $f(x + 2\pi) = f(x)$. It sufficed to draw the graph for $(0, \pi)$. Since $\sin(x + \frac{\pi}{2}) = \sin(-x + \frac{\pi}{2})$, $f(x)$ is even with respect to $x = \frac{\pi}{2}$.

(1) $0 < \sin x \leq 1$, hence $\frac{1}{\sin x} \geq 1$ and $\log(\frac{1}{\sin x}) \geq 0$.

(2) The domain is $(0, \pi)$, so we need to check $\{0, \pi\}$. $\lim_{x\to 0} f(x) = \lim_{x\to\pi} f(x) = \infty$. The asymptotes are $x = 0, \pi$.

(3) $f'(x) = -\frac{\cos x}{\sin x}$. $f'(x) = 0 \Leftrightarrow \cos x = 0 \Leftrightarrow x = \frac{\pi}{2}$. $f(\frac{\pi}{2}) = 0$. $f'(x) < 0$ if $x \in (0, \frac{\pi}{2})$, and $f'(x) > 0$ if $x \in (\frac{\pi}{2}, \pi)$.

(4) $f''(x) = \frac{1}{\sin^2 x} > 0$.

x		$\frac{\pi}{2}$	
$f'(x)$	$-$	0	$+$
$f''(x)$		$+$	
$f(x)$	$\searrow$	0	$\nearrow$

Solutions to equations

We can draw the graphs of $f(x) = 1 - x^2$ and $g(x) = e^x - 1$, and prove that there are two solutions of the equation $f(x) = g(x)$.

Indeed, let us consider the function $h(x) = g(x) - f(x) = e^x + x^2 - 2$ and it suffices to find all x such that $h(x) = g(x) - f(x) = 0$. We have $\lim_{x\to\pm\infty} g(x) - f(x) = \infty$ and $g(0) - f(0) = (1-1) - 1 = -1$. By the intermediate value theorem, there are solutions in $x > 0$ e $x < 0$. Moreover, $h'(x) = e^x + 2x$, hence there is only one stationary point (because in $x > 0$ $h'(x)$ is positive and it is negative for sufficiently small x, while $g''(x) - f''(x) = e^x + 2$ is positive, therefore, $g'(x) - f'(x)$ is monotonically increasing). Therefore, $h(x) = g(x) - f(x)$ is decreasing in a negative half line and is increasing in the rest, hence it can have only two points x where $h(x) = 0$.

Some applications of the minimum/maximum finding

If one can express a problem as a problem of finding the maximum or the minimum of a function, we can solve it using derivatives and graphs.

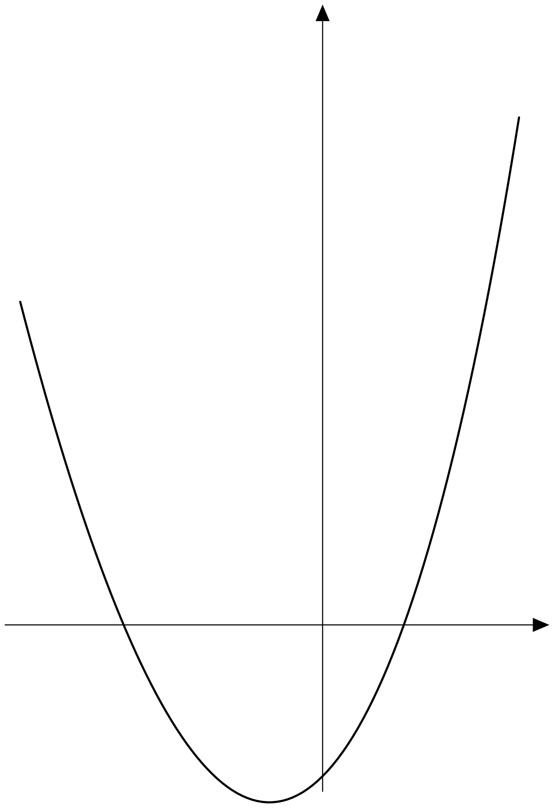


Figure 60: The graph of $h(x) = e^x + x^2 - 2$. It crosses the x-axis twice and only twice.

- Among all rectangles of given perimeter $2r$, which one has the largest area? Let the vertical side x, then $0 \le x \le r$ and the other side is $r - x$, hence the area is $x(r - x)$. We need to find the maximum of $f(x) = x(r - x)$ on the domain $\{x : 0 < x < r\}$. We have $\lim_{x\to 0} f(x) = \lim_{x\to r} f(x) = 0$, while $f'(x) = r - 2x$, and hence there is a stationary point at $x = \frac{r}{2}$, and $f''(x) = -2$, hence this is a local maximum. There is no other stationary points, and $f(0) = f(r) = 0$, hence this is the maximum.
- The geometric mean $\sqrt{ab}$ is smaller than or equal to the arithmetic mean $\frac{a+b}{2}$. Let us fix $P = \sqrt{ab}$ and put $a = x$, then $b = \frac{P^2}{x}$ and $0 < x$. Let us find the minimum of $f(x) = \frac{x+\frac{P^2}{x}}{2}$. This tends to ∞ as $x \to 0$ or $x \to \infty$. On the other hand, $f'(x) = \frac{1}{2}(1-\frac{P^2}{x^2})$, and hence there is only one stationary point at $x = P$, and $f''(x) = \frac{2P^2}{x^3}$, hence this is a local minimum, and is the minimum. At $x = P$, we have $f(P) = P$. Hence we have $P \le \frac{x+\frac{P^2}{x}}{2}$.

Nov. 10 (14:00). Theorem of Bernoulli-de l'Hôpital

Let us recall the mean value theorem of Cauchy: let f, g be continuous in $[a, b]$ and differentiable in (a, b). Then there is $x_0 \in (a, b)$ such that

$$f'(y)(g(b) - g(a)) = g'(y)(f(b) - f(a)).$$

(Bernoulli-)de l'Hôpital rule is a useful tool to compute limits of the type $\frac{0}{0}$ or $\frac{\infty}{\infty}$.

Theorem 122 (Bernoulli-de l'Hôpital, case 1)**.** *Let $a < x_0$, f, g differentiable in (a, x_0) such that $g'(x) \ne 0$ for x sufficiently close to $x_0, x \ne x_0$, $\lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^-} g(x) = 0$, $\lim_{x\to x_0^-} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}$. Then $g(x) \ne 0$ for x close to $x_0, x \ne x_0$ and $\lim_{x\to x_0^-} \frac{f(x)}{g(x)} = L$.*

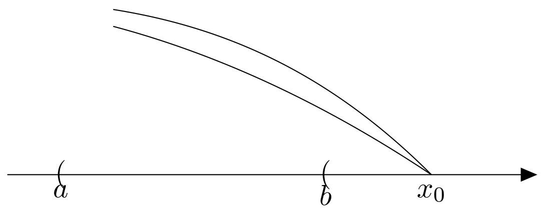


Figure 61: Theorem of de l'Hôpital. The limit $\lim_{x\to x_0^-}\frac{f(x)}{g(x)}$ is determined by $\lim_{x\to x_0}\frac{f'(x)}{g'(x)}$.

Proof. We can extend f, g to $(a, x_0]$ by putting $f(x_0) = g(x_0) = 0$, such that they are continuous. By the hypothesis we may assume that $g'(x) \neq 0$ in (b, x_0). Let $x \in (b, x_0)$, by Lagrange's mean value theorem, there is $y \in (x, x_0)$ such that $g(x) = g(x) - g(x_0) = g'(y)(x - x_0) \neq 0$, in particular, $g(x) \neq 0$.

By Cauchy's mean value theorem, for x above, there is $y \in (x, x_0)$ such that $f'(y)(g(x) - g(x_0)) = g'(y)(f(x) - f(x_0))$, that is,

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(y)}{g'(y)}.$$

If $x \to x_0$, such y tends to x_0. Because $\lim_{y\to x_0^-}\frac{f'(y)}{g'(y)} = L$ by the hypothesis it holds that $\lim_{x\to x_0^-}\frac{f(x)}{g(y)} = L$. □

A similar result holds for right limits.

Example 123.
- Consider $\frac{e^x-1}{\sin(2x)}$. The limit $x \to 0$ is of the form $\frac{0}{0}$. It holds that $(\sin(2x))' = 2\cos(2x) \neq 0$ as $x \to 0$. In addition $(e^x - 1)' = e^x$. Therefore, $\lim_{x\to 0}\frac{e^x-1}{\sin x} = \lim_{x\to 0}\frac{e^x}{2\cos(2x)} = \frac{1}{2}$.
- $\lim_{x\to 0}\frac{x}{e^x-1} = \frac{1}{e^0} = 1$.
- $\lim_{x\to 0}\frac{x^2}{\cos x-1} = \lim_{x\to 0}\frac{2x}{-\sin x} = \frac{2}{-\cos 0} = -2$.

Theorem 124 (Bernoulli-de l'Hôpital, case 2)**.** *Let f, g differentiable in (a, ∞) such that $g'(x) \neq 0$ for x sufficiently large, $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = 0$, $\lim_{x\to\infty}\frac{f'(x)}{g'(x)} = L$. Then $g(x) \neq 0$ for x sufficiently large and $\lim_{x\to\infty}\frac{f(x)}{g(x)} = L$.*

Proof. Let $F(x) = f(\frac{1}{x}), G(x) = g(\frac{1}{x})$. Note that, as $x \to \infty$, we have $\frac{1}{x} \to 0^+$, and $F'(x) = -\frac{1}{x^2}f'(\frac{1}{x}), G'(x) = -\frac{1}{x^2}g'(\frac{1}{x})$. Then for sufficiently small x, $G'(x) \neq 0$ because $g'(\frac{1}{x}) \neq 0$ for such x. By applying case 1, we obtain

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = \lim_{x\to 0^+}\frac{F(x)}{G(x)} = \lim_{x\to 0^+}\frac{F'(x)}{G'(x)} = \lim_{x\to 0^+}\frac{-x^2 f'(\frac{1}{x})}{-x^2 g'(\frac{1}{x})} = \lim_{x\to\infty}\frac{f'(x)}{g'(x)}$$

as desired. □

Example 125.
- $\lim_{x\to\infty}\frac{\sin(\frac{1}{x^2})}{\frac{1}{x^2}} = \lim_{x\to\infty}\frac{-\frac{2}{x^3}\cos(\frac{1}{x^2})}{-\frac{2}{x^3}} = 1$.

Theorem 126 (Bernoulli-de l'Hôpital, case 3)**.** *Let $a < x_0$, f, g differentiable in (a, x_0) such that $g'(x) \neq 0$ for x sufficiently close to x_0, $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = +\infty$, $\lim_{x\to x_0}\frac{f'(x)}{g'(x)} = L$. Then $g(x) \neq 0$ for x sufficiently close to x_0 and $\lim_{x\to x_0}\frac{f(x)}{g(x)} = L$.*

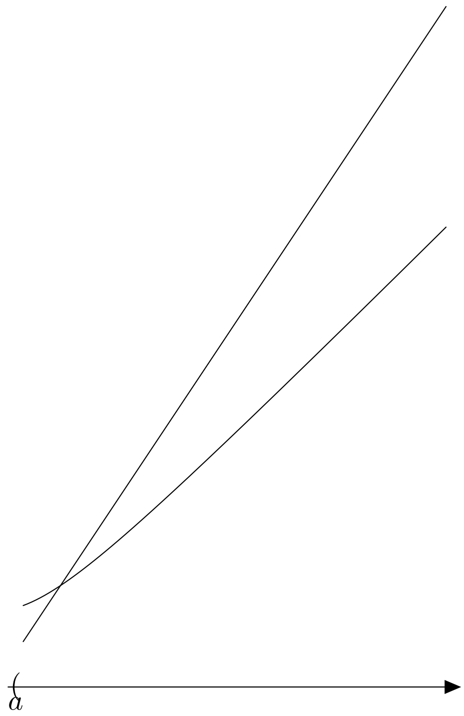


Figure 62: Theorem of de l'Hôpital. The limit $\lim_{x\to\infty}\frac{f(x)}{g(x)}$ is determined by $\lim_{x\to\infty}\frac{f'(x)}{g'(x)}$.

Proof. Let $\varepsilon > 0$. By the hypothesis, there is b such that $\left|\frac{f'(y)}{g'(y)} - L\right| < \frac{\varepsilon}{3}$ for $y \in (b, x_0)$. In addition, there is $\tilde{b}$ such that $b < \tilde{b} < x_0$ and in $(\tilde{b}, x_0)$ $f(x) > 2f(b) > 0, g(x) > 2g(b) > 0$. Then the function $h(x) = \frac{1-\frac{g(b)}{g(x)}}{1-\frac{f(b)}{f(x)}}$ is continuous on $(\tilde{b}, x_0]$ and its value at x_0 is 1 . Furthermore, it holds that

$$\frac{f(x)-f(b)}{g(x)-g(b)} \cdot h(x) = \frac{f(x)-f(b)}{g(x)-g(b)} \cdot \frac{1-\frac{g(b)}{g(x)}}{1-\frac{f(b)}{f(x)}} = \frac{f(x)}{g(x)}.$$

Let $\tilde{\tilde{b}}$ such that $|h(x)-1| < \frac{\varepsilon}{3L+1}$ for $x \in (\tilde{\tilde{b}}, x_0)$. By Cauchy's mean value theorem, there is $y \in (b, x)$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(b)}{g(x)-g(b)} \cdot h(x) = \frac{f'(y)}{g'(y)}h(x).$$

Now $\left|\frac{f(x)}{g(x)} - L\right| = \left|\frac{f'(y)}{g'(y)}h(x) - L\right| < \left|\frac{f'(y)}{g'(y)} - L\right|(1+\frac{\varepsilon}{3L+1}) + L|h(x)-1| < \frac{\varepsilon}{3} + \frac{\varepsilon^2}{3} + \frac{\varepsilon}{3} < \varepsilon$. □

Theorem 127 (Bernoulli-de l'Hôpital, case 4)**.** *Let f, g be differentiable (a, ∞) such that $g'(x) \neq 0$ as $x \to \infty$, $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = +\infty$, $\lim_{x\to\infty}\frac{f'(x)}{g'(x)} = L$. Then $g(x) \neq 0$ for x sufficiently large and* $\lim_{x\to\infty}\frac{f(x)}{g(x)} = L$.

Proof. Consider $F(y) = f(\frac{1}{y}), G(y) = g(\frac{1}{y})$. Since $\frac{1}{y} \to \infty$ as $y \to 0^+$, and $D(F(y)) = \frac{Df(\frac{1}{y})}{-y^2}, D(G(y)) = \frac{Dg(\frac{1}{y})}{-y^2}$ we can apply case 3 and obtain $L = \lim_{y\to 0^+}\frac{DF(y)}{DG(y)} = \lim_{y\to 0^+}\frac{F(y)}{G(y)} = \lim_{x\to\infty}\frac{f(x)}{g(x)}$. □

Example 128.

- Let us compute $\lim_{x\to\infty}\frac{x^2}{e^x}$. If the limit $\lim_{x\to\infty}\frac{2x}{e^x}$ exists, then by the de l'Hôpital rule, they should coincide. The latter exists if $\lim_{x\to\infty}\frac{2}{e^x}$ exists, and it does: it is 0. Therefore, the second limit exists and it is 0, and hence the first limit exists and it is 0.

- $\lim_{x\to 0} \frac{\sin 2x}{\sin x} = \lim_{x\to 0} \frac{2\cos 2x}{\cos x} = 2$.
- $\lim_{x\to 0} \frac{\log x}{1/\tan x} = \lim_{x\to 0} \frac{1/x}{1/\sin^2 x} = 0$.
- $\lim_{x\to 0} \frac{\log(\sin x)}{\log x} = \lim_{x\to 0} \frac{\frac{\cos x}{\sin x}}{x} = 1$.
- $\lim_{x\to\infty} \frac{x^n}{e^x} = 0$.
- $\lim_{x\to\infty} \frac{\log\cosh x}{x} = \lim_{x\to\infty} \frac{\sinh x/\cosh x}{1} = 1$.

Nov. 11. Landau's symbols, Taylor's formula.

Definition 129. Let I be an open interval, $f, f_1, f_2, g : I \to \mathbb{R}$, $x_0 \in I$ and suppose that $g(x) \neq 0$ in an neighbourhood of $x_0, x \neq x_0$. We write:

- $f(x) = O(g(x))$ (as $x \to x_0$) if there is $M > 0$ such that $|f(x)| \leq M|g(x)|$ in an neighbourhood of x_0.
- $f(x) = o(g(x))$ (as $x \to x_0$) if $\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0$.
- $f_1(x) = f_2(x) + O(g(x))$ ($f_1(x) = f_2(x) + o(g(x))$, respectively) if $f_1(x) - f_2(x) = O(g(x))$ ($= o(g(x))$, respectively).

Similarly, let $f, g : (a, \infty) \to \mathbb{R}$, and suppose that $g(x) \neq 0$ for sufficiently large x (that is, there is $X > 0$ such that $g(x) \neq 0$ if $x > X$). We write:

- $f(x) = O(g(x))$ (as $x \to \infty$) if there is $M > 0$ such that $|f(x)| \leq M|g(x)|$ for sufficiently large x.
- $f(x) = o(g(x))$ (as $x \to \infty$) if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.

The cases for $(-\infty, a)$, or $f(x) \to 0$ and the cases in I but $f(x) \to 0$ are analogous.

Example 130.
- If $n > 1$, $x^n = o(x)$ as $x \to 0$ (because $\lim_{x\to 0} \frac{x^n}{x} \to 0$).
- $x^n = o(x^m)$ as $x \to 0$ if $n > m$ (because $\lim_{x\to 0} \frac{x^n}{x^m} \to 0$).
- $x^m = o(x^n)$ as $x \to \infty$ se $n > m$ (because $\lim_{x\to\infty} \frac{x^m}{x^n} \to 0$).
- $\log x = o(x)$ as $x \to \infty$ (because $\lim_{x\to\infty} \frac{\log x}{x} \to 0$).
- $\log x = o(\frac{1}{x})$ as $x \to 0$ (because $\lim_{x\to 0} x\log x \to 0$).
- $\sin x = O(x)$ as $x \to 0$ (because $\lim_{x\to 0} \frac{\sin x}{x} \to 1$).
- $\sin x = o(x)$ as $x \to \infty$ (because $\lim_{x\to\infty} \frac{\sin x}{x} \to 0$).
- $\cos x = O(1)$ as $x \to 0$ (because $\lim_{x\to 0} \cos x \to 1$).
- $e^x - 1 = O(x)$ as $x \to 0$ (because $\lim_{x\to 0} \frac{e^x-1}{x} \to 1$).

Lemma 131. *Let us consider the behaviour $x \to x_0 = 0$ (other cases are analogous).*

(a) Let $a, b \in \mathbb{R}$. If $f(x) = O(h(x)), g(x) = O(h(x))$, then $af(x) + bg(x) = O(h(x))$.

(b) Let $a, b \in \mathbb{R}$. If $f(x) = o(h(x)), g(x) = o(h(x))$, then $af(x) + bg(x) = o(h(x))$.

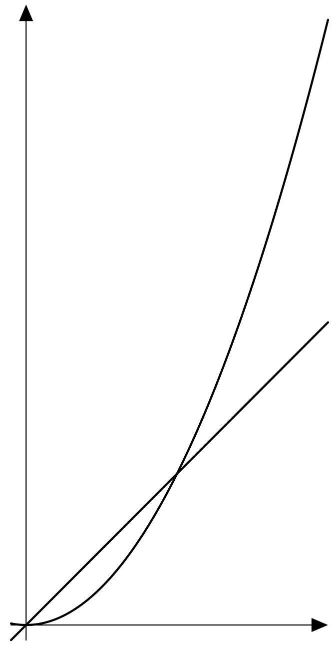


Figure 63: Landau's symbol. $x^2 = o(x)$ as $x \to 0$, but $x = o(x^2)$ as $x \to \infty$.

(c) If $g(x) = o(h(x))$, then $f(x)g(x) = o(f(x)h(x))$.
(Similarly, if $g(x) = O(h(x))$, then $f(x)g(x) = O(f(x)h(x))$)

(d) If $f(x) = o(h(x))$, then $f(x) = O(h(x))$.

(e) Let $f(x) = o(h(x))$ and $f(0) = 0, \lim_{x\to 0} g(x) = 0$. Then $f(g(x)) = o(h(g(x)))$.
(Similarly, if $f(x) = O(h(x))$, then $f(g(x)) = O(h(g(x)))$)

Proof. (a) We have $|f(x)| \leq M_1|h(x)|, |g(x)| \leq M_2|h(x)|$, hence $|af(x) + bg(x)| \leq |a||f(x)| + |b||g(x)| \leq (|a|M_1 + |b|M_2)|h(x)|$.

(b) Analogous.

(c) If $\lim_{x\to 0} \frac{g(x)}{h(x)} = 0$, then $\lim_{x\to 0} \frac{f(x)g(x)}{f(x)h(x)} = \lim_{x\to 0} \frac{g(x)}{h(x)} = 0$.

(d) If $\lim_{x\to 0} \frac{f(x)}{h(x)} \to 0$, then $\left|\frac{f(x)}{h(x)}\right| < M$ for x close enough to 0, hence $|f(x)| < M|h(x)|$.

(e) Let us define

$$u(k) = \begin{cases} \frac{f(k)}{h(k)} & \text{if } k \neq 0 \\ 0 & \text{if } k = 0. \end{cases}$$

Then $u(k)$ is continuous at $k = 0$ because $\frac{f(k)}{h(k)} \to 0$ as $k \to 0$. We have $f(g(x)) = h(g(x))u(g(x))$, and Altogether, $\lim_{x\to 0} \left|\frac{f(g(x))}{h(g(x))}\right| = \lim_{x\to 0} \left|\frac{h(g(x))u(g(x))}{h(g(x))}\right| = 1 \cdot 0 = 0$. The other claim is analogous.

□

Example 132. As $x \to 0$,

- $\sin(x^2) = O(x^2)$, because $\sin(y) = O(y)$ and we put $y = x^2$.
- $e^{x^3} - 1 = O(x^3)$, because $e^y - 1 = O(y)$, and we put $y = x^3$.
- $\sin^2(x) = O(x^2)$, because $\sin x = O(x)$, and hence $\sin^2 x = O(x \sin x) = O(x^2)$.

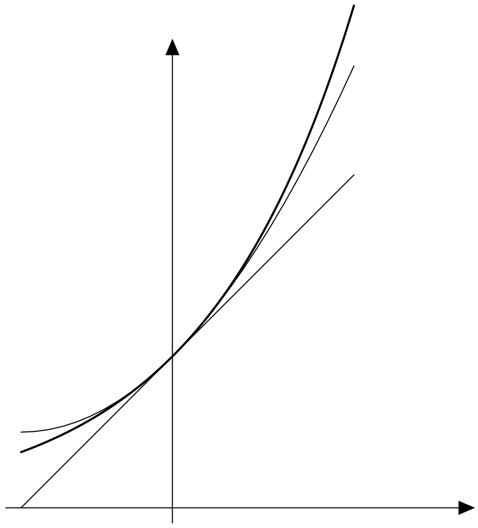


Figure 64: The second order Taylor formula. We approximate a general function by a second order polynomial.

Second order Taylor Formula

We have defined derivative by $\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$. If f is differentiable at x_0, then we have $\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} = f'(x_0)$, or equivalently,

$$\lim_{x\to x_0}\left(\frac{f(x)-f(x_0)}{x-x_0} - \frac{f'(x_0)(x-x_0)}{x-x_0}\right) = \lim_{x\to x_0}\frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x-x_0} = 0,$$

therefore, $f(x)-f(x_0)-f'(x_0)(x-x_0) = o(x-x_0)$. This means that we can approximate f **to the first order** by $f(x_0)+f'(x_0)(x-x_0)$. This is indeed called the first order Taylor formula.

The Taylor formula can be extended to higher order.

Proposition 133 (Second order Taylor formula)**.** *Let f be differentiable in (a,b) and twice differentiable at $x_0 \in (a,b)$. Then $f(x) = f(x_0)+(x-x_0)f'(x_0)+\frac{1}{2}(x-x_0)^2f''(x_0)+o((x-x_0)^2)$ as $x \to x_0$.*

Proof. Let us put $P_2(x) = f(x_0)+(x-x_0)f'(x_0)+\frac{1}{2}(x-x_0)^2f''(x_0)$. Then $P_2'(x) = f'(x_0)+(x-x_0)f''(x_0)$. Furthermore, the first order Talylor formula holds for f': $f'(x) = f'(x_0)+(x-x_0)f''(x_0)+o(x-x_0)$ as $x \to x_0$. That is,

$$\lim_{x\to x_0}\frac{D(f(x)-P_2(x))}{D((x-x_0)^2)} = \lim_{x\to x_0}\frac{f'(x)-f'(x_0)-(x-x_0)f''(x_0)}{2(x-x_0)} = \frac{1}{2}(f''(x_0)-f''(x_0)) = 0.$$

By the Bernoulli-de l'Hôpital theorem,

$$\lim_{x\to x_0}\frac{f(x)-P_2(x)}{(x-x_0)^2} = 0.$$

that is, $f(x) = P_2(x)+o((x-x_0)^2)$. □

Example 134. As $x \to 0$,

- $e^x = 1+x+\frac{x^2}{2}+o(x^2)$
- $\log(1+x) = x-\frac{x^2}{2}+o(x^2)$.
- $\sin(x) = x+o(x^2)$.
- $\cos(x) = 1-\frac{x^2}{2}+o(x^2)$.

Nov. 15. Higher order Taylor formula and more examples.

Higher order Taylor(-Peano) Formula

For f n-times differentiable, the following holds (as we prove later).

With the convention $f^{(0)}(x) = f(x)$,

$$\begin{aligned} f(x) &= f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + \cdots + \frac{1}{n!}(x - x_0)^n f^{(n)}(x_0) + o((x - x_0)^n) \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + o((x - x_0)^n) \end{aligned}$$

The part $\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$ is called the Taylor polynomial of f.

Lemma 135. *Let f, g differentiable n times in (a, b) and $x_0 \in (a, b)$. Suppose that $g^{(k)}(x) \neq 0$ for $x \neq x_0, 0 \leq k \leq n$ but $f^{(k)}(x_0) = g^{(k)}(x_0) = 0$ for $0 \leq k \leq n - 1$. Then for any $x \neq x_0, x \in (a, b)$ there is ξ between x, x_0 such that $\frac{f(x)}{g(x)} = \frac{f^{(n)}(\xi)}{g^{(n)}(\xi)}$.*

Proof. By Proposition 111,

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(\xi_1)}{g'(\xi_1)} = \frac{f'(\xi_1) - f'(x_0)}{g'(\xi_1) - g'(x_0)} = \frac{f^{(2)}(\xi_2)}{g^{(2)}(\xi_2)} = \cdots = \frac{f^{(n)}(\xi_n)}{g^{(n)}(\xi_n)}$$

and we put $\xi = \xi_n$. □

Proposition 136. *Let F be differentiable n times at $x_0 \in (a, b)$. Then, $F(x) = o((x - x_0)^n)$ as $x \to x_0$ if and only if $F^{(k)}(x_0) = 0$ for $0 \leq k \leq n$.*

Proof. We know this for $n = 0$ by definition. Let us prove the general case by induction, by assuming that it is true for n.

Let $F(x) = o((x - x_0)^{n+1})$. Then $F^{(k)}(x_0) = 0$ for $0 \leq k \leq n$ by the hypothesis of induction. The assumption is $0 = \lim_{x \to x_0} \frac{F(x)}{(x - x_0)^{n+1}}$. On the other hand, $\frac{F(x)}{(x - x_0)^{n+1}} = \frac{F^{(n)}(\xi)}{(n+1)!(\xi - x_0)}$ for some ξ. If $x \to x_0$, $\xi \to x_0$, that is, $0 = \lim_{x \to x_0} \frac{F(x)}{(x - x_0)^{n+1}} = \lim_{\xi \to x_0} \frac{F^{(n)}(\xi)}{(n+1)!(\xi - x_0)} = \lim_{\xi \to x_0} \frac{F^{(n)}(\xi) - F^{(n)}(x_0)}{(n+1)!(\xi - x_0)} = \frac{F^{(n+1)}(x_0)}{(n+1)!}$, hence $F^{(n+1)}(x_0) = 0$.

Let $F^{(k)}(x_0) = 0$ for $0 \leq k \leq n + 1$. Then by the Bernoulli-de l'Hôpital theorem,

$$\begin{aligned} 0 = \frac{F^{(n+1)}(x_0)}{(n+1)!} &= \lim_{x \to x_0} \frac{F^{(n)}(x) - F^{(n)}(x_0)}{(n+1)!(x - x_0)} \\ &= \lim_{x \to x_0} \frac{F^{(n)}(x)}{(n+1)!(x - x_0)} = \lim_{x \to x_0} \frac{F^{(n-1)}(x) - F^{(n-1)}(x_0)}{\frac{(n+1)!}{2}(x - x_0)^2} \\ &= \lim_{x \to x_0} \frac{F^{(n-1)}(x)}{\frac{(n+1)!}{2}(x - x_0)^2} = \lim_{x \to x_0} \frac{F^{(n-2)}(x) - F^{(n-2)}(x_0)}{\frac{(n+1)!}{3!}(x - x_0)^3} \\ &\cdots = \lim_{x \to x_0} \frac{F(x)}{(x - x_0)^{n+1}}. \end{aligned}$$

That is, $F(x) = o((x - x_0)^{n+1})$. □

Corollary 137. *Let $f(x)$ differentiable n times at $x_0 \in (a, b)$. Then with*

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k,$$

we have $f(x) = P_n(x) + o((x - x_0)^n)$.

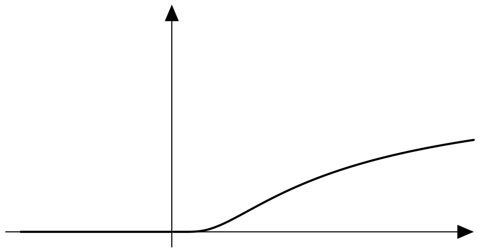


Figure 65: The graph of a function whose Taylor series converges but not to itself.

Proof. $D^k(f(x_0) - P_n(x_0)) = 0$ for $0 \le k \le n$. By Proposition 136, $f(x_0) = P_n(x_0) + o((x - x_0)^n)$. □

Example 138.
- $f(x) = e^x$. As $f^{(n)}(x) = e^x$, we have $f^{(n)}(0) = 1$ and hence $e^x = \sum_{k=0}^n \frac{x^n}{n!} + o(x^n)$. That is, $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n)$ as $x \to x_0 = 0$.
- $f(x) = \sin x$. As $f^{(4n)}(x) = \sin x$, $f^{(4n+1)}(x) = \cos x$, $f^{(4n+2)}(x) = -\sin x$, $f^{(4n+3)}(x) = -\cos x$, we have $f^{(4n)}(0) = 0$, $f^{(4n+1)}(x) = 1$, $f^{(4n+2)}(x) = 0$, $f^{(4n+3)}(x) = -1$, and hence $\sin x = \sum_{k=0}^n \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o(x^{2n+1})$. That is, $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + o(x^{2n+1})$ as $x \to x_0 = 0$.

Very often, the Taylor series converges to the original function $f(x)$, that is,

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

holds for some functions $(e^x, \sin x, \cos x, \log(1+x))$ and for some x. One example we can show easily such convergence is $f(x) = \frac{1}{1-x}$. Indeed, $f'(x) = \frac{1}{(1-x)^2}, f^{(n)}(x) = \frac{n!}{(1-x)^n}$. And hence the Taylor series around $x = 0$ is

$$\sum_{k=0}^{n} \frac{n! x^n}{n!} = \sum_{k=0}^{n} x^n,$$

and we know that this partial sum is $\frac{1-x^n}{1-x}$, which converges to $\frac{1}{1-x}$ for $|x| < 1$. But the series does not converge for if $|x| \ge 0$.

There are functions whose Taylor series converges but not to the original function. For example, if we take

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{if } x > 0 \\ 0 & \text{if } x \ge 0 \end{cases}$$

then $f^{(n)}(0) = 0$ for all n, hence the Taylor polynomial is identically 0, but the original function $f(x)$ is not identically 0.

The question of for which function the Taylor series converges to the origial function will be studied in Mathematical Analysis II.

Applications to certain limits

Taylor's formula can be used to compute certain indefinite limits.

Example 139.
- $$\lim_{x \to 0} \frac{e^x - 1 - x}{\sin(x^2)}$$

  As $x \to x_0 = 0$, we have

  - $e^x = 1 + x + \frac{x^2}{2} + o(x^2)$

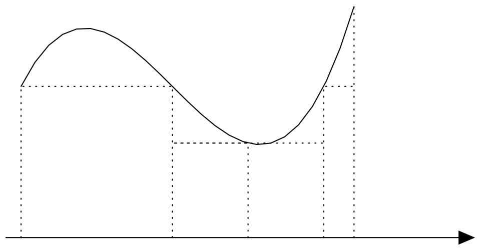


Figure 66: Approximating the area surrounded by $f(x)$ by rectangles.

- $\sin y = y + o(y)$
- $\sin(x^2) = x^2 + o(x^2)$

Then it holds, as $x \to 0$,

$$\frac{e^x - 1 - x}{\sin(x^2)} = \frac{\frac{x^2}{2} + o(x^2)}{x^2 + o(x^2)}$$

hence $\lim_{x\to 0} \frac{e^x-1-x}{\sin(x^2)} = \frac{1}{2}$.

-
$$\lim_{x\to 0} \frac{x - \ln(1-x) - 2x\sqrt{1+x}}{\sin(x) - xe^x}$$

As $x \to x_0 = 0$, we have

- $\ln(1-x) = -x - \frac{1}{2!}x^2 + o(x^2)$
- $x\sqrt{1+x} = x(1 + \frac{1}{2}x - o(x)) = x + \frac{1}{2}x^2 + o(x^2)$
- $\sin x = x + 0 \cdot x^2 + o(x^2)$
- $xe^x = x(1 + x + o(x)) = x + x^2 + o(x^2)$

Then it holds, as $x \to 0$,

$$\frac{x - \ln(1-x) - 2x\sqrt{1+x}}{\sin(x) - xe^x} = \frac{-\frac{1}{2}x^2 + o(x^2)}{-x^2 + o(x^2)}$$

hence $\lim_{x\to 0} \frac{x-\ln(1-x)-2x\sqrt{1+x}}{\sin(x)-xe^x} = \frac{1}{2}$.

Nov. 17. Definite integral.

Given a function f, we consider (Riemann) integral. This is a concept that extends the area of familiar figures such as triangles and disks. If $f(t)$ represents the velocity of a car at time t, then the integral of f gives the distance the car travels in a time interval. If f is the density of a piece of iron, the integral gives the weight.

The area of a region defined by a function can be approximated by rectangles. We know that the area of a rectangle with sides a, b is ab.

For an interval $I = (a, b)$ or $(a, b]$ etc., we define $|I| = b - a$.

Definition 140. (i) Let I be a bounded interval in $\mathbb{R}$. A partition of I is a finite set of disjoint intervals $P = \{I_j : 1 \le j \le n\}$ such that $\bigcup_{j=1}^n I_j = I$.

(ii) $\mathrm{diam}(P) = \max\{|I_j| : 1 \le j \le n\}$.

(iii) A partition P' is called a refinement of P if every interval of P admits a partizione formed by intervals in P'. That is, every $I_j \in P$ can be written as $I_j = \bigcup_{k=1}^{n_j} I_{jk}, I_{jk} \in P'$.

We denote this by $P' \succ P$.

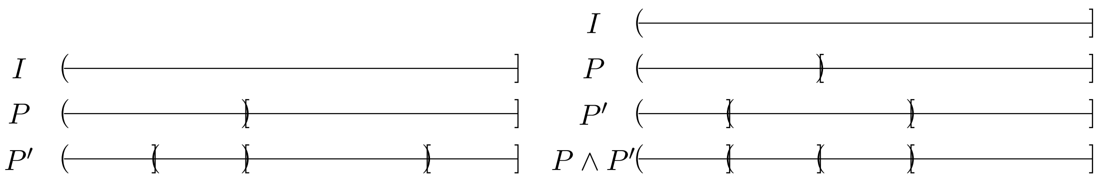


Figure 67: Left: a partition P of an interval I and a refinement $P' \succ P$. Right: two partitions P, P' of I and $P \wedge P'$.

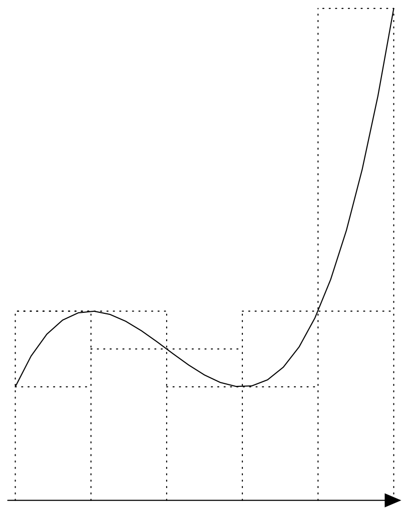


Figure 68: $\overline{S}_I(f,P)$ and $\underline{S}_I(f,P)$ for a given a partition P of I.

(iv) If P, P' are partitions of I, we define $P \wedge P' = \{I \cap I' : I \in P, I \in P', I \cap I' \neq \emptyset\}$. We have $P \wedge P' \succ P, P'$.

If P is a partition of I, then $|I| = \sum_{j=1}^n |I_j|$.
For a partition P of I and a bounded function $f : I \to \mathbb{R}$, we put

$$\underline{S}_I(f,P) := \sum_{j=1}^n (\inf_{I_j} f)|I_j|, \text{ (“lower sum”)}$$

$$\overline{S}_I(f,P) := \sum_{j=1}^n (\sup_{I_j} f)|I_j| \text{ (“upper sum”)}.$$

We have $\underline{S}_I(f,P) \leq \overline{S}_I(f,P)$.

Example 141. $I = [0,1]$. $P_n = \{[0, \frac{1}{n}), [\frac{1}{n}, \frac{2}{n}), \cdots, [\frac{n-1}{n}, 1]\}$. If $n' = mn$ for $m \in \mathbb{N}$, then $P_{n'} \succ P_n$. $\mathrm{diam}(P_n) = \frac{1}{n}$.

- If $f(x) = a$, then $\underline{S}_I(f,P_n) = \overline{S}_I(f,P_n) = a$.
- Let $f(x) = x$.

$$\overline{S}_I(f,P_n) = \sum_{j=1}^n \frac{j}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \cdot \frac{n(n+1)}{2}$$

Analogously,

$$\underline{S}_I(f,P_n) = \sum_{j=1}^n \frac{j-1}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \cdot \left(\frac{n(n+1)}{2} - n\right) = \frac{1}{n^2} \frac{n(n-1)}{2}.$$

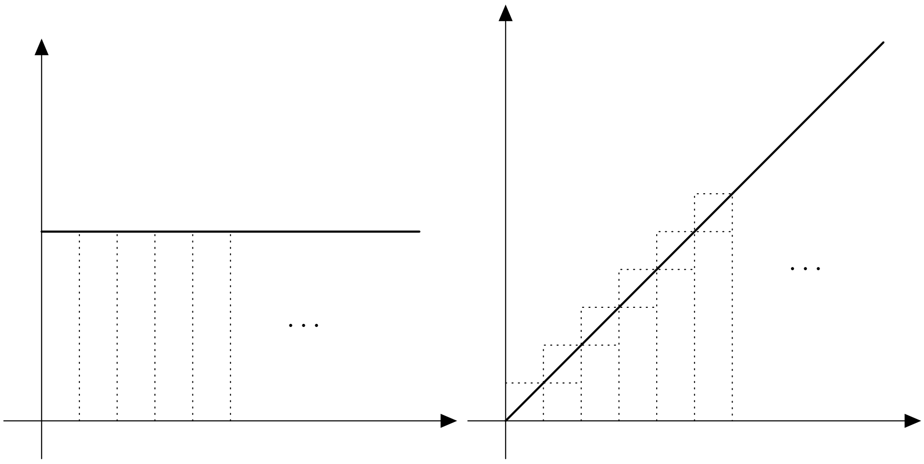


Figure 69: The upper and lower sum for $f(x) = a$ (constant) and $f(x) = x$.

Therefore, by taking $n \to \infty$, we obtain $\lim_{n\to\infty} \overline{S}_I(f, P_n) = \lim_{n\to\infty} \underline{S}_I(f, P_n) = \frac{1}{2}$, which is the area of the triangle $\{(x, y) : x \in [0, 1], 0 \le y \le x = f(x)\}$.

Lemma 142. *Let I be a bounded interval, $f : I \to \mathbb{R}$ a bounded function. If P, P' are two partitions of I, then*

(i) If $P \prec P'$, then

$$\underline{S}_I(f, P) \le \underline{S}_I(f, P') \le \overline{S}_I(f, P') \le \overline{S}_I(f, P).$$

(ii) $\underline{S}_I(f, P) \le \overline{S}_I(f, P')$

Proof. (i) If $P \prec P'$, then we can take $P = \{I_j : 1 \le j \le n\}$ and $P' = \{I_{jk} : 1 \le j \le n, 1 \le j \le n_j\}$ such that $\cup_{k=1}^{n_j} I_{jk} = I_j$. Then, for every j, k, $\inf_{I_j} f \le \inf_{I_{jk}} f$, $\sup_{I_j} f \ge \sup_{I_{jk}} f$. It follows that

$$\begin{aligned}
\underline{S}_I(f, P) &= \sum_{j=1}^{n} (\inf_{I_j} f)|I_j| = \sum_{j=1}^{n} (\inf_{I_j} f) \sum_{k=1}^{n_k} |I_{jk}| = \sum_{j=1}^{n} \sum_{k=1}^{n_k} (\inf_{I_j} f)|I_{jk}| \\
&\le \sum_{j=1}^{n} \sum_{k=1}^{n_k} (\inf_{I_{jk}} f)|I_{jk}| = \underline{S}_I(f, P'). \\
\overline{S}_I(f, P) &= \sum_{j=1}^{n} (\sup_{I_j} f)|I_j| = \sum_{j=1}^{n} (\sup_{I_j} f) \sum_{k=1}^{n_k} |I_{jk}| = \sum_{j=1}^{n} \sum_{k=1}^{n_k} (\sup_{I_j} f)|I_{jk}| \\
&\ge \sum_{j=1}^{n} (\sup_{I_{jk}} f) \sum_{k=1}^{n_k} |I_{jk}| = \overline{S}_I(f, P').
\end{aligned}$$

Note that $\underline{S}_I(f, Q) \le \overline{S}_I(f, Q)$ for any partition Q.

(ii) Since $P \prec P \wedge P', P' \prec P \wedge P'$, it follows from the previous point $\underline{S}_I(f, P) \le \underline{S}_I(f, P \wedge P') \le \overline{S}_I(f, P \wedge P') \le \overline{S}_I(f, P')$.

□

Definition 143. Let $I = (a, b)$ or $[a, b]$ etc. a bounded interval and $f : I \to \mathbb{R}$ bounded. f is said to be **integrable on** I if

$$\underbrace{\sup_P \underline{S}_I(f, P)}_{\text{“lower integral”}} = \underbrace{\inf_P \overline{S}_I(f, P)}_{\text{“upper integral”}},$$

where $\inf_P$ and $\sup_P$ are taken over all possible partitions of P of I and in this case we denote this number by

$$\int_I f(x)dx = \int_a^b f(x)dx.$$

x does not have any meaning, and one can also write thsi as $\int_I f(t)dt$.

Example 144.

- $\int_0^1 adx = a$. Indeed, for all partitions $\underline{S}_I(f,P) = \overline{S}_I(f,P) = a$.
- $\int_0^1 xdx = \frac{1}{2}$. Indeed, with $f(x) = x$, we have found P_n such that $\underline{S}_I(f,P) = \frac{n(n-1)}{2n^2}$ and $\overline{S}_I(f,P) = \frac{n(n+1)}{2n^2}$, hence the sup and the inf coincide and it is $\frac{1}{2}$.

In general, it is difficult to show integrability by definition. Fortunately, we can prove that continuous functions on a closed bounded interval are integrable, and we also have the fundamental theorems of calculus, that let us calculate integrals with the knowledge of derivatives.

Nov. 18. Integrability of continuous functions, fundamental theorems of calculus.

Recall that any continuous function f on a closed interval $I = [a,b]$ is uniformly continuous, that is, for given $\epsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for $x, y \in I$ with $|x - y| < \delta$.

Theorem 145. *Let $I = [a,b]$ be a closed bounded interval. Then any continuous function f on I is integrable.*

Proof. By Theorem 90, for $\frac{\epsilon}{2(b-a)}$ there is δ such that for $x, y \in I, |x-y| < \delta$ it holds that $|f(x) - f(y)| < \frac{\epsilon}{2(b-a)}$. Now, for any partition $P = \{I_j\}_{j=1}^n$ with $\operatorname{diam} P = \max\{|I_j| : 1 \le j \le n\} < \delta$, we have

$$\overline{S}_I(f,P) - \underline{S}_I(f,P) = \sum_{j=1}^n (\sup_{I_j} f - \inf_{I_j} f)|I_j| < \sum_{j=1}^n \frac{\epsilon}{2(b-a)}|I_j| = \frac{\epsilon}{2} < \epsilon.$$

Therefore, f is integrable. □

Example 146. (of a nonintegrable function) If $f(x) = \begin{cases} 0 & x \text{ rational} \\ 1 & x \text{ irrational} \end{cases}$, $f(x)$ is not integrable: $\overline{S}_I(f,P) = 1, \underline{S}_I(f,P) = 0$, because any interval I contains both a rational number and an irrational number.

Proposition 147. *Let I be a bounded interval, f, g bounded and integrable on I.*

(i) If $c, d \in \mathbb{R}$, then $cf + dg$ is integrable on I and $\int_I (cf(t) + dg(t))dt = c\int_I f(t)dt + d\int_I g(t)dt$.

(ii) If $f \le g$, then $\int_I f(t)dt \le \int_I g(t)dt$.

(iii) If $\underline{I} \subset I$, then f is integrable on $\underline{I}$. If $P = \{I_j : 1 \le j \le n\}$ is a partition of I, then $\int_I f(t)dt = \sum_{j=1}^n \int_{I_j} f(t)dt$.

Proof. (i) Integrability of cf is easy: $\underline{S}_I(cf,P) = c\underline{S}_I(f,P)$ if $c \ge 0$ and $\underline{S}_I(cf,P) = c\overline{S}_I(f,P)$ if $c < 0$. If $c = 0$ everything becomes 0, and if $c > 0$ one obtains the limit directly. If $c < 0$ sup and inf are exchanged.

Let f, g be integrable. We have $\inf_{I_j} f + \inf_{I_j} g \le \inf_{I_j}(f+g)$, hence for any partition P, $\underline{S}_I(f,P) + \underline{S}_I(g,P) \le \underline{S}_I(f+g,P)$. Analogously, $\overline{S}_I(f+g,P) \le \overline{S}_I(f,P) + \overline{S}_I(g,P)$. By taking inf and sup with respect to P, we obtain integrability of $f+g$ and the equality $\int_I (cf(t) + df(t))dt = c\int_I f(t)dt + d\int_I g(t)dt$.

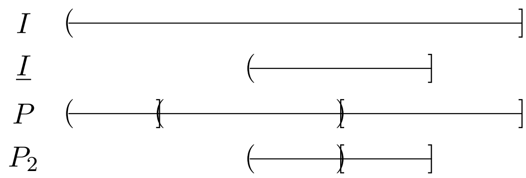


Figure 70: The patition P_2 of a subinterval $\underline{I}$ of I obtained from a partition P of I.

(ii) If $f \le g$, then $\overline{S}_I(f,P) \le \overline{S}_I(g,P)$ for any P. Similarly $\underline{S}_I(f,P) \le \underline{S}_I(g,P)$.

(iii) Let $\underline{I} \subset I$. Then, for any partition P of I, we can take a refinement P' which consists of the intervals of the form $I_j \cap \underline{I}$ and $I_j \setminus \underline{I}$ (the latter may be a union of two intervals).

Let $P_2 = \{I_j \cap \underline{I} : 1 \le j \le n\}$. Then

$$\begin{aligned}\overline{S}_{\underline{I}}(f,P_2) - \underline{S}_{\underline{I}}(f,P_2) &\le \overline{S}_{\underline{I}}(f,P_2) - \underline{S}_{\underline{I}}(f,P_2) + \sum_{j=1}^{n}(\sup_{I_j\setminus\underline{I}} f - \inf_{I_j\setminus\underline{I}} f)|I_j \setminus \underline{I}| \\ &= \overline{S}_{\underline{I}}(f,P') - \underline{S}_{\underline{I}}(f,P')\end{aligned}$$

As f is integrable, there is P' such that $\overline{S}_{\underline{I}}(f,P') - \underline{S}_{\underline{I}}(f,P') < \epsilon$, then $\overline{S}_{\underline{I}}(f,P_2) - \underline{S}_{\underline{I}}(f,P_2) < \epsilon$.

Let us consider the case $n = 2$, that is $I = I_1 \cup I_2$. For P_1, P_2 partitions of I_1, I_2, $P = P_1 \cup P_2$ is a partition of I. Then

$$\underline{S}_{I_1}(f,P_1) + \underline{S}_{I_2}(f,P_2) = \underline{S}_I(f,P_1 \cup P_2) \le \overline{S}_I(f,P_1 \cup P_2) = \overline{S}_{I_1}(f,P_1) + \overline{S}_{I_2}(f,P_2).$$

If $\overline{S}_{I_1}(f,P_1) - \underline{S}_{I_1}(f,P_1) < \frac{\varepsilon}{2}$ and $\overline{S}_{I_2}(f,P_2) - \underline{S}_{I_2}(f,P_2) < \frac{\varepsilon}{2}$, then $\overline{S}_I(f,P_1 \cup P_2) - \underline{S}_I(f,P_1 \cup P_2) < \varepsilon$ and the limits coincide: $\int_I f(x)dx = \int_{I_1} f(x)dx + \int_{I_2} f(x)dx$.

□

Corollary 148. *If f is continuous, then $|f|$ is continuous and by Theorem 147(ii), $\left|\int_I f(x)dx\right| \le \int_I |f(x)|dx$.*

Definition 149. If $a < b$, then we put $\int_b^a f(x)dx = -\int_a^b f(x)dx$.

Lemma 150. *It holds that $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$ for all $a, b, c \in \mathbb{R}$.*

Proof. If $a < c < b$, then this follows from 147(iii). If $a < b < c$, then

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^b f(x)dx - \int_c^b f(x)dx.$$

The other cases are analogous. □

Theorem 151 (Fundamental theorem of calculus 1)**.** *Let $I = [a,b]$ a bounded closed interval and $f : I \to \mathbb{R}$ continuous. Then the function of x on I defined by*

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

is differentiable and $F'(x) = f(x)$.

Proof. We have

$$\frac{F(x+h)-F(x)}{h} = \frac{1}{h}\left(\int_a^{x+h} f(t)dt - \int_a^x f(t)dt\right) = \frac{1}{h}\int_x^{x+h} f(t)dt$$

As f is continuous, $\left|\int_x^{x+h} f(t)dt - f(x)h\right| = \left|\int_x^{x+h} f(t)dt - \int_x^{x+h} f(x)dt\right| \leq \int_x^{x+h} |f(t)-f(x)|dt$ and for any $\epsilon > 0$ there is $\delta > 0$ such that if $|t-x| < \delta$, then $|f(t)-f(x)| < \epsilon$, therefore, $\int_x^{x+h} |f(t)-f(x)|dt < h\epsilon$. Then, for such h, $\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| = \left|\frac{1}{h}\int_x^{x+h} f(t)dt - f(x)\right| < \epsilon$. Since $\epsilon > 0$ is arbitrary, $\lim_{h\to 0}\frac{F(x+h)-F(x)}{h} = f(x)$. □

Theorem 152 (Fundamental theorem of calculus 2). *Let $I = [a,b]$ be a closed bounded interval and $f : I \to \mathbb{R}$ differentiable in (a,b) and f' is continuous and extends to a continuous function on Then*

$$f(x) - f(a) = \int_a^x f'(t)dt.$$

Proof. $D(f(x)-f(a)) = f'(x)$, while $D(\int_a^x f'(t)dt) = f'(x)$ by Theorem 151. Therefore, $D(f(x)-f(a)-\int_a^x f'(t)dt) = 0$, and by Corollary 106, $f(x)-f(a)-\int_a^x f'(t)dt$ is constant, but with $x = a$, $f(a)-f(a)-\int_a^a f'(t)dt = 0$, hence $f(x)-f(a)-\int_a^x f'(t)dt = 0$. □

Theorem 152 allows us to compute integrals of certain functions.

- We know that $D(x^{n+1}) = (n+1)x^n$, or $D(\frac{x^{n+1}}{n+1}) = x^n$. Hence $\int_a^x t^n dt = \frac{1}{n+1}(x^{n+1} - a^{n+1})$. We know that $D(e^x) = e^x$, hence
- $\int_a^x e^t dt = (e^x - e^a)$.

Nov. 22. Primitive and examples of integral calculus.

Definition 153. Let $f : I \to \mathbb{R}$. If there is a function $F : I \to \mathbb{R}$ such that $F' = f$, then F is called a **primitive of** f.

By Corollary 106, if there are two primitives F, G of f, then $F(x) - G(x)$ is a constant.
By Theorem 145 and Theorem 151 there is a primitive if f is continuous: $F(x) = \int_a^x f(t)dt$ is a primitive of f.

Corollary 154. *Let f be a continuous function on a closed bounded interval $I = [a,b]$, and F be a primitive of f. Then $\int_a^b f(t)dt = F(b) - F(a)$.*

Proof. Let $\tilde{F}(x) = \int_a^x f(t)dt$. Then $\tilde{F}$ is a primitive of f, and hence $\tilde{F}(x) - F(x) = c$ (constant). By Theorem 152, $\int_a^b f(x)dx = \tilde{F}(b) - \tilde{F}(a) = (F(b)+c) - (F(a)+c) = F(b) - F(a)$. □

This tells us a way to compute the integral $\int_a^b f(x)dx$: we only have to find a primitive $F(x)$ of $f(x)$ and take the difference $F(b) - F(a)$. This is denoted by $[F(x)]_a^b$. Namely,

$$\int_a^b f(x)dx = [F(x)]_a^b.$$

A primitive of f is also written by $\int f(x)dx$ (up to a constant) and it is called the **indefinite integral** of f. With a generic constanti it is written, for example $\int xdx = \frac{x^2}{2} + C$. In contrast, $\int_a^b f(x)dx$ (the integral in the interval $[a,b]$) is called a **definite integral**.

- Let $f(x) = x^n, n \in \mathbb{N}$. Then $F(x) = \frac{x^{n+1}}{n+1}$ is a primitive of $f(x)$. That is, $\int x^n dx = \frac{x^{n+1}}{n+1} + C$.

- Let $f(x) = \sin x$. Then $F(x) = -\cos x$ a primitive of $f(x)$. That is, $\int \sin x dx = -\cos x + C$.
- Let $f(x) = \cos x$. Then $F(x) = \sin x$ a primitive of $f(x)$. That is, $\int \cos x dx = \sin x + C$.
- Let $f(x) = e^x$. Then $F(x) = e^x$ a primitive of $f(x)$. That is, $\int e^x dx = e^x + C$.
- Let $f(x) = \frac{1}{x}$. Then $F(x) = \log x$ a primitive of $f(x)$. That is, $\int \frac{1}{x} dx = \log x + C$.
- Let $f(x) = \frac{1}{x^{n+1}}, n \in \mathbb{N}$. Then $F(x) = -\frac{1}{nx^n}$ a primitive of $f(x)$. That is, $\int \frac{1}{x^{n+1}} dx = -\frac{1}{nx^n} + C$.

With this knowledge of primitives, we can compute definite integrals.

- $\int_0^1 x^2 dx = [\frac{x^3}{3}]_0^1 = \frac{1}{3} - \frac{0}{3} = \frac{1}{3}$.
- $\int_{-1}^2 x^4 dx = [\frac{x^5}{5}]_{-1}^2 = \frac{32}{5} - (-\frac{1}{5}) = \frac{33}{5}$.
- $\int_0^\pi \sin x dx = [-\cos x]_0^\pi = -(-1) - (-1) = 2$.
- $\int_0^\pi \cos x dx = [\sin x]_0^\pi = 0 - 0 = 0$.
- $\int_1^2 e^x dx = [e^x]_1^2 = e^2 - e$.
- $\int_1^2 \frac{1}{x} dx = [\log x]_1^2 = \log 2 - \log 1 = \log 2$.
- $\int_1^2 \frac{1}{x^2} dx = [-\frac{1}{x}]_1^2 = -\frac{1}{2} - (-\frac{1}{1}) = \frac{1}{2}$.

Note that $\int_{-1}^2 \frac{1}{x} dx$ is not integrable!

Lemma 155. *We have the folowing.*

- $\int (f(x) + g(x)) dx = \int f(x) dx + \int g(x) dx + C$.
- $\int af(x) dx = a \int f(x) dx + C$ *for* $a \in \mathbb{R}$.
- *If* $\int f(x) dx = F(x) + C$, *then* $\int f(x-a) dx = F(x-a) + C$ *for* $a \in \mathbb{R}$.

Proof. All these follow from the rules for derivatives.

- If $DF(x) = f(x), DG(x) = g(x)$, then $D(F(x) + G(x)) = f(x) + g(x)$.
- If $DF(x) = f(x)$, then $D(aF(x)) = af(x)$.
- If $DF(x) = f(x)$, then $D(F(x-a)) = f(x-a)$ by the chain rule.

□

With this, we can compute also definite integrals.

Lemma 156. *We have the folowing.*

- $\int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x)$.
- $\int_a^b cf(x) dx = c \int_a^b f(x) dx$ *for* $c \in \mathbb{R}$.

Proof. This follows immediately from Lemma 155. □

- $\int_0^1 (x-1)^2 dx = [\frac{(x-1)^3}{3}]_0^1 = \frac{0}{3} - \frac{(-1)^3}{3} = \frac{1}{3}$.
- $\int_{-1}^2 x^2(x-2) dx = \int_{-1}^2 x^3 - 2x^2 dx = [\frac{x^4}{4} - \frac{2x^3}{3}]_{-1}^2 = (4 - \frac{16}{3}) - (\frac{1}{4} - (-\frac{2}{3})) = -\frac{9}{4}$.

- $\int_0^\pi \sin(x - \frac{\pi}{3})dx = [-\cos(x - \frac{\pi}{3})]_0^\pi = -\cos\frac{2\pi}{3} - (-\cos(-\frac{\pi}{3})) = \frac{1}{2} + \frac{1}{2} = 1.$
- $\int_1^2 \frac{2}{x+1}dx = 2[\log(x+1)]_1^2 = 2(\log 3 - \log 2).$

Theorem 157. *Let f be a continuous function on $I = [a,b]$. Then there is $c \in (a,b)$ such that $\int_a^b f(x)dx = f(c)(b-a)$.*

Proof. Note that $F(x) = \int_a^x f(t)dt$ is differentiable and $F'(x) = f(x)$. By Lagrange's mean value theorem, there is $c \in (a,b)$ such that $F'(c) = \frac{F(b)-F(a)}{b-a}$, that is, $f(c)(b-a) = \int_a^b f(x)dx$. □

Some applications

Let us imagine a car travelling at the speed $v(t)$ at time t. Then the distance travelled from time a and b is $\int_a^b v(t)dt$. Indeed, if $x(t)$ is the place of the car at time t, then we have $x'(t) = v(t)$ by definition. By Theorem 152, we have $x(b) - x(a) = \int_a^b v(t)dt$.

As another example from physics, consider the situation where someone is pushing up vertically a mass m (kg) to a certain height h. The work done by this motion is, as far as the gravitational force is constant, mgh, where g is the gravitational acceleration. mg is called the weight, which is the downward force. If one is pushing a mass in a changing gravitational field $g(x)$ (like a rocket carrying a payload), the work done by this motion is $\int_{h_1}^{h_2} mg(x)dx$.

Nov. 24. Integral calculus.

Indefinfite integral of elementary functions

Note that, for $x < 0$, $D(\log|x|) = D(\log(-x)) = -\frac{1}{-x} = \frac{1}{x}$. Altogether,

$f(x)$	$f'(x)$	$\int f(x)dx$	
c (constant)	0	$cx + C$	
x^α	$\alpha x^{\alpha-1}$	$\frac{x^{\alpha+1}}{\alpha+1} + C$	for $\alpha \neq 0, -1$,$x \neq 0$ for negative power
x^{-1}	$-\frac{1}{x^2}$	$\log\|x\| + C$	$x \neq 0$
$\frac{1}{x^2+1}$	$-\frac{2x}{(x^2+1)^2}$	$\arctan x + C$	
$\frac{1}{\sqrt{1-x^2}}$	$\frac{x}{(1-x^2)^{\frac{3}{2}}}$	$\arcsin x + C$	$-1 < x < 1$
e^x	e^x	$e^x + C$	
$\log\|x\|$	$\frac{1}{x}$	$x\log\|x\| - x + C$	see below, $x \neq 0$
$\sin x$	$\cos x$	$-\cos x + C$	
$\cos x$	$-\sin x$	$\sin x + C$	
$\sinh x$	$\cosh x$	$\sinh x + C$	
$\cosh x$	$\sinh x$	$\cosh x + C$	

Integration by parts

Recall that, if f, g are differentiable, then it holds that $D(f(x)g(x)) = Df(x)g(x) + f(x)Dg(x)$. By writing this as $Df(x)g(x) = D(f(x)g(x)) - f(x)Dg(x)$, we can find a primitive of $Df(x)g(x)$ if we know a primitive of $f(x)Dg(x)$. Schematically,

$$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx + C.$$

This is called **integration by parts**.

Example 158.
- Consider $\int x\cos x dx$. With $f(x) = \sin x, g(x) = x$, this is of the form $f'(x)g(x)$, because $f'(x) = \cos x$. By integration by parts, with $g'(x) = 1$, we obtain

$$\int x\cos x dx = x\sin x - \int \sin x \cdot 1 dx + C = x\sin x + \cos x + C.$$

We can check this results by taking the derivative: $D(x \sin x + \cos x) = \sin x + x \cos x - \sin x = x \cos x$.

- Consider $\int \log x dx$. We can see this as $1 \cdot \log x$, and $1 = D(x)$. Therfore, with $f(x) = x, g(x) = \log x$ and $g'(x) = \frac{1}{x}$, we have

$$\int \log x dx = x \log x - \int x \cdot \frac{1}{x} dx + C = x \log x - \int 1 dx + C = x \log x - x + C.$$

- Consider $\int x^2 \sin x dx$. This cannot be integrated by one step, but by successive applications of integration by parts. By noting that $\sin x = D(-\cos x)$ and $\cos x = D(\sin x)$,

$$\begin{aligned}\int x^2 \sin x dx &= x^2(-\cos x) - \int 2x(-\cos x) dx + C \\ &= -x^2 \cos x + 2x \sin x - \int 2 \sin x dx + C \\ &= -x^2 \cos x + 2x \sin x + 2 \cos x + C.\end{aligned}$$

As for indefinite integral, we do not have to find the whole indefinite integral, but we can give values to parts. Let us recall that $f(b) - f(a) = \int_a^b f'(x) dx$.

Lemma 159. *If f, g are differentiable and f', g' are continuous, then*

$$\int_a^b f'(x) g(x) dx = [f(x)g(x)]_a^b - \int_a^b f(x) g'(x) dx.$$

Proof. $(fg)' = f'g + fg'$, hence $\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx$ (integration by parts) and this follows from Theorem 152, that is, If $H(x)$ is a primitive of $h(x)$, then $\int_a^b h(x)dx = H(b) - H(a)$. Note that with $h(x) = f(x)g'(x)$, we have we can take $H(x) = \int_a^x h(x)dx$ and $H(b) - H(a) = \int_a^b h(x)dx - \int_a^a h(x)dx = \int_a^b h(x)dx$. ☐

Example 160. $\int_0^1 xe^{2x} dx = \frac{1}{2}[xe^{2x}]_0^1 - \int_0^1 \frac{1}{2} e^{2x} dx = \frac{1}{2}(e^2 - 0) - \frac{1}{4}[e^{2x}]_0^1 = \frac{e^2}{2} - \frac{1}{4}(e^2 - 1) = \frac{e^2}{4} + \frac{1}{4}$.

Substitution

Next, let us consider the case where the integral is of the form $\int \varphi'(x) f'(\varphi(x)) dx$. We know that $D(f(\varphi(x))) = \varphi'(x) f'(\varphi(x))$ by the chain rule, hence in this case we have

$$\int \varphi'(x) f'(\varphi(x)) dx = f(\varphi(x)) + C.$$

This is called **substitution**.

Example 161.

- Consider $\int 2x \sin(x^2) dx$. Note that $2x = D(x^2)$ and $\sin(y) = D(-\cos y)$, hence

$$\int 2x \sin(x^2) dx = -\cos(x^2) + C.$$

Indeed, by the chain rule, $D(-\cos(x^2)) = -(2x(-\sin(x^2))) = 2x \sin(x^2)$.

- Consider $\int \frac{x}{x^2+1} dx$. Note that $2x = D(x^2)$, and hence

$$\int \frac{x}{x^2+1} dx = \frac{1}{2} \int \frac{2x}{x^2+1} dx = \frac{1}{2} \int \frac{D(x^2)}{x^2+1} dx = \frac{1}{2} \log(x^2 + 1).$$

- Consider $\int \tan x dx$. Recall that $\tan x = \frac{\sin x}{\cos x}$ and note that $D(\cos x) = -\sin x$. Hence

$$\int \tan x dx = -\int D(\cos x) \cdot \frac{1}{\cos x} dx + C = -\log|\cos x| + C.$$

Lemma 162. *If f, φ are differentiable and f', φ' is continuous, then*

$$\int_a^b \varphi'(x) f'(\varphi(x)) dx = [f(\varphi(x))]_a^b = [f(y)]_{\varphi(a)}^{\varphi(b)} = f(\varphi(b)) - f(\varphi(a)).$$

Proof. This follows immediately because $f(\varphi(x))$ is a primitive of $\varphi'(x)f'(\varphi(x))$. □

Example 163.

$$\begin{aligned}\int_0^\pi \sin^3 x dx &= -\int_0^\pi (\cos^2 x - 1) \sin x dx \\ &= \int_0^\pi ((\cos x)^2 - 1) D(\cos x) dx = \left[\frac{\cos^3 x}{3} - \cos x\right]_0^\pi \\ &= \left(\frac{(-1)^3}{3} - (-1) - (\frac{1^3}{3} - 1)\right) = \frac{4}{3}.\end{aligned}$$

Nov. 25. Integral calculus.

Rational functions

We know that

- $\int \frac{1}{(x-a)} dx = \log|x-a|$.
- for $n \in \mathbb{N}, n \geq 2$, $\int \frac{1}{(x-a)^n} dx = \frac{-1}{(n-1)(x-a)^{n-1}}$.
- $\int \frac{1}{x^2+1} dx = \arctan x$, $\int \frac{1}{x^2+a^2} dx = \frac{1}{a} \arctan(\frac{x}{a})$.
- $\int \frac{1}{(x-b)^2+a^2} dx = \frac{1}{a} \arctan(\frac{(x-b)}{a})$.

We also have

$$\begin{aligned}\int \frac{1}{(x^2+1)^2} dx &= \int \frac{1+x^2-x^2}{(x^2+1)^2} dx = \int \frac{1}{(x^2+1)} + \int \frac{-2x \cdot x}{2(x^2+1)^2} dx \\ &= \arctan x + \frac{x}{2(x^2+1)} - \int \frac{1}{2(x^2+1)} = \arctan x + \frac{x}{2(x^2+1)} - \frac{1}{2} \arctan x \\ &= \frac{1}{2} \arctan x + \frac{x}{2(x^2+1)}.\end{aligned}$$

Indeed, by taking the derivative,

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

In general, the derivative of the primitive F of f must be the original function f. We can check that the primitive in this way. This is often useful because the calculus of primitive is often complicated, while derivative is mechanical.

Example 164. • $\int \frac{x^3-1}{4x^3-x}dx = \int \frac{\frac{1}{4}(4x^3-x)+\frac{x}{4}-1}{4x^3-x}dx$. Note that $\frac{x-4}{4x^3-x} = \frac{x-4}{x(2x-1)(2x+1)} = \frac{4}{x} + \frac{-\frac{7}{2}}{2x-1} + \frac{-\frac{9}{2}}{2x+1}$ (see below) and hence

$$\begin{aligned}\int \frac{x^3-1}{4x^3-x}dx &= \frac{x}{4} + \frac{1}{4}\int \left(\frac{4}{x} + \frac{-\frac{7}{2}}{2x-1} + \frac{-\frac{9}{2}}{2x+1}\right)dx \\ &= \frac{x}{4} + \log|x| - \frac{7}{16}\log|2x-1| - \frac{9}{16}\log|2x+1|.\end{aligned}$$

• Using $\frac{1}{(x-1)^2(x^2+1)} = \frac{Ax+B}{(x-1)^2} + \frac{Cx+D}{x^2+1} = \frac{-\frac{x}{2}+1}{(x-1)^2} + \frac{\frac{x}{2}}{x^2+1}$ we get

$$\begin{aligned}\int \frac{1}{(x-1)^2(x^2+1)}dx &= \int \frac{-\frac{x}{2}+1}{(x-1)^2}dx + \int \frac{\frac{x}{2}}{x^2+1}dx \\ &= \int \frac{-\frac{(x-1)}{2}+\frac{1}{2}}{(x-1)^2}dx + \int \frac{\frac{x}{2}}{x^2+1}dx \\ &= -\frac{1}{2}\log|x-1| - \frac{1}{2(x-1)} + \frac{1}{4}\log(x^2+1).\end{aligned}$$

In general, if $P(x)$ and $Q(x)$ are polynomials, $\frac{P(x)}{Q(x)}$ can be written as a sum of $\frac{P_1(x)}{(x-a)^n}$ or $\frac{P_2(x)}{((x-b)^2+a^2)^n}$ (with different polynomials P_1, P_2), and for each of them one can find a primitive.

Example 165. • $\frac{1}{(x-1)(x+1)}$. We put $\frac{1}{(x-1)(x+1)} = \frac{A}{x-1} + \frac{B}{x+1}$. Then

$$\frac{1}{(x-1)(x+1)} = \frac{A(x+1)+B(x-1)}{(x-1)(x+1)}$$

and $1 = (A+B)x + (A-B)$. This means that $A+B=0$ and $1 = A-B$, therfore, $A = \frac{1}{2}, B = -\frac{1}{2}$.

• $\frac{1}{(x+1)(x^2+1)^2}$. We put $\frac{1}{(x+1)(x^2+1)^2} = \frac{A}{x+1} + \frac{Bx^3+Cx^2+Dx+E}{(x^2+1)^2}$.
Then

$$\frac{1}{(x+1)(x^2+1)^2} = \frac{A(x^2+1)^2 + (Bx^3+Cx^2+Dx+E)(x+1)}{(x+1)(x^2+1)^2}$$

hence $1 = A(x^4+2x^2+1) + (Bx^4 + (B+C)x^3 + (C+D)x^2 + (D+E)x + E)$, and this means $A+B=0, B+C=0, 2A+C+D=0, D+E=0, A+E=1$. To solve this, we observe $A=-B$, $C=-B, D=-E$ and hence $-3B-E=0, -B+E=1$, hence $B=-\frac{1}{4}, E=\frac{3}{4}, A=\frac{1}{4}, C=\frac{1}{4}, D=-\frac{3}{4}$. Altogether,

$$\frac{1}{(x+1)(x^2+1)^2} = \frac{\frac{1}{4}}{x+1} + \frac{-\frac{1}{4}x^3+\frac{1}{4}x^2-\frac{3}{4}x+\frac{3}{4}}{(x^2+1)^2} = \frac{\frac{1}{4}}{x+1} + \frac{-\frac{x}{4}+\frac{1}{4}}{x^2+1} + \frac{-\frac{1}{2}x+\frac{1}{2}}{(x^2+1)^2}$$

Change of variables

Let $F(x)$ be a primitive of $f(x)$, that is $\int f(x)dx = F(x)$. If it is difficult to find F directly, one may consider a change of variables $x = \varphi(t)$. By the chain rule, $\frac{d}{dt}F(\varphi(t)) = f(\varphi(t))\varphi'(t)$. If $G(t)$ is a primitive of $f(\varphi(t))\varphi'(t)$, then $F(x) = G(\varphi^{-1}(t))$.

In order to recall the rule, it is useful to write

$$\int f(x)dx = \int f(\varphi(t))\frac{dx}{dt}dt,$$

even if this is only formal.

Example 166. • $f(x) = \frac{1}{e^x+1}$. With $t = e^x + 1, x = \varphi(t) = \log(t-1), \varphi'(t) = \frac{1}{t-1}$,

$$\int \frac{1}{e^x+1} dx = \int \frac{1}{t(t-1)} dt = \int \left(\frac{1}{t-1} - \frac{1}{t} \right) dt = \log \left| \frac{t-1}{t} \right|$$

and with $t = e^x + 1$, $\int \frac{1}{e^x+1} dx = \log \frac{e^x}{e^x+1}$.

• $f(x) = \frac{x^2}{\sqrt{1-x^2}}$. With $x = \varphi(t) = \sin t, \varphi'(t) = \cos t, t = \arcsin x$, if $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$,

$$\int \frac{x^2}{\sqrt{1-x^2}} dx = \int \frac{\sin^2 t}{\cos t} \cos t dt = \int \sin^2 t dt = \int \frac{1 - \cos 2t}{2} dt = \frac{t}{2} - \frac{\sin 2t}{4}$$

and with $t = \arcsin x, \sin 2t = 2 \sin t \cos t = 2x\sqrt{1-x^2}$, we obtain $\int \frac{x^2}{\sqrt{1-x^2}} dx = \frac{1}{2} \arcsin x - \frac{1}{2} x\sqrt{1-x^2}$.

Nov. 26. Integral calculus, Taylor's formula with remainder, log derivative, improper integral

When the function contains $\sin x$ and $\cos x$, it is often useful to do the change of variable $x = \varphi(t) = 2 \arctan t$, or $t = \tan \frac{x}{2}$. Indeed, we have $\varphi'(t) = \frac{2}{1+t^2}$, while $\frac{1}{\cos^2 \frac{x}{2}} = \frac{\cos^2 \frac{x}{2} + \sin^2 \frac{x}{2}}{\cos^2 \frac{x}{2}} = 1 + t^2$ and $\sin x = \sin(2 \cdot \frac{x}{2}) = 2 \sin \frac{x}{2} \cos \frac{x}{2} = \frac{2t}{1+t^2}$ and $\cos x = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2} = \frac{1-t^2}{1+t^2}$.

For example,

$$\int \frac{1}{\sin x} dx = \int \frac{t^2+1}{2t} \cdot \frac{2}{1+t^2} dt = \log |t| + C = \log \left| \tan \frac{x}{2} \right| + C.$$

Definite integral by change of variables

Corollary 167. *Let f be contiunous on $[a,b]$, φ differentiable and φ' continuous on $[\alpha, \beta]$, and $\varphi([\alpha,\beta]) \subset [a,b]$, $\varphi(\alpha) = a, \varphi(\beta) = b$. Then*

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)dt$$

Proof. Let $F(x) = \int_a^x f(s)ds$. Since $\frac{d}{dt}(F(\varphi(t))) = f(\varphi(t)) \cdot \varphi'(t)$,

$$\int_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)dt = [F(\varphi(t))]_\alpha^\beta = [F(x)]_a^b = \int_a^b f(x)dx.$$

□

Example 168. • Note that $\sqrt{1 - \sin^2 t} = |\cos t|$ and this is equal to $\cos t$ if $|t| < \frac{\pi}{2}$, hence with $x = \sin t$,

$$\begin{aligned}\int_0^1 \sqrt{1-x^2} dx &= \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 t} \cos t \, dt = \int_0^{\frac{\pi}{2}} \cos^2 t \, dt \\ &= \int_0^{\frac{\pi}{2}} \frac{\cos(2t)+1}{2} dt = \left[\frac{\sin(2t)}{4} + \frac{t}{2} \right]_0^{\frac{\pi}{2}} \\ &= \frac{\pi}{4}.\end{aligned}$$

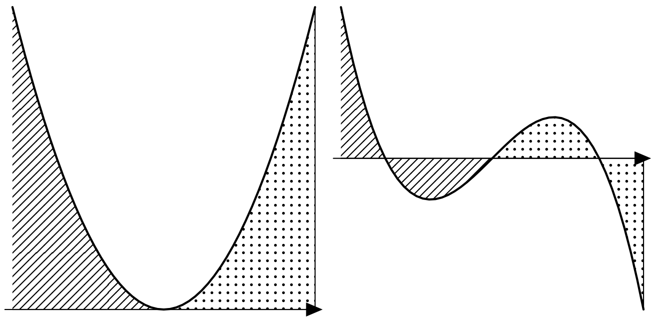


Figure 71: Integral of symmetric and antisymmetric functions.

Some remarks

- If $f(x) = f(-x)$, then by the change of variables $x = -t$, $\int_{-a}^0 f(x)dx = \int_a^0 f(-t)(-t)'dt = \int_0^a f(t)dt$, hence $\int_{-a}^a f(x)dx = 2\int_0^a f(x)dx$.

  For example, $\int_{-1}^1 \sqrt{1-x^2}dx = 2\int_0^1 \sqrt{1-x^2}dx = \frac{\pi}{2}$.

- If $f(x) = -f(-x)$, then by the change of variables $x = -t$, $\int_{-a}^0 f(x)dx = \int_a^0 f(-t)(-t)'dt = -\int_0^a f(t)dt$, hence $\int_{-a}^a f(x)dx = 0$.

  For example, $\int_{-1}^1 e^{x^2} \sin x dx = 0$.

- Logarithmic differentiation: If $f(x)$ is difficult to differentiate but $\log f(x)$ is easy, then we have $D(\log f(x)) = \frac{f'(x)}{f(x)}$, hence we have $f'(x) = f(x)D(\log f(x))$.

  For example, $f(x) = x^x$ (for $x > 0$) is not a simple product or a composition. But $\log f(x) = x \log x$, hence $D(\log f(x)) = \log x + 1$, hence $f'(x) = x^x(\log x + 1)$.

Taylor's formula with remainder

Proposition 169. *If f is differentiable $n+1$ times in an neighbourhood of x_0 with continuous derivative, then for x in that neighbourhood,*

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x, x_0),$$

where $R_n(x, x_0) = \frac{1}{n!}\int_{x_0}^x f^{(n+1)}(y)(x-y)^n dy$.

Proof. This is true for $n = 0$, because

$$f(x_0) + \int_{x_0}^x f'(y)dy = f(x_0) + [f(y)]_{x_0}^x = f(x).$$

To prove the formula by induction, assume the claim for n and let f be $n+2$ times differentiable, then $f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x, x_0)$ and

$$\begin{aligned} R_n(x, x_0) &= \frac{1}{n!}\int_{x_0}^x f^{(n+1)}(y)(x-y)^n dy \\ &= -\frac{1}{(n+1)!}\left[f^{(n+1)}(y)(x-y)^{n+1}\right]_{x_0}^x + \frac{1}{(n+1)!}\int_{x_0}^x f^{(n+2)}(y)(x-y)^{n+1}dy \\ &= \frac{1}{(n+1)!}f^{(n+1)}(x_0)(x-x_0)^{n+1} + \frac{1}{(n+1)!}\int_{x_0}^x f^{(n+2)}(y)(x-y)^{n+1}dy \end{aligned}$$

□

This is interesting, because for some functions, we can prove that the Taylor series *converges* to the original function. Let us take $x_0 = 0$ and consider the interval $(-R, R)$.

- $e^x = \sum_{k=0}^n \frac{x^n}{n!} + \frac{1}{(n+1)!}\int_0^x e^y(y-x)^n dy$. As $|x| < R$, we have $e^y < e^R$ and $|(y-x)^n| < R^n$. Altogether, the remainder term is

$$\frac{1}{(n+1)!}\left|\int_0^x e^y(y-x)^n dy\right| \le \frac{1}{(n+1)!}\left|\int_0^x e^R R^n dy\right| \le \frac{e^R R^{n+1}}{(n+1)!}.$$

  Note that for any R, $\frac{R^{n+1}}{(n+1)!} \to 0$ because for sufficiently large n we have $n > 2R$, hence from that point the sequence decreases more than by $\frac{1}{2}$. This means that $e^x - \sum_{k=0}^n \frac{x^k}{k!} \to 0$, that is, the Taylor series converges to e^x for $x \in (-R, R)$, and this is denoted by

$$e^x = \sum_{n=0}^\infty \frac{x^n}{n!}.$$

  Furthermore, R was arbitrary, hence this holds for any x.

- The same argument holds for $\sin x, \cos x$, because $|D^n(\sin x)| \le 1, |D^n(\cos x)| \le 1$. That is,

$$\sin x = \sum_{n=0}^\infty \frac{(-1)^n x^{2n+1}}{(2n+1)!},$$
$$\cos x = \sum_{n=0}^\infty \frac{(-1)^n x^{2n}}{(2n)!},$$

  in the sense that for each x the series converges to the original function.

More properties of Taylor *series* will be studied in Mathematical Analysis II.

The value of Napier's number

We can also find the approximate value of $e = e^1$ using the Taylor formula with remainder. We know that $e < (1 + \frac{1}{n})^{n+1}$ for any n. In particular, $e < 4$. Therefore, by

$$e^1 = \sum_{k=0}^{13} \frac{1^k}{k!} + \frac{1}{13!}\int_0^1 e^y(1-y)^{13} dy$$

and the error term satisfies $0 < \frac{1}{13!}\int_0^1 e^y(1-y)^{12}dy < \frac{4}{13!} < 0.0000002$. Therefore, the approximation of e,

$$\sum_{k=0}^{12} \frac{1^k}{k!} \cong 2.71828182,$$

is correct up to the 7-th digit.

Improper integral

We can define integral for (some) funtions that are not bounded and on an interval not bounded.

Definition 170. Let (a, b) be an interval, $a \in \mathbb{R}$ or $a = -\infty$ and $b \in \mathbb{R}$ or $b = +\infty$. Let f be a function integrable on all $[\alpha, \beta]$, where $a < \alpha < \beta < b, \alpha, \beta \in \mathbb{R}$. If there exists the limit $\lim_{\alpha \to a} \int_\alpha^\beta f(x)dx$, then we denote it by

$$\int_a^\beta f(x)dx = \lim_{\alpha \to a} \int_\alpha^\beta f(x)dx.$$

It also holds that $\int_a^\gamma f(x)dx = \int_a^\beta f(x)dx + \int_\beta^\gamma f(x)dx$ for $\gamma \in (a,b)$. Analogously if there exists the limit $\lim_{\beta\to b}\int_\alpha^\beta f(x)dx$, then we write $\int_\alpha^b f(x)dx = \lim_{\beta\to b}\int_\alpha^\beta f(x)dx$. If both limits exist, then we denote $\int_a^b f(x)dx = \int_a^{x_0} f(x)dx + \int_{x_0}^b f(x)dx$ for some $x_0 \in (a,b)$.

This definition does not depend on $x_0 \in (a,b)$. Indeed,

$$\int_a^{x_0} f(x)dx + \int_{x_0}^b f(x)dx = \int_a^{x_0} f(x)dx + \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_0} f(x)dx + \int_{x_0}^b f(x)dx$$
$$= \int_a^{x_1} f(x)dx + \int_{x_1}^b f(x)dx$$

Example 171.

- Consider $(0,1)$ and the function $f(x) = x^\alpha, \alpha \in \mathbb{R}$. For $\varepsilon > 0$, if $\alpha \neq -1$,

$$\int_\varepsilon^1 x^\alpha dx = \frac{1}{\alpha+1}[x^{\alpha+1}]_\varepsilon^1 = \frac{1}{\alpha+1}(1-\varepsilon^{\alpha+1}),$$

  and as $\varepsilon \to +0$, this tends to $\frac{1}{\alpha+1}$ if $\alpha > -1$, and diverges if $\alpha < -1$. If $\alpha = -1$,

$$\int_\varepsilon^1 x^{-1}dx = [\log x]_\varepsilon^1 = -\log\varepsilon,$$

  and this tends to ∞ as $\varepsilon \to 0+$. Therefore, for $\alpha > -1$, $\int_0^1 x^\alpha dx = \frac{1}{\alpha+1}$.

- Consider $(1,\infty)$ and the function $f(x) = x^\alpha, \alpha \in \mathbb{R}$. For $\beta > 1$, if $\alpha \neq -1$,

$$\int_1^\beta x^\alpha dx = \frac{1}{\alpha+1}[x^{\alpha+1}]_1^\beta = \frac{1}{\alpha+1}(\beta^{\alpha+1}-1),$$

  and as $\beta \to +\infty$, this tends to $-\frac{1}{\alpha+1} = \frac{1}{|\alpha+1|}$ if $\alpha < -1$, and diverges if $\alpha > -1$. For $\alpha = -1$,

$$\int_1^\beta x^{-1}dx = [\log x]_1^\beta = \log\beta,$$

  and this tends to ∞ as $\beta \to +\infty$. Therefore, $\int_1^\infty f(x)dx = \frac{1}{|\alpha+1|}$ for $\alpha < -1$.

- Consider $(-\infty,\infty)$.

$$\int_\alpha^\beta xe^{-x^2}dx = \frac{1}{2}[e^{-x^2}]_\alpha^\beta = \frac{1}{2}(e^{-\beta^2} - e^{-\alpha^2}).$$

  Then both limits $\lim_{\alpha\to-\infty}, \lim_{\beta\to\infty}$ exist. Furthermore, $\int_{-\infty}^\infty xe^{-x^2}dx = \frac{1}{2}([e^{-x^2}]_{-\infty}^0 + [e^{-x^2}]_0^\infty) = 0$.

Nov. 30. Some properties of improper integral

Let us recall that we introduced a proper integral for an unbounded function or on an unbounded interval by

$$\int_a^\beta f(x)dx = \lim_{\alpha\to a}\int_\alpha^\beta f(x)dx,$$

where $a < \alpha$, and the function is bounded and integrable on all bounded intervals $[\alpha,\beta]$. Similarly, $\int_\alpha^b f(x)dx = \lim_{\beta\to b}\int_\alpha^\beta f(x)dx$ and $\int_a^b f(x)dx = \int_a^{x_0} f(x)dx + \int_{x_0}^b f(x)dx$, where $x_0 \in (a,b)$. When these limits exist, we say that the improper integral **converges**, and otherwise **does not converge**, or **diverges** if the limit tends to ∞ or $-\infty$.

Example 172. • The improper integral $\int_0^\infty \sin x dx$ does not converge. Indeed, it holds that $\int_0^\beta \sin x dx = [-\cos x]_0^\beta = -\cos\beta + 1$, and as $\beta \to \infty$, $-\cos\beta$ oscillates and does not converge to any value.

- Consider $\int_0^\infty e^{-x}dx$. For $\beta > 0$, we have $\int_0^\beta e^{-x}dx = [-e^{-x}]_0^\beta = -e^{-\beta} - (-1) = 1 - e^{-\beta}$, hence as $\beta \to \infty$, this tends to 1. That is, $\int_0^\infty e^{-x}dx = 1$.

- Consider $(-\infty, \infty)$.

$$\int_\alpha^\beta xe^{-x^2}dx = -\frac{1}{2}[e^{-x^2}]_\alpha^\beta = -\frac{1}{2}(e^{-\beta^2} - e^{-\alpha^2}).$$

and both limits $\lim_{\alpha\to-\infty}, \lim_{\beta\to\infty}$ exist. Furthermore, $\int_{-\infty}^\infty xe^{-x^2}dx = -\frac{1}{2}([e^{-x^2}]_{-\infty}^0 + [e^{-x^2}]_0^\infty) = 0$.

Proposition 173. *Let f, g be integrable on all $[\alpha, \beta] \subset (a, b)$.*

(i) Let $0 \le f \le g$. If $\int_a^b g(x)dx$ converges, then so does $\int_a^b f(x)dx$. If $\int_a^b f(x)dx$ diverges, then so does $\int_a^b g(x)dx$.

(ii) If $f \ge 0$ and there is $M > 0$ such that $\int_\alpha^\beta f(x)dx < M$ for all $a < \alpha < \beta < b$, then $\int_a^b f(x)dx$ converges.

(iii) Let $0 < g$, and $\lim_{x\to b-} \frac{f(x)}{g(x)} = c \neq 0$. Then $\int_\alpha^b f(x)dx$ exists if and only if $\int_\alpha^b g(x)dx$ exists.

(iv) Let $f > 0$ be decreasing on $[\alpha, \infty)$. $\int_\alpha^\infty f(x)dx$ converges if and only if $\sum_{n=N}^\infty f(n)$ converges for some N.

(v) $\left|\int_a^b f(x)dx\right| \le \int_a^b |f(x)|dx$.

Proof. (i) On any interval $[\alpha, \beta]$ it holds that $\int_\alpha^\beta f(x)dx \le \int_\alpha^\beta g(x)dx$, hence $\int_\alpha^\beta f(x)dx$ is bounded and increases as α, β tend to a, b.

(ii) As $f(x) \ge 0$, when $\alpha \to a$ (and $\beta \to b$), the integral $\int_\alpha^\beta f(x)dx$ increases. But as it is bounded by M, it must converge to a certain number $\int_a^b f(x)dx \le M$.

(iii) Let $c > 0$ (the other case is analogous). For x_0 close enough to b, it holds that $\frac{c}{2}g(x) \le f \le 2cg(x)$. Hence the claim follows from (i).

(iv) We have $f(n+1) \le \int_n^{n+1} f(x)dx \le f(n)$, therefore, $\sum_{n=N}^M f(n) \le \int_N^{M+1} f(x)dx \le \sum_{n=N}^{M+1} f(n)$.

(v) This follows from $-|f| \le f \le |f|$ and (i) for an interval $[\alpha, \beta]$, then by taking the limits. □

Now we can show that $\sum_{n=1}^\infty \frac{1}{n}$ is divergent using integral. Indeed, $\int_1^N \frac{1}{x}dx \le \sum_{n=1}^{N-1} \frac{1}{n}$, but $\int_1^N \frac{1}{x}dx = [\log x]_1^N = \log N - 0 \to \infty$, therefore, also $\sum_{n=1}^{N-1} \frac{1}{n} \to \infty$ as $N \to \infty$.

When the improper integral $\int_a^b |f(x)|dx$, then we say that the improper integral $\int_a^b f(x)dx$ **converges absolutely**.

Example 174. • $\int_1^\infty \frac{\cos x}{x^2}dx$ converges. Indeed, $|\frac{\cos x}{x^2}| \le \frac{1}{x^2}$ and $\int_1^\alpha \frac{1}{x^2}dx = [-\frac{1}{x}]_1^\alpha = 1 - \frac{1}{\alpha}$ which tends to 1 as $\alpha \to \infty$.

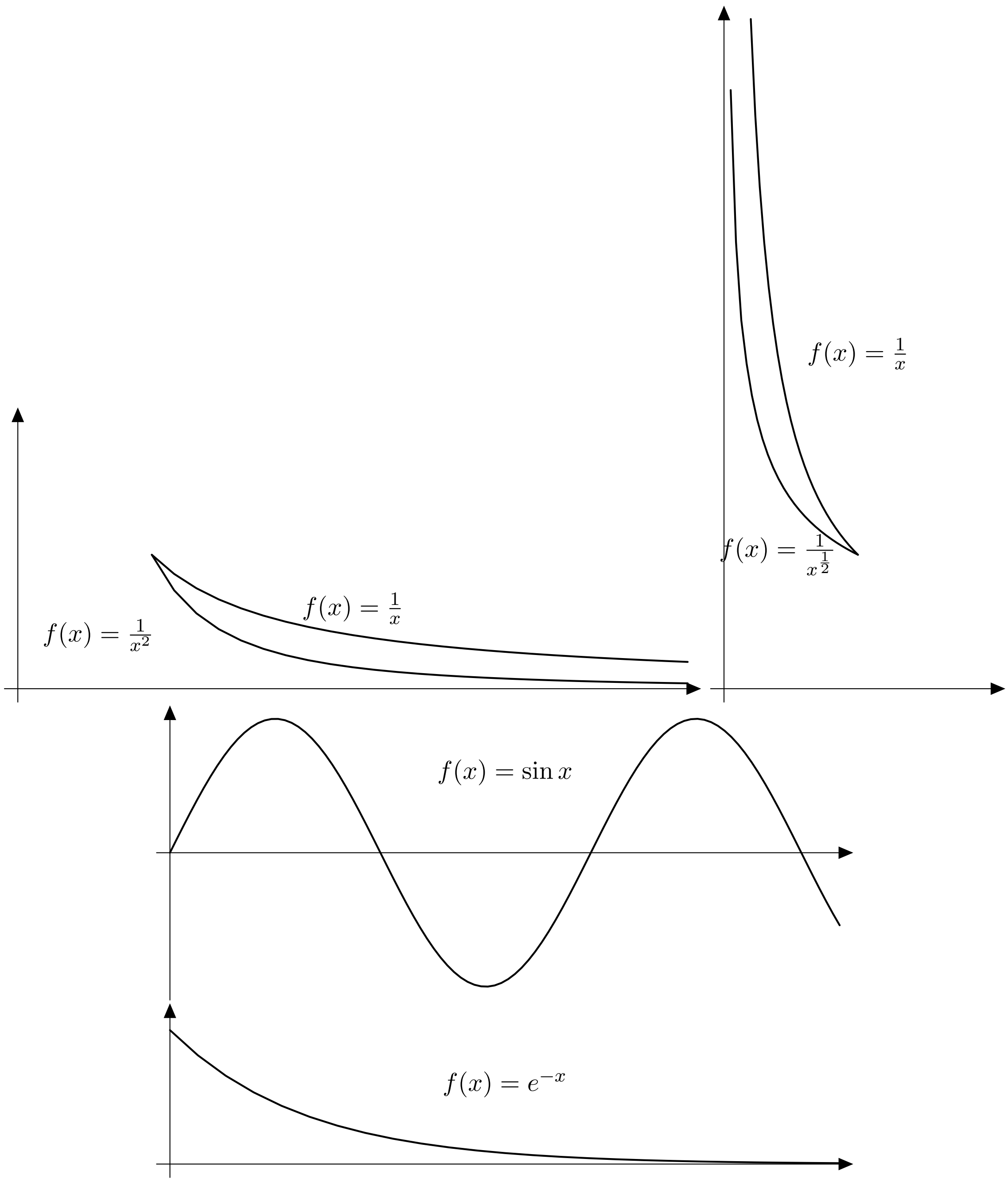


Figure 72: Convergent and non convergent improper integrals.

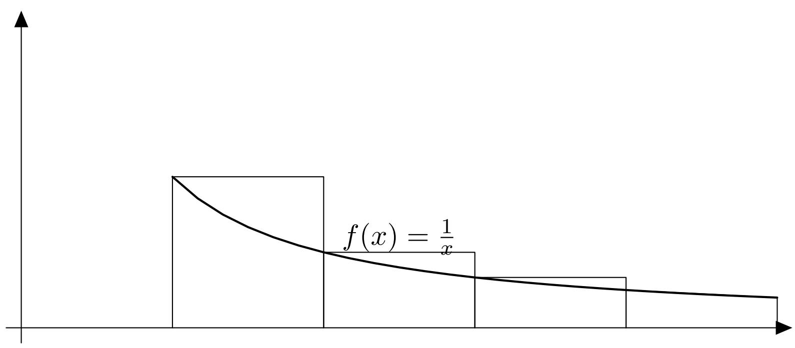


Figure 73: A graphical proof that $\sum_{n=1}^{N} \frac{1}{n}$ diverges as $N \to \infty$.

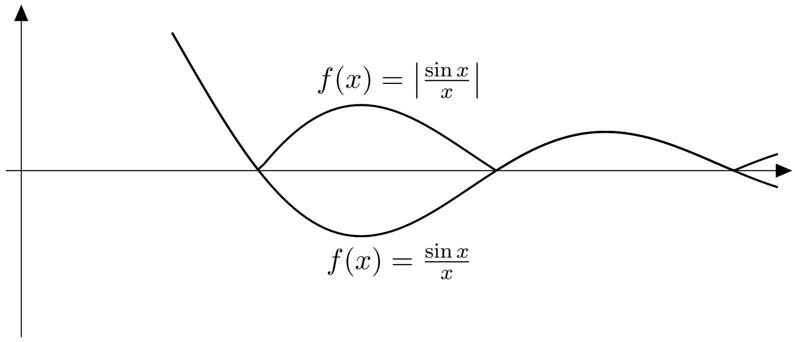


Figure 74: $f(x) = \frac{\sin x}{x}$ and $f(x) = \left|\frac{\sin x}{x}\right|$. The improper integral of the former in $[1, \infty)$ is convergent, while the latter is not.

- $\int_1^\infty \frac{\sin x}{x} dx$ converges. Indeed, by integration by parts,
$$\int_1^\alpha \frac{\sin x}{x} dx = \left[\frac{-\cos x}{x}\right]_1^\alpha + \int_1^\alpha \frac{\cos x}{x^2} dx = \cos 1 - \frac{\cos\alpha}{\alpha} + \int_1^\alpha \frac{\cos}{x^2} dx$$
The first two terms tend to $\cos 1$ while the last one is convergent.

- $\int_1^\infty |\frac{\sin x}{x}| dx$ diverges. Indeed, $\int_{n\pi}^{(n+1)\pi} |\frac{\sin x}{x}| \geq \frac{1}{(n+1)\pi}\int_0^\pi \sin x dx = \frac{2}{(n+1)\pi}$ and hence we have $\int_1^\alpha |\frac{\sin x}{x}| dx \geq \sum_{n=2}^{[\alpha]} \frac{2}{\pi(n+1)} \to \infty$.

- $\int_0^\infty \frac{1}{\sqrt{x^3+1}} dx$ is convergent. Indeed, it is enough to consider $\int_1^\infty \frac{1}{\sqrt{x^3+1}} dx$, and since $\frac{1}{\sqrt{x^3+1}} \leq \frac{1}{\sqrt{x^3}} = \frac{1}{x^{\frac{3}{2}}}$, we have $\int_1^\beta \frac{1}{\sqrt{x^3+1}} dx \leq \int_1^\beta x^{-\frac{3}{2}} dx$, where the right-hand side is convergent.

- $\int_{-\infty}^\infty \frac{1}{\sqrt{x^4+1}} dx$ is convergent. Indeed, $\frac{1}{\sqrt{x^4+1}} \leq \frac{1}{\sqrt{x^4}} = \frac{1}{x^2}$ and $\int_1^\beta \frac{1}{x^2} dx = \int_{-\beta}^{-1} \frac{1}{x^2} dx = [-x^{-1}]_1^\beta = 1 - \frac{1}{\beta} \to 1$ as $\beta \to \infty$.

Dec. 02. Area and length

Area

We know the area of rectangles, triangles and disks Let us define the area of a more general region.

Definition 175. • Let $f \geq g$ be two integrable functions on an interval I. The area of the region between g, f is defined by the following:
$$D_{g,f} := \{(x, y) \in \mathbb{R}^2 : x \in I, g(x) \leq y \leq f(x)\}$$
$$\text{area}(D_{g,f}) := \int_I (f(x) - g(x)) dx.$$

- Even if I is not bounded, if the improper integral $\int_I (f(x) - g(x)) dx$ exists, then we define the area of the region $D_{g,f} = \{(x, y) \in \mathbb{R}^2 : x \in I, g(x) \leq y \leq f(x)\}$ by the same formula.

- If D is the disjoint union of such regions, then area(D) is the sum of the areas of such regions.

Example 176. • Rectancles. $D = \{(x, y) \in \mathbb{R}^2 : x \in I, a \leq y \leq b\}$, with the length $|I|$ and width $b - a$, then area$(D) = \int_I (b - a) dx = (b - a)|I|$.

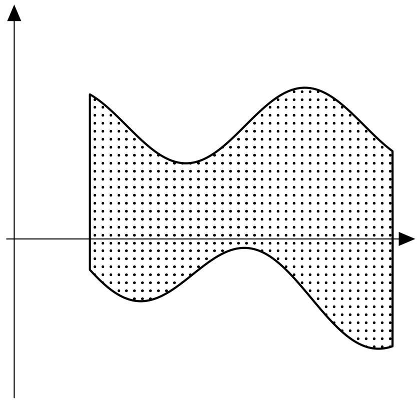


Figure 75: The area of the region between two functions.

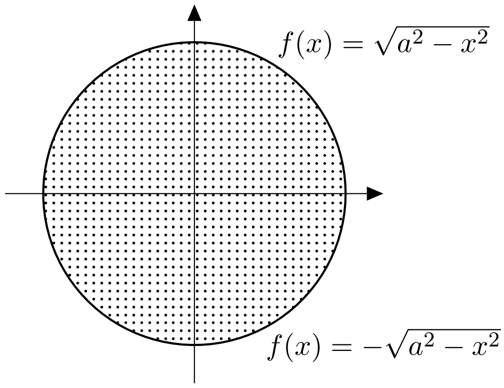


Figure 76: The area of a disk.

- Triangles. $D = \{(x,y) \in \mathbb{R}^2 : x \in [0,a], 0 \leq y \leq \frac{b}{a}x\}$, with length a and width b, then area$(D) = \int_0^a \frac{b}{a}xdx = [\frac{b}{2a}x^2]_0^a = \frac{ab}{2}$.
- Disks. $D = \{(x,y) \in \mathbb{R}^2 : \sqrt{x^2+y^2} \leq a\}$, with radius a, then D can be also written as

$$y^2 \leq a^2 - x^2 \Longleftrightarrow -\sqrt{a^2-x^2} \leq y \leq \sqrt{a^2-x^2}.$$

Furthermore, $-a \leq x \leq a$ because, if $x > a$ then there is no y such that $x^2 + y^2 \leq a^2$. Therefore,

$$D = \{(x,y) \in \mathbb{R}^2 : -a \leq x \leq a, -\sqrt{a^2-x^2} \leq y \leq \sqrt{a^2-x^2}\}$$

for which we can compute the area.

By our definition,

$$\begin{aligned} \text{area}(D) &= \int_{-a}^{a} (\sqrt{a^2-x^2} - (-\sqrt{a^2-x^2}))dx \\ &= 2\int_{-a}^{a} \sqrt{a^2-x^2}dx. \end{aligned}$$

By change of variables $x = a\sin\theta$ with $\frac{dx}{d\theta} = a\cos\theta$, this integral corresponds to that on the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$ because $a\sin\frac{\pi}{2} = a, a\sin(-\frac{\pi}{2}) = -a$,

$$\int_{-a}^{a} \sqrt{a^2-x^2}dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{a^2 - a^2\sin^2\theta}a\cos\theta d\theta.$$

Using $\cos^2\theta = \frac{\cos 2\theta + 1}{2}$,

$$\begin{aligned}\text{area}(D) &= 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{a^2 - a^2\sin^2\theta}\, a\cos\theta d\theta \\ &= 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a^2\cos^2\theta d\theta \\ &= a^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos 2\theta + 1) d\theta \\ &= a^2\left[\frac{\sin 2\theta}{2} + \theta\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\ &= a^2\pi.\end{aligned}$$

Length

A curve can be, at least partially, described using a function.

- segment. $\{(x, y) : x \in I, y = ax + b\}$.
- semicircle. $\{(x, y) : x \in (-a, a), y = \sqrt{a^2 - x^2}\}$.
- parabola. $\{(x, y) : x \in \mathbb{R}, y = x^2\}$.
- hyperbola. $\{(x, y) : x \in \mathbb{R}, y = \sqrt{x^2 + 1}\}$.

As we defined the area of a general region using integral, we can define the length of a curve with integral.

Definition 177. For a curve represented by $G_f := \{(x, y) : x \in [a, b], y = f(x)\}$, where f is differentiable and f' is continuous, we define the length by

$$\ell(G_f) := \int_a^b \sqrt{1 + f'(x)^2} dx.$$

If G_f is a union of such graphs, then $\ell(G_f)$ is defined to be the sum of the lengths of the partial graphs.

Let us see that this coincides with the case of segment: a segment that goes by a horizontally and b vertically is represented by $\{(x, y) : x \in [0, a], y = \frac{b}{a}x\}$. Hence $f(x) = \frac{b}{a}x, f'(x) = \frac{b}{a}$. By definition, $\ell(G_f) = \int_0^a \sqrt{1 + (\frac{b}{a})^2} dx = a\sqrt{1 + (\frac{b}{a})^2} = \sqrt{a^2 + b^2}$, which coincides with the length of the segment by the theorem of Pytagoras.

If a curve is a union of different parts, each of which is represented by a function f_j, then the length of the curve is the sum of the lengths of the parts.

Another possibile definition is to approximate a curve by segments: let $f(x)$ be a function on $I = [a, b]$ and take a parition P by $a = x_0 < x_1 < \cdots < x_n = b$. Correspondingly, we consider the attached segments $P_f(\{x_k\})$: $(x_0, f(x_0)), (x_1, f(x_1)), \cdots, (x_n, f(x_n))$. Let us recall $|P| = \max_{1\le k\le n-1}\{x_{k+1} - x_k\}$.

Proposition 178. *Let f be differentiable and f' be continuous. Then for any ϵ there is δ such that if $a = x_0 < x_1 < \cdots < x_n = b$ is a partition P with $|P| < \delta$, then $|\ell(G_f) - \ell(P_f(\{x_k\}))| < \epsilon$.*

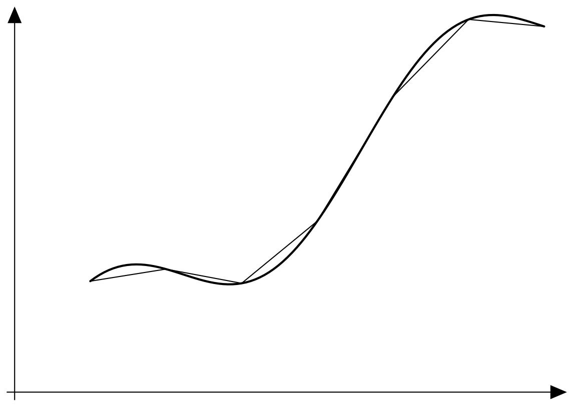


Figure 77: The length and segments.

Proof. By the mean value theorem, there are $x_k \le \xi_k \le x_{k+1}$ such that $f(x_k) - f(x_{k+1}) = f'(\xi_k)(x_k - x_{k+1})$. Since $P_f(\{x_k\})$ is the union of segments,

$$\begin{aligned}\ell(P_f(\{x_k\}) &= \sum_{k=1}^{n-1} \sqrt{(x_{k+1} - x_k)^2 + (f(x_{k+1}) - f(x_k))^2} \\ &= \sum_{k=1}^{n-1} \sqrt{(x_{k+1} - x_k)^2 + f'(\xi_k)^2 (x_{k+1} - x_k)^2} \\ &= \sum_{k=1}^{n-1} \sqrt{1 + f'(\xi_k)^2}(x_{k+1} - x_k).\end{aligned}$$

On the other hand, $\sqrt{1 + f'(x)^2}$ is continous and hence integrable. By uniform continuity, there is δ such that $|\sqrt{1 + f'(x)^2} - \sqrt{1 + f'(y)^2}| < \frac{\epsilon}{b-a}$ if $|x - y| < \delta$. With such a parition, we have $\underline{S}_I(\sqrt{1 + f'(x)^2}, P) \le \ell(P_f(\{x_k\}) \le \overline{S}_I(\sqrt{1 + f'(x)^2}, P)$.

If $|P|$ is small, the difference between these sides are smaller than ϵ, and $\underline{S}_I(\sqrt{1 + f'(x)^2}, P) \le \ell(G_f) \le \overline{S}_I(\sqrt{1 + f'(x)^2}, P)$. Therefore, $|\ell(G_f) - \ell(P_f(\{x_k\}))| < \epsilon$. □

Example 179. Semicircle. $I = [-1, 1], f(x) = \sqrt{1 - x^2}, f'(x) = \frac{x}{\sqrt{1-x^2}}$.

$$\begin{aligned}\ell(G_f) &= \int_{-1}^{1} \sqrt{1 + f'(x)^2} dx \\ &= \int_{-1}^{1} \frac{1}{\sqrt{1 - x^2}} dx = [\arcsin x]_{-1}^{1} = \pi\end{aligned}$$

(note that this is an improper integral). That is, the length of the circle is 2π.

Dec. 06. Series

Zeno's paradox

The Achilles and Tortoise paradox goes as follows: Achilles (ancient Greek hero, runs very fast) running behind a tortoise (walks very slowly). At the beginning, Achilles is 10 meter behind the tortoise. In the next moment, Achilles arrives at the position where the tortouse was there, but in the meantime it moves by 1 meter. Then Achilles arrives at the position where the tortoise was there in the previous moment, but in the meantime it moves by 0.1 meter. Then Achilles arrives...

So how can we be sure that Achilles catches up with the tortoise?

Series and convergence

Let us recall that we have considered **sequences** of numbers $a_1, a_2, \cdots, a_n, \cdots$, and the **series** $\sum_{k=1}^{n} a_k$, that is a new sequence

$$a_1, a_1 + a_2, a_1 + a_2 + a_3, \cdots, \sum_{k=1}^{n} a_k, \cdots$$

As this is a new sequence, we can consider its convergence or divergence.

That is, we say that a series $\sum_{k=1}^{n} a_k$ **converges** to $S \in \mathbb{R}$ if for each $\epsilon > 0$ there is $N \in \mathbb{N}$ such that for $n > N$ it holds that $|S - \sum_{k=1}^{n} a_k| < \epsilon$. We say taht the series **diverges** if for any $R > 0$ there is $N \in \mathbb{N}$ such that for $n > N$ it holds that $|\sum_{k=1}^{n} a_k| > R$. In other cases we just say that the series does not converge.

If a series converges, we denote the limit by $\sum_{k=1}^{\infty} a_k$. Sometimes we just write $\sum_n a_n$ a general term in a series.

Example 180.
- $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$. This diverges.
- $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$. This diverges.
- $\sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}$ (if $a \neq 1$) This converges if and only if $|a| < 1$.

For these examples, we know the exact form of the n-th sum. For other series, it is difficult to compute such general term, but still we may be able to say whether the series converges or not.

For example, let us take $a_n = \frac{1}{n}$ and consider $\sum_{k=1}^{n} \frac{1}{k}$. This is called the **harmonic series**. As we have seen, this sum is larger than the integral of $\frac{1}{x}$ on $[1, n+1]$:

$$\int_1^{n+1} \frac{1}{x} dx \leq \sum_{k=1}^{n} \frac{1}{k}.$$

On the other hand, we can calculate the left-hand side and we obtain $\int_1^{n+1} \frac{1}{x} dx = [\log x]_1^{n+1} = \log(n+1)$, and this diverges as $n \to \infty$. Therefore, $\sum_{k=1}^{n} \frac{1}{k}$ diverges as well.

Lemma 181. *If $\sum_{k=1}^{\infty} a_k$ converges, then $a_k \to 0$ as $k \to \infty$.*

Proof. As $\sum_{k=1}^{\infty} a_k$ is convergent to S, for any $\epsilon > 0$ there is N such that if $n > N$ then $|\sum_{k=1}^{n} a_k - S| < \frac{\epsilon}{2}$. In particular, we have $S - \frac{\epsilon}{2} < \sum_{k=1}^{n} a_k < S + \frac{\epsilon}{2}$ and $S - \frac{\epsilon}{2} < \sum_{k=1}^{n+1} a_k < S + \frac{\epsilon}{2}$.

From this it follows that $-\epsilon < \sum_{k=1}^{n+1} a_k - \sum_{k=1}^{n} a_k < \epsilon$, that is, $|\sum_{k=1}^{n+1} a_k - \sum_{k=1}^{n} a_k| = |a_{n+1}| < \epsilon$. This means that $a_n \to 0$. □

Example 182.
- $\sum_{k=1}^{n} k$ does not converge because $a_k = k$ diverges.
- $\sum_{k=1}^{n} (\frac{1}{2})^k$ converges to 1, and in this case indeed $(\frac{1}{2})^k$ converges to 0.
- $\sum_{k=1}^{n} \frac{1}{k}$ diverges, although in this case indeed $\frac{1}{k}$ converges to 0.

Theorem 183.
(i) Let $\sum_{k=1}^{\infty} a_k$ be convergent. Then for any $c \in \mathbb{R}$, $\sum_{k=1}^{\infty} c a_k$ is also convergent and $\sum_{k=1}^{\infty} c a_k = c \sum_{k=1}^{\infty} a_k$.

(ii) Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be convergent. Then $\sum_{k=1}^{\infty} (a_k + b_k)$ is also convergent and $\sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} (a_k + b_k)$

(iii) Let $\sum_{k=1}^{\infty} a_k$ convergent and $\sum_{k=1}^{\infty} b_k$ be divergent. Then $\sum_{k=1}^{\infty} (a_k + b_k)$ is divergent.

Proof. (i)(ii) These follow from the properties of sequences.

(iii) Suppose the contrary that $\sum_{k=1}^{\infty}(a_k + b_k)$ converges. Then by (ii) $\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty}(a_k + b_k) - a_k$ would converge, which contradicts the assumption that $\sum_{k=1}^{\infty} a_k + b_k$ diverges.

□

Example 184.

- The series $\sum_n (\frac{1}{n} + \frac{1}{2^n})$ diverges, because $\sum_n \frac{1}{n}$ diverges while $\sum_n \frac{1}{2^n}$ converges.
- The series $\sum_n 1$ and $\sum_n -1$ both diverge, but the sum $\sum_n (1-1) = \sum_n 0$ converges to 0.

Let us consider some cases where the sum converges.

Example 185.

- Telescopic series. Let b_n a sequence and $a_n = b_{n+1} - b_n$. We call such $\sum_n a_n$ a **telescopic series**. (Any series can be written in the form of telescopic series, but we are interested in the case where b_n is simpler than a_n)

  Then we have

  $$\sum_{k=1}^{n} a_n = (b_2 - b_1) + (b_3 - b_2) + \cdots + (b_{n+1} - b_n) = b_{n+1} - b_1.$$

  From this we infer that $\sum_n a_n$ is convergent if and only if b_n is convergent.

  For example, consider $a_n = \frac{1}{n(n+1)}$. This can be written as

  $$a_n = \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1},$$

  hence with $b_n = -\frac{1}{n}$, this is a telescopic series. By the argument above, we see that $\sum_{k=1}^{n} \frac{1}{n(n+1)} = b_{n+1} - b_1 = 1 - \frac{1}{n+1}$, and $\sum_{k=1}^{\infty} a_n = 1$.

  Next let us take $a_n = \log(n/(n+1))$. Then it holds that $a_n = \log n - \log(n+1)$, hence with $b_n = -\log n$ this is a telescopic series. As $b_n \to -\infty$, the series $\sum_n a_n$ diverges.

- Geometric series. Let us take $x \in \mathbb{R}, x \neq 1$. We know that $\sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}$. It is clear that the series converges to $\frac{1}{1-a}$ if $|a| < 1$, and diverges if $|a| > 1$. If $a = 1$, then the series is simply $\sum_{k=1}^{n} 1 = n + 1$ and diverges as well.

Geometric series can be seen as a function of x: For a given number $x \in \mathbb{R}$, we consider a sequence $a_n(x) = x^n$ and it holds for $|x| < 1$ that

$$\sum_{n=0}^{\infty} a_n(x) = \frac{1}{1-x}.$$

The right-hand side is again a function of x. In this sense, a convergent series which depends on x defines a new function.

We have seen other examples of this type:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} \cdots$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots$$

In all these cases, for a fixed $x \in \mathbb{R}$, we have seen that the right-hand converges by the Taylor formula with remainder.

In a similar way, we can define many other useful functions by series.

Dec. 9. Convergence criteria for positive series

Let$\{a_n\} \subset \mathbb{R}$ be a sequence. We have considered a series $\sum_{k=1}^n a_k$, which is a new sequence and its convergence or divergence.

When all the term are non-negative: $a_n \geq 0$, there are some criteria that can be often used to determine the convergence or divergence.

Theorem 186. *We have the following.*

- *Let $\{a_n\}, \{b_n\}$ be two sequences, $a_n \geq 0, b_n \geq 0$ such that there is $c > 0$ and $a_n \leq cb_n$ for n sufficiently large. In this case, if $\sum b_n$ converges, then so does $\sum a_n$. If $\sum a_n$ diverges, so does $\sum b_n$.*
- *Let $\{a_n\}, \{b_n\}$ be two sequences, $a_n \geq 0, b_n \geq 0$ such that $a_n/b_n \to c, c \neq 0, \infty$. Then $\sum a_n$ converges if and only if $\sum b_n$ converges.*

Proof.

- As $a_n \geq 0, b_n \geq 0$, the series $\sum_n a_n, \sum_n b_n$ are increasing. If $\sum_n b_n$ is convergent and $a_n \leq cb_n$, then $\sum_{k=1}^n a_k \leq \sum_{k=1}^n cb_k \leq \sum_{k=1}^\infty cb_k$, The other case is analogous. hence the former is bounded and increasing, therefore, it must converge.
- If $a_n/b_n \to\neq 0$, then this implies that $\frac{c}{2}b_n \leq a_n \leq 2cb_n$ for sufficiently large n, hence the previous point applies.

□

We have seen that, for $0 < a < 1$, $\sum_{k=1}^n a^k$ converges to $\frac{1}{1-a}$. We can use this fact to show the convergence of some other series.

Theorem 187 (root test)**.** *Let $a_n > 0$ be a sequence.*

(i) If $a_n^{\frac{1}{n}} \leq \theta < 1$ for n sufficiently large, then $\sum_n a_n$ converges.

(ii) If $a_n^{\frac{1}{n}} \geq \theta > 1$ for n sufficiently large, then $\sum_n a_n$ diverges.

(iii) Let $a_n^{\frac{1}{n}} \to a$. If $a < 1$, $\sum_n a_n$ converges. If $a > 1$, $\sum_n a_n$ diverges.

Proof. The series $\sum_n \theta^n$ converges if $\theta < 1$ (the geometric series) and diverges if $\theta > 1$ (θ^n does not tend to 0). By Theorem 186, and $a_n^{\frac{1}{n}} < \theta$ or $a_n^{\frac{1}{n}} > \theta$, the first two claims follow.

If $a_n^{\frac{1}{n}} \to a < 1$, then we can take θ such that $a_n^{\frac{1}{n}} < \theta < 1$ for n sufficiently large. The case $a_n^{\frac{1}{n}} \to a > 1$ is analogous. □

Example 188.

- $\sum_n \frac{1}{n3^n}$ is convergent.
- $\sum_n \frac{n2^n}{3^n}$ is convergent.
- $\sum_n \frac{n}{2^n}$ is convergent.

When $\lim_n a_n^{\frac{1}{n}} = 1$, this criterion does not give information. Indeed, $\sum_n \frac{1}{n}$ is divergent, but $\sum_n \frac{1}{n^2}$ is convergent (compare it with $\sum \frac{1}{n(n-1)}$), while in both cases $\lim(\frac{1}{n})^{\frac{1}{n}} = \lim(\frac{1}{n^2})^{\frac{1}{n}} = 1$.

Proposition 189 (ratio test)**.** *Let $a_n > 0$ be a sequence.*

(i) If $\frac{a_{n+1}}{a_n} \leq \theta < 1$ for n sufficiently large, then $\sum_n a_n$ converges.

(ii) If $\frac{a_{n+1}}{a_n} \geq \theta > 1$ for n sufficiently large, then $\sum_n a_n$ diverges.

(iii) Let $\frac{a_{n+1}}{a_n} \to \theta$. If $\theta < 1$, $\sum_n a_n$ converges. If $\theta > 1$, $\sum_n a_n$ diverges.

Proof. Let $\frac{a_{n+1}}{a_n} \le \theta < 1$ for $n \ge N$. Then,

$$a_{N+m} \le a_{N+m-1}\theta \le a_{N+m-2}\theta^2 \le \cdots \le a_N\theta^m.$$

Now $\sum_n a_N\theta^m$ is convergent, hence by Theorem 186, $\sum_n a_n$ is convergent. If $\frac{a_{n+1}}{a_n} \ge \theta > 1$, then a_n is increasing and does not convergent to 0.

If $\frac{a_{n+1}}{a_n} \to a < 1$ or > 1, then $\frac{a_{n+1}}{a_n} \le \theta < 1$ or > 1 for n sufficiently large, hence the claim follow from (i), (ii). □

Example 190.

- $\sum \frac{n}{2^n}$ is convergent.
- $\sum \frac{n^2}{n!}$ is convergent.
- $\sum \frac{(n!)^2}{2^{n^2}}$ is convergent.

When $a_{n+1}/a_n \to 1$ or $a_n^{\frac{1}{n}} \to 1$, we need to study the series better.

Lemma 191 (integral test)**.** *Let $\{a_n\}$ be a decreasing sequence of positive numbers and assume that there is a positive decreasing function $f(x)$ defined on $[1,\infty)$. If $a_n \le f(n)$ and $\int_1^\infty f(x)dx$ converges, then $\sum_n a_n$ converges. If $a_n \ge f(n)$ and $\int_1^\infty f(x)dx$ diverges, then $\sum_n a_n$ diverges.*

Proof. For the first case, we have $\sum_{k=2}^n a_n \le \int_1^n f(x)dx$, and the later converges, hence so does the former.

For the first case, we have $\sum_{k=1}^n a_n \ge \int_1^{n+1} f(x)dx$, and the later diverges, hence so does the former. □

Example 192. Let us fix $s \in \mathbb{R}$ and consider $\sum_{n=1}^\infty \frac{1}{n^s}$. We can compare this with $f_s(x) = \frac{1}{n^s}$. We know that $\int_1^\infty f_s(x)dx$ converges if and only if $s > 1$.

$\zeta(s)$ is called the Riemann zeta function.

Lemma 193 (condensation principle)**.** *Let $\{a_n\}$ be a decreasing sequence of positive numbers. Then $\sum a_n$ converges if and only if so does $\sum 2^n a_{2^n}$.*

Proof. Since a_n is decreasing and positive, $a_{2^n} \ge a_{2^n+1} \ge \cdots \ge a_{2^{n+1}}$, hence

$$2^n a_{2^n} \ge \sum_{k=2^n}^{2^{n+1}-1} a_k \ge 2^n a_{2^{n+1}}.$$

By summing this with respect to n,

$$\sum_{n=1}^N 2^n a_{2^n} \ge \sum_{n=1}^N \sum_{k=2^n}^{2^{n+1}-1} a_k = \sum_{n=1}^{2^{N+1}-1} a_k \ge \sum_{n=1}^N 2^n a_{2^{n+1}} = \frac{1}{2}\sum_{n=1}^N 2^{n+1}a_{2^{n+1}} = \frac{1}{2}\left(\sum_{n=1}^{N+1} 2^n a_{2^n} - a_1\right).$$

Therefore, $\sum_n a_n$ converges if and only if $\sum_n 2^n a_{2^n}$ converges by Theorem 186. □

Example 194. $\sum \frac{1}{n(\log n)^\alpha}$. By condensation principle, it is enough to study $\sum 2^n \frac{1}{2^n(\log 2^n)^\alpha} = \sum \frac{1}{(n\log 2)^\alpha} = \frac{1}{\log 2^\alpha}\sum \frac{1}{n^\alpha}$. Hence this converges if and only if $\alpha > 1$.

Dec. 10. Convergence criteria for general series

Let $\{a_n\}$ be a sequence, not necessarily positive. We say that the series $\sum_n a_n$ **converges absolutely** if $\sum_n |a_n|$ converges.

Lemma 195. *If $\sum_n |a_n|$ converges, then so does $\sum_n a_n$.*

Proof. As $\sum_n |a_n|$ converges, for $\epsilon > 0$ there is $N \in \mathbb{N}$ such that for $m > n > N$ it holds that $\sum_{k=n}^m |a_n| < \epsilon$. Then it holds that $|\sum_{k=n}^m a_k| < \epsilon$ by the triangle inequality. This means that $\sum_{k=1}^n a_k$ is a Cauchy sequence, hence it converges. □

This Lemma, combined with the criteria for positive series, enables us to show convergence of many series.

Example 196.

- $\sum_{n=1}^\infty \frac{1}{(-1)^n n^s}$ converges absolutely if $s > 1$. Indeed, with $a_n = \frac{1}{(-1)^n n^s}$, $|a_n| = \frac{1}{n^s}$, and we know that $\sum_{n=1}^\infty \frac{1}{n^s}$ converges if and only if $s > 1$.
- $\sum_{n=1}^\infty \frac{n}{(-1)^n 2^n}$ converges absolutely. Indeed, with $a_n = \frac{n}{(-1)^n 2^n}$, $|a_n| = \frac{n}{2^n}$ and by root test,

$$\left(\frac{n}{2^n}\right)^{\frac{1}{n}} = \frac{n^{\frac{1}{n}}}{2} \to \frac{1}{2} < 1.$$

  Therefore, $\sum_n |a_n|$ converges.

If a series $\sum_n a_n$ converges absolutely, its limit does not depend on the order: indeed, as it is absolutely convergent, then its positive elements $b_n = \begin{cases} a_n & \text{if } a_n > 0 \\ 0 & \text{if } a_n \leq 0 \end{cases}$ and the negative elements $c_n = a_n - b_n$ are both convergent. Therefore, even if we sum first the positive elements $\sum_n b_n$ and then the negative elements $\sum_n c_n$, the result is the same: $\sum_n b_n + \sum_n c_n = \sum_n a_n$.

On the other hand, if a series $\sum_n a_n$ converges but not absolutely, their positive and negative parts both diverges (because otherwise it would be absolute convergence). By rearranging the sum, one can make it diverge to ∞ (by taking many elements of b_n) or to $-\infty$ (by taking many elements of c_n).

A series whose terms change sign at each stem is called an **alternating series**. That is, for $a_n > 0$, it is given by

$$\sum_{n=1}^\infty (-1)^{n-1} a_n = a_1 - a_2 + a_3 - a_4 + \cdots,$$

Lemma 197 (Leibniz criterion)**.** *Let $\{a_n\}$ be a decreasing sequence with positive terms and assume that $a_n \to 0$. Then $\sum (-1)^{n-1} a_n$ converges.*

Proof. Let $s_n = \sum_{k=1}^n (-1)^k a_k$. Then $s_{2n} = \sum_{k=1}^n (-a_{2k-1} + a_{2k})$ is decreasing. Analogously $s_{2n+1} = -a_1 + \sum_{k=1}^n (a_{2k} - a_{2k+1})$ is increasing. In addition, $s_{2n} - s_{2n+1} = -(-a_{2n+1}) = a_{2n+1} \geq 0$. Hence s_{2n} and s_{2n+1} converge to $\overline{s}$ and $\underline{s}$, respectively, and $\overline{s} \geq \underline{s}$. But $s_{2n+1} \leq \underline{s} \leq \overline{s} \leq s_{2n}$ and $s_{2n} - s_{2n+1} = a_{2n+1} \to 0$, hence $\overline{s} = \underline{s}$. □

Example 198.

- $\sum_n (-1)^{n-1} \frac{1}{n}$ is convergent. Note that this series is not absolutely convergent, indeed, $\sum_n \frac{1}{n}$ is divergent.
- $\sum_n (-1)^{n-1} \frac{1}{\log(n+1)}$ is convergent. Note that this series is not absolutely convergent, indeed, $\sum_n \frac{1}{\log(n+1)}$ is divergent.

- The sequence $\sum_{k=1}^n \frac{1}{k} - \log n$ converges. Indeed, this can be seen as

$$\begin{aligned}
&1+\frac{1}{2}+\frac{1}{3}\cdots+\frac{1}{n}-\int_1^2\frac{1}{x}dx-\int_2^3\frac{1}{x}dx+\cdots-\int_{n-1}^n\frac{1}{x}dx\\
&=\frac{1}{1}-\int_1^2\frac{1}{x}dx+\frac{1}{2}-\int_2^3\frac{1}{x}dx+\cdots-\int_{n-1}^n\frac{1}{x}dx+\frac{1}{n},
\end{aligned}$$

and the last expression is an alternating series. Note that $\frac{1}{k} > \int_k^{k+1}\frac{1}{x}dx > \frac{1}{k+1}$. Furthermore, $\frac{1}{k} \to 0$ as $k \to \infty$. Therefore, the Leibniz criterion applies and the series converges (to a number known as the Euler-Mascheroni constant).

Lemma 199 (Abel's partial summation formula)**.** *Let* $\{a_n\}, \{b_n\}$ *be two sequences, and let* $A_n = \sum_{k=1}^n a_k$*. Then we have the identity* $\sum_{k=1}^n a_k b_k = A_n b_{n+1} + \sum_{k=1}^n A_k(b_k - b_{k+1})$*.*

Proof. Let us define $A_0 = 0$. By computing the right-hand side,

$$\begin{aligned}
A_n b_{n+1} + \sum_{k=1}^n A_k(b_k - b_{k+1}) &= A_n b_{n+1} + \sum_{k=1}^n A_k b_k - \sum_{k=2}^{n+1} A_{k-1} b_k\\
&= A_n b_{n+1} + \sum_{k=1}^n A_k b_k - \sum_{k=1}^{n+1} A_{k-1} b_k\\
&= A_n b_{n+1} + \sum_{k=1}^n a_k b_k - A_n b_{n+1}\\
&= \sum_{k=1}^n a_k b_k.
\end{aligned}$$

□

Theorem 200 (Dirichlet's test)**.** *Let* $\sum_n a_n$ *be a series and assume that* $A_n = \sum_{k=1}^n a_k$ *is a bounded sequence. Let* $b_n > 0$ *be a decreasing sequence and* $b_n \to 0$*. Then the series* $\sum_n a_n b_n$ *converges.*

Proof. By Lemma 199, we have $\sum_{k=1}^n a_n b_n = A_n b_{n+1} + \sum_{k=1}^n A_k(b_k - b_{k+1})$. As b_k is decreasing, we have $b_k - b_{k+1} > 0$ and since A_k is bounded, say by C, we have $\sum_{k=1}^n |A_k(b_k - b_{k+1})| \le \sum_{k=1}^n C(b_k - b_{k+1})$. The latter is a telescopic series and it is equal to $C(b_1 - b_{n+1})$, which converges (to Cb_1). Therefore, $\sum_{k=1}^n A_k(b_k - b_{k+1})$ is absolutely convergent.

On the other hand, $A_n b_{n+1}$ tends to 0 because $|A_n| < C$ and $b_{n+1} \to 0$. Altogether, the series $\sum_{k=1}^n a_n b_n$ converges. □

Theorem 201 (Abel's test)**.** *Let* $\sum_n a_n$ *be a convergent series and assume* b_n *is monotonic and convergent. Then the series* $\sum_n a_n b_n$ *converges.*

Proof. We may assume that b_n is decreasing (otherwise we can consider $-b_n$). In this case, $A_n = \sum_n a_n$ is bounded (because it is convergent), and with $B = \lim b_n$, $b_n - B$ is decreasing and converges to 0.

Hence we can apply Dirichlet's test to the first term of $\sum_n a_n b_n = \sum_n a_n(b_n - b) + \sum_n a_n b$, where the last term is convergent because $\sum_n a_n$ is convergent. □

Example 202. Let us consider the series $\sum_n \frac{\sin n\pi\theta}{n^s}$, where $\theta = \frac{p}{q}$ is a rational number and $s > 0$. This is not absolutely convergent, nor alternating (if $q \neq 2$). On the other hand, the sequence $\sin n\pi\theta$ is periodic, because $\sin x$ is 2π-periodic, that is, $\sin\frac{n\pi p}{q} = \sin\frac{(n+2q)\pi p}{q}$. Let us assume q is even and p is odd. Then $\sin\frac{(n+q)\pi p}{q} = -\sin\frac{n\pi p}{q}$, and hence also the sum $\sum_n \sin n\pi\theta$ is periodic, and in particular bounded.

In this case, as $\frac{1}{n^s}$ is monotonically decreasing and tends to 0, we can apply Dirichlet's test and conclude that $\sum_{n=1}^{\infty} \frac{\sin n\pi\theta}{n^s}$ is convergent. Similarly, $\sum_{n=1}^{\infty} \frac{\cos n\pi\theta}{n^s}$

The same holds for q odd, and actually for any $\theta \in \mathbb{R}$.

Dec. 14. Ordinary differential equations

Many scientific questions are expressed in terms of differential equation (equation about functions and their derivatives).

- The equation of motion in a gravitational field $m\frac{d^2x}{dt^2} = -\frac{mMG}{x^2}$
- The heat equation $\frac{\partial u}{\partial t} = \alpha \frac{\partial u^2}{\partial x^2}$ (this is **partial differential equation**, because it contains partial derivatives, studied in Mathematical Analysis II) .
- The SIR model in epidemiology $\frac{dS}{dt} = -\frac{\beta S(t)I(t)}{N}$, $\frac{dI}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t)$, $\frac{dR}{dt} = \gamma I(t)$

This is because the rate of change (the derivative) is often determined by the current status (the function). For example, in the equation of motion, the gravitational force $-\frac{mMG}{x^2}$ depends on the place of a particle $x(t)$, while the force determines the the rate of change of the speed (the acceleration), and the speed is $x'(t)$, hence the second derivative appears on the left-hand side.

Once the scientific problem is written in the form of differential equation, it is a mathematical problem to solve it, that is, to find functions that satisfies the given equation.

In the following, $y(x)$ will be a function of x and the derivatives are denoted by $y'(x), y''(x)$ and so on. Some more examples of differential equation are

- $y'(x) = y(x)$
- $y'(x) = x^3y(x) + \sin(xy''(x))$
- Sometimes we just write this as $y' = x^3y + \sin(xy'')$, keeping in mind that y is a function of x.

In a differential equation, certain higher derivative of y may appear. The highest order of the derivative of y is called the **order** of the differential equation. For example,

- $y'(x) = 2y(x)$ is a first-order differential equation.
- $y'(x) = x^3y(x) + \sin(xy''(x))$ is a second-order differential equation.

We need to find functions $y(x)$ that satisfy the given equation. This is why it is called a differential equation. Compare it with an algebraic equation $x^2+3x-4 = 0$, where we need to find numbers that satisfy this equation.

Let us consider first-order differential equations. In an abstract form, we can write it as

$$y' = f(x, y),$$

where f is explicitly written in examples, while y is the unknown functions which we need to find. In the example $y'(x) = 2y(x)$, we take $f(x, y) = 2y$. A **solution** of a differential equation is a (differentiable) function that satisfies this equation. For example, by taking $y(x) = Ce^{2x}$, we can check that this is a solution:

$$y'(x) = 2Ce^{2x} = 2y(x).$$

Some first-order differential equations

The simplest case is where f does not depend on y: that is,

$$y'(x) = f(x).$$

This means that f is the derivative of y, or y is a primitive of f. Therefore, y can be obtained by integrating f: $y(x) = \int f(x)dx + C$. Indeed, this y satisfies the given equation for any $C \in \mathbb{R}$, and there is no other solution.

Example 203. When a ball falls freely without drag, the speed $-gx$ is proportional to the time x. As the speed is the derivative of the position y, we have the equation

$$y'(x) = -gx.$$

This can be solved by integration, that is $y(x) = \int(-gx)dx = -\frac{gx^2}{2} + C$. The constant C depends on the position where the ball starts to fall.

As we see in this example, a differential equation may have many solutions. In practice, we are interested in one of them which satisfies additional conditions, the **initial conditions** or **boundary conditions**, that give the value of y, y' at a given time x.

Next, let us consider again the simplest differential equation $y' = f(x, y)$ where f depends on y.

Theorem 204. *Let $a, C \in \mathbb{R}$. Then there is only one (differentiable) function y such that $y'(x) = ay(x)$ and $y(0) = C$.*

Proof. We know that there is one such function: $y(x) = Ce^{ax}$. Indeed, we can check that $y'(x) = aCe^{ax} = ay(x)$ and $y(0) = Ce^0 = C$.

Suppose that there is $g(x)$ with the same condition. Let $h(x) = e^{-ax}g(x)$, then $h'(x) = -ae^{-ax}g(x) + e^{-ax}g'(x) = -ae^{-ax}g(x) + ae^{-ax}g(x) = 0$ for all $x \in \mathbb{R}$, hence $h(x)$ must be a constant. As $h(0) = e^0 g(0) = C$, $h(x) = C$ hence $g(x) = Ce^{ax}$. □

Let us consider when we see the equation $y' = ay$.

- A very typical example is radioactive atoms. Let $y(x)$ be the number of a single species of radioactive atoms at time x. It is known that each atom decays, independently from other atoms, in a certain time period by a certain probability. This means that, at each moment, the rate of decrease in numbers $y(x)$ is proportional to $y(x)$. With a constant a, we can write this as

  $$y'(x) = -ay(x).$$

  If there are C atoms at time $x = 0$, we know that the solution is $y(x) = Ce^{-ax}$, hence the number of radioactive atoms decays exponentially. This can be written more conveniently as $y(x) = Ce^{-ax} = C2^{-ax/\log 2}$. Then with $T = \frac{\log 2}{a}$, we have $y(x) = C2^{-x/T}$, and it is clear that the number of atoms halves in time T. T is called the **half life** of this particular species of atom.

- Another instance is the SIR model in epidemiology. We consider the total population N, the numbers of $S(t)$ (succeptible), $I(t)$ (infected) and $R(t)$ (removed/recovered). It is assumed that each infected people has contact with a certain number of people in each day, hence this number is proportional to $\frac{S(t)}{I(t)}$, and assume that in each such contact transmission

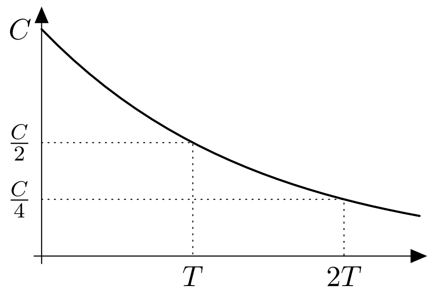


Figure 78: The exponential decay $y(x) = Ce^{-ax} = C2^{-ax/\log 2}$ The half-life is $T = \log 2/a$.

occurs by the rate β. On the other hand, each infected people lose infectivity by the rate γ.

$$\frac{dS}{dt} = -\frac{\beta S(t)I(t)}{N}$$
$$\frac{dI}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t)$$
$$\frac{dR}{dt} = \gamma I(t)$$

It is difficult to solve this set of equations. Yet, we can understand the behaviour when there are few infected people $I(t)$ compared to the total number N. When $I(t)$ is small, then $R(t)$ is also small and $S(t) = N - I(t) - R(t)$ is close to N. By putting $S(t) = N$, we have

$$\frac{dI}{dt} = (\beta - \gamma)I(t).$$

As a function of t, we know that $I(t) = Ce^{(\beta-\gamma)t}$, where C is the number of infected at day $t = 0$. This epidemic grows when $\beta - \gamma > 0$, and decays when $\beta - \gamma < 0$. $R_0 = \frac{\beta}{\gamma}$ is called the **basic reproduction number in the SIR model**. When $R_0 > 1$ the epidemic grows and when R_0 it decays.

Dec. 16. Ordinary differential equations

Linear equations with constant coefficients

Let us next consider a linear differential equation of the form

$$y'' + P(x)y' + Q(x)y = R(x).$$

This contains the second derivative y'', hence it is a second-order differential equation. Here, $P(x), Q(x), R(x)$ are a known function and we have to find a function $y(x)$ which satisfies this equation. Such an equation is called **linear homogeneous differential equation** of second-order. (Linear means that there is no term containing $y^2, y^3, (y')^2$ etc., homogeneous means that the right-hand side (the term which does not depend on y) is 0.)

If y_1, y_2 are two solutions of a homogeneous equation, then then so is $y_3(x) = ay_1(x) + by_2(x)$, because

$$\begin{aligned}
&y_3''(x) + P(x)y_3'(x) + Q(x)y_3(x)\\
&= ay_1''(x) + by_2''(x) + P(x)(ay_1'(x) + by_2(x)) + Q(x)(ay_1(x) + by_2(x))\\
&= a(y_1''(x) + P(x)y_1'(x) + Q(x)y_1''(x)) + b(y_2''(x) + P(x)y_2'(x) + Q(x)y_2''(x)) = 0
\end{aligned}$$

If y_1, y_2 are two solutions of a linear equation, then the difference $y_1 - y_2$ is a solution of the equation where $R(x)$ is set to 0.

A simplest of such equations is one where $P(x), Q(x), R(x)$ are constant:

$$y'' + ay' + by = 0.$$

As we will see, such equations appear naturally in physics.

An even simpler case is where $a = 0$:

$$y'' + by = 0.$$

Let us start with solutions of this type.

- Case 1. $b = 0$. In this case, we have $y'' = 0$. This means that $y' = C_1$ (constant) and further $y = C_1x + C_2$. It is easy to see that any solution is of this form.
- Case 2. $b < 0$. In this case, the equation can be written as $y'' = k^2y$ where $b = -k^2$ and we can take easily check that $y(x) = C_1e^{kx} + C_2e^{-kx}$ is a solution for any constant C_1, C_2. Indeed, $y'(x) = kC_1e^{kx} - kC_2e^{-kx}$ and $y''(x) = k^2C_1e^{kx} + (-k)^2C_2e^{-kx} = k^2(C_1e^{kx} + C_2e^{-kx}) = k^2y(x)$.
- Case 3. $b > 0$. In this case, the equation can be written as $y'' = -k^2y$ where $b = -k^2$. There are solutions of the form $y(x) = C_1 \sin(kx) + C_2 \cos(kx)$ is a solution for any constant C_1, C_2.

These solutions are all, and no other solutions (we present later the general uniqueness theorem).

Note that, in all these cases, there are two constants C_1, C_2. If we require an initial condition

- $y(a) = b_1$
- $y'(a) = b_2$

these constants are fixed.

For example, in Case 2 with $a = 0$, $y(x) = C_1e^{kx} + C_2e^{-kx}$, we should have $y(0) = C_1 + C_2 = b_1$ and $y'(0) = kC_1 - kC_2 = b_2$, hence $C_1 = \frac{1}{2}(b_1 + \frac{b_2}{k}), C_2 = \frac{1}{2}(b_1 - \frac{b_2}{k})$.

Let us consider the general case $y'' + ay' + by = 0$. This can be reduced to the special case above as follows. We write $y(x) = u(x)e^{-\frac{ax}{2}}$, then

$$\begin{aligned}
y'(x) &= u'(x)e^{-\frac{ax}{2}} - \frac{a}{2}u(x)e^{-\frac{ax}{2}} = u'(x)e^{-\frac{ax}{2}} - \frac{a}{2}y(x) \\
y''(x) &= u''(x)e^{-\frac{ax}{2}} - \frac{a}{2}u'(x)e^{-\frac{ax}{2}} - \frac{a}{2}u'(x)e^{-\frac{ax}{2}} + \frac{a^2}{4}u(x)e^{-\frac{ax}{2}} \\
&= u''(x)e^{-\frac{ax}{2}} - au'(x)e^{-\frac{ax}{2}} + \frac{a^2}{4}u(x)e^{-\frac{ax}{2}}
\end{aligned}$$

Therefore, if y is a solution of this equation, it must hold that

$$\begin{aligned}
0 &= (u''(x)e^{-\frac{ax}{2}} - au'(x)e^{-\frac{ax}{2}} + \frac{a^2}{4}u(x)e^{-\frac{ax}{2}}) + a(u'(x)e^{-\frac{ax}{2}} - \frac{a}{2}u(x)e^{-\frac{ax}{2}}) + bu(x)e^{-\frac{ax}{2}} \\
&= u''(x)e^{-\frac{ax}{2}} + \left(b - \frac{a^2}{4}\right)u(x)e^{-\frac{ax}{2}}
\end{aligned}$$

hence if u safisfies $u''(x) + \left(b - \frac{a^2}{4}\right)u(x) = 0$, then $y(x) = u(x)e^{-\frac{ax}{2}}$ satisfies $y'' + ay' + by = 0$. We know how to solve the former, hence so the latter.

Example 205. Consider the equation $y'' + y' - y = 0$. Then we can write $y = ue^{-\frac{x}{2}}$ and then u should satisfy $u'' - \frac{5}{4}u = 0$. We know that $u(x) = C_1e^{\frac{\sqrt{5}}{2}x} + C_2e^{-\frac{\sqrt{5}}{2}x}$ is a solution of this, hence $y = C_1e^{\frac{\sqrt{5}-1}{2}x} + C_2e^{-\frac{\sqrt{5}+1}{2}x}$.

Next, let us consider the inhomogeneous case, that is $y'' + ay' + by = R(x)$. In some cases we can find solutions.

Example 206. Take $R(x) = x^2$. Then $y(x) = \frac{1}{b}(x^2 - \frac{2ax}{b} + \frac{2a^2-2b}{b^2})$ is a solution. Indeed, $y'(x) = \frac{1}{b}(2x - \frac{2a}{b}), y''(x) = \frac{2}{b}$.

A general solution can be obtained by adding a solution of the homogeneous version $y'' + ay' + by = 0$ to this solution.

Physical examples

- Simple harmonic motion. Consider a mass m which is attached to a spring. Let us call $x(t)$ the position of the mass. When a spring is stretched by the distance r, then it pulls back the mass by the force kr. Similarly, when a spring is pressed by the distance r (hence the mass is displaced to $-r$), then it pushes back the mass by the force kr. Together with the direction of the force, it can be written as $-kx$.

  The equation of motion is about the variable $x(t)$ and the acceleration is $a(t) = x''(t)$, hence $F(x) = ma = mx''$ becomes

  $$mx'' = F(x) = -kx.$$

  That is, $x'' + \frac{k}{m}x = 0$. The general solution of this is

  $$x(t) = C_1 \sin\sqrt{\frac{k}{m}}t + C_2 \cos\sqrt{\frac{k}{m}}t.$$

  If we pull the mass to a and leave quietly at time $t = 0$, then the solution should have $x(0) = a, x'(0) = 0$. That is, $C_2 = a$ and $C_1 = 0$, and the special solutions is

  $$x(t) = a\cos\sqrt{\frac{k}{m}}t.$$

  This means that the mass oscilates between $-a$ and a.s

  In general, if we specify the values $x(0)$ and $x'(0)$, then there is only one solution. These values are called the **initial conditions**.

- In addition to the previous example, let us consider the case where the mass lies on a floor hence receives the friction. The friction is proportional to the velocity and in the converse direction. Therefore, the equation of motion is

  $$mx''(t) = -kx(t) - cx'(t),$$

  or $x'' + \frac{c}{m}x + \frac{k}{m}x = 0$.

  The solution is given by solving $u'' + (\frac{k}{m} - \frac{1}{4}(\frac{c}{m})^2)u = 0$. If $-s^2 = \frac{k}{m} - \frac{1}{4}(\frac{c}{m})^2 < 0$, then we have $x(t) = C_1 e^{-(\frac{c}{2m}-s)t} + C_2 e^{-(\frac{c}{2m}+s)t}$.

  If $x(0) = a, x'(0) = 0$, then $C_1 + C_2 = a, -(\frac{c}{2m}-s)C_1 + -(\frac{c}{2m}+s)C_2 = 0$ hence $C_1 = C_2 = \frac{a}{2}$, and

  $$x(t) = \frac{a}{2}\left(e^{-(\frac{c}{2m}-s)t} + e^{-(\frac{c}{2m}+s)t}\right)$$

  Note that $\frac{c}{2m} > s$, hence this decays exponentially. This means that the mass arrives at 0 without going back and forth.

  We leave the remaining case $\frac{k}{m} - \frac{1}{4}(\frac{c}{m})^2 \geq 0$ as exercises.

Dec. 17. Ordinary differential equations

General remarks

Many interesting differential equations are nonlinear: just for example, the motion in a gravitational field is given by

$$mx'' = -\frac{mMG}{x^2}$$

(in one-dimension). And it is difficult to solve such nonlinear equations.

Let us consider a first-order differential equation

$$y' = f(x, y),$$

that is, f is a given function of two variables and the question is to find a function $y(x)$ such that $y'(x) = f(x, y(x))$ for x in a certain interval.

If from the differential equation $y' = f(x, y)$ we can derive a relation between x, y of the form

$$F(x, y, C) = 0,$$

where $F(x, y, C)$ is another two-variable function with a parameter C (hence 3-variables), then we say that the differential equation is solved, or integrated. This is because the relation $F(x, y, C) = 0$ for a fixed number C defines a function $y(x)$ implicitly: recall that, if $F(x, y) = x^2 + y^2 - C^2$, then it defines the function(s) $y(x) = \pm\sqrt{C^2 - x^2}$.

Separable differential equations

We call a first-order differential equation $y' = f(x, y)$ **separable** if it can be written in the form $y' = Q(x)R(y)$, where $Q(x)$ is a function of x alone and $R(y)$ is a function of y alone. For example,

- $y' = x^3$
- $y' = yx^2$
- $y' = \sin y \log x$.

When $R(y) \neq 0$, we can write this in the form $A(y)y' = Q(x)$.

Theorem 207. *Let $G(y)$ be a primitive of $A(y)$ and $H(x)$ be a primitive of $Q(x)$. Then any differentiable function $y(x)$ which satisfies*

$$G(y(x)) = H(x)$$

satisfies the differential equation $A(y(x))y'(x) = Q(x)$, and conversely, any solution $y(x)$ satisfies this equation for certain $H(x)$.

Proof. Let $y(x)$ satisfies the equation above, then by the chain rule, we have $\frac{d}{dx}G(y(x)) = y'(x)A(y(x))$, while $\frac{d}{dx}H(x) = Q(x)$, hence we obtain $A(y(x))y'(x) = Q(x)$ by differentiating $G(y(x)) = H(x)$.

Conversely, if $y(x)$ is the solution of the differential equation, then by integrating both sides of $A(y(x))y'(x) = Q(x)$ by substitution, we have $G(y(x)) = H(x) + C$ for some constant C. Note that $H(x) + C$ is a primitive of $Q(x)$, hence we proved the claim. □

Example 208.

- Consider $y' = y^2 x$. This can be written as $\frac{y'}{y^2} = x$. Each sides can be integrated, and we obtain

$$-\frac{1}{y} = \frac{x^2}{2} + C,$$

or $y = -\frac{1}{\frac{x^2}{2} + C}$.

- Consider $y' = \frac{x}{y}$. This can be written as $yy' = x$. Each sides can be integrated, and we obtain

$$\frac{y^2}{2} = \frac{x^2}{2} + C,$$

or $y = \sqrt{x^2 + 2C}$.

- Consider $xy' + y = y^2$. This can be written as $y' = \frac{y(y-1)}{x}$, or $\frac{y'}{y(y-1)} = \frac{1}{x}$. Each sides can be integrated, and we obtain

$$\log\left|1 - \frac{1}{y}\right| = \log x + C,$$

or $|1 - \frac{1}{y}| = C'x$ for some constant $C' = e^C$.

Not many equations are separable, but some can be reduced to a separable equation. For example, consider the case where $y' = f(x, y)$ and

$$f(tx, ty) = f(x, y)$$

for any $t \neq 0$. In this case, we can introduce $v = \frac{y}{x}$, or $y = vx$ and hence $y' = v'x + v$. Therefore, if y' is the solution of the equation above, then it must hold that

$$v'x + v = y' = f(x, y) = f(1, y/x) = f(1, v).$$

This can be written as $v' = (f(1, v) - v)\frac{1}{x}$, hence is separable. Once v is obtained as a function of x, we can recover $y = vx$.

Example 209. Consider $y' = \frac{y-x}{x+y}$. $f(tx, ty) = \frac{ty-tx}{tx+ty} = \frac{y-x}{x+y} = f(x, y)$, hence this is can be solved by introducing $y = vx$.

We have $v' = (f(1, v) - v)\frac{1}{x} = (\frac{v-1}{1+v} - v)\frac{1}{x} = -\frac{1+v^2}{1+v}\frac{1}{x}$, and we have

$$\int \frac{1+v}{1+v^2} dv = \arctan v + \frac{1}{2}\log(1 + v^2)$$
$$-\int \frac{1}{x} dx = -\log x + C$$

By bringing back $y = vx$, we have $\arctan \frac{y}{x} + \frac{1}{2}\log(x^2 + y^2) = C$.

Concrete applications

- Consider a falling body in a resisting medium. For example, we drop a ball from a window. The gravitational force is constant g when the body moves the distance much shorter than the radius of the Earth. In addition, the ball is resisted by the air and the resistance is proportional to the velocity. To express this in a differential equation, let $v(t)$ be the velocity of the ball at time t, we leave it at time $t = 0$ from the height 0. Then

$$mv' = mg - kv,$$

or $\frac{mv'}{mg-kv} = 1$. This is separable. We have $-\frac{m}{k}\log(mg - kv) = t + C$, or $\frac{mg}{k}e^{-\frac{k}{m}t+C} = v(t)$.

- Let us consider a small particle in a large medium. If the temperature of the particle and that of the medium is different, then the changing rate of the temperature is proportional to the difference of the temperature. As the medium is large, we may assume that only the temperature $y(t)$ changes with $y(0) = T$, while the medium remain in the same temperature M. In a differential equation,

$$y'(t) = k(M - y(t)).$$

This is again separable.

Dec. 20. Complex numbers

Consider the equation $x^2 + 1 = 0$. There is no real solution x that satisfies this equation because $x^2 > 0$, hence $x^2 + 1 > 0$. Yet, it is possible to *extend* the set of real numbers in such a way to include solutions to this equation. One of such solutions is denoted by i, that means $i^2 = -1$.

Complex numbers as ordered pairs

We expect that i behaves in a way similar to the real numbers. We can consider $a + ib, a, b \in \mathbb{R}$, it should hold that $(a + ib)(c + id) = ac + i^2bd + i(ad + bc) = ac - bd + i(ad + bc)$.

This can be formulated as follows: A **complex number** is a pair (a, b) of real numbers, and we define

- (Equality) $(a, b) = (c, d)$ as complex numbers if and only if $a = c$ and $b = d$.
- (Sum) $(a, b) + (c, d) = (a + c, b + d)$.
- (Product) $(a, b)(c, d) = (ac - bd, ad + bc)$.

This is a formal definition, and it is customary to denote (a, b) by $a + ib$. a is called the **real part** and b is called the **imaginary part**.

The set of complex numbers is denoted by $\mathbb{C}$.

Theorem 210. *For $x, y, z \in \mathbb{C}$, we have the following.*

- *(Commutativity) $x + y = y + x, xy = yx$.*
- *(Associativity) $x + (y + z) = (x + y) + z, x(yz) = (xy)z$.*
- *(Distributive law) $x(y + z) = xy + xz$.*

Proof. All these properties can be shown by calculating both sides and using the properties of real numbers. For example, for $x = a + ib = (a, b), y = c + id = (c, d)$, $x + y = (a, b) + (c, d) = (a + c, b + d) = (c + a, d + b) = y + x$, $xy = (a, b)(c, d) = (ac - bd, ad + bc) = (ca - db, da + cb) = (c, d)(a, b) = yx$.

Other properties are left as exercises. □

Real numbers in complex numbers

Let us consider the complex number $(0, 0)$. For any other complex number (a, b), it holds that $(0, 0)(a, b) = (0, 0)$ and $(0, 0) + (a, b) = (a, b)$. That is, $(0, 0)$ plays the same role of $0 \in \mathbb{R}$.

Next, we consider $(1, 0)$. For any other complex number (a, b), it holds that $(1, 0)(a, b) = (a, b)$. That is, $(1, 0)$ plays the same role of $1 \in \mathbb{R}$.

For any $(a, b) \in \mathbb{C}$, it holds that $(a, b) + (-a, -b) = (0, 0)$. We write $-(a, b)$ for $(-a, -b)$.

For any $(a, b) \in \mathbb{C}, (a, b) \neq (0, 0)$, it holds that $(a, b)(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}) = (\frac{a^2+b^2}{a^2+b^2}, \frac{-ab+ba}{a^2+b^2}) = (1, 0)$. We write $(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$ for $1/(a, b)$, or $(a, b)^{-1}$.

All these properties tell that $\mathbb{C}$ is an object called a field.

For any number $a \in \mathbb{R}$, the complex number $(a, 0)$ behaves exactly as $a \in \mathbb{R}$. Indeed, $(a, 0) + (b, 0) = (a + b, 0), (a, 0)(b, 0) = (ab, 0)$.

Imaginary unit

The complex number $(0, 1)$ satisfies $(0, 1)^2 = (-1, 0)$. Indeed, $(0, 1)^2 = (0 - 1, 0 + 0) = (-1, 0)$. With the understanding that $(-1, 0) = -1 \in \mathbb{R}$, we can interpret this as $(0, 1) = i \in \mathbb{C}$.

For a real number a and a complex number (b, c), we have $a \cdot (b, c) = (a, 0)(b, c) = (ab, ac)$.

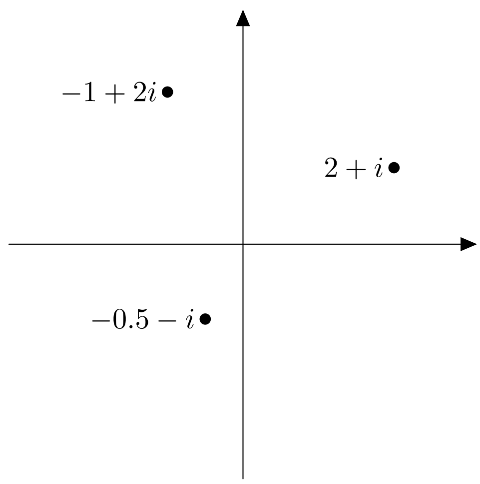


Figure 79: Various complex numbers on the plane

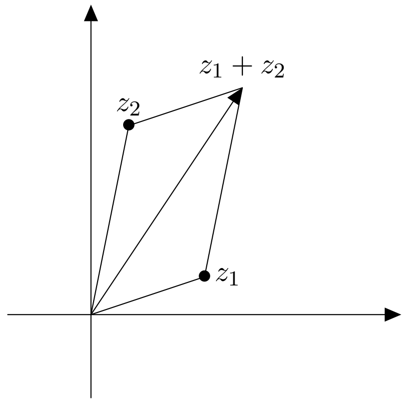


Figure 80: Sum of two complex numbers

With this understanding, any complex number (a, b) can be written as $a + ib$, where $a = (a, 0), b = (b, 0), i = (0, 1)$. We can perform all the usual computations using $i^2 = -1$, for example,

$$(a + ib)(c + id) = ac + iad + ibc + i^2bd = ac - bd + i(ad + bc),$$
$$\frac{1}{a + ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2}.$$

Geometric interpretation

As we represented a real number as a point on the line, we can represent a complex number on the plane:

This is helpful especially when we consider various operations on complex numbers. For example, any complex number (a, b) can be considered as a segment from $(0, 0)$. Then the sum can be found by forming a parallelogram.

As we identify a complex number with a point on the plane, for each complex number (a, b) the length of the segment $(0, 0)$–(a, b) is $\sqrt{a^2 + b^2}$ and the angle from the horizontal axis (the real numbers), and we can write this as $(a, b) = (r \cos\theta, r \sin\theta)$.

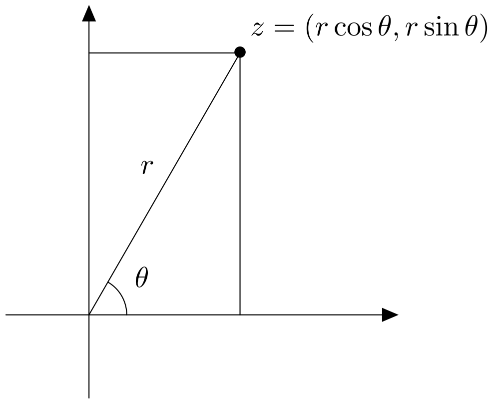


Figure 81: Complex number represented by radius and angle

With two complex numbers $(r_1\cos\theta_1, r_1\sin\theta_1), (r_2\cos\theta_2, r_2\sin\theta_2)$, we have

$$\begin{aligned}&(r_1\cos\theta_1, r_1\sin\theta_1)(r_2\cos\theta_2, r_2\sin\theta_2)\\&=(r_1r_2(\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2), r_1r_2(\cos\theta_1\sin\theta_2+\sin\theta_1\cos\theta_2))\\&=(r_1r_2\cos(\theta_1+\theta_2), r_1r_2\sin(\theta_1+\theta_2))\end{aligned}$$

Therefore, the product has the length r_1r_2 and the angle $\theta_1+\theta_2$. In particular, if we take $z=(r\cos\theta, r\sin\theta)$, then we have $z^n=(r^n\cos n\theta, r^n\sin n\theta)$.

From this, we can conlude that for any $z\in\mathbb{C}$ there is the n-th root in $\mathbb{C}$. Indeed, if $z=(r\cos\theta, r\sin\theta)$, then we can take $w=(r^{\frac{1}{n}}\cos\frac{\theta}{n}, r^{\frac{1}{n}}\sin\frac{\theta}{n})$.

Fundamental theorem of algebra

In complex numbers, we can solve the second degree equation $ax^2+bx+c=0$. Indeed, we can use the usual formula

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a},$$

where the square root is interpreted as a complex number, which always exists.

In general, a polynomial equation $a_nx^n+\cdots a_1x+a_0=0$ has n solution in $\mathbb{C}$. This is called the fundamental theorem of algebra.

Dec. 22. Complex valued functions

Let us consider $x=a+ib\in\mathbb{C}$. We denote by $|x|=\sqrt{a^2+b^2}$ the **absolute value** of x. For two complex numbers x, y, the **disntance** between x, y is $|x-y|$. This distance coincide with the distance on the plane.

We have the triangle inequality

$$|x-y|\le|x-z|+|z-y|.$$

This is literally a triangle inequality, in the sense that $|x-y|, |y-z|, |z-x|$ are the lengths of the sides of the triangle formed by x, y, z in $\mathbb{C}$.

For $x=a+ib$, let $\operatorname{Re}x=a, \operatorname{Im}x=b$. We have $|x|\ge|\operatorname{Re}x|, |\operatorname{Im}x|$, while $|x|\le|\operatorname{Re}x|+|\operatorname{Im}x|$.

Complex sequences

Let us consider a sequence of complex numbers $\{x_n\}$ and $x\in\mathbb{C}$. We say that x_n converges to x (and write $x_n\to x$) if $|x_n-x|\to 0$. Note that $\{|x_n-x|\}$ is a sequence of real numbers, hence we can use the definition and results there.

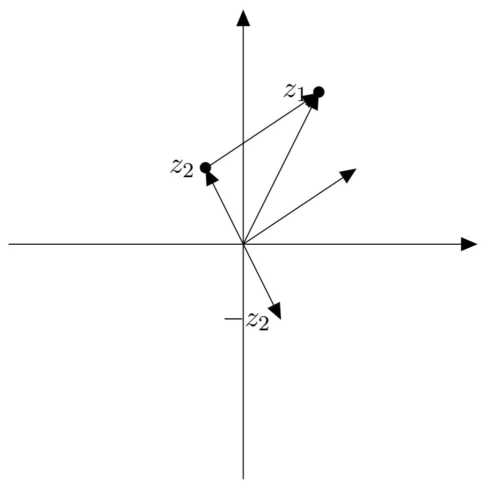


Figure 82: The distance of two complex numbers

Lemma 211. *x_n converges to x if and only if $\operatorname{Re} x_n$ and $\operatorname{Im} x_n$ converge to $\operatorname{Re} x$ and $\operatorname{Im} x$, respectively.*

Proof. If $|x_n - x| \to 0$, then $|\operatorname{Re} x_n - \operatorname{Re} x| = |\operatorname{Re}(x_n - x)| \leq |x_n - x|$ hence $\operatorname{Re} x_n \to \operatorname{Re} x$. Similarly for $\operatorname{Im} x$.

If $|\operatorname{Re} x_n - \operatorname{Re} x|, |\operatorname{Im} x_n - \operatorname{Im} x| \to 0$, then $|x_n - x| \leq |\operatorname{Re} x_n - \operatorname{Re} x| + |\operatorname{Im} x_n - \operatorname{Im} x| \to 0$. □

In particular, we say that $\{x_n\}$ is a Cauchy sequence if for any $\epsilon > 0$ there is N such that $|x_m - x_n| < \epsilon$ if $m, n > N$. In this case, $\{x_n\}$ is convergent.

Complex series

If $\{z_n\}$ are complex numbers, we can also consider the series

$$\sum_{k=0}^{n} z_k = z_0 + z_1 + \cdots + z_n.$$

We can define the convergence of the series as the convergence of the sequence $\sum_{k=0}^{n} z_k$ just as with real numbers.

We say that the series $\sum_{k=0}^{n} z_k$ **converges absolutely** if $\sum_{k=0}^{n} |z_k|$ converges.

Lemma 212. *If $\sum_{k=0}^{n} z_k$ converges absolutely, then the series $\sum_{k=0}^{n} z_k$ converges.*

Proof. If $\sum_{k=0}^{n} z_k$ converges absolutely, then $\sum_{k=0}^{n} |z_k|$ is a Cauchy sequence, and hence $\sum_{k=0}^{n} z_k$ is a Cauchy sequence (in the complex sense as above) because

$$\left|\sum_{k=0}^{n} z_k - \sum_{k=0}^{m} z_k\right| = \left|\sum_{k=m+1}^{n} z_k\right| \leq \sum_{k=m+1}^{n} |z_k| < \epsilon$$

for sufficiently large m, n, by triangle inequality. This implies that $\operatorname{Re} \sum_{k=0}^{n} z_k$ and $\operatorname{Im} \sum_{k=0}^{n} z_k$ are Cauchy, hence converge. Therefore, $\sum_{k=0}^{n} z_k$ converges by Lemma 211. □

Recall that, for real number x, we have proved

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$
$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!},$$
$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.$$

We can extend these functions by replacing x by a complex number z. Indeed, the series

$$\sum_{n=0}^{N} \frac{z^n}{n!}$$

is absolutely convergent for any $z \in \mathbb{C}$, because

$$\sum_{n=0}^{N} \frac{|z|^n}{n!}$$

is convergent by the ratio test: with $a_n = \frac{|z|^n}{n!}$, $\frac{a_{n+1}}{a_n} = \frac{|z|}{n+1} \to 0$.

Therefore, we can *define* the complex exponential function e^z by $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$. Let us see what happens if $z = i\theta, \theta \in \mathbb{R}$.

$$\begin{aligned} e^{i\theta} = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} &= \sum_{n=0}^{\infty} \frac{(i\theta)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(i\theta)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} \\ &= \cos\theta + i \sin\theta. \end{aligned}$$

This last formula is called the Euler formula. In particlar with $\theta = \pi$, we have $e^{i\pi} = -1$, or $e^{i\pi} + 1 = 0$.

We can also extend $\cos z, \sin z$ to complex variables (convergence is proven again by ratio test). As we have

$$\begin{aligned}
e^{\theta} &= \sum_{n=0}^{\infty} \frac{\theta^n}{n!}, \\
e^{-\theta} &= \sum_{n=0}^{\infty} \frac{(-1)^n\theta^n}{n!}, \\
\cos(i\theta) &= \sum_{n=0}^{\infty} \frac{(-1)^n(i\theta)^{2n}}{(2n)!} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n(-1)^n\theta^{2n}}{(2n)!} \\
&= \sum_{n=0}^{\infty} \frac{\theta^{2n}}{(2n)!} = \frac{1}{2}(e^{\theta}+e^{-\theta}) = \cosh\theta, \\
\sin(i\theta) &= \sum_{n=0}^{\infty} \frac{(-1)^n(i\theta)^{2n+1}}{(2n+1)!} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n i(-1)^n\theta^{2n+1}}{(2n+1)!} \\
&= i\sum_{n=0}^{\infty} \frac{\theta^{2n+1}}{(2n+1)!} = \frac{i}{2}(e^{\theta}-e^{-\theta}) = i\sinh\theta,
\end{aligned}$$

Furthermore, this explains why the differential equation $y'' + y = 0$ has a general solution $y(x) = C_1 \sin x + C_2 \cos x$. By formally applying (this can be justified by the material in Mathematical Analysis II) the chain rule, we have $D(e^{ix}) = ie^{ix}, D^2(e^{ix}) = -e^{ix}$ and $D(e^{-ix}) = -ie^{-ix}, D^2(e^{-ix}) = -e^{-ix}$, hence they are formally two solutions of the equation $y'' + y$. Hence their linear combinations $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$ and $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$ satisfy the same equation.

Jan. 11. Ordinary differential equations

Integral curves

As in previous example, it is typical that, by solving a differential equation, we obtain an implicit equation $F(x, y, C) = 0$. This means that, for each value of C, we have a relation between x, y, and in certain cases, it defines a function y of x. As this function $y(x)$ satisfies the differential equation $y'(x) = f(x, y(x))$, $y'(x)$ should mean the slope of the curve $y(x)$ at the point $(x, y(x))$.

For example, consider the equation $y' = x$. This can be integrated and $y = \frac{x^2}{2} + C$, and depending on the value of C, we have different parabolas. On the other hand, at each point in the xy-plane, we can draw an arrow which goes from (x, y) to $(x + \epsilon, y + y'(x)\epsilon)$. These arrows are tangent to the curve which represents the solution.

This plot of arrows is called a vector field, and a solution is obtained by "connecting" these arrows.

(One can visualize the arrows by a command `VectorPlot[1,f(x,y)]` on Wolfram Alpha, and the stream by `StreamPlot[1,f(x,y)]`, where we took $\epsilon = 1$).

Existence and uniqueness of solution

We have considered ordinary differential equations $y' = f(x, y)$ and found solutions to some of them. Yet, some differential equation does not have a solution for a given initial condition, and

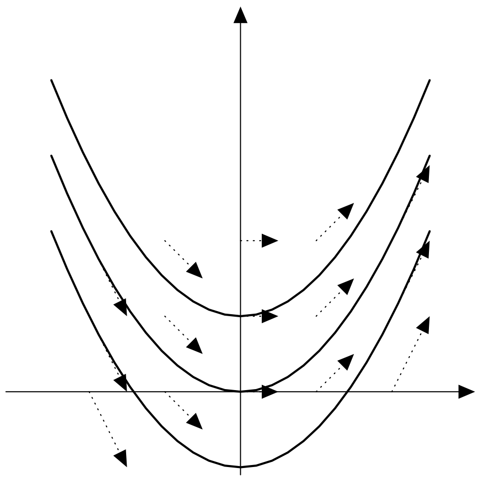


Figure 83: The integral curves of $y' = x$.

others have many solutions.

- Consider $(y')^2 - xy' + y + 1 = 0$: no solution with $y(0) = 0$, because then $y'(0)^2 + 1 = 0$, which is impossible because $y'(0)$ should be a real number.

- Consider $y' = 3y^{\frac{2}{3}}$: the initial condition $y(0) = 0$ has two solutions $y = 0$ and $y = x^3$.

Yet, as we have seen, a differential equation gives a vector field as in Figure 83, and it should be enough to "chase the arrows". For this to be possible, $f(x, t)$ should have certain nice properties. We only state the theorem, and leave the proof to a more advanced book.

For this purpose, we need the following concept: Let $f(x, y)$ be a function of two variables, that is, f gives a number for a given pair of numbers (x, y). For each fixed y, we can think of $f(x, y)$ as a function of x, and hence take the derivative with respect to x. This is called the **partial derivative** with respect to x, and denoted by $\frac{\partial f}{\partial x}$.

Example 213.
- Let $f(x, y) = x^2 + y^2$. Then $\frac{\partial f}{\partial x} = 2x$.
- Let $f(x, y) = xy$. Then $\frac{\partial f}{\partial x} = y$.
- Let $f(x, y) = \sin(xy^2)$. Then $\frac{\partial f}{\partial x} = y^2 \cos(xy^2)$.

It is also possible to consider $\frac{\partial f}{\partial y}$. The detail will be explained in Mathematical Analysis II.

Theorem 214. *Suppose that $f(x, y)$ and $\frac{\partial f}{\partial x}$ are continuous in a rectangle*

$$R = \{(x, y) : x_0 - \delta < x < x_0 + \delta, y_0 - \epsilon < y < y_0 + \epsilon\}.$$

Then there is δ_1 such that the equation $y' = f(x, y)$ has a unique solution $y(x)$ with initial condition $y(x_0) = y_0$ for $x_0 - \delta_1 < x < x_0 + \delta_1$.

Euler method

The proof of this theorem amounts to construct approximate solutions. At the end, for applications in science and engineering, we are satisfied with having sufficiently good approximate solutions.

There are many methods to obtain a numerical solution of a differential equation. One of the simplest of them is called the Euler's method, and it literally chase the vector field as follows.

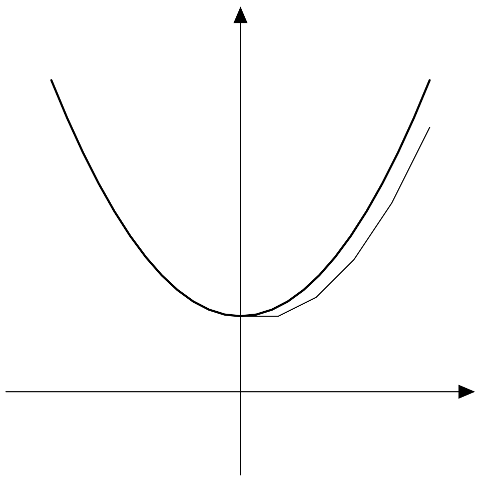


Figure 84: The Euler method to solve $y' = x$ with $(x_0, y_0) = (0, 1)$ with $\epsilon = 0.5$.

Let us consider the differential equation $y' = f(x, y)$ with the inizial condition $y(x_0) = y_0$, where $x_0, y_0 \in \mathbb{R}$. This means that the solution $y(x)$ passes the point (x_0, y_0). Furthermore, by "chasing the arrows", the slope of the curve $y(x)$ at the point (x_0, y_0) is $f(x_0, y_0)$. That is, if we take a small step ϵ, then the next point on the curve is close to $(x_0 + \epsilon, y_0 + f(x_0, y_0)\epsilon) = (x_1, y_1)$. Then, again at the point (x_1, y_1), the slope of the curve is $f(x_1, y_1)$, hence the next point on the curve is close to $(x_1 + \epsilon, y_1 + f(x_1, y_1)\epsilon) = (x_2, y_2)$, and so on. In this way, we obtain a union of segments which approximates the solution.

If we take smaller ϵ, the approximation gets better, while we need do more computations.

A few codes in Python

In the language Python, it is very easy to write a code to solve a differential equation. Let us see some examples[3]

The following code solves the equation $y' = x, x_0 = 0, y_0 = 1$.

```
import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt

x = np.linspace(0, 5, 500)

# parameter k
k = 0.5

# equation y' = k y

def deriv(y, x):
    return k*y

# the initial condition
y0 = 1

# solve the equation numerically
```

[3]The plot part is taken from the book by Christian Hill, https://scipython.com/book/chapter-8-scipy/additional-examples/the-sir-epidemic-model/.

```
ret = odeint(deriv, y0, x)
y = ret


# plot the graph
plt.plot(x,y, label='Solution of y\'(x)='+str(k)+'y with y(0)='+str(y0))
legend = plt.legend()
plt.show()
```

To solve a second-order differential equation $y''+b=0$, use the trick of doubling the variables: $y_0' = y_1$ and $y_1' + b = 0$, which means $y_0'' + b = 0$.

```
import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt

x = np.linspace(0, 5, 500)

# parameter k
k = -10

# equation y'' = ky, that is, y_0' = y_1 and y_1' = - k y_0

def deriv(y, x):
    return [y[1], k*y[0]]

# the initial condition
y0 = [1,0.1]

# solve the equation numerically
ret = odeint(deriv, y0, x)
y_0,y_1 = ret.T

# plot the graph
plt.plot(x,y_0, label='Solution of y\'\'(x)='+str(k)+'y with y(0)='+str(y0[0])
legend = plt.legend()
plt.show()
```

It is also possible to solve a differential equation with more variables, for example the SIR model

$$\frac{dS}{dt} = -\frac{\beta S(t)I(t)}{N},$$
$$\frac{dI}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t),$$
$$\frac{dR}{dt} = \gamma I(t),$$

see for example this page.