Paweł Góra

Notes: Analysis I

Concordia University, Montreal

Updated: April 11, 2026

These are notes for one semester course in Analysis I. They are obviously based on a number of previous books, usually not referenced.

Contents

1. Short Introduction to The Absolute Basics of Math Logic

Logic is useful for everybody. For mathematicians it is the basic tool used in their every professional activity. What is presented here is the very minimum of mathematical logic every math student should know and enjoy using.

1.1. **Elementary sentences and calculus of sentences.** A "sentence" is a primary notion, i.e., something which is not defined but everybody knows what it is. It is a notion like a point in geometry or a horse in an old Polish encyclopedia: "Horse, what it is, everybody knows". For a sentence to be a mathematical sentence we need to be able to say (at least in principle) whether it is true or false. Thus, "The sky is blue" is a mathematical sentence, while "Let's go to the movies" is not. It does not matter if the sentence is false or true, as long as its "logical value" can be assigned it is a mathematical sentence.

Other examples of mathematical sentences:

There is 23 students in this class.

It will be raining tomorrow.

I have 231456 hairs on my head.

Some English sentences are not mathematical sentences:

Is this blackboard white?

Why is it raining right now?

This sentence is a lie.

The last example "This sentence is a lie" is particularly interesting. It seems that it says something so it should be either true or false. If we assume that it is true, then it is a lie, i.e., false and we obtain a contradiction. If we assume that it is false, then it is not a lie, i.e., it is true and again we obtain a contradiction. Thus, it is not true neither false, and for this reason it is not a mathematical sentence. Sentences like this one are called self referencing sentences and often cause problems.

The sentences considered above were "simple" sentences. Using them and logical operators we can build more complex sentences. Most popular logical operators are:

"and", "or", "not" and "implies". There are many others, and everybody can define her/his own new operators. Traditionally, we denote them by

and $\wedge$;

or $\vee$;

not $\neg$;

implies $\Rightarrow$.

Note, that they are similar to the set operation symbols: intersection $\cap$, union $\cup$, set subtraction $\setminus$. The symbol for implication should be similar for the inclusion symbol $\subset$ but it is not.

We define logical operators by giving their value tables, i.e., for all possible logical values of sentences α and β we specify the value of the sentence αoperatorβ. We will denote value of "true" by 1 and value of "false" by 0. (Some computer languages use opposite notation.) Here are the tables:

α	β	$\alpha \wedge \beta$	$\alpha \vee \beta$	$\neg\alpha$	$\alpha \Rightarrow \beta$
0	0	0	0	1	1
0	1	0	1	1	1
1	0	0	1	0	0
1	1	1	1	0	1

For example, for $\alpha = 1$ and $\beta = 1$ we have $(\alpha \vee \beta) = 1$. This is slightly different from everyday use of "or". Also, for $\alpha = 0$ and $\beta = 1$ we have $(\alpha \Rightarrow \beta) = 1$. This may seem incorrect but we will understand it better later.

A sentence which is always true is called a theorem (or a tautology). The simplest way to check if a given sentence is a theorem is just to check its logical value for all possible combinations of arguments values. Let us consider a sentence

$$\neg(\alpha \wedge \beta) \Longleftrightarrow (\neg\alpha \vee \neg\beta)\,. \tag{1}$$

$\gamma \iff \delta$ means that γ and δ have the same logical value, i.e., are equivalent. It is the same as $(\gamma \Rightarrow \delta) \wedge (\delta \Rightarrow \gamma)$ (You can prove this). We have

α	β	$\alpha \wedge \beta$	$\neg(\alpha \wedge \beta)$	**LHS**	$\neg\alpha$	$\neg\beta$	$\neg\alpha \vee \neg\beta$	**RHS**
0	0	0	1	**1**	1	1	1	**1**
0	1	0	1	**1**	1	0	1	**1**
1	0	0	1	**1**	0	1	1	**1**
1	1	1	0	**0**	0	0	0	**0**

Thus, the sentences on both sides of (1) are equivalent and sentence (1) is a theorem. With its sibling

(2) $$\neg(\alpha \vee \beta) \iff (\neg\alpha \wedge \neg\beta)\,,$$

they are called De Morgan's laws, named after Augustus De Morgan (1806-71), British mathematician.

Let us consider a sentence $\neg(\alpha \implies \beta)$. By common sense it should be true only when α is true and β is false (Do you agree?). Let us check if the following sentence is a theorem:

(3) $$\neg(\alpha \implies \beta) \iff (\alpha \wedge \neg\beta)\,.$$

α	β	$\alpha \implies \beta$	$\neg(\alpha \implies \beta)$	**LHS**	$\neg\beta$	$\alpha \wedge \neg\beta$	**RHS**
0	0	1	0	**0**	1	0	**0**
0	1	1	0	**0**	0	0	**0**
1	0	0	1	**1**	1	1	**1**
1	1	1	0	**0**	0	0	**0**

Thus, the sentences on both sides of (3) are equivalent and sentence (3) is a theorem. In particular, we can see that for $\alpha = 0$ and $\beta = 1$ we have $(\alpha \Rightarrow \beta) = 1$, which in a way justifies the definition of implication. (The definitions do not need justification, but it is nicer to have them reasonable.)

1.2. **Proofs by contraposition and by contradiction:** The following sentences are theorems, i.e., are true for all choices of sentences α, β, γ. You can prove this using the truth tables.

This one is the pattern of the proof by contraposition:

$$(a) \qquad (\alpha \Rightarrow \beta) \Longleftrightarrow (\neg\beta \Rightarrow \neg\alpha),$$

This one is the pattern of the proof by contradiction:

$$(b) \qquad [(\alpha \wedge \neg\beta) \Rightarrow (\gamma \wedge \neg\gamma)] \Rightarrow (\alpha \Rightarrow \beta).$$

Now we will use these patterns to prove the simple theorem: If n is divisible by 6, then n is divisible by 2.

(a) First by contraposition: the theorem is written as an implication: If n is divisible by 6, then n is divisible by 2. We can denote "n is divisible by 6" by α and "n is divisible by 2" by β. Then the theorem is $\alpha \Rightarrow \beta$. Using pattern (a) we see that instead we can prove $\neg\beta \Rightarrow \neg\alpha$. This means we need to prove "If n is not divisible by 2, then n is not divisible by 6".

If n is not divisible by 2, then $n = 2k + 1$. Now, k is of one of the three forms: $k = 3s$, $k = 3s + 1$ or $k = 3s + 2$ (depending on the remainder when k is divided by 3). Then, we have $n = 2 \cdot 3s + 1$, $n = 2 \cdot 3s + 2 + 1$, or $n = 2 \cdot 3s + 4 + 1$. In other words $n = 6s + 1$, $n = 6s + 3$, or $n = 6s + 5$. Thus, n is not divisible by 6. We proved $\neg\beta \Rightarrow \neg\alpha$. Using pattern (a) this implies $\alpha \Rightarrow \beta$. We proved the theorem by contraposition.

(b) Now, by contradiction: we again see that the theorem is written as an implication and denote "n is divisible by 6" by α and "n is divisible by 2" by β. Then the theorem is $\alpha \Rightarrow \beta$. Using pattern (b) we write the negation of the theorem $\alpha \wedge \neg\beta$, i.e., "n is divisible by 6 and n is not divisible by 2".

Now, we will get a contradiction: since n is divisible by 6, we have $n = 6k = 2(3k)$ and we see that n is divisible by 2. At the same time our assumption says " and n is not divisible by 2". If we denote by γ the sentence "n is divisible by 2", we proved $\gamma \wedge \neg\gamma$. According to pattern (b) this proves the theorem.

I understand that this was an unnecessary complication of the trivial proof of a trivial theorem but we did this to present an example of how the patterns (a) and

(b) are used. We will use them many times in the future for much more challenging proofs.

1.3. **Sentences and set operations.** Again, we do not define the notion of a set. It has to be understood. Let X denote out space, i.e., the set which contains all sets we are going to consider. Considering sets "in general", i.e., without deciding on some space containing all of them, leads to contradictions. The basic operations on sets are defined using sentences:

$A \cap B = \{x \in X : (x \in A) \wedge (x \in B)\}$, intersection of A and B;

$A \cup B = \{x \in X : (x \in A) \vee (x \in B)\}$, union of A and B;

$A^c = X \setminus A = \{x \in X : \neg(x \in A)\}$, complement of A.

We say that $A \subset B$ (A is contained in B, or A is a subset of B) if and only if $(x \in A) \implies (x \in B)$.

We can use this correspondence to prove theorems about sets. For example De Morgan's laws for set operations are

$(A \cap B)^c = A^c \cup B^c$, corresponding to (1) and

$(A \cup B)^c = A^c \cap B^c$, corresponding to (2).

Other set identities also can be "translated" into sentences equivalences and proved this way. For example:

$(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$ corresponds to

$(\alpha \vee \beta) \wedge \neg\gamma \iff (\alpha \wedge \neg\gamma) \vee (\beta \wedge \neg\gamma)$ and can be proved using a table of logical values.

1.4. **Quantifiers.** Let $P(x)$ be a sentence depending on variable x (say, $x^2 \geq 0$), which is true for all x in some domain X. Instead of saying: $P(x_1)$ and $P(x_2)$ and $P(x_3)$ and ... we just say: for all $x \in X$ $P(x)$, or symbolically

$$\forall_{\, x \in X} \; P(x) \,.$$

Quantifier $\forall$ is called a general quantifier. For example, sentence $\forall_{\, x \in \mathbb{R}} \; x^2 \geq 0$ is true and the sentence $\forall_{\, x \in \mathbb{R}} \; x \geq 3$ is false.

Similarly, the multiple (possibly infinite) or are expressed through existential quantifier "there exists", for example

$$\exists_{\, x\in\mathbb{R}}\; x > 3\,,$$

is a true sentence.

In some books "for all" is denoted by $\bigwedge$ and "there exists" by $\bigvee$ to highlight the fact that they are generalizations of "and" and "or", respectively.

The negations of sentences with quantifiers are obtained using de Morgan's Laws:

$$\neg\forall_{\, x\in X}\; P(x) \Longleftrightarrow \exists_{\, x\in X}\; \neg P(x)\,,$$

$$\neg\exists_{\, x\in X}\; P(x) \Longleftrightarrow \forall_{\, x\in X}\; \neg P(x)\,,$$

which seems to actually be in agreement with our common sense.

Example: below we write the definition of a continuous function $f:\mathbb{R}\to\mathbb{R}$:

$$\forall_{\, x\in\mathbb{R}}\;\forall_{\, y\in\mathbb{R}}\;\forall_{\, \varepsilon>0}\;\exists_{\, \delta>0}\;|y-x|<\varepsilon \Rightarrow |f(y)-f(x)|<\varepsilon.$$

To obtain a definition of the function which is not continuous we will negate the above sentence and simplify it step by step using de Morgan's laws:

$$\neg\left[\forall_{\, x\in\mathbb{R}}\;\forall_{\, y\in\mathbb{R}}\;\forall_{\, \varepsilon>0}\;\exists_{\, \delta>0}\;|y-x|<\varepsilon \Rightarrow |f(y)-f(x)|<\varepsilon\right] \Longleftrightarrow$$

$$\exists_{\, x\in\mathbb{R}}\;\neg\left[\forall_{\, y\in\mathbb{R}}\;\forall_{\, \varepsilon>0}\;\exists_{\, \delta>0}\;|y-x|<\varepsilon \Rightarrow |f(y)-f(x)|<\varepsilon\right] \Longleftrightarrow$$

$$\exists_{\, x\in\mathbb{R}}\;\exists_{\, y\in\mathbb{R}}\;\neg\left[\forall_{\, \varepsilon>0}\;\exists_{\, \delta>0}\;|y-x|<\varepsilon \Rightarrow |f(y)-f(x)|<\varepsilon\right] \Longleftrightarrow$$

$$\exists_{\, x\in\mathbb{R}}\;\exists_{\, y\in\mathbb{R}}\;\exists_{\, \varepsilon>0}\;\neg\left[\exists_{\, \delta>0}\;|y-x|<\varepsilon \Rightarrow |f(y)-f(x)|<\varepsilon\right] \Longleftrightarrow$$

$$\exists_{\, x\in\mathbb{R}}\;\exists_{\, y\in\mathbb{R}}\;\exists_{\, \varepsilon>0}\;\forall_{\, \delta>0}\;\neg\left[|y-x|<\varepsilon \Rightarrow |f(y)-f(x)|<\varepsilon\right] \Longleftrightarrow$$

$$\exists_{\, x\in\mathbb{R}}\;\exists_{\, y\in\mathbb{R}}\;\exists_{\, \varepsilon>0}\;\forall_{\, \delta>0}\;|y-x|<\varepsilon \wedge |f(y)-f(x)|\geq\varepsilon\,.$$

We used the fact that $\neg(\alpha\Rightarrow\beta) \Longleftrightarrow\ \alpha\wedge\neg\,\beta$.

Quantifiers correspond to generalized (possibly infinite) operations on sets: if $\{A_s\}_{s\in S}$ are subsets of some space X, then we define their intersection and union as follows:

$$\bigcap_{s\in S} A_s = \{x\in X : \forall_{\, s\in S}\; x\in A_s\}\,,$$

$$\bigcup_{s\in S} A_s = \{x \in X : \exists_{\ s\in S}\ x \in A_s\} .$$

The Morgan's laws "translate" for set operations as

$$\left(\bigcap_{s\in S} A_s\right)^c = \bigcup_{s\in S}(A_s)^c ,$$

$$\left(\bigcup_{s\in S} A_s\right)^c = \bigcap_{s\in S}(A_s)^c .$$

2. Representation of numbers in different bases

We usually use number written in decimal notation, which means that

$$2345678 = 2\cdot 10^6 + 3\cdot 10^5 + 4\cdot 10^4 + 5\cdot 10^3 + 6\cdot 10^2 + 7\cdot 10^1 + 8\cdot 10^0 ,$$

and

$$0.2345678 = 2\cdot 10^{-1} + 3\cdot 10^{-2} + 4\cdot 10^{-3} + 5\cdot 10^{-4} + 6\cdot 10^{-5} + 7\cdot 10^{-6} + 8\cdot 10^{-7} .$$

This notation allows to conveniently express arbitrarily large and arbitrarily small numbers using 10 digits, 0,1,2,3,4,5,6,7,8 and 9. We say that the base of this system is 10. Popularity of number ten comes probably from the fact that an average person has ten fingers naturally used to count objects. Other civilizations used (or may be using) different representations of numbers and/or different bases. You can view examples following the link
(http://mypage.concordia.ca/mathstat/pgora/Number representations.html).

More generally, let us assume that our base is an integer $n > 1$. We use digits $0, 1, 2, \ldots, n-1$. We have, for $0 \le a_k, a_{k-1}, a_{k-2}, \ldots, a_4, a_3, a_2, a_1, a_0 \le n-1$,

$$a_k a_{k-1} a_{k-2} \ldots a_4 a_3 a_2 a_1 a_0 = a_k n^k + a_{k-1} n^{k-1} + a_{k-2} n^{k-2} + \cdots + a_4 n^4 + a_3 n^3 + a_2 n^2 + a_1 n^1 + a_0 n^0 ,$$

and

$$0.a_1 a_2 a_3 \ldots a_{k-2} a_{k-1} a_k = a_1 n^{-1} + a_2 n^{-2} + a_3 n^{-3} + \cdots + a_{k-2} n^{-(k-2)} + a_{k-1} n^{-(k-1)} + a_k n^{-k} .$$

Examples: Base is denoted by the number in brackets at the bottom.

$$457_{(8)} = 4\cdot 8^2 + 5\cdot 8 + 7 = 111_{(10)}\ ,\ 10101_{(2)} = 1\cdot 2^4 + 1\cdot 2^2 + 1 = 21_{(10)} .$$

$$0.121_{(3)} = 1/3 + 2/9 + 1/27 = 16/27\ ,\ 0.1001_{(2)} = 1/2 + 1/16 = 7/16\,,$$
$$0.AF_{(16)} = 10/16 + 15/256 = 175/256\,.$$

The last example is in hexadecimal system (base 16) using A=10, B=11, C=12, D=13, E=14 and F=15 as additional digits. This system finds a variety of computer related uses (check at Wikipedia). As everybody knows (?) computers translate everything into binary (base 2) representations to operate on them.

The formulas above show how to change numbers represented in other bases into decimal numbers. How to perform an inverse operation, i.e., represent a decimal number in say base 3? First we have to think how we really make decimal representation of numbers. (We are so used to it that we never think about this process.) We do it performing consecutive division by 10 and writing down the remainders:

$$\begin{aligned} 736 : 10 &= 73 \text{ remainder } 6 \\ 73 : 10 &= 7 \text{ remainder } \ 3 \\ 7 : 10 &= 0 \text{ remainder } \ 7\,. \end{aligned}$$

Similarly:

$$\begin{aligned} 736 : 3 &= 245 \text{ remainder } 1 \\ 245 : 3 &= 81 \text{ remainder } \ 2 \\ 81 : 3 &= 27 \text{ remainder } \ 0 \\ 27 : 3 &= 9 \text{ remainder } \ \ 0 \\ 9 : 3 &= 3 \text{ remainder } \ \ 0 \\ 3 : 3 &= 1 \text{ remainder } \ \ 0 \\ 1 : 3 &= 0 \text{ remainder } \ \ 1\ , \end{aligned}$$

and $736_{(10)} = 1000021_{(3)}$. To check this we write $1000021_{(3)} = 1 \cdot 3^6 + 2 \cdot 3 + 1 = 736_{(10)}$. We also have

$$\begin{aligned} 736 : 16 &= 46 \text{ remainder } 0 \\ 46 : 16 &= 2 \text{ remainder } \ 14 \\ 2 : 16 &= 0 \text{ remainder } \ 2\,, \end{aligned}$$

and $736_{(10)} = 2E0_{(3)}$. To check this we write $2E0_{(16)} = 2 \cdot 16^2 + 14 \cdot 16 = 736_{(10)}$.

Now, we will try to represent a usual fraction $\frac{3}{7}$ as a "decimal" fraction in different bases. First, in base 10:

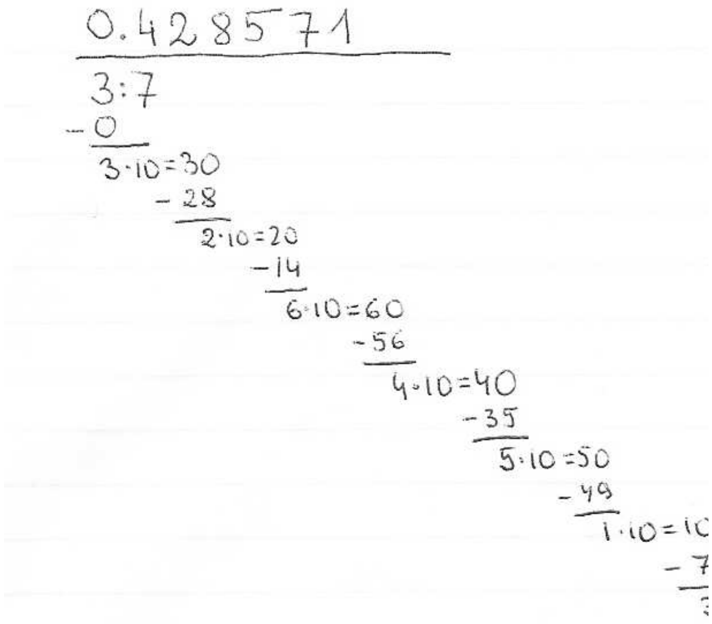

FIGURE 1. Long division 3:7 in base 10

The important fact we used is that each unit makes 10 units of the lower order. To check correctness of our result we calculate

$$0.428571428571\cdots = \frac{428571}{10^6}\left(\frac{1}{1-1/10^6}\right) = \frac{3}{7}\,.$$

We used the formula for the sum of geometric series

$$1 + q + q^2 + q^3 + \cdots = \frac{1}{1-\frac{1}{q}} \text{ for } |q| < 1\,.$$

Similarly, we do this for base 2 and base 16, remembering that for these cases the units make 2 and 16 units of lower order, correspondingly.

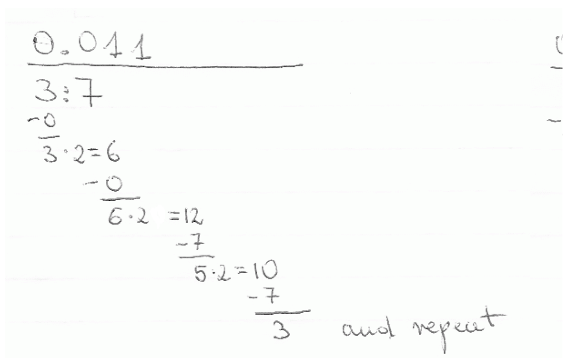

Figure 2. Division 3:7 in base 2 and base 16

To check correctness we calculate:

$$0.011011011\ldots_{(2)} = \left(\frac{1}{4}+\frac{1}{8}\right)\left(\frac{1}{1-\frac{1}{8}}\right) = \frac{3}{7}\,,$$

and

$$0.6(13)(11)6(13)(11)6(13)(11)\ldots_{(16)} = \left(\frac{6}{16}+\frac{13}{16^2}+\frac{11}{16^2}\right)\left(\frac{1}{1-\frac{1}{16^3}}\right) = \frac{3}{7}\,.$$

Of course, we should have used D instead of (13) and B instead of (11).

An easy corollary of the presented algorithm is the theorem:

Theorem 1. *A rational number (a usual fraction) has eventually periodic digital representation in any base. An eventually periodic digital fraction represents a rational number.*

Proof of the first claim uses the fact that in base n there is only n possible remainders so the digital representation has to repeat after at most n steps. The second follows from the calculation.

The last question in this section is changing a decimal fraction into digital fraction in other base. First we will do a short geometric introduction into representation of fractions in general. Let us consider base $n = 4$. We will use 4 digits $0, 1, 2, 3$. Let us divide interval $[0, 1)$ into 4 equal subintervals:

$$[0,1) = \left[0, \frac{1}{4}\right) \cup \left[\frac{1}{4}, \frac{2}{4}\right) \cup \left[\frac{2}{4}, \frac{3}{4}\right) \cup \left[\frac{3}{4}, 1\right) ,$$

see Figure 3 a). The points in the first interval have first digit 0 (since they are between 0 and 1/4), the points in the second interval have first digit 1 (since they are between 1/4 and 2/4), the points in the third interval have first digit 2 (since they are between 2/4 and 3/4) and the points in the fourth interval have first digit 3 (since they are between 3/4 and 4/4). To assign the second digit we perform the same operation on each of the intervals of the first partition.

In Figure 3 b) we zoomed up the first interval corresponding to the the first digit 0. We divided it into 4 equal subintervals and the digits of the points in these subintervals are:

00 (since they are between 0 and 1/16),

01 (since they are between 1/16 and 2/16),

02 (since they are between 2/16 and 3/16),

03 (since they are between 3/16 and 4/16).

In Figure 3 c) we zoomed up the third interval of the second generation corresponding to the the first digits 02. We divided it into 4 equal subintervals and the digits of the points in these subintervals are:

020 (since they are between $0/4 + 2/16 + 0/16^2$ and $0/4 + 2/16 + 1/16^2$),

021 (since they are between $0/4 + 2/16 + 1/16^2$ and $0/4 + 2/16 + 2/16^2$),

022 (since they are between $0/4 + 2/16 + 2/16^2$ and $0/4 + 2/16 + 3/16^2$),

023 (since they are between $0/4 + 2/16 + 3/16^2$ and $0/4 + 2/16 + 4/16^2$).

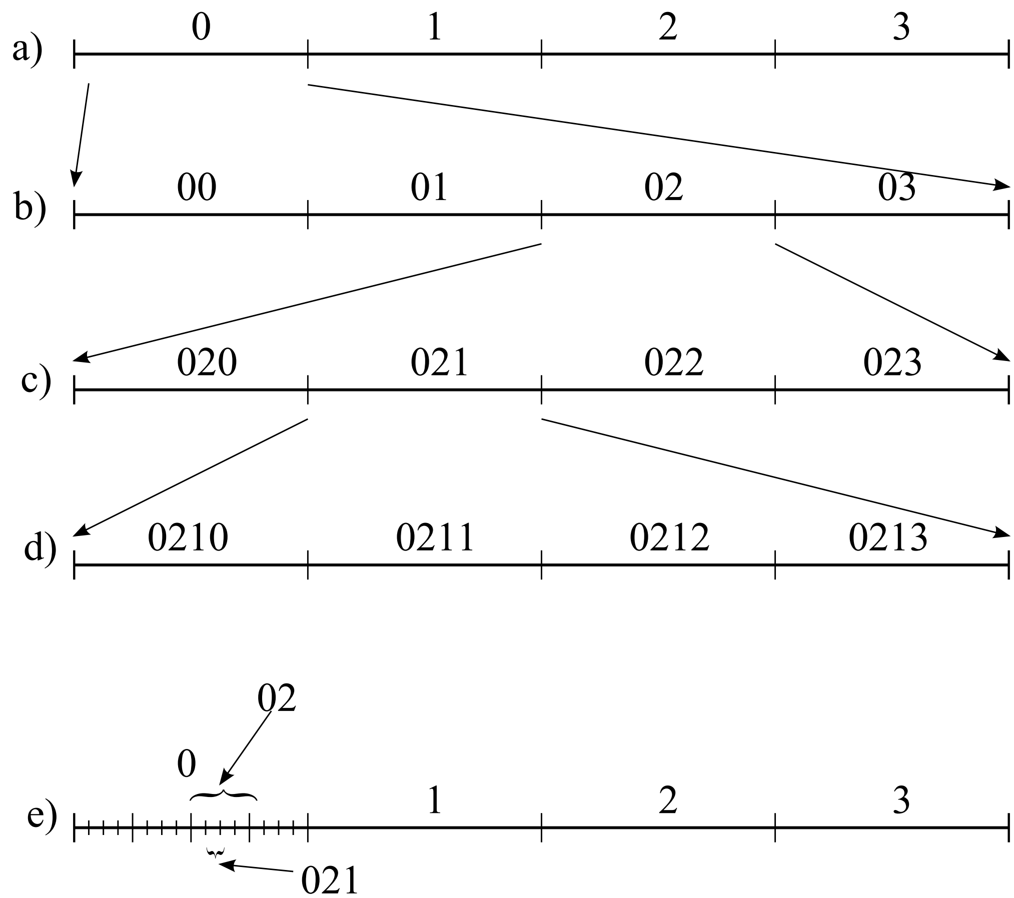

FIGURE 3.

In Figure 3 d) we zoomed up the second interval of the the third generation corresponding to the the first digits 021. We divided it into 4 equal subintervals and the digits of the points in these subintervals are:

0210 (since they are between $0/4 + 2/16 + 1/16^2 + 0/16^3$ and $0/4 + 2/16 + 1/16^2 + 1/16^3$),

0211 (since they are between $0/4 + 2/16 + 1/16^2 + 1/16^3$ and $0/4 + 2/16 + 1/16^2 + 2/16^3$),

0212 (since they are between $0/4 + 2/16 + 1/16^2 + 2/16^3$ and $0/4 + 2/16 + 1/16^2 + 3/16^3$), 0213 (since they are between $0/4+2/16+1/16^2+3/16^3$ and $0/4+2/16+1/16^2+4/16^3$).

These operations are repeated infinitely and every point receives it expansion in base 4.

In Figure 3 e) we showed intervals corresponding to digits 02 and 021 without zooming.

Now we will represent a decimal fraction $x = 0.123457$ as a fraction in base 4. We want to represent x in the form $\frac{d_1}{4}+\frac{d_2}{4^2}+\frac{d_3}{4^3}+\frac{d_4}{4^4}+\frac{d_5}{4^5}+\ldots$. How to get d_1: it is integer part of $4\cdot x$ as $4\cdot x = d_1+\frac{d_2}{4^1}+\frac{d_3}{4^2}+\frac{d_4}{4^3}+\frac{d_5}{4^4}+\ldots$. We have $4\cdot 0.123457 = 0.493828$ so $d_1 = 0$. This shows that x is in the first ("0") interval of the first generation in geometric construction above.

To find d_2 we remove d_1, i.e., consider $x_1 = 4x - E(4x) = \frac{d_2}{4^1}+\frac{d_3}{4^2}+\frac{d_4}{4^3}+\frac{d_5}{4^4}+\ldots$, where $E(t)$ denotes the integer part of t. Now, $d_2 = E(4x_1)$. We have $4\cdot 0.493828 = 1.975312$ so $d_2 = 1$. This locates x more precisely in $[0,1)$. x is in the interval corresponding to 01 in the partition of second generation. And so on:

$x_2 = 4x_1 - E(4x_1) = 0.975312$ and $4x_2 = 3.901248$ so $d_3 = 3$ and

$x_3 = 4x_2 - E(4x_2) = 0.901248$ and $4x_3 = 3.604992$ so $d_4 = 3$ and

$x_4 = 4x_3 - E(4x_3) = 0.604992$ and $4x_4 = 2.419968$ so $d_5 = 2$ and so on. We have

$$0.123457_{(10)} = 0.01332\ldots_{(4)}\ .$$

To check we write

$$0.01332\ldots_{(4)} = 0/4+1/4^2+3/4^3+3/4^4+2/4^5 = 0.1230468750\,,$$

which agrees with what we expected on the first three places. Precision in base 4 in much worse than in base 10, i.e., one needs much more digits to represent a number to the same precision.

3. Cantor set.

The Cantor set is a very popular source of examples and counterexamples in Analysis.

Construction:

We start with the closed interval

$$A_0 = [0, 1].$$

On the step one we divide A_0 into three subintervals of the equal lengths $A_0 = B_{0,0} \cup B_{0,1} \cup B_{0,2}$, where $B_{0,0}$ and $B_{0,2}$ are closed and $B_{0,1}$ is open, and remove the middle open interval $B_{0,1}$. We have

$$A_1 = B_{0,0} \cup B_{0,2}.$$

Note that the length of the removed interval is $m(B_{0,1}) = \frac{1}{3}$. Removing $B_{0,1}$ from $[0, 1]$ we removed all numbers x which have the first digit $d_1 = 1$ in the ternary expansion

$$x = \sum_{k=1}^{\infty} \frac{d_k}{3^k}, \ d_k \in \{0, 1, 2\}.$$

FIGURE 4. Sets A_0, A_1, A_2 and A_3 in the construction of the Cantor set

Step two: We divide each of the remaining intervals $B_{0,0}$, $B_{0,2}$ into three subintervals of equal lengths, side intervals closed and the middle one open,

$$B_{0,0} = B_{0,0,0} \cup B_{0,0,1} \cup B_{0,0,2},$$

$$B_{0,2} = B_{0,2,0} \cup B_{0,2,1} \cup B_{0,2,2},$$

and remove from each the middle subinterval. We have

$$A_2 = B_{0,0,0} \cup B_{0,0,2} \cup B_{0,2,0} \cup B_{0,2,2}.$$

On step two we removed 2 intervals each of the length $\frac{1}{9} = \frac{1}{3^2}$. Removing $B_{0,0,1}$ and $B_{0,2,1}$ from $[0,1]$ we removed all remaining numbers x which have the second digit $d_2 = 1$ in the ternary expansion of x.

Let us assume that after step n we have

$$A_n = \bigcup B_{0,d_1,d_2,\ldots,d_n},\ d_1, d_2, \ldots, d_n \in \{0,2\}.$$

A_n is the union of 2^n disjoint closed intervals of length $\frac{1}{3^n}$ each. Each interval $B_{0,d_1,d_2,\ldots,d_n}$ consists of numbers x with the the beginning of the ternary expansion is

$$x = \frac{d_1}{3} + \frac{d_2}{3^2} + \frac{d_3}{3^3} + \cdots + \frac{d_n}{3^n} + \ldots.$$

Step $n+1$: We divide each of the remaining intervals $B_{0,d_1,d_2,\ldots,d_n}$ into three subintervals of equal lengths, side intervals closed and the middle one open,

$$B_{0,d_1,d_2,\ldots,d_n} = B_{0,d_1,d_2,\ldots,d_n,0} \cup B_{0,d_1,d_2,\ldots,d_n,1} \cup B_{0,d_1,d_2,\ldots,d_n,2},\ d_1, d_2, \ldots, d_n \in \{0,2\},$$

and remove all middle subintervals. We have

$$A_{n+1} = \bigcup B_{0,d_1,d_2,\ldots,d_n,d_{n+1}},\ d_1, d_2, \ldots, d_n, d_{n+1} \in \{0,2\},$$

the union of 2^{n+1} disjoint closed intervals of the length $\frac{1}{3^{n+1}}$ each. On the $n+1$ step we removed 2^n disjoint intervals $B_{0,d_1,d_2,\ldots,d_n,1}$, $d_1, d_2, \ldots, d_n \in \{0,2\}$, of the length $\frac{1}{3^{n+1}}$ each. Removing these intervals we removed from $[0,1]$ remaining numbers x with the $n+1$ digit $d_{n+1} = 1$ in the ternary expansion.

This describes the inductive procedure to construct the Cantor set. We define the Cantor set

$$\mathfrak{c} = \bigcap_{n=0} A_n.$$

Some properties of Cantor set:

(i) Cantor set $\mathfrak{c}$ is closed (as an intersection of closed sets). Actually, it is compact, which is a stronger property.

(ii) The "length" of Cantor set is $m(\mathfrak{c}) = 0$. On step n we removed 2^{n-1} disjoint intervals of length $\frac{1}{3^n}$ each. Thus, the construction removes intervals of total length

$$\sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = \frac{1/3}{1-2/3} = 1.$$

Thus, no "length " is left.

(iii) Every point of $\mathfrak{c}$ is the limit of other points of $\mathfrak{c}$. The easiest way to see is to see that the endpoints of the removed open intervals are dense $\mathfrak{c}$.

(iv) Cantor set is the set

$$\mathfrak{c} = \{x \in [0,1] : x = \sum_{k=1}^{\infty} \frac{d_k}{3^k},\ d_k \in \{0,2\},$$

i.e., set of numbers in $[0,1]$ without the digit 1 in their ternary expansions. This shows that the Cantor set is of cardinality continuum, $\operatorname{Card}(\mathfrak{c}) = \mathfrak{c}$.

(v) The function

$$f\left(\sum_{k=1}^{\infty} \frac{d_k}{3^k}\right) = \frac{1}{2} \sum_{k=1}^{\infty} \frac{d_k}{2^k},$$

transforms Cantor set onto the interval $[0,1]$. It can be proven that f is a continuous function.

4. Supremum, Infimum, Induction

4.1. **Supremum, Infimum.**

Definition 4.1. *Let A be a non-empty subset of real line$\mathbb{R}$. We say that α is an upper bound for A $\iff$ for any $a \in A$ we have $a \leq \alpha$. We say that α is a supremum of A, $\alpha = \sup A$, $\iff$ α is an upper bound for A and for any upper bound γ for A we have $\alpha \leq \gamma$, i.e., α is the least upper bound for A.*

Definition 4.2. *We say that β is a lower bound for A $\iff$ for any $a \in A$we have $\beta \leq a$. We say that β is an infimum of A, $\beta = \inf A$ $\iff$ β is a lower bound for A and for any lower bound γ for A we have $\gamma \leq \beta$, i.e., β is the greatest lower bound for A.*

Theorem 2. *Let A be a non-empty bounded subset of $\mathbb{R}$. Then, (i) $\alpha = \sup A \iff \alpha$ is an upper bound for A and $\forall_{\varepsilon>0} \exists_{a\in A}$ such that $\alpha - \varepsilon < a$. (ii) $\beta = \inf A \iff \beta$ is a lower bound for A and $\forall_{\varepsilon>0} \exists_{a\in A}$ such that $a < \beta + \varepsilon$.*

Proof. We prove (i). Part (ii) is proven similarly.

To prove equivalence we need to prove two implications:

(1) $\alpha = \sup A \implies \alpha$ is an upper bound for A and $\forall_{\varepsilon>0} \exists_{a\in A}$ such that $\alpha - \varepsilon < a$.

Let us assume that $\alpha = \sup A$ and negation of $\forall_{\varepsilon>0} \exists_{a\in A}$ such that $\alpha - \varepsilon < a$., i.e., $\exists_{\varepsilon>0} \forall_{a\in A} \alpha - \varepsilon \geq a$. The last statement means that $\alpha - \varepsilon < \alpha$ is an upper bound for A, so α is not the $\sup A$. Contradiction.

(2) $\alpha = \sup A \impliedby \alpha$ is an upper bound for A and $\forall_{\varepsilon>0} \exists_{a\in A}$ such that $\alpha - \varepsilon < a$. Let us assume the RHS statement and that α is not the least upper bound for A. Then, there exists $\gamma < \alpha$ such that $a \leq \gamma$ for all $a \in A$. Let $\varepsilon = (\alpha - \gamma)/2$. Then, $\gamma < \alpha - \varepsilon$ and we cannot find $a \in A$ with $\alpha - \varepsilon < a$. Contradiction with the assumed RHS statement.

□

Example 1: Let $A = \{a_k = \frac{1+2k}{1+k^2} : k = 1, 2, 3, \dots\}$. Find $\sup A$ and $\inf A$.

First, we list a few first elements of A: $a_1 = 3/2$, $a_2 = 5/5 = 1$, $a_3 = 7/10$, $a_4 = 9/17$, etc.

(i) $\sup A$: We see that the largest element of A is $a_1 = 3/2$. We will prove that for $k \geq 2$ we have $a_k \leq 1$. We have

$$\frac{1+2k}{1+k^2} \leq 1 \Leftrightarrow 1 + 2k \leq 1 + k^2 \Leftrightarrow 2k \leq k^2 \Leftrightarrow 2 \leq k,$$

so our claim is proven. Since if A has the largest element then this element is the supremum of A, we obtain $\sup A = 3/2$.

(ii) $\inf A$: We see that for all $k \geq 1$ we have $0 < a_k$, so 0 is a lower bound for A. Let us assume that there exists a larger lower bound, $\varepsilon > 0$. We will find a k such that $\frac{1+2k}{1+k^2} < \varepsilon$. We have

$$\frac{1+2k}{1+k^2} \leq \frac{k+2k}{1+k^2} \leq \frac{3k}{k^2} = \frac{3}{k},$$

so it is enough to find k such that $3/k < \varepsilon$. Assume that for all $k \geq 1$ we have $3/k \geq \varepsilon$. Then, $3/\varepsilon \geq k$ for all $k \geq 1$ which is impossible. Thus, there exists a $k \geq 1$ such that $\frac{1+2k}{1+k^2} < \varepsilon$ and $\varepsilon > 0$ is not a lower bound for the set A. Thus, $\inf A = 0$.

Example 4.1. ***Example 2:*** *Let*

$$A = \left\{ \left| \frac{\sin t}{t} \right| : t \in (0, +\infty) \right\}.$$

Find $\sup A$ *and* $\inf A$.

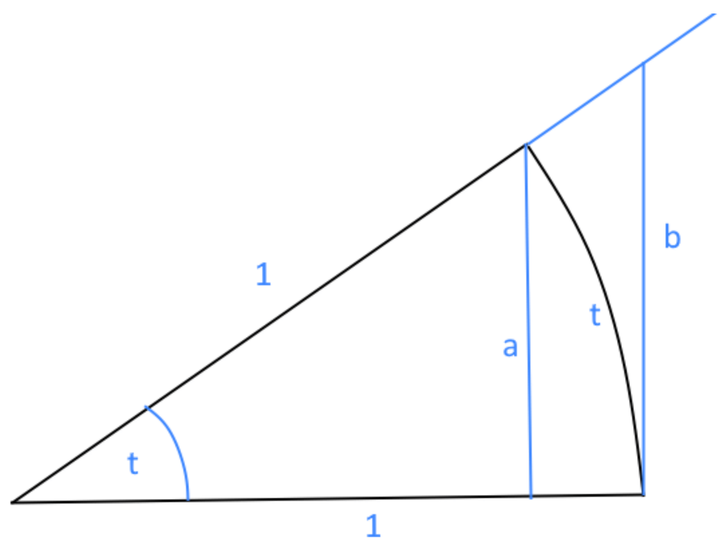

FIGURE 5.

In the Figure 1 we see a part of the unit disk corresponding to the central angle t (in radians). Then, the length of the corresponding arc is also t. We also have $a = \sin t$ and $b = \tan t$. From the picture we see that

$$(*) \sin t \leq t \leq \tan t.$$

We have $0 \leq \left|\frac{\sin t}{t}\right|$ for all t and $\left|\frac{\sin \pi}{\pi}\right| = 0$ so $\inf A = 0$. From the first inequality we have $\left|\frac{\sin t}{t}\right| \leq 1$ for all $t \neq 0$ so 1 is an upper bound for A. We will show that $1 = \sup A$. The second part of inequality $(*)$ gives

$$|\cos t| \leq \left| \frac{\sin t}{t} \right|, \quad t \in (0, +\infty).$$

Take an $\varepsilon > 0$. We can find a t such that $1 - \varepsilon < |\cos t| \leq 1$. Then, we also have $1 - \varepsilon < \left|\frac{\sin t}{t}\right| \leq 1$. This proves that $\sup A = 1$.

Some inequalities for Suprema and Infima:

Let A and B be non-empty bounded subsets of $\mathbb{R}$.

$$\text{(i)}\quad \sup\{a + b : a \in A, b \in B\} = \sup A + \sup B.$$

Proof: For any $a \in A$, $b \in B$ we have $a + b \leq \sup A + \sup B$, so it is an upper bound for $A + B = \{a + b : a \in A, b \in B\}$. Let us take $\varepsilon > 0$. Then, we can find an element $a \in A$ such that $\sup A - \varepsilon/2 < a$ and an element $b \in B$ such that $\sup B - \varepsilon/2 < b$.

Then, we have $\sup A + \sup B - \varepsilon < a + b$ which proves that $\sup A + \sup B$ is the supremum of $A + B$.

Is the equality $\inf\{a + b : a \in A, b \in B\} = \inf A + \inf B$, always true?

Is the inequality $\sup\{a \cdot b : a \in A, b \in B\} \leq \sup A \cdot \sup B$, always true? Add a condition ensuring that it holds.

$$\text{(ii)}\quad \sup\{-a : a \in A\} = -\inf A.$$

Proof: For any $a \in A$ we have $-a \leq -\inf A$, so it is an upper bound for $-A = \{-a : a \in A\}$. Let us take $\varepsilon > 0$. Then, we can find an element $a \in A$ such that $a < \inf A + \varepsilon$. Then, $-a > -\inf A - \varepsilon$ which shows that $-\inf A$ is the supremum of $-A$.

Completeness Axiom (Axiom # 8):

Let $A \subset \mathbb{R}$ be a non-empty subset of $\mathbb{R}$. If A is bounded above, then the real number $\alpha = \sup A$ exists. If A is bounded below, then the real number $\beta = \inf A$ exists.

The real number $\sqrt{3}$ exists by the Completeness Axiom.

We will show that $\sqrt{3}$, the number a with property $a^2 = 3$ exists. We know it cannot be a rational number.

Let us consider the set $A = \{x \in \mathbb{R} : x^2 < 3\}$. A is not empty ($1 \in A$) and bounded above for example by 2. The completeness Axiom says that $\alpha = \sup A$ is areal number. We will show that $\alpha^2 = 3$.

(i) Assume $\alpha^2 < 3$. Take $0 < \varepsilon < 1$. Then $\varepsilon^2 < \varepsilon$. Also, we have $\alpha \leq 2$. Then, we have $(\alpha+\varepsilon)^2 = \alpha^2 + 2\alpha\varepsilon + \varepsilon^2 \leq \alpha^2 + 5\varepsilon$. We see that as long as $5\varepsilon < 3 - \alpha^2$ we have $(\alpha+\varepsilon)^2 < 3$. This contradicts the assumption that $\alpha = \sup A$.

(ii) Assume $\alpha^2 > 3$. Since $\alpha = \sup A$ for any $\varepsilon > 0$ there exists $a \in A$ with $\alpha - \varepsilon < a$. Then, $a^2 > (\alpha-\varepsilon)^2 = \alpha^2 - 2\alpha\varepsilon + \varepsilon^2 > \alpha^2 - 2\alpha\varepsilon \geq \alpha^2 - 4\varepsilon$. As long as $4\varepsilon < \alpha^2 - 3$ we have $(\alpha-\varepsilon)^2 > 3$ so we found $a \in A$ with $a^2 > 3$, which is a contradiction.

Since (i) and (ii) are both impossible, we have $\alpha^2 = 3$. QED

Arithmetic-Geometric Means Inequality:

Lemma: Assume that $a_1, a_2, \ldots, a_n \geq 0$. If $a_1 \cdot a_2 \cdots a_n \geq 1$, then $a_1 + a_2 + \cdots + a_n \geq n$.

Proof: We will prove this by induction. For $n = 1$ we have statement: If $a_1 \geq 0$ and $a_1 \geq 1$ then $a_1 \geq 1$ which is obviously true. To make the further proof more understandable we will do case $n = 2$ which in principle is not necessary.

Let $n = 2$. $a_1, a_2 \geq 0$ and $a_1 \cdot a_2 \geq 1$. We want to show $a_1 + a_2 \geq 2$. If both a_1 and a_2 are greater than 1, there is nothing to show. We assume $a_1 \leq 1$ and $a_2 \geq 1$. We have

$$a_1 + a_2 \geq a_1 + a_2 + 1 - a_1 \cdot a_2 + 1 - 1 = 2 + a_1 - 1 - a_2(a_1 - 1) = 2 + (a_1 - 1)(1 - a_2) \geq 2,$$

since $(a_1 - 1)(1 - a_2) \geq 0$.

We assume that the Lemma holds for n. We will prove it for $n+1$. Let $0 \leq a_1 \leq a_2 \leq \cdots \leq a_n \leq a_{n+1}$ and $a_1 \cdot a_2 \cdots a_n \cdot a_{n+1} \geq 1$. We want to prove $a_1 + a_2 + \cdots + a_n + a_{n+1} \geq n+1$. If $a_1 \geq 1$ there is nothing to prove. We assume $a_1 \leq 1$ and $a_{n+1} \geq 1$. We also have $a_2 \cdot a_3 \cdots a_n \cdot (a_1 \cdot a_{n+1}) \geq 1$ so by induction assumption $a_2 + a_3 + \cdots + a_n + (a_1 \cdot a_{n+1}) \geq n$.

This implies $a_1 + a_2 + a_3 + \cdots + a_n + a_{n+1} \geq n + a_1 + a_{n+1} - (a_1 \cdot a_{n+1})$. Now,

$n+a_1+a_{n+1}-(a_1 \cdot a_{n+1}) = n+a_{n+1}(1-a_1)+a_1-1+1 = n+1+(a_{n+1}-1)(1-a_1) \geq n+1.$

QED

Corollary 1: Arithmetic-Geometric Mean Inequality. Let $x_1, x_2, \ldots, x_n > 0$, Then

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

Proof: Set

$$a_i = \frac{x_i}{\sqrt[n]{x_1 \cdot x_2 \cdots x_n}}, \quad i = 1, \ldots n.$$

Then, $a_1 \cdot a_2 \cdots a_n = 1$ so by the Lemma $a_1 + a_2 + \ldots a_n \geq n$.

Corollary 2: For any $n \geq 1$ we have

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

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

This means, that

$$a_n = \left(1 + \frac{1}{n}\right)^n, \text{ is increasing,}$$

and

$$b_n = \left(1 + \frac{1}{n}\right)^{n+1}, \text{ is decreasing.}$$

Proof: By Arithmetic-Geometric Mean Inequality, we have

$$\sqrt[n+1]{\left(1 + \frac{1}{n}\right)^n} < \frac{1}{n+1}\left(1 + 1 + \frac{1}{n} + 1 + \frac{1}{n} + \cdots + 1 + \frac{1}{n}\right) = 1 + \frac{1}{n+1}\left(n \cdot \frac{1}{n}\right) = 1 + \frac{1}{n+1},$$

which is equivalent to our claim (1).

The proof of (2) is similar, although not obviously similar.

“e” - the base of the natural logarithms:

If we set $a_n = \left(1+\frac{1}{n}\right)^n$ and $b_n = \left(1+\frac{1}{n}\right)^{n+1}$, then the intervals $I_n = [a_n, b_n]$ form a nested sequence. Since $a_n \le b_1 = 4$ we have

$$b_n - a_n = \frac{1}{n}\left(1+\frac{1}{n}\right)^n \le \frac{4}{n} \to 0\ ,\ \text{as } n \to +\infty.$$

Thus, the intersection $\bigcap_{\ n\ge 1} I_n$ is one point, the number called "e", the base of the natural logarithms. It has many beautiful properties, the most known probably is

$$(e^x)' = e^x.$$

We will prove this property much later.

Proposition 4.1. *We constructed the number e and we have the two-sided bound for it*

$$\left(1+\frac{1}{n}\right)^n \le e \le \left(1+\frac{1}{n}\right)^{n+1}\ ,\ n \ge 1.$$

These are not very good approximations, for $n = 1000$ we obtain

$$2.716923932 < e < 2.719640856.$$

4.2. **Principle of Mathematical Induction:** The principle of Mathematical Induction gives a very useful method of proving theorems which hold for natural numbers.

Principle of Mathematical Induction: Let P be a statement depending on natural numbers $n \in \mathbb{N}$.

If

(i) $P(1)$ is true, and

(ii) the implication $P(n) \implies P(n+1)$ is true,

then $P(n)$ holds for all $n \in \mathbb{N}$.

Example 1: Show that $6|(n^3+5n)$ (n^3+5n is divisible by 6) for any $n \in \mathbb{N}$.

Proof: We check $P(1):\ 6|6$ - true.

Now, we assume $P(n)$, i.e., $6|(n^3+5n)$. We will prove $P(n+1)$, i.e., $6|((n+1)^3+5(n+1))$.

We have $(n+1)^3+5(n+1) = n^3+3n^2+3n+1+5n+5 = (n^3+5n)+3(n^2+n)+6$. The first summand n^3+5n divides by 6 by Inductive assumption, the last one is 6 and also divides by 6. Now, $3(n^2+n) = 3n(n+1)$. It divides by 3 and one of the numbers $n, n+1$ is even so divides by 2. The sum of expressions divisible by 6 is divisible by 6 and we proved $P(n+1)$. By the Principle of Mathematical Induction we proved $6|(n^3+5n)$ for any $n \in N$.

Example 1: Show that

$$(*)\quad 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} \geq \sqrt{n},$$

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

Proof: We check $P(1) : 1 \geq 1$ - true. Now, we assume $P(n)$, i.e., the inequality $(*)$. We will prove $P(n+1)$, i.e.,

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n+1}} \geq \sqrt{n+1}.$$

We start with assumed inequality $(*)$ and add $\frac{1}{\sqrt{n+1}}$ to both sides. Now, it is enough to show that

$$\sqrt{n} + \frac{1}{\sqrt{n+1}} \geq \sqrt{n+1}.$$

It is equivalent to

$$\sqrt{n(n+1)} + 1 \geq n+1 \iff \sqrt{n(n+1)} \geq n.$$

Squaring both sides (they are positive), we obtain $n(n+1) \geq n^2$, which is true. We proved $P(n+1)$ and by the Principle of Mathematical Induction we proved inequality $(*)$ for any $n \in \mathbb{N}$.

Example 3: Theorem SM : Every non-empty subset of $\mathbb{N}$ has the smallest element. This property is often formulated differently: The set $\mathbb{N}$ is well ordered (for every element we can point out the next element in the sense of the relation $<$).

Proof: First, we will prove Theorem SM for finite subsets of $\mathbb{N}$. We use Mathematical Induction over the number of elements in the subset. We check $P(1)$: The set with 1 element has the smallest element. True.

Now, we assume $P(n)$: Any subset of $\mathbb{N}$ with n elements has the smallest element. Consider a subset A with $n+1$ elements. Let $n_0 \in A$. Then, $A \setminus \{n_0\}$ has n elements, so it has the smallest element, say n_1. If $n_1 < n_0$ then n_1 is the smallest element of A.

If $n_0 < n_1$ then n_0 is the smallest element of A. We proved $P(n+1)$. This completes the proof for finite subsets of $\mathbb{N}$.

Let B be a non-empty subset of $\mathbb{N}$. Since B is nonempty, it contains an element, say n_0. The set $B \cap \{1, 2, 3, ..., n_0 - 1, n_0\}$ is a finite non-empty subset of $\mathbb{N}$ so it contains the smallest element, say n_1. Then, n_1 is the smallest element of B.

We proved Theorem SM. QED

Now, we will prove that Theorem SM implies the Principle of Mathematical Induction. We will use a proof by contradiction. We assume Theorem SM and that PMI does not hold, i.e., there exists a theorem P such that $P(1)$ is true, $P(n) \implies P(n+1)$ is true and there exists $n_0 \in \mathbb{N}$ such that $P(n_0)$ is false. Let $A = \{n \in \mathbb{N} : P(n) \text{ is false } \}$. A is non-empty ($n_0 \in A$), by Theorem SM A has the smallest element, say n_1. Since $P(1)$ is true, $n_1 > 1$. Consider $P(n_1 - 1)$.

If it is true, then $P(n_1)$ is also true since we know $P(n) \implies P(n+1)$ is true. Thus, $P(n_1 - 1)$ cannot be true.

If it is false, then $n_1 - 1 \in A$ which contradicts n_1 being the smallest element of A.

We obtained a contradiction, which proves our claim. QED

Example 4: Pigeon Hole Principle : If there are $n+1$ pigeons in n holes, there is a hole with at least two pigeons. Or more "scientifically": If $n \in \mathbb{N}$, there is no injection of the set with $n+1$ elements into a set with n elements.

A function $f : X \to Y$ is called an injection $\iff$ if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$ or equivalently if $f(x_1) = f(x_2)$ then $x_1 = x_2$.

We will prove PHP by induction: Below by A_n we denote a set with n elements.

We check $P(1)$: There is no injection from $A_2 = a, b$ into $A_1 = c$. If $f : A_2 \to A_1$, then $f(a) = c$ and $f(b) = c$, so f is not an injection.

We assume $P(n)$: there is no injection from A_{n+1} into A_n. We will prove $P(n+1)$ by contrapositive. Assume that $P(n+1)$ is not true, i.e., there is an injection $f : A_{n+2} \to A_{n+1}$. Take an element $a_0 \in A_{n+2}$. Since f is an injection, the only point which goes onto $f(a_0)$ is a_0. This means that $f : A_{n+2} \setminus \{a_0\} \to A_{n+1} \setminus \{f(a_0)\}$ is also an injection. $A_{n+2} \setminus \{a_0\}$ has $n+1$ elements and $A_{n+1} \setminus \{f(a_0)\}$ has n elements, so this shows that $P(n)$ does not hold. This contradiction completes the proof.

Now, we will prove that the PHP implies the PMI. We use the proof by contrapositive. Let us assume PHP and assume that PMI does not hold, i.e., there exists a theorem P such that $P(1)$ is true, $P(n) \implies P(n+1)$ is true and there exists $n_0 \in \mathbb{N}$ such that $P(n_0)$ is false. Let $A = \{n \in \mathbb{N} : P(n) \text{ is true }\}$ and consider the set $B = A \cap \{1, 2, 3, ..., n_0 - 1, n_0\}$. Assume B has n elements. It is non-empty as $1 \in B$. We will define injection $f : B \to B \setminus \{1\}$. Let $f(n) = n + 1$, for $n \in B$. We have $f(n) \in B$ since if $n \in B$ then $n + 1 \in B$ and $n_0 \notin B$. f is injective. Element 1 is not a value of f. We showed that PHP does not hold. This completes the proof of our claim. QED

5. Cardinality

5.1. **Relations:** If X and Y are non-empty sets then their Cartesian product is the set of pairs

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

Remember that $(x_1, y_1) = (x_2, y_2) \iff x_1 = x_2$ and $y_1 = y_2$, i.e., the order in the pair counts. A **relation** is a subset $R \subset X \times Y$. We write $x \sim_R y$, x is in relation with y, if and only is $(x, y) \in R$.

A relation $R \subset X \times X$ is called:

(1) **Reflexive** $\iff x \sim_R x, \ \forall_{x \in X}$;

(2) **Symmetric** $\iff x \sim_R y \implies y \sim_R x, \ \forall_{x,y \in X}$;

(3) **Transitive** $\iff x \sim_R y$ and $y \sim_R z \iff x \sim_R z, \ \forall_{x,y \in X}$.

For example, relation $<$ on $\mathbb{R}$ is transitive but it is neither reflexive nor symmetric.

A relation R on X which is transitive, reflexive and symmetric is called an **equivalence relation**. It divides the set X into disjoint "equivalence classes", subsets A_α such that $x, y \in A_\alpha$ if and only if $x \sim_R y$. If $x \in A_{\alpha_1}$ and $y \in A_{\alpha_2}$ with $\alpha_1 \neq \alpha_2$, then $x \not\sim_R y$. The equivalence class containing an element x is denoted by $[x]$. In particular, $x \sim_R y \iff [x] = [y]$.

Example: Let $X = \mathbb{N}$, $n \sim_R m \iff$ there exists an integer k such that both $n/2^k$ and $m/2^k$ are odd numbers. For example $6 \sim_R 14$ (k=1) and $6 \not\sim_R 9$. It is easy to see that this relation is an equivalence relation. It divides the set $\mathbb{N}$ into disjoint equivalence classes. We have

$$[1] \; = \; \text{odd numbers } (k = 0);$$
$$[2] \; = \; 2\cdot\text{odd numbers } (k = 1);$$
$$[2^k] \; = \; 2^k\cdot\text{odd numbers } (k = k).$$

We have $\mathbb{N} = \bigcup_{k=0}^{\infty}\{[2^k]\}$ and $\mathbb{N}/R = \{2^k : k = 0, 1, 2, \ldots\}$. The last one is called a factor space obtained by dividing $\mathbb{N}$ by the equivalence relation R.

5.2. **Functions:** A **function** $f : X \to Y$ is a relation on $X \times Y$ with special properties. We write $f(x) = y$ instead of $x \sim_f y$. We assume

(1) for any $x \in X$ there exists an $y \in Y$ such that $f(x) = y$;

(2) if $f(x) = y_1$ and $f(x) = y_2$ then $y_1 = y_2$ (the value $f(x)$ is uniquely determined).

A function $f : X \to Y$ is called **injection** $\iff$ if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$, or equivalently if $f(x_1) = f(x_2)$ then $x_1 = x_2$.

A function $f : X \to Y$ is called **surjection** or "onto function" $\iff$ for any $y \in Y$ there exists $x \in X$ with $f(x) = y$.

A function $f : X \to Y$ is called **bijection** or "one-to-one function" $\iff$ f is both an injection and a surjection.

Examples: Function $f : [0, = \infty) \to \mathbb{R}$, $f(x) = x^2$ is an injection, but not a surjection. Function $f : \mathbb{R} \to [0, +\infty)$, $f(x) = x^2$ is a surjection, but not an injection. Function $f : (-\pi/2, \pi/2) \to \mathbb{R}$, $f(x) = \tan x$ is a bijection.

If $f : X \to Y$ is a bijection $x \overset{f}{\mapsto} y$, then the **inverse function** is $f^{-1} : Y \to X$, $y \overset{f^{-1}}{\longmapsto} x$. For any $x \in X$ we have $f^{-1}(f(x)) = x$ and for every $y \in Y$ we have $f(f^{-1}(y)) = y$.

5.3. **Cardinality:** Cardinality of a set is a notion generalizing the number of elements in a set. For finite sets the cardinality equals the number of elements.

There are two basic ways to compare the number of elements in two sets. For example, consider acrowd of students in front of a classroom and the set of chairs inside. We do not know if we have enough chairs for all the students. How to check?

(1) We can count the number of students, the number of chairs and compare. This works well for practical, finite cases.

(2) We can let the students inside and ask them to sit on the chairs. This creates a function from the set of students to the set of chairs. If this function is an injection, the number of students is less than or equal to the number of students. If it is a bijection the numbers are equal. If there is not enough chairs we cannot create a proper function.

To define cardinality of general sets we use the second method.

We say that the sets A and B are equivalent (in the sense of cardinality), $A \sim_C B$, or have the same cardinality, $\mathrm{Card(A)} = \mathrm{Card(B)}$, $\iff$ the exist a bijection $f : A \to B$. It is easy to see that relation $\sim_C$ is an equivalence relation (identity is bijection, inverse of a bijection is a bijection, composition of two bijections is a bijection). The equivalence classes of this relation are **cardinal numbers**. For example: 5 corresponds to the class of all sets equivalent to $\{1, 2, 3, 4, 5\}$, all sets equivalent to $\mathbb{N}$ have cardinality $\aleph_0$ (Hebrew letter aleph), all sets equivalent to $\mathbb{R}$ have cardinality $\mathfrak{c}$ (Gothic small c for Georg Cantor, who introduced cardinality). There infinitely many different infinite cardinal numbers. We will prove this later.

We can define addition and multiplication of cardinal numbers.

Let α_1 and α_2 be cardinal numbers. If A_1 and A_2 are disjoint and $\mathrm{Card(A_1)} = \alpha_1$ and $\mathrm{Card(A_2)} = \alpha_2$, then $\alpha_1 + \alpha_2$ is defined as a cardinality of the set $A_1 \cup A_2$. $\alpha_1 \cdot \alpha_2$ is defined as a cardinality of the set $A_1 \times A_2$, the Cartesian product.

If there is an injection $f : A \to B$, then we have $\mathrm{Card(A)} \leq \mathrm{Card(B)}$. If there is a surjection $f : A \to B$, then we have $\mathrm{Card(B)} \leq \mathrm{Card(A)}$.

The following theorem is often useful

Theorem 3. ***Cantor-Bernstein-Schröder Theorem:*** *If* Card(A) $\leq$ Card(B) *and* Card(B) $\leq$ Card(A)*, then* Card(A) = Card(B)*. In other words: If there is an injection $f : A \to B$ and an injection $g : B \to A$, then there exists a bijection $h : A \to B$.*

Proof from Larsen's book:

Let us assume that f is not a bijection.

Let $B_1 = B \setminus f(A)$. If $B_k \subset B$ is defined for some $k \in \mathbb{N}$, let $A_k = g(B_k)$ and $B_{k+1} = f(A_k)$. This inductively defines A_k and B_k for all $k \in \mathbb{N}$. Use these sets to define $\tilde{A} = \bigcup_{k \geq 1} A_k$ and $h : A \to B$ as

$$h(x) = \begin{cases} g^{-1}(x), & x \in \tilde{A}; \\ f(x), & x \in A \setminus \tilde{A}. \end{cases}$$

It must be shown that h is well-defined, injective and surjective. To show h is well-defined, let $x \in A$. If $x \in A \setminus \tilde{A}$, then it is clear $h(x) = f(x)$ is well defined. On the other hand, if $x \in A$, then $x \in A_k$ for some k. Since $x \in A_k = g(B_k)$, we see $h(x)g^{-1}(x)$ is well defined. Therefore, h is well-defined.

To show h is injective, let $x, y \in A$ with $x \neq y$. If both $x, y \in A$ or $x, y \in A \setminus \tilde{A}$, then the assumptions that f and g are injective, respectively, imply $h(x) \neq h(y)$.

The remaining case is when $x \in A$ and $y \in A \setminus \tilde{A}$. Suppose $x \in A_k$ and $h(x) = h(y)$. If $k = 1$, then $h(x) = g^{-1}(x) \in B_1$ and $h(y) = f(y) \in f(A) \setminus B_1$. This is clearly incompatible with the assumption that $h(x) = h(y)$. Now, suppose $k > 1$. Then there is an $x_1 \in B_1$ such that

$$x = \underbrace{g \circ f \circ g \circ f \circ g \circ \cdots \circ f \circ g}_{\text{k-1 } f\text{'s and k } g\text{'s}}(x_1).$$

This implies

$$h(x) = g^{-1}(x) = \underbrace{f \circ g \circ f \circ g \circ \cdots \circ f \circ g}_{\text{k-1 } f\text{'s and k-1 } g\text{'s}}(x_1) = f(y).$$

so that

$$y = \underbrace{g \circ f \circ g \circ f \circ g \circ \cdots \circ f \circ g}_{\text{k-2 } f\text{'s and k-1 } g\text{'s}}(x_1) \in A_{k-1} \subset \tilde{A}.$$

This contradiction shows that$h(x) \neq h(y)$. We conclude h is injective.

To show h is surjective, let $y \in B$. If $y \in B_k$ for some k, then $h(A_k) = g^{-1}(A_k) = B_k$ shows $y \in h(A)$. If $y \notin B_k$ for any k, $y \in f(A)$ because $B_1 = B \setminus f(A)$, and $g(y) \notin \tilde{A}$, so $y = h(x) = f(x)$ for some $x \in A$. This shows h is surjective.

QED

As we said before $\mathrm{Card}(\mathbb{N}) = \aleph_0$. Any set A with $\mathrm{Card(A)} \leq \aleph_0$ is called a **countable set**. Another way to say that $\mathrm{Card(A)} = \aleph_0$ is to say that the elements of A can be ordered in a sequence.

Consider $\{0\} \cup \mathbb{N}$. The function $f(n) = n + 1$ is a bijection $f : \{0\} \cup \mathbb{N} \to \mathbb{N}$. This shows that $\mathrm{Card}(\{0\} \cup \mathbb{N}) = \mathrm{Card}(\mathbb{N})$ or

$$\aleph_0 + 1 = \aleph_0. \tag{4}$$

As we see, the arithmetic of cardinal numbers is somewhat surprising. Equality (4) implies, by induction

$$\aleph_0 + n = \aleph_0,\ n \in \mathbb{N}. \tag{5}$$

The equations (4) and (5) are often presented as an anegdote about “Hilbert’s hotel”. Hotel “Hilbert” is an unusual one. It has $\aleph_0$ rooms numbered $H1, H2, H3, \dots$. One night all rooms are taken. The weather is awful. Around midnight one more guest comes and asks for a room. “No problem” says the night receptionist. He moves the guest from room H1 to room H2, the guest from room H2 to room H3, the guest from room H3 to room H4, etc., leaving room H1 free for a new guest.

If instead of one guest a group of n guests arrives and asks for n rooms he proceeds similarly: $H1 \mapsto H(n+1)$, $H2 \mapsto H(n+2)$, $H3 \mapsto H(n+3)$,etc, leaving n first rooms free.

Now, a more serious trouble arises: a whole bus full of extra guests in need of separate rooms comes. And the bus itself has infinite number of seats: $B1, B2, B3, B4, \dots$. The receptionist is still undefeated: he moves the old and new guests as follows:

$H1 \mapsto H1$, $H2 \mapsto H3$, $H3 \mapsto H5$,..., $Hn \mapsto H(2n-1)$, $n = 1, 2, 3, \dots$.

$B1 \mapsto H2$, $B2 \mapsto H4$, $B3 \mapsto H6, \ldots$, $Bn \mapsto H(2n)$, $n = 1, 2, 3, \ldots$.

This way old and new guest have their own separate rooms.

The last example proves

$$\aleph_0 + \aleph_0 = \aleph_0.$$

This, by induction, implies that any finite sum of $\aleph_0$'s is again $\aleph_0$, or that the finite sum of countable sets is countable. We can write it as

$$n \cdot \aleph_0 = \aleph_0 \ , \ n \in \mathbb{N}.$$

In particular it proves that for the set of integers $\mathbb{Z} = \mathbb{N} \cup -\mathbb{N} \cup \{0\}$ we have $\mathrm{Card}(\mathbb{Z}) = \aleph_0$.

Now, we will show that

$$\aleph_0 \cdot \aleph_0 = \aleph_0.$$

We need to show that $\mathrm{Card}(\mathbb{N} \times \mathbb{N}) = \aleph_0$. The set $\mathbb{N} \times \mathbb{N}$ can identified with $\mathbb{Q}^+ = \{\frac{n}{m} : n, m \in \mathbb{N}\}$.

Fast proof: Obviously $\mathbb{N} \subset \mathbb{Q}$ so $\mathrm{Card}(\mathbb{N}\times\mathbb{N}) \geq \aleph_0$. Now, the function $f(\frac{n}{m}) = 2^n 3^m$ is an injection of $\mathbb{Q}$ into $\mathbb{N}$. Thus, $\mathrm{Card}(\mathbb{Q}) \leq \aleph_0$. Using Cantor-Bernstein-Schröder Theorem we obtain $\mathrm{Card}(\mathbb{Q}) = \aleph_0$.

We present a direct proof constructing a bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$. In Figure 6 we show the set $\mathbb{N} \times \mathbb{N}$ in the form of doubly infinite matrix. The red arrows show how we can order all elements of the set into a sequence, defining the bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$. Thus, without invoking any theorems, $\mathrm{Card}(\mathbb{N} \times \mathbb{N}) = \aleph_0$, and

$$\aleph_0 \cdot \aleph_0 = \aleph_0. \tag{6}$$

This proof is called a "snake proof" for the shape of the red arrows path. By induction we obtain $\aleph_0^n = \aleph_0$, for any $n \in \mathbb{N}$.

Corollary 5.1. *Equality 6 implies that a countable union of countable sets is countable.*

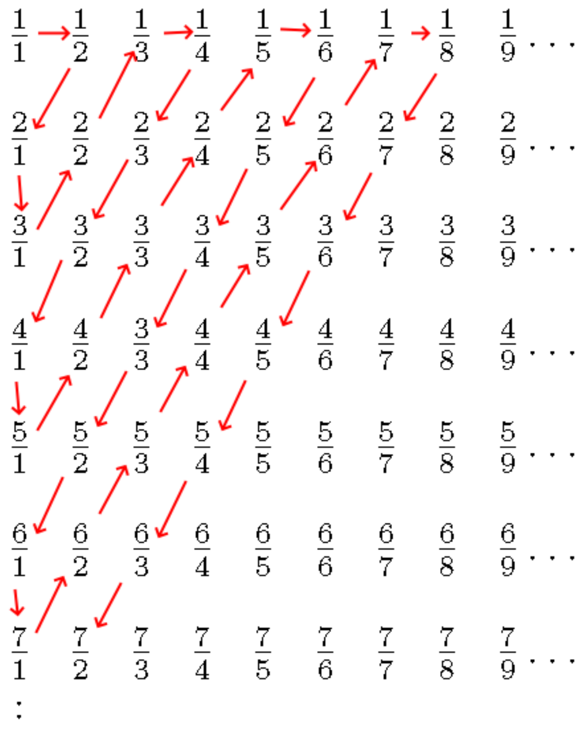

FIGURE 6. Bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$

Example 5.1.

Consider a set $\mathcal{J}$ of all solutions of the equations $\sin x = r$, where r is a rational number. In other words $\mathcal{J}$ is the set of all $x \in \mathbb{R}$ such that $\sin x$ is a rational number. Show that

$$\text{Card}\mathcal{J} = \aleph_0 \ .$$

Is $\mathcal{J}$ dense in $\mathbb{R}$?

Consider the graph of function $\sin(x)$ for $x \in [-\pi/2, \pi/2]$.

The function sin is a bijection $\sin : [-\pi/2, \pi/2] \to [-1, 1]$ and it inverse function $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ is also a bijection. We have $\arcsin(\mathbb{Q} \cap [-1, 1]) = J \cap [-\pi/2, \pi/2]$, so the set $J \cap [-\pi/2, \pi/2]$ is countable. At the same time we see

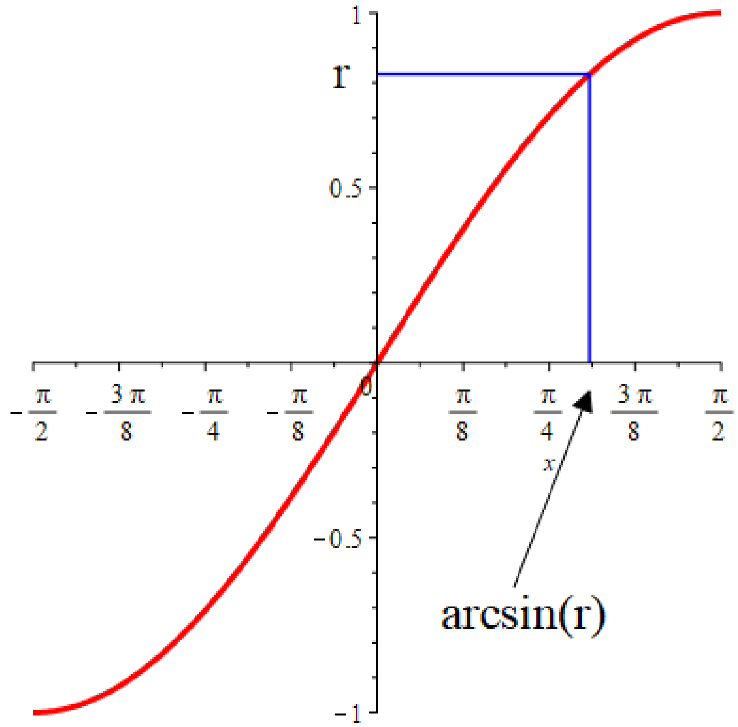

Figure 7. Graph of $\sin(x)$, $x \in [-\pi/2, \pi/2]$

that $\arcsin(\mathbb{Q} \cap [-1,1])$ is dense in $[-\pi/2, \pi/2]$, since $\mathbb{Q} \cap [-1,1]$ is dense in $[-1,1]$ and arcsin is a continuous function. In the same way we can show that each set $J_n = J \cap ([-\pi/2, \pi/2] + n \cdot \pi)$, $n = 0, \pm 1, \pm 2, \pm 3, \ldots$ is countable and dense in $[-\pi/2, \pi/2] + n \cdot \pi$. Since a countable union of countable sets is countable we obtain that J is countable. We also proved that J is dense in $\mathbb{R} = \bigcup_{n \in \mathbb{Z}}([-\pi/2, \pi/2] + n \cdot \pi)$.

Since $\mathbb{Q} \subset \mathbb{R}$ we have $\aleph_0 \leq \text{Card}(\mathbb{R})$. We will prove that $\mathbb{R}$ is not countable, so $\text{Card}(\mathbb{R}) = \mathfrak{c} > \aleph_0$. We will actually show that the interval $(0, 1)$ is not countable, but this is enough. Let us assume that all $x \in (0, 1)$ can be ordered into a sequence: $x_1, x_2, x_3, \ldots$. We can expand any x_k into its binary expansion
$x_k = 0.b_1^{(k)} b_2^{(k)} b_2^{(k)} b_4^{(k)} b_5^{(k)} b_6^{(k)} \ldots$

We put the sequence $x_1, x_2, x_3, \ldots$ into a table for visualization:

$$
\begin{aligned}
x_1 &= 0.b_1^{(1)} b_2^{(1)} b_3^{(1)} b_4^{(1)} b_5^{(1)} b_6^{(1)} \ldots \\
x_2 &= 0.b_1^{(2)} b_2^{(2)} b_3^{(2)} b_4^{(2)} b_5^{(2)} b_6^{(2)} \ldots \\
x_3 &= 0.b_1^{(3)} b_2^{(3)} b_3^{(3)} b_4^{(3)} b_5^{(3)} b_6^{(3)} \ldots \\
x_4 &= 0.b_1^{(4)} b_2^{(4)} b_3^{(4)} b_4^{(4)} b_5^{(4)} b_6^{(4)} \ldots \\
x_5 &= 0.b_1^{(5)} b_2^{(5)} b_3^{(5)} b_4^{(5)} b_5^{(5)} b_6^{(5)} \ldots \\
x_6 &= 0.b_1^{(6)} b_2^{(6)} b_3^{(6)} b_4^{(6)} b_5^{(6)} b_6^{(6)} \ldots \\
&\vdots
\end{aligned}
$$

If $b = 0$ then $1 - b = 1$. If $b = 1$ then $1 - b = 0$. We construct points

$$\bar{x} = 0.(1 - b_1^{(1)})(1 - b_2^{(2)})(1 - b_3^{(3)})(1 - b_4^{(4)})(1 - b_5^{(5)})(1 - b_6^{(6)}) \ldots$$

The point $\bar{x}$ is not in the sequence: $\bar{x}_1 \neq b_1^{(1)}$, so $\bar{x} \neq x_1$; $\bar{x}_2 \neq b_2^{(2)}$, so $\bar{x} \neq x_2$; $\bar{x}_3 \neq b_3^{(3)}$, so $\bar{x} \neq x_3$, etc. We proved that $(0, 1)$ is not countable. Since we used the digits on the diagonal of the table this proof is called a "diagonal proof".

We can give another, more geometric proof. Now we will prove that the interval $[0, 1]$ is not countable. Again, let us assume that all $x \in [0, 1]$ can be ordered into a sequence: $x_1, x_2, x_3, \ldots$. Let $I_0 = [0, 1]$. We divide I_0 into three subintervals of equal length 1/3: $I_1^{(0)}$, $I_2^{(0)}$ and $I_3^{(0)}$. We make the middle subinterval open, both side subintervals closed. If $x_1 \in I_1^{(0)} \cup I_2^{(0)}$, we define $I_1 = I_3^{(0)}$. If $x_1 \in I_2^{(0)} \cup I_3^{(0)}$, we define $I_1 = I_1^{(0)}$.

We have $I_1 \subset I_0$, length of I_1 is 1/3 and $x_1 \notin I_1$.

Now, we iterate this procedure: We divide I_1 into three subintervals of equal length $1/9$: $I_1^{(1)}$, $I_2^{(1)}$ and $I_3^{(1)}$. We make the middle subinterval open, both side subintervals closed. If $x_2 \in I_1^{(1)} \cup I_2^{(1)}$, we define $I_2 = I_3^{(1)}$. If $x_2 \in I_2^{(1)} \cup I_3^{(1)}$, we define $I_2 = I_1^{(1)}$.

We have $I_2 \subset I_1 \subset I_0$, length of I_2 is $1/9$ and $x_1, x_2 \notin I_2$.

We continue by induction: Let us assume that we have intervals $I_n \subset I_{n-1} \subset \cdots \subset I_1 \subset I_0$, length of I_n is $1/3^n$ and $x_1, x_2, \ldots, x_{n-1}, x_n \notin I_n$. We divide I_n into three subintervals of equal length $1/3^{n+1}$: $I_1^{(n)}$, $I_2^{(n)}$ and $I_3^{(n)}$. We make the middle subinterval open, both side subintervals closed. If $x_{n+1} \in I_1^{(n)} \cup I_2^{(n)}$, we define $I_{n+1} = I_3^{(1)}$. If $x_{n+1} \in I_2^{(1)} \cup I_3^{(1)}$, we define $I_{n+1} = I_1^{(1)}$.

We have intervals $I_{n+1} \subset I_n \subset I_{n-1} \subset \cdots \subset I_1 \subset I_0$, length of I_n is $1/3^{n+1}$ and $x_1, x_2, \ldots, x_{n-1}, x_n, x_{n+1} \notin I_{n+1}$.

The intervals I_n, $n \geq 1$, form a nested sequence of closed intervals and their lengths go to 0. Thus, the intersection $\bigcap_{n\geq 1}$ is one point, call it $\bar{x}$. This point is not in the sequence since $\bar{x} \in I_k$ for all $k \geq 1$ and $x_k \notin I_k$.

We proved that $[0,1]$ is not countable, once more showing that $\aleph_0 < \mathfrak{c}$.

The countable infinite Cartesian product of finite sets is not countable (it is of cardinality $\mathfrak{c}$). We will prove that

$$\{0,1\}^{\mathbb{N}} = \prod_{n=1}^{\infty} \{0,1\},$$

is of cardinality $\mathfrak{c}$, which shows $2^{\aleph_0} = \mathfrak{c}$. The same proof shows $n^{\aleph_0} = \mathfrak{c}$.

Proof: The element of $\{0,1\}^{\mathbb{N}}$ is a sequence of 0's and 1's : $\bar{x} = (b_1, b_2, b_3, \ldots)$, $b_k \in \{0,1\}$ for any $k = 1, 2, \ldots$ We define $f : \{0,1\}^{\mathbb{N}} \to [0,1]$, $f(\bar{x}) = 0.b_1 b_2 b_3 \ldots$, where $0.b_1 b_2 b_3 \ldots$ is a real number in binary expansion. f is a surjection, and f is a bijection if we ignore the real numbers with double binary expansions (like $0.10000000\cdots = 0.01111111111\ldots$). There are only countably many such numbers so they do not matter in this consideration.

For the proof of $n^{\aleph_0} = \mathfrak{c}$ we use expansions in base n. QED

The countable Cartesian product of sets of cardinality $\mathfrak{c}$ is of cardinality $\mathfrak{c}$.

We will prove that $[0,1]^{\mathbb{N}}$ is of cardinality $\mathfrak{c}$.

An element of $[0,1]^{\mathbb{N}}$ is a sequence of numbers $\bar{x} = (x_1, x_2, x_3, \dots)$, $x_k \in [0,1]$, k=1,2,.... We can represent every fraction $x-k$ as its binary expansion. Then an element $\bar{x} = (x_1, x_2, x_3, \dots)$ of $[0,1]^{\mathbb{N}}$ can be presented as the table

$$\begin{aligned}
x_1 &= 0.b_1^{(1)}\ b_2^{(1)}\ b_3^{(1)}\ b_4^{(1)}\ b_5^{(1)}\ b_6^{(1)}\dots\\
x_2 &= 0.b_1^{(2)}\ b_2^{(2)}\ b_3^{(2)}\ b_4^{(2)}\ b_5^{(2)}\ b_6^{(2)}\dots\\
x_3 &= 0.b_1^{(3)}\ b_2^{(3)}\ b_3^{(3)}\ b_4^{(3)}\ b_5^{(3)}\ b_6^{(3)}\dots\\
x_4 &= 0.b_1^{(4)}\ b_2^{(4)}\ b_3^{(4)}\ b_4^{(4)}\ b_5^{(4)}\ b_6^{(4)}\dots\\
x_5 &= 0.b_1^{(5)}\ b_2^{(5)}\ b_3^{(5)}\ b_4^{(5)}\ b_5^{(5)}\ b_6^{(5)}\dots\\
x_6 &= 0.b_1^{(6)}\ b_2^{(6)}\ b_3^{(6)}\ b_4^{(6)}\ b_5^{(6)}\ b_6^{(6)}\dots\\
&\ \vdots
\end{aligned}$$

Now, we can order all digits of the table into one sequence using the “snake method’ used in the proof that $\mathbb{Q}$ is countable. The beginning of the sequence is $b_1^{(1)}b_2^{(1)}b_1^{(2)}b_1^{(3)}b_2^{(2)}b_3^{(1)}b_4^{(2)}b_3^{(2)}b_2^{(3)}b_1^{(4)}b_1^{(5)}b_2^{(4)}\dots$ The map

$$f(\bar{x}) = 0.b_1^{(1)}b_2^{(1)}b_1^{(2)}b_1^{(3)}b_2^{(2)}b_3^{(1)}b_4^{(2)}b_3^{(2)}b_2^{(3)}b_1^{(4)}b_1^{(5)}b_2^{(4)}\dots$$

is a bijection of $[0,1]^{\mathbb{N}}$ onto interval $[0,1]$ if we ignore possibility of double binary representations. It is possible only for countably many elements of $[0,1]$ so it is unimportant in this consideration.

This shows that $\mathfrak{c}^{\aleph_0} = \mathfrak{c}$. QED

The countable infinite Cartesian product of countable sets is not countable (it is of cardinality $\mathfrak{c}$).

In other words $\aleph_0^{\aleph_0} = \mathfrak{c}$. To show this we will use Cantor-Bernstein-Schröder theorem. We have

$$\{0,1\}^{\mathbb{N}} \subset \mathbb{N}^{\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}} \quad \text{so} \quad \mathfrak{c} \le \aleph_0^{\aleph_0} \le \mathfrak{c}.$$

Thus, $\aleph_0^{\aleph_0} = \mathfrak{c}$. QED

Algebraic numbers are countable. The algebraic numbers are the roots (zeros) of polynomials with integer coefficients. $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0$, $a_n, a_{n-1}, \ldots, a_2, a_1, a_0 \in \mathbb{Z}$ is an example of such a polynomial. Its order is n. Since $\mathrm{Card}(\mathbb{Z}) = \aleph_0$ and finite union of countable sets is countable there is $\aleph_0$ different polynomials of order n. A polynomial of order n can have at most n real roots (exactly n complex roots) so the number of roots of order n is $n \cdot \aleph_0 = \aleph_0$. Now we take a union of the sets of roots of order n over natural number and since $\aleph_0 \cdot \aleph_0 = \aleph_0$ the union is countable. QED

Cardinality of a power set: The power set $\mathcal{P}(A)$ (or 2^A) of a set A is the set of all subsets of A. Note that if A has n elements, then $\mathcal{P}(A)$ has 2^n elements. This is the origin of the second notation. Easy way to prove this is order the elements of A and for any subset $B \subset A$ consider the vector $V_B = (v_1, v_2, \ldots, v_n)$ with $v_k = 0$ if $a_k \notin B$ and $v_k = 1$ if $a_k \in B$. Obviously, we have exactly 2^n such vectors. Vector V_B corresponds to characteristic (or indicator) function χ_B defined as

$$\chi_B(x) = \begin{cases} 1 \text{ if } x \in B; \\ 0 \text{ if } x \notin B. \end{cases}$$

This definition works for general subsets, not necessarily finite.

Theorem 4. ***Theorem: Cantor:*** $\mathrm{Card}(2^{\mathrm{A}}) > \mathrm{Card}(\mathrm{A})$.

Proof: If $A = \emptyset$, then $\mathrm{Card}(\mathrm{A}) = 0$ and $\mathrm{Card}(2^{\mathrm{A}}) = 1$ as $\emptyset$ is the only subset of A.

Assume that $A \neq \emptyset$. We will show that there is no surjection $f : A \to 2^A$. Assume there is one. Let us consider the set $B = \{x \in A : x \notin f(x)\}$. Since f is a surjection there is an $x_0 \in A$ such that $f(x_0) = B$.

Then: If $x_0 \in B = f(x_0)$, then by the definition of B, $x_0 \notin f(x_0) = B$. Contradiction.

If $x_0 \notin B = f(x_0)$, then by the definition of B, $x_0 \in f(x_0) = B$. Contradiction.

This proves the theorem. QED

6. Sequences of Numbers

6.1. Absolute Value (Modulus) of a Number.

Recall that for $x \in \mathbb{R}$ we have

$$|x| = \begin{cases} x, \text{ for } x \geq 0; \\ -x, \text{ for } x < 0. \end{cases}$$

The function $x \mapsto |x|$ is called absolute value or modulus of x. We have

$$|x \pm y| \leq |x| + |y| \text{ , “triangle inequality”;}$$

$$|x - y| \geq ||x| - |y||.$$

Both are easy to prove. Note that

$$|x - A| < \varepsilon \iff A - \varepsilon < x < A + \varepsilon.$$

6.2. **Definition of a Sequence.**

A sequence of real numbers is a function $a : \mathbb{N} \to \mathbb{R}$. Usually, instead of writing $a(n)$ we write a_n, $n = 1, 2, \dots$.

A sequence can be defined explicitly, for example

$$a_n = \frac{1}{n} \quad , \quad n = 1, 2, \dots ,$$

or inductively, for example

$$a_1 = 1, a_{n+1} = \sqrt{1 + a_n} \quad , \quad n = 1, 2, \dots .$$

The last sequence is $a_1 = 1$, $a_2 = \sqrt{2}$, $a_3 = \sqrt{1 + \sqrt{2}}$, $a_4 = \sqrt{1 + \sqrt{1 + \sqrt{2}}}$, etc.

Fibonacci sequence: Let $a_1 = 1$, $a_2 = 1$, $a_n = a_{n-1} + a_{n-2}$, $n = 3, 4, 5, \dots$.

Then, the sequence is: $1, 1, 2, 3, 5, 8, 13, 21, \dots$. It is the famous Fibonacci sequence found in many features of nature and useful in many mathematical theories.

6.3. **Limit of a Sequence.**

Definition 6.1. *We say that a sequence $(a_n)_{n=1}^{\infty}$ converges to a limit L* $\iff$

$$\forall_{\,\varepsilon>0}\ \exists_{\,N\geq 1}\ \forall_{\,n\geq N}\ |a_n - L| < \varepsilon.$$

Example: Let $a_n = \frac{1+3n^2}{1+5n+2n^3}$. We will prove that $\lim_{n\to\infty} a_n = 0$. We need to prove $\forall_{\,\varepsilon>0}\ \exists_{\,N\geq 1}\ \forall_{\,n\geq N}\ |a_n| < \varepsilon$. Let us fix an $\varepsilon > 0$. We have

$$\frac{1 + 3n^2}{1 + 5n + 2n^3} < \varepsilon \iff \frac{3n^2 + 3n^2}{2n^3} < \varepsilon \iff \frac{3}{n} < \varepsilon \iff n > \frac{3}{\varepsilon}.$$

We set $N = \mathrm{Int}(\frac{3}{\varepsilon}) + 1$, where $\mathrm{Int}(t)$ denotes the integer part (floor) of the number t. For $n > N$ we have $\frac{1+3n^2}{1+5n+2n^3} < \varepsilon$ and we proved our claim.

Proposition 6.1. *The limit of a sequence is unique, i.e, if $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} a_n = B$, then $A = B$.*

Proof. Assume that $A < B$ are the limits of the sequence $(a_n)_{n=1}^{\infty}$.
Take $\varepsilon = (B - A)/3 > 0$. Then, there exist $N_A \geq 1$ such that $|a_n - A| < \varepsilon \iff A - \varepsilon < a_n < A + \varepsilon$, for $n \geq N_A$ and $B - \varepsilon < a_n < B + \varepsilon$, for $n \geq N_B$. Then, for $n \geq \max(N_A, N_B)$ we have

$$a_n < A + \varepsilon < B - \varepsilon < a_n,$$

which is impossible. The contradiction proves Proposition 6.1. □

Proposition 6.2. *A convergent sequence is bounded.*

Proof. Assume that $\lim_{n\to\infty} a_n = A$. Then, for $\varepsilon = 1$ we can find an $N \geq 1$ such that

$$A - 1 < a_n < A + 1, \quad n \geq N.$$

The set $\{a_1, a_2, \ldots, a_{N-2}, a_{N-1}\}$ is finite so it bounded between some numbers m and M. Then, we have

$$\min(A - 1, m) \leq a_n \leq \max(A + 1, M), \quad n \geq 1,$$

and $(a_n)_{n=1}^{\infty}$ is bounded. □

Proposition 6.3. *If* $\lim_{n\to\infty} a_n = A > 0$*, then the sequence* $(a_n)_{n=1}^{\infty}$ *is eventually positive, i.e.,*

$$\exists_{\, B>0} \, \exists_{\, N\geq 1} \, \forall_{\, n\geq N} \; a_n \geq B.$$

Proof. Take $B = A/2$. Since $\varepsilon = A - B = A/2 > 0$ there exists $N \geq 1$ such that for all $n \geq N$ we have

$$B = A - \varepsilon < a_n < A + \varepsilon.$$

□

Proposition 6.4. ***Arithmetic of Limits:*** *Assume that* $\lim_{n\to\infty} a_n = A$ *and* $\lim_{n\to\infty} b_n = B$*. Then,*

(a) $\lim_{n\to\infty}(a_n + b_n) = A + B$*.*

(b) $\lim_{n\to\infty}(c \cdot a_n) = c \cdot A$*.*

(c) $\lim_{n\to\infty}(a_n \cdot b_n) = A \cdot B$*.*

(d) If in addition $b_n \neq 0$*,* $n = 1, 2, \ldots$ *and* $\lim_{n\to\infty} b_n \neq 0$*, then* $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{A}{B}$*.*

Proof. We will prove (c). By assumptions $\forall_{\varepsilon>0}\ \exists_{N_A\geq 1}\ \forall_{n\geq N_A}\ |a_n - A| < \varepsilon$ and $\forall_{\varepsilon>0}\ \exists_{N_B\geq 1}\ \forall_{n\geq N_B}\ |b_n - B| < \varepsilon$. By Proposition 6.2, we know that both $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are bounded, say $|a_n| \leq M_A$ and $|b_n| \leq M_B$, for all $n \geq 1$. Let us fix an $\varepsilon > 0$ and let N_A, N_B be the constants corresponding to this ε. Then for $n \geq \max(N_A, N_B)$ we have

$$\begin{aligned}|a_n \cdot b_n - A \cdot B| &= |a_n \cdot b_n - a_n \cdot B + a_n \cdot B - A \cdot B| \\ &\leq |a_n \cdot b_n - a_n \cdot B| + |a_n \cdot B - A \cdot B| \\ &\leq |a_n||b_n - B| + |a_n - A||B| < \varepsilon(M_A + |B|).\end{aligned}$$

Since ε is arbitrarily small this proves (c). □

Proposition 6.5. *If the sequence* $(a_n)_{n=1}^{\infty}$ *is convergent,* $\lim_{n\to\infty} a_n = A$*, and* $a_n \in [a, b]$ *for all* $n \geq 1$*, then* $A \in [a, b]$*.*

Proof. Assume that $b < A$ and take $\varepsilon = (A - b)/2$. Then, there exists an $N \geq 1$ such that for all $n \geq N$ we have

$$b < A - \varepsilon < a_n,$$

which contradicts the assumption $a_n \in [a, b]$ for all $n \geq 1$. □

Proposition 6.6. ***Three Sequences Theorem or Squeeze Theorem:*** *Let us assume that* $a_n \leq c_n \leq b_n$ *for all* $n \geq 1$*, and* $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = A$*. Then,* $\lim_{n\to\infty} c_n = A$*, as well.*

Proof. Fix an $\varepsilon > 0$. By definition of the limit there are an $N_a \geq 1$ and an $N_b \geq 1$ such that $|a_n - A| < \varepsilon$ and $|b_n - A| < \varepsilon$ for $n \geq N_a$ and for $n \geq N_b$, respectively. Let $n \geq \max(N_a, N_b)$. Then, we have

$$A - \varepsilon < a_n \leq c_n \leq b_n < A + \varepsilon \implies A - \varepsilon < c_n < A + \varepsilon,$$

which shows $\lim_{n\to\infty} c_n = A$. □

6.4. **Increasing and Decreasing Sequences.**

Definition 6.2. *We say that the sequence* $(a_n)_{n=1}^{\infty}$ *is* ***increasing*** $\iff$ *we have*

$$a_n \leq a_{n+1}, \; n = 1, 2, 3, \ldots$$

We say that the sequence $(a_n)_{n=1}^{\infty}$ *is* ***decreasing*** $\iff$ *we have*

$$a_n \geq a_{n+1}, \; n = 1, 2, 3, \ldots$$

We say the sequence is ***strictly increasing*** *or* ***strictly decreasing*** *if the inequalities in the above definition are strict.*

Proposition 6.7. *(a) An increasing and bounded above sequence* $(a_n)_{n=1}^{\infty}$ *is convergent to the* $\sup\{a_n : n \geq 1\}$.

(b) A decreasing and bounded below sequence $(a_n)_{n=1}^{\infty}$ *is convergent to the* $\inf\{a_n : n \geq 1\}$.

Proof. We prove (a). Let $\alpha = \sup\{a_n : n \geq 1\}$. We need to show that

$$\forall_{\,\varepsilon>0} \; \exists_{\,N\geq 1} \; \forall_{\,n\geq N} \; |a_n - \alpha| < \varepsilon.$$

Let us fix an $\varepsilon > 0$. Since $\alpha = \sup\{a_n : n \geq 1\}$ we can find a_N such that $\alpha - \varepsilon < a_N$. Since $(a_n)_{n=1}^{\infty}$ is increasing we have $\alpha - \varepsilon < a_n$, for all $n \geq N$. On the other hand we have $a_n \leq \alpha$ for all $n \geq 1$. Thus

$$\alpha - \varepsilon < a_n \leq \alpha,$$

for all $n \geq N$. This shows (a). □

Example 6.1.

Let $a_1 = 1$ and $a_{n+1} = \sqrt{1+a_n}$ for $n \geq 1$. We will show that $(a_n)_{n=1}^{\infty}$ is increasing and bounded above and we will find its limit.

$$a_n \leq a_{n+1} \iff a_n \leq \sqrt{1+a_n} \iff a_n^2 \leq 1 + a_n.$$

Consider inequality $t^2 - t - 1 \leq 0$. The numbers $\frac{1}{2}(1-\sqrt{5})$ and $\frac{1}{2}(1+\sqrt{5})$ are the roots. Thus the inequality holds for $\frac{1}{2}(1-\sqrt{5}) \leq t \leq \frac{1}{2}(1+\sqrt{5})$. We have $0 < a_1 < \frac{1}{2}(1+\sqrt{5})$. We will show by induction that $a_n < \frac{1}{2}(1+\sqrt{5})$, for all $n \geq 1$. Let us assume that $a_n < \frac{1}{2}(1+\sqrt{5})$, for some n. Then $a_{n+1} = \sqrt{1+a_n} < \sqrt{1+\frac{1}{2}(1+\sqrt{5})}$. We have

$$\sqrt{1+\frac{1}{2}(1+\sqrt{5})} \leq \frac{1}{2}(1+\sqrt{5}) \iff 1+\frac{1}{2}(1+\sqrt{5}) \leq \left(\frac{1}{2}(1+\sqrt{5})\right)^2$$

$$\iff 4+2+2\sqrt{5} \leq 1+2\sqrt{5}+5,$$

which holds. This shows that the sequence $(a_n)_{n=1}^{\infty}$ is bounded above and at the same time that it is increasing. Thus, it converges to its supremum. Let us say that $\alpha = \sup\{a_n : n \geq 1\}$. For any $n \geq 1$ we have $a_{n+1} = \sqrt{1+a_n}$. We let $a_n \to \alpha$ and obtain

$$\alpha = \sqrt{1+\alpha}.$$

We already solved this equation. Since all a_n are positive we have

$$\lim_{n\to\infty} a_n = \frac{1}{2}(1+\sqrt{5}).$$

6.5. **Interesting Limits and Other Useful Statements.**

Proposition 6.8. ***Bernoulli Inequality:*** *If $-1 < x$, then*

$$(1+x)^n \geq 1+nx \ , \ n = 1, 2, \ldots.$$

Proof. We prove this by induction: for $n = 1$ we have $(1+x) \geq 1+x$, which obviously holds. Assume $(1+x)^n \geq 1+nx$, for some $n \geq 1$. Since $x > -1$, we have $(1+x) > 0$ and

$$(1+x)^n \cdot (1+x) \geq (1+nx)(1+x) \iff (1+x)^{n+1} \geq 1+(n+1)x+nx^2$$
$$\implies (1+x)^{n+1} \geq 1+(n+1)x.$$

By induction, the inequality holds for all $n \geq 1$. □

Proposition 6.9. ***Generalized Bernoulli Inequality:*** *If $-1 < x_k$, and all x_k's are of the same sign, $k = 1, 2, \ldots, n$, $n = 1, 2, \ldots,$ then*

$$(1+x_1)(1+x_2)\ldots(1+x_n) \geq 1+(x_1+x_2+\cdots+x_n)\ .$$

Proof. The proof is the same as above. □

Proposition 6.10. ***Newton's Binomial formula:*** *For any $a, b \in \mathbb{R}$ and any $n \geq 1$ we have*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

We are not proving it here. It can be proven by induction or by straight forward combinatorial reasoning.

Proposition 6.11. ***Limit 1:*** *Let $a > 0$. Then* $\lim_{n\to\infty} \sqrt[n]{a} = 1$.

Proof. First, let us assume that $a > 1$. We can write $\sqrt[n]{a} = 1+\varepsilon_n$. We have

$$a = (1+\varepsilon_n)^n \geq 1+n\varepsilon_n,$$

by Bernoulli inequality. This implies that

$$\varepsilon_n \leq \frac{a-1}{n} \underset{n\to\infty}{\longrightarrow} 0.$$

This proves our claim for $a > 1$.

Let $a < 1$. Then $1/a > 1$ and we have

$$\lim_{n\to\infty} \sqrt[n]{a} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{\frac{1}{a}}} = \frac{1}{1} = 1.$$

□

Proposition 6.12. ***Limit 2:*** $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Proof. Similarly as before we write $\sqrt[n]{n} = 1 + \varepsilon_n$. Then,

$$n = (1+\varepsilon_n)^n = \sum_{k=0}^{n} \binom{n}{k} \varepsilon_n^k 1^{n-k} \geq \binom{n}{0} \cdot 1 + \binom{n}{1} \varepsilon_n^1 + \binom{n}{2} \varepsilon_n^2 \geq \binom{n}{2} \varepsilon_n^2 = \frac{n!}{2 \cdot (n-2)!} \varepsilon_n^2.$$

This implies that

$$\varepsilon_n^2 \leq \frac{2}{n-1} \underset{n\to\infty}{\longrightarrow} 0,$$

which proves the claim. □

Proposition 6.13. ***Limit 3:*** $\lim_{n\to\infty} \sqrt[n]{6^n + 3^n} = 6$.

Proof. We will use the Three Sequences Theorem: we have

$$6 = \sqrt[n]{6^n} \leq \sqrt[n]{6^n + 3^n} \leq \sqrt[n]{6^n + 6^n} = \sqrt[n]{2 \cdot 6^n} = \sqrt[n]{2}\,\sqrt[n]{6^n} \underset{n\to\infty}{\longrightarrow} 1 \cdot 6.$$

Invoking the Three Sequences Theorem completes the proof. □

Proposition 6.14. ***Limit 4:*** $\lim_{n\to\infty} \frac{1}{n} \ln(6^n - 3^n) = ln 6$.

Proof. We will again use the Three Sequences Theorem: we have

$$\frac{1}{2} 6^n = 6^n - \frac{1}{2} 6^n \leq 6^n - 3^n \leq 6^n,$$

since $\frac{1}{2} 6^n \geq 3^n$ for all $n \geq 1$ (easy to prove). Then,

$$\frac{1}{n} \ln(\frac{1}{2} \cdot 6^n) \leq \frac{1}{n} \ln(6^n - 3^n) \leq \frac{1}{n} \ln(6^n) = \ln 6.$$

We have

$$\frac{1}{n} \ln(\frac{1}{2} \cdot 6^n) = \frac{1}{n} \ln(\frac{1}{2}) + \frac{1}{n} \ln(6^n) \underset{n\to\infty}{\longrightarrow} 0 + \ln 6.$$

Invoking the Three Sequences Theorem completes the proof. □

Proposition 6.15. ***Limit 5:*** $\lim_{n\to\infty} \sqrt{n^2+n} - n = \frac{1}{2}$.

Proof. Using $(a+b)(a-b) = a^2 - b^2$ we obtain

$$\sqrt{n^2+n} - n = \frac{n^2+n-n^2}{\sqrt{n^2+n}+n} = \frac{n}{\sqrt{n^2+n}+n} = \frac{1}{\sqrt{1+1/n}+1} \underset{n\to\infty}{\longrightarrow} \frac{1}{2}.$$

□

Limit 6:

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

Proof. Using the Nested Intervals Lemma we proved that

$$\left(1+\frac{1}{n}\right)^n \le e \le \left(1+\frac{1}{n}\right)^{n+1}, \tag{7}$$

and that the sequences on both sides converge to e. We use this to write

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

□

Example 6.2. ***Limit 7:*** $\lim_{n\to\infty} \frac{e^{1/n}-1}{1/n} = 1$.

Proof. We use inequalities from the previous example to obtain:

$$\left(1+\frac{1}{n}\right) \le e^{\frac{1}{n}} \le \left(1+\frac{1}{n-1}\right) \iff \frac{1}{n} \le e^{\frac{1}{n}} - 1 \le \frac{1}{n-1} \iff \frac{\frac{1}{n}}{\frac{1}{n}} \le \frac{e^{\frac{1}{n}}-1}{\frac{1}{n}} \le \frac{\frac{1}{n-1}}{\frac{1}{n}}.$$

The Three Sequences Theorem completes the proof. □

Limit 7A: Consider the sequences

$$a_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n$$

and

$$b_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-2} + \frac{1}{n-1} - \ln n.$$

$(a_n)_{n\geq 1}$ is strictly decreasing and bounded below so convergent. $(b_n)_{n\geq 1}$ is strictly increasing and bounded above so convergent. The common limit is called Euler's constant and denoted by γ. We have $\gamma \sim 0.577$.

Proof. From double inequality (7) we obtain

$$\frac{1}{n+1} < \ln\left(1+\frac{1}{n}\right) < \frac{1}{n}. \tag{8}$$

To have $a_n > a_{n+1}$ we need

$$-\ln n > \frac{1}{n+1} - \ln(n+1) \text{ or } \ln(1+\frac{1}{n}) > \frac{1}{n+1}.$$

To have $b_n < b_{n+1}$ we need

$$-\ln n < \frac{1}{n} - \ln(n+1) \text{ or } \ln(1+\frac{1}{n}) < \frac{1}{n}.$$

Both hold by inequality (8). We have $a_n - b_n = \frac{1}{n}$, so both sequences converge to a common limit $b_n < \gamma < a_n$. For $n = 5000$ we obtain $0.577315662 < \gamma < 0.577115662$. □

Proposition 6.16. ***Limit 7B:*** *In Proposition 4.1 we proved that*

$$\left(1+\frac{1}{n}\right)^n \to e, \; n \to \infty.$$

Similar proof shows that

$$\left(1+\frac{x}{n}\right)^n \to e^x, \; n \to \infty, \; x \in \mathbb{R}.$$

Now, we will prove that the sequence

$$a_n = 1 + \frac{x}{1} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!},$$

also converges to e.

Proof. By Newton's Binomial formula we have

$$\left(1+\frac{x}{n}\right)^n = \sum_{k=0}^{n} \frac{n!}{(n-k)!n^k}\frac{x^k}{k!} = \sum_{k=0}^{n} \frac{(n-k+1)}{n}\frac{(n-k+2)}{n}\frac{(n-k+k-1)}{n}\frac{x^k}{k!}$$

$$= \sum_{k=0}^{n}\left(1-\frac{k-1}{n}\right)\left(1-\frac{k-2}{n}\right)\cdots\left(1-\frac{1}{n}\right)\frac{x^k}{k!}$$

This shows that $|(1+x/n)|^n < |a_n|$ and using Proposition 6.9 we have

$$\left|a_n - \left(1+\frac{x}{n}\right)^n\right| \le \sum_{k=0}^{n}\left(\frac{1}{n}\sum_{j=1}^{k-1} j\right)\frac{|x|^k}{k!} = \frac{1}{n}\sum_{k=0}^{n}\frac{k(k-1)}{2}\frac{|x|^k}{k!} \underset{n\to\infty}{\longrightarrow} 0,$$

as the sequence

$$\sum_{k=0}^{n}\frac{k(k-1)}{2}\frac{|x|^k}{k!}$$

is bounded above.

□

Proposition 6.17. ***Limit 8:*** *If* $\lim_{n\to\infty} a_n = 0$*, then*

$$\lim_{n\to\infty}\frac{\sin a_n}{a_n} = 1.$$

Proof. We proved that $\sin t \le t \le \tan t$. This implies

$$\cos t \le \frac{\sin t}{t} \le 1,$$

so for our sequence we obtain

$$\cos a_n \le \frac{\sin a_n}{a_n} \le 1.$$

Since $\lim_{a_n\to 0}\cos a_n = 1$, the Three Sequences Theorem again completes the proof.

□

Limit 9: Let $a > 1$. Then, for any $k \ge 1$ we have

$$\lim_{n\to\infty}\frac{n^k}{a^n} = 0.$$

This says that a geometric growth is much faster than any polynomial growth.

Proof. Fix $k \ge 1$. We can write $a = 1 + b$ with $b > 0$. Then, for $n > k+1$,

$$a^n = (1+b)^n = \sum_{j=0}^{n}\binom{n}{j}b^j \ge \binom{n}{k+1}b^{k+1}.$$

Since

$$\binom{n}{k+1} = \frac{n!}{(k+1)!(n-k-1)!} = \frac{1}{(k+1)!}n(n-1)(n-2)\cdots(n-k),$$

for $n > k+1$ we have

$$a^n \geq \frac{1}{(k+1)!}[n^{k+1} + c_k n^k + c_{k-1}n^{k-1} + \cdots + c_1 x + c_0],$$

where $c_k, c_{k-1}, \ldots, c_1, c_0 \in \mathbb{R}$ are fixed. We have

$$\begin{aligned}\frac{n^k}{a^n} &\leq \frac{n^k}{\frac{1}{(k+1)!}[n^{k+1} + c_k n^k + c_{k-1}n^{k-1} + \cdots + c_1 x + c_0]} \\ &= \frac{1}{n} \cdot \frac{(k+1)!}{1 + c_k \frac{1}{n} + c_{k-1}\frac{1}{n^2} + \cdots + c_1 \frac{1}{n^k} + c_0 \frac{1}{n^{k+1}}} \underset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

□

Limit 10:

$$\lim_{n\to\infty} \frac{\ln n}{n} = 0.$$

Proof. The simplest way to show this is

$$\frac{\ln n}{n} = \ln(n^{1/n}) \underset{n\to\infty}{\longrightarrow} 0,$$

since we know that $n^{1/n} \underset{n\to\infty}{\longrightarrow} 1$. This method uses some knowledge we do not have yet, like continuity of the ln function. We will present another proof.

For any sufficiently large $n \in \mathbb{N}$ we can find $k_n \in \mathbb{N}$, such that $e^{k_n} \leq n \leq e^{k_n+1}$.

Then, we can write

$$\frac{\ln e^{k_n}}{e^{k_n+1}} \leq \frac{\ln n}{n} \leq \frac{\ln e^{k_n+1}}{e^{k_n}},$$

or

$$\frac{k_n}{e^{k_n+1}} \leq \frac{\ln n}{n} \leq \frac{k_n+1}{e^{k_n}}.$$

When $n \to +\infty$, $k_n \to +\infty$ as well. By Limit 9, we know that as $k_n \to +\infty$ both fractions on the sides converge to 0. This shows our claim. □

Corollary 6.1.

$$\lim_{n\to\infty} \frac{\ln n}{n^a} = 0, \;\; a > 0.$$

In particular

$$\lim_{n\to\infty} \frac{\ln n}{\sqrt{n}} = 0\ .$$

Limit 11:

$$\lim_{n\to\infty} \frac{\sin(e^n + 999^{\sqrt{n}})}{n} = 0.$$

Proof. Since $-1 \leq \sin t \leq 1$, we have

$$\frac{-1}{n} \leq \frac{\sin(e^n + 999^{\sqrt{n}})}{n} \leq \frac{1}{n}.$$

Invoking the Three Sequences Theorem completes the proof.

□

6.6. **Infinite Limits.**

Definition 6.3. *We say that* $\lim_{n\to\infty} a_n = +\infty \iff$

$$\forall_{M\in\mathbb{R}} \, \exists_{N\geq 1} \, \forall_{n\geq N} \, M < a_n.$$

We say that $\lim_{n\to\infty} a_n = -\infty \iff$

$$\forall_{M\in\mathbb{R}} \, \exists_{N\geq 1} \, \forall_{n\geq N} \, a_n < M.$$

It seems that the definition of a limit equal $\pm\infty$ is different from the definition of a finite limit. We can unify both definitions using a bit more advanced language.

Definition 6.4. ***Neighbourhood:***

If $A \in \mathbb{R}$, we say that any open interval $(A-\varepsilon, A+\varepsilon)$, $\varepsilon > 0$ is a neighbourhhood of A.

If $A = +\infty$, we say that any open interval $(M, +\infty)$, $M \in \mathbb{R}$ is a neighbourhhood of A.

If $A = -\infty$, we say that any open interval $(-\infty, M)$, $M \in \mathbb{R}$ is a neighbourhhood of A.

The unified definition says:

Definition 6.5. *We say that* $\lim_{n\to\infty} a_n = A$, $A \in \mathbb{R}$ *or* $A = \pm\infty \iff$ *for any neighbourhood U of A the sequence $(a_n)_{n=1}^{\infty}$ is eventually contained in U, i.e., there exists $N \geq 1$ such that for all $n \geq N$ we have $a_n \in U$.*

Example: Show that $\lim_{n\to\infty} \sqrt[n]{n!} = +\infty$. It is enough to show this for even n, i.e., to show $\lim_{n\to\infty} \sqrt[2n]{(2n)!} = +\infty$. Since

$$\underbrace{n \cdot n \cdots n \cdot n}_{n-\text{times}} \leq 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n \cdot (n+1) \cdots (2n-1) \cdot (2n),$$

we have

$$\sqrt{n} = \sqrt[2n]{n^n} \leq \sqrt[n]{(2n)!}.$$

Let $M \in \mathbb{R}$. Then, there we can find an $N > M^2$ (otherwise the set $\mathbb{N}$ would be bounded). Then, for $n \geq N$ we have $a_n = \sqrt[2n]{(2n)!} > \sqrt{n} > M$ which proves our claim.

Now, to justify doing this only for even n, we show:

$$\sqrt[2n]{(2n)!} \leq \sqrt[2n+1]{(2n+1)!} \iff ((2n)!)^{2n+1} \leq ((2n+1)!)^{2n}$$

$$\iff ((2n)!)^{2n}(2n)! \leq ((2n)!)^{2n}(2n+1)^{2n} \iff (2n)! \leq (2n+1)^{2n},$$

which holds.

6.7. **Subsequences.**

Definition 6.6. *Let $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$ be two sequences of real numbers. Then, $(y_n)_{n\geq 1}$ is a subsequence of $(x_n)_{n\geq 1}$ $\iff$ there is a* ***strictly increasing*** *function $h : \mathbb{N} \to \mathbb{N}$, such that $n < m \implies h(n) < h(m)$ and $y_n = x_{h(n)}$, for all $n \in \mathbb{N}$.*

In short $y_k = x_{n_k}$ with $k_1 < k_2 \implies n_{k_1} < n_{k_2}$.

For example the sequence $x_1, x_2, x_3, x_3, x_4, x_5, \ldots$ is not a subsequence of $x_1, x_2, x_3, x_4, x_5, x_6, \ldots$. Also, $x_1, x_2, x_4, x_3, x_5, \ldots$ is not a subsequence of $x_1, x_2, x_3, x_4, x_5, x_6, \ldots$.

Proposition 6.18. *If $x_n \underset{n\to\infty}{\longrightarrow} x$, then any subsequence of $(x_n)_{n\geq 1}$ also converges to x.*

Proof. Let us consider a subsequence $(x_{n_k})_{k\geq 1}$ of $(x_n)_{n\geq 1}$. Fix an $\varepsilon > 0$. We need to show that there exists a $K \geq 1$ such that for any $k \geq K$ we have $|x_{n_k} - x| < \varepsilon$. We know $\exists_{N\geq 1} \forall_{n\geq N} |x_n - x| < \varepsilon$. Since $n_k \underset{k\to\infty}{\longrightarrow} +\infty$ (it is a strictly increasing unbounded above sequence), we can find a $K \geq 1$ such that $N \leq n_K$. Since the function $k \mapsto n_k$ is strictly increasing, we have $N \leq n_k$ for all $k \geq K$. Thus, $|x_{n_k} - x| < \varepsilon$ for $k \geq K$. $\square$

Proposition 6.19. *$x_n \not\to x \iff \exists_{\,\varepsilon>0}$ and a subsequence $(x_{n_k})_{k\geq 1}$ of $(x_n)_{n\geq 1}$ such that $|x_{n_k} - x| \geq \varepsilon$ for all $k \geq 1$.*

Proof. The definition of the limit

$$\forall_{\,\varepsilon>0}\ \exists_{\,N\geq 1}\ \forall_{\,n\geq N}\ |x_n - x| < \varepsilon.$$

The negation is

$$(*)\quad \exists_{\,\varepsilon>0}\ \forall_{\,N\geq 1}\ \exists_{\,n\geq N}\ |x_n - x| \geq \varepsilon.$$

The statement $(*)$ gives us an $\varepsilon > 0$. For this ε we will construct the consecutive n_k, $k = 1, 2, 3, \ldots$

Let $N_1 = 1$. By $(*)$ there exists $n_1 \geq N_1 = 1$ with $|x_{n_1} - x| \geq \varepsilon$.

Let $N_2 = n_1 + 1$. By $(*)$ there exists $n_2 \geq N_2 = n_1 + 1 > n_1$ with $|x_{n_2} - x| \geq \varepsilon$.

Let $N_3 = n_2 + 1$. By $(*)$ there exists $n_3 \geq N_3 = n_2 + 1 > n_2$ with $|x_{n_3} - x| \geq \varepsilon$.

In this way we construct the indices n_k, $k = 1, 2, 3, \ldots$ By construction the sequence $(n_k)_{k\geq 1}$ is strictly increasing so $(x_{n_k})_{k\geq 1}$ is a subsequence of $(x_n)_{n\geq 1}$. □

Theorem 5. ***Sunset Theorem:*** *Every sequence contains a monotonic subsequence.*

Proof. Let $(x_n)_{n\geq 1}$ be a sequence of real numbers.

Element x_m is called a peak $\iff \forall_{\,n>m}\ x_m > x_n$. This is illustrated in Figure 8. Imagine a mountain range with the heights of consecutive mountains equal to x_n. The peaks of the sequence are the points from which one can see the sunset.

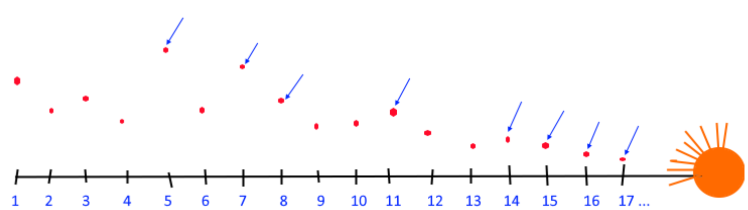

FIGURE 8. Peaks are pointed out by arrows.

There are two possibilities:

(i) The number of peaks is finite. Then, for any m there exists $n > m$ such that $x_n \geq x_m$.

Let x_{n_1} be the last peak. There exists $n_2 > n_1$ with $x_{n_2} \geq x_{n_1}$;

Then, there exists $n_3 > n_2$ with $x_{n_3} \geq x_{n_2}$, and so on.

The constructed subsequence $(x_{n_k})_{k\geq 1}$ is increasing.

(ii) The number of peaks is infinite. Then, the peaks with increasing indices form a decreasing subsequence since for the peaks x_{n_k} and $x_{n_{k+1}}$ with $n_k \leq n_{k+1}$ we have $x_{n_k} > x_{n_{k+1}}$.

□

Theorem 6. ***Bolzano-Weierstass Theorem:*** *A bounded sequence in $\mathbb{R}$ contains a convergent subsequence.*

Proof. Let $(x_n)_{n\geq 1}$ be a bounded sequence of real numbers. By Theorem 5 it contains a monotonic subsequence. By Proposition 6.7 this subsequence is convergent. □

Example 6.3.

Let the sequence $(a_n)_{n\geq 1}$ be unbounded above. Prove that it contains an increasing subsequence divergent to $+\infty$.

If $(a_n)_{n\geq 1}$ is bounded above, then $\exists_{\, M\in\mathbb{R}}\ \forall\, n \in \mathbb{N}\ a_n \leq M$. Thus, if $(a_n)_{n\geq 1}$ be unbounded above, we have

$$(*)\quad \forall_{\, M\in\mathbb{R}}\ \exists\, n \in \mathbb{N}\ \ a_n > M.$$

Let $a_{n_1} = a_1$.

Using $(*)$ with $M = \max\{a_{n_1}, 2\}$ we can find $n_2 > n_1$ such that $a_{n_1} < a_{n_2}$ and $2 < a_{n_2}$.

Using $(*)$ with $M = \max\{a_{n_2}, 3\}$ we can find $n_3 > n_2$ such that $a_{n_2} < a_{n_3}$ and $3 < a_{n_3}$.

And we proceed by induction. Assume that we have a_{n_k} such that $a_{n_{k-1}} < a_{n_k}$ and $k < a_{n_k}$. Using $(*)$ with $M = \max\{a_{n_k}, k+1\}$ we can find $n_{k+1} > n_k$ such that $a_{n_k} < a_{n_{k+1}}$ and $k+1 < a_{n_{k+1}}$.

This construction produces a subsequence which is increasing ($a_{n_k} < a_{n_{k+1}}$ for each k) and convergent to $+\infty$ ($k < a_{n_k}$ for each k).

6.8. **Cauchy Sequences.**

Definition: A sequence $\{x_n\}_{n=1}^{\infty}$ is Cauchy if and only if

$$\forall_{\varepsilon>0} \exists_{N\geq 1} \forall_{n,m\geq N} |x_n - x_m| < \varepsilon.$$

Example: Let

$$x_n = \sum_{k=0}^{n} \frac{1}{2^k}.$$

We will show it is Cauchy. Let $N \leq n < m$. Then,

$$x_n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{2^n}\ ,\ x_m = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{2^n} + \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^m},$$

and

$$|x_n - x_m| = \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^m} \leq \frac{1}{2^{n+1}}(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots) = \frac{1}{2^{n+1}}\frac{1}{1-\frac{1}{2}} = \frac{1}{2^n} \leq \frac{1}{2^N}.$$

Since $\frac{1}{2^N}$ can be made smaller than any positive ε the sequence is a Cauchy sequence.
A few theorems:

Theorem 7. ***Theorem CS1:*** *A convergent sequence is Cauchy.*

Proof. Assume that $x_n \to L$, as $n \to \infty$. Fix an *eps* > 0. For this ε we can find an $N \geq 1$ such that for any $n \geq N$ we have

$$|x_n - L| < \varepsilon/2.$$

Take $n, m \geq N$. Then

$$|x_n - x_m| \leq |x_n - L| + |L - x_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

which shows that $\{x_n\}$ is a Cauchy sequence. □

Theorem 8. ***Theorem CS2:*** *A Cauchy sequence is bounded.*

Proof. Let $\varepsilon = 1$ and $\{x_n\}$ be a Cauchy sequence. There exists an $N \geq 1$ such that for any $n, m \geq N$ we have $|x_n - x_m| < 1$. In particular, for any $n \geq N$ we have $|x_n - x_N| < 1$. This means that all elements $x_N, x_{N+1}, x_{N+2}, \ldots$ are in the interval $I_1 = [x_N - 1, x_N + 1]$. The remaining elements $x_1, x_2, x_3, \ldots, x_{N-1}$ are a finite number of elements so they fit into some interval $I_2 = [-M, M]$. We can find an interval $[-K, K]$ such that both $I_1, I_2 \subset [-K, K]$. Then the sequence $\{x_n\}$ is bounded (in modulus) by K. ☐

Theorem 9. ***Theorem CS3:*** *A Cauchy sequence containing a convergent subsequence is convergent (to the same limit).*

Proof. Let $\{x_n\}$ be a Cauchy sequence and its subsequence $x_{n_k} \to L$ as $k \to \infty$. Fix an $\varepsilon > 0$. We can find a $K \geq 1$ such that for any $k \geq K$ we have

$$|x_{n_k} - L| < \varepsilon/2.$$

We can also find an $N \geq 1$ such that for any $n, m \geq N$ we have

$$|x_n - x_m| < \varepsilon/2.$$

We can find $k_0 \geq K$ such that $n_{k_0} \geq N$. Let $m \geq N$. The, we have

$$|x_m - L| \leq |x_{n_{k_0}} - L| + |x_{n_{k_0}} - x_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

This shows that $x_m \to L$ as $m \to \infty$. ☐

Theorem 10. ***Theorem CS4:*** *A Cauchy sequence in $\mathbb{R}$ is convergent.*

Proof. Let $\{x_n\}$ be a Cauchy sequence. By Th. CS2 sequence $\{x_n\}$ is bounded. By Bolzano-Weierstrass Th. $\{x_n\}$ contains a convergent subsequence. By Th. CS3 the sequence $\{x_n\}$ is convergent. ☐

Example: The theorem CS4 depends on the space in which the sequence is considered. Let consider a space $X = (0, 1]$. The sequence $x_n = 1/n$, $n = 1, 2, \dots$ is Cauchy. For any $n, m \geq N$ we have $|1/n - 1/m| \leq 1/N$ so it is Cauchy as $1/N$ can be made smaller than any positive epsilon. The sequence is convergent in $\mathbb{R}$ to 0, but in the space X the 0 does not exist so the sequence is not convergent.

Contraction maps: A map (function) $f : \mathbb{R} \to \mathbb{R}$ is called a contraction if

$$|f(x) - f(y)| \leq \alpha|x - y| \ , \ 0 \leq \alpha < 1 \ , \ \forall \ x, y \in \mathbb{R}.$$

Theorem 11. ***(Banach Contraction Principle):*** *An contraction map $f : \mathbb{R} \to \mathbb{R}$ has a unique fixed point, i.e., there exist unique point $p \in \mathbb{R}$ with $f(p) = p$.*

Proof. Let $x_0 \in \mathbb{R}$. we consider the sequence x_0, $x_1 = f(x_0)$, $x_2 = f(x_1)$, $x_3 = f(x_2)$, and in general $x_{n+1} = f(x_n)$, for all n. We will show that $\{x_n\}$ is a Cauchy sequence.

We have

$$|x_2 - x_1| = |f(x_1) - f(x_0)| \leq \alpha|x_1 - x_0|.$$

$$|x_3 - x_2| = |f(x_2) - f(x_1)| \leq \alpha|x_2 - x_1| \leq \alpha^2|x_1 - x_0|.$$

$$|x_4 - x_3| = |f(x_3) - f(x_2)| \leq \alpha|x_3 - x_2| \leq \alpha^3|x_1 - x_0|.$$

In general

$$|x_n - x_{n-1}| = |f(x_{n-1}) - f(x_{n-2})| \leq \alpha|x_{n-1} - x_{n-2}| \leq \alpha^{n-1}|x_1 - x_0|.$$

(Can proved by induction if three dots do not convince You.)

Now, we will estimate $|x_n - x_m|$ for $n, m \geq N$ and $n < m$.

$$|x_n - x_m| \leq |x_n - x_{n+1}| + |x_{n+1} - x_{n+2}| + |x_{n+2} - x_{n+3}|$$
$$+ \cdots + |x_{m-3} - x_{m-2}| + |x_{m-2} - x_{m-1}| + |x_{m-1} - x_m|$$
$$\leq (\alpha^n + \alpha^{n+1} + \alpha^{n+2} + \cdots + \alpha^{m-3} + \alpha^{m-2} + \alpha^{m-1})|x_1 - x_0|$$
$$\leq |x_1 - x_0|\alpha^n \frac{1}{1 - \alpha} \leq |x_1 - x_0|\alpha^N \frac{1}{1 - \alpha}.$$

Since $|x_1 - x_0|\frac{1}{1-\alpha}$ is a fixed constant and α^N can be made arbitrarily small, we proved that $\{x_n\}$ is a Cauchy sequence. By Th. CS4 $\{x_n\}$ is convergent, say $x_n \to p$ as $n \to \infty$. We also have

$$|f(x_n) - f(p)| \leq \alpha |x_n - p|,$$

so $f(x_n) \to f(p)$ as as $n \to \infty$.

We write $x_{n+1} = f(x_n)$ and go to the limit with n. We obtain $p = f(p)$, so we proved the existence of the fixed point.

Uniqueness: Assume there are two fixed points $p_1 = f(p_1)$ and $p_2 = f(p_2)$. Then,

$$|p_1 - p_2| = |f(p_1) - f(p_2)| \leq \alpha |p_1 - p_2|.$$

Since $\alpha < 1$ this implies $p_1 = p_2$.

□

Corollary: In the proof we obtained:

$$|x_n - x_m| \leq |x_1 - x_0| \frac{\alpha^n}{1-\alpha}.$$

Going with m to infinity we obtain

$$|x_n - p| \leq |x_1 - x_0| \frac{\alpha^n}{1-\alpha},$$

which is an useful estimate on the error of approximation to p.

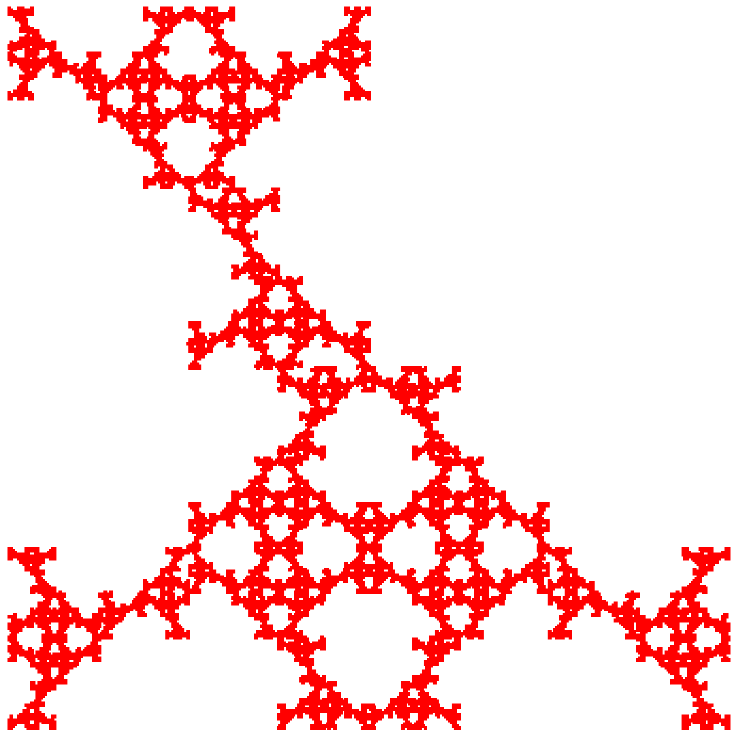

FIGURE 9. Fixed point of a contraction on the space of sets.

Newton's Algorithm for solving equations:

To solve an equation $f(x) = 0$ we can use Newton's method. Starting with any point x_0, hopefully close to the solution, we iterate

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

This iteration, under some assumptions, gives a solution to the equation. For different starting points we may obtain different solutions.

We will relate this method to the Banach contraction principle. Consider the function

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

If $\bar{x}$ is a zero of f, ($f(\bar{x}) = 0$), then $\bar{x}$ is the fixed point of F, i.e., $F(\bar{x}) = \bar{x}$. We have

$$F'(x) = 1 - \frac{f'(x)f'(x) - f''(x)f(x)}{f'(x)f'(x)}.$$

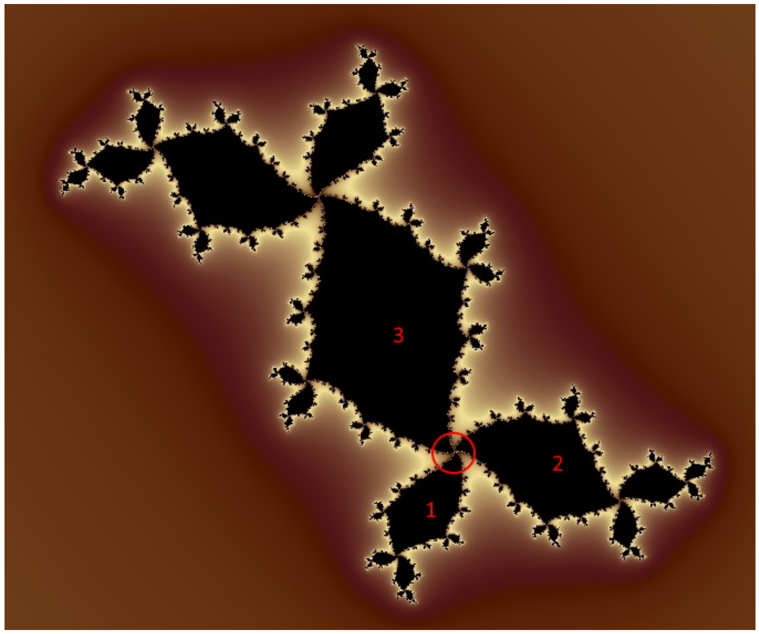

FIGURE 10. Fixed point of another contraction on the space of sets.

If $f(\bar{x}) = 0$ then $F'(\bar{x}) = 0$, which means that F is the strong contraction in a neighbourhood of the point $\bar{x}$. This explains the convergence of Newton's algorithm. This also shows that if there are multiple roots of the equation, we need to start iteration close to the root to find it.

In Figure 11 we see the regions of starting points which give different solutions to the equation $z^4 = 1$. Solution $z = 1$ color yellow, $z = -1$ color blue, $z = i$ color red and $z = -i$ color green. The points on the boundaries between the different "basins" do not give any solutions.

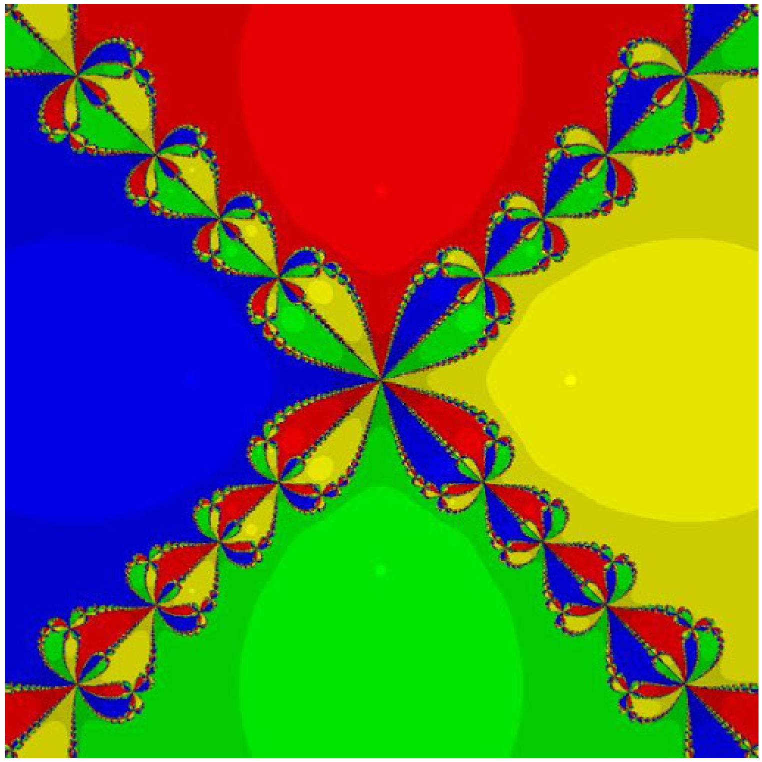

FIGURE 11. Attracting regions for Newton's algorithm for equation $z^4 = 1$.

6.9. **Partial Limits, Limit Superior, Limit Inferior.**

Definition 6.7. ***Partial Limits, Limit Points*** *If $(x_{n_k})_{k\geq 1}$ is a convergent subsequence of $(x_n)_{n\geq 1}$, $\lim_{k\to\infty} x_{n_k} = x$, then the point x is called a* ***partial limit*** *or a* ***limit point*** *of the sequence $(x_n)_{n\geq 1}$.*

Example 6.4. Let $x_n = (-1)^n$, $n = 1, 2, \ldots$. Then, the points $1, -1$ are the partial limits of the sequence as $\lim_{k\to\infty} x_{2k} = 1$ and $\lim_{k\to\infty} x_{2k+1} = -1$. The sequence $(x_n)_{n\geq 1}$ does not have any other limit points. Note, the sequence with more than one limit point is not convergent.

Definition 6.8. ***Limsup, Liminf*** *Let* $(x_n)_{n\geq 1}$ *be a sequence of real numbers and let* $P = \{a : a \text{ is a partial limit of } (x_n)_{n\geq 1}\}$. *Then,*

$$\limsup_{n\to\infty} x_n = \sup P,$$

$$\liminf_{n\to\infty} x_n = \inf P,$$

i.e., $\limsup_{n\to\infty} x_n$ *is the supremum of all limit points of* $(x_n)_{n\geq 1}$ *and* $\liminf_{n\to\infty} x_n$ *is the infimum of all limit points of* $(x_n)_{n\geq 1}$.

Why is the set P non-empty for any sequence $(x_n)_{n\geq 1}$?

Example 6.5. For the sequence $(x_n)_{n\geq 1}$ of Example 6.4 we have $\limsup_{n\to\infty} x_n = 1$, and $\liminf_{n\to\infty} x_n = -1$.

Example 6.6. For the sequence $x_n = \sin(n\pi/2)$, $n \geq 1$ we have $P = \{-1, 0, 1\}$ and again $\limsup_{n\to\infty} x_n = 1$, $\liminf_{n\to\infty} x_n = -1$.

Example 6.7. Let the sequence $(x_n)_{n\geq 1}$ be formed of all positive rational numbers. Since the rational numbers are dense in $\mathbb{R}$ we have $P = [0, +\infty)$ and $\limsup_{n\to\infty} x_n = +\infty$, $\liminf_{n\to\infty} x_n = 0$.

Proposition 6.20. *Let* $(x_n)_{n\geq 1}$ *be a sequence of real numbers. Then,*

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

Proof. $\Longrightarrow$: By Proposition 6.18 we know that any subsequence of $(x_n)_{n\geq 1}$ converges to a. Thus, $P = \{a\}$ and our claim follows.

$\Longleftarrow$: If $\lim_{n\to\infty} x_n \neq a$ then By Proposition 6.19 we know that there is a subsequence $(x_{n_k})_{k\geq 1}$ of $(x_n)_{n\geq 1}$ which stays away from a. The $\limsup_{k\to\infty} x_{n_k}$ and $\liminf_{k\to\infty} x_{n_k}$ are the elements of P and at least one of them is different from a. This contradicts the assumption that P consists of one point. $\square$

Theorem 12. *For a sequence $(x_n)_{n\geq 1}$ there exist subsequences $(x_{n_k})_{k\geq 1}$ and $(x_{n_\ell})_{\ell\geq 1}$ such that*

$$\lim_{k\to\infty} x_{n_k} = \limsup_{n\to\infty} x_n \quad , \quad \lim_{\ell\to\infty} x_{n_\ell} = \liminf_{n\to\infty} x_n \, .$$

Proof. Let $\alpha = \limsup_{n\to\infty} x_n$. We want to show that there exists a subsequence $\{x_{n_k}\}$ such that $\lim_{k\to\infty} x_{n_k} = \alpha$. We will consider finite and infinite cases separately.

(1) α is finite. We want to show that

$$\forall_{\,\varepsilon>0} \; \forall_{\,N\geq 1} \; \exists_{\,n\geq N} \; |x_n - \alpha| < \varepsilon.$$

Let us assume that this does not hold, i.e.,

$$(*) \qquad \exists_{\,\varepsilon>0} \; \exists_{\,N\geq 1} \; \forall_{\,n\geq N} \; |x_n - \alpha| \geq \varepsilon.$$

Let us fix such an $\varepsilon_0 > 0$ and the corresponding $N_0 \geq 1$. Since $\alpha = \sup A$, where A is the set of all partial limits of (x_n), we can find a partial limit a such that $\alpha - \varepsilon_0/3 < a \leq \alpha$. There is a subsequence $\{x_{n_\ell}\}$ converging to a, i.e., satisfying

$$\forall_{\,\varepsilon>0} \; \exists_{\,M\geq 1} \; \exists_{\,\ell\geq M} \; |x_{n_\ell} - a| < \varepsilon/3.$$

In particular, there exists an element x_{n_ℓ} with $n_\ell > N_0$ satisfying $|x_{n_\ell} - a| < \varepsilon_0/3$. Then,

$$|x_{n_\ell} - \alpha| < |x_{n_\ell} - a| + |a - \alpha| < \varepsilon_0/3 + \varepsilon_0/3 < \varepsilon_0,$$

which contradicts $(*)$.

(2) $\alpha = +\infty$. The idea of the proof is exactly the same but formally it looks different.

We want to show that

$$\forall_{\,K>0} \; \forall_{\,N\geq 1} \; \exists_{\,n\geq N} \; x_n > K.$$

Let us assume that this does not hold, i.e.,

$$(**) \qquad \exists_{\,K>0} \; \exists_{\,N\geq 1} \; \forall_{\,n\geq N} \; x_n \leq K.$$

Let us fix such an $K_0 > 0$ and the corresponding $N_0 \geq 1$. Since $+\infty = \sup A$, where A is the set of all partial limits of (x_n), we can find a partial limit a (assume that it is finite, if not there is nothing to prove) such that $a > 3K_0$. There is a subsequence $\{x_{n_\ell}\}$ converging to a, i.e., satisfying

$$\forall_{\varepsilon>0} \exists_{M\geq 1} \exists_{\ell\geq M} |x_{n_\ell} - a| < \varepsilon.$$

In particular, for $\varepsilon = K_0$, there exists an element x_{n_ℓ} with $n_\ell > N_0$ satisfying $a - K_0 < x_{n_\ell}$. Then,

$$x_{n_\ell} > a - K_0 > 3K_0 - K_0 > K_0,$$

which contradicts $(**)$.

□

Now, we will present another proof of Theorem 12. It is more intuitive but heavier on notation. We will do only case (1) $\alpha = \limsup x_n$ is finite.

Proof. We have $\alpha = \sup P$, where P is the set of partial limits of the sequence $(x_n)_{n\geq 1}$. If P has the largest element $a = \alpha$, then the subsequence convergent to a converges to α as well. If Then, there exists a sequence of partial limits

□

Proposition 6.21. *Let $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$ be sequences of real numbers.*

(a) We assume that if $\limsup_{n\to\infty} x_n = +\infty$, then $\limsup_{n\to\infty} y_n \neq -\infty$ and if $\limsup_{n\to\infty} x_n = -\infty$, then $\limsup_{n\to\infty} y_n \neq +\infty$. Then,

$$\limsup_{n\to\infty}(x_n + y_n) \leq \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n.$$

(b) We assume that if $\liminf_{n\to\infty} x_n = +\infty$, then $\liminf_{n\to\infty} y_n \neq -\infty$ and if $\liminf_{n\to\infty} x_n = -\infty$, then $\liminf_{n\to\infty} y_n \neq +\infty$. Then,

$$\liminf_{n\to\infty} x_n + \liminf_{n\to\infty} y_n \leq \liminf_{n\to\infty}(x_n + y_n).$$

Proof. We prove (a). First, we deal with infinite cases. If at least one of limsup on the RHS is $+\infty$, say $\limsup_{n\to\infty} x_n = +\infty$, then $\limsup_{n\to\infty} y_n \neq -\infty$ and the inequality holds whatever there is on the LHS.

If at least one of limsup on the RHS is $-\infty$, say $\limsup_{n\to\infty} x_n = -\infty$ (it means that $\lim_{n\to\infty} x_n = -\infty$), then $\limsup_{n\to\infty} y_n \neq +\infty$ which means that $(y_n)_{n\geq 1}$ is bounded above. Then, we have $\lim_{n\to\infty}(x_n + y_n) = -\infty$ and the inequality holds whatever there is on the RHS.

Now, we can assume that both sequences are bounded. By Theorem 12 we know there is a subsequence $(x_{n_k} + y_{n_k})_{k\geq 1}$ of $(x_n + y_n)_{n\geq 1}$ with $\lim_{k\to\infty}(x_{n_k} + y_{n_k}) = \limsup_{n\to\infty}(x_n + y_n)$. The subsequence $(x_{n_k})_{k\geq 1}$ is bounded so by Theorem 5 it contains a convergent further subsequence $(x_{n_{k_\ell}})_{\ell\geq 1}$. The corresponding subsequence $(y_{n_{k_\ell}})_{\ell\geq 1}$ of $(y_{n_k})_{k\geq 1}$ is bounded so again by Theorem 5 it contains a convergent further subsequence $(y_{n_{k_{\ell_j}}})_{j\geq 1}$. Now, both subsequences $(x_{n_{k_{\ell_j}}})_{j\geq 1}$ and $(y_{n_{k_{\ell_j}}})_{j\geq 1}$ are convergent and we have

$$\limsup_{n\to\infty}(x_n+y_n) = \lim_{j\to\infty}(x_{n_{k_{\ell_j}}}+y_{n_{k_{\ell_j}}}) = \lim_{j\to\infty} x_{n_{k_{\ell_j}}} + \lim_{j\to\infty} y_{n_{k_{\ell_j}}} \leq \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n.$$

We used Proposition 6.18 and Proposition 6.4. □

Proposition 6.22. ***Alternate Definition of limsup, liminf:*** *Let $(x_n)_{n\geq 1}$ be a sequence of real numbers. Then,*

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k : k\geq n\} = \inf_{n\geq 1}\left(\sup\{x_k : k\geq n\}\right).$$
$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k : k\geq n\} = \sup_{n\geq 1}\left(\inf\{x_k : k\geq n\}\right).$$

Proof. We prove the first statement. The sequence $s_n = \sup\{x_k : k\geq n\}$, $n\geq 1$ is decreasing (suprema of smaller and smaller sets), so $\lim_{n\to\infty} s_n = \inf_{n\geq 1}\{s_n\}$.

Let $A = \limsup_{n\to\infty} x_n$ and $B = \inf_{n\geq 1}\{s_n\}$. Assume that $A < B$ and let $\varepsilon = (B-A)/3 > 0$. Only a finite number of elements of the sequence $(x_n)_{n\geq 1}$ can be larger than $A+\varepsilon$. Let x_N be such the element with the largest index. We have $B \leq s_{N+1} = \sup\{x_k : k\geq N+1\}$. Thus, there exists an x_k with $k\geq N+1$ with $B-\varepsilon \leq s_{N+1}-\varepsilon < x_k$. Then

$$A+\varepsilon < B-\varepsilon \leq s_{N+1}-\varepsilon < x_k,$$

a contradiction.

Now, assume that $B < A$ and let again $\varepsilon = (B - A)/3 > 0$.

By Theorem 12 there exists a subsequence $(x_{n_k})_{k\geq 1}$ with $\lim_{k\to\infty} x_{n_k} = A$. Then, for $k \geq K$ we have $A - \varepsilon < x_{n_k}$.

Since $B = \inf_{n\geq 1}\{s_n\}$ there exists an $N \geq 1$ with $s_N < B+\varepsilon$. $s_N = \sup\{x_j : j \geq N\}$ so $x_j \leq s_N$ for all $j \geq N$. For elements of the subsequence $(x_{n_k})_{k\geq 1}$ with $k \geq K$ and $n_k \geq N$ we have

$$x_{n_k} \leq s_N < B + \varepsilon \ < \ A - \varepsilon < x_{n_k}$$

again a contradiction.

□

Corollary 6.2. Using Proposition 6.22 we can give another proof of Proposition 6.21. Again we prove only part (a). We have

$$(\sup\{x_k + y_k : k \geq n\}) \leq (\sup\{x_k : k \geq n\}) + (\sup\{y_k : k \geq n\}) .$$

Then, for any $n \geq 1$ we have

$$\inf_{n\geq 1} (\sup\{x_k + y_k : k \geq n\}) \leq (\sup\{x_k : k \geq n\}) + (\sup\{y_k : k \geq n\}) .$$

Now we can take separately infima on the RHS:

$$\begin{aligned} \inf_{n\geq 1} (\sup\{x_k + y_k : k \geq n\}) &\leq \inf_{n\geq 1} (\sup\{x_k : k \geq n\}) + \sup\{y_k : k \geq n\} \\ &\leq \inf_{n\geq 1} (\sup\{x_k : k \geq n\}) + \inf_{n\geq 1} (\sup\{y_k : k \geq n\}) . \end{aligned}$$

By Proposition 6.22 this is equivalent to part (a) of Proposition 6.21.

6.10. **Other Interesting Topics.**

Theorem 13. ***Subadditive Sequence*** *A sequence $(x_n)_{n\geq 1}$ of real numbers is called* ***subadditive*** $\iff$ *we have*

$$x_{n+m} \leq x_n + x_m, \ \forall \ n, m \in \mathbb{N}.$$

Then, the sequence

$$\frac{x_n}{n} \underset{n\to\infty}{\longrightarrow} \inf_{n\geq 1} \left\{ \frac{x_n}{n} \right\} .$$

Proof. Let $\beta = \inf_{n\geq 1}\left\{\frac{x_n}{n}\right\}$. We consider two cases:

(i) β is finite. Let us fix an $\varepsilon > 0$. There exists an $N \geq 1$ with $\frac{x_N}{N} < \beta + \varepsilon/2$. Any $n \in \mathbb{N}$ can be written as $n = kN + r$ with $0 \leq r \leq N-1$. We have

$$\beta \leq \frac{x_n}{n} = \frac{x_{kN+r}}{kN+r} \leq \frac{kx_N + x_r}{kN+r} = \frac{kx_N}{kN}\frac{kN}{kN+r} + \frac{x_r}{kN+r} \leq \beta + \frac{\varepsilon}{2} + \frac{\max\{x_1, x_2, \ldots, x_{N-1}\}}{n}.$$

We can find an $K \geq 1$ such that for $n \geq K$ we have

$$\frac{\max\{x_1, x_2, \ldots, x_{N-1}\}}{n} < \frac{\varepsilon}{2}.$$

this completes the proof of (i).

(ii) $\beta = -\infty$. As in the proof of (i) for any $M \in \mathbb{R}$ we can find $N \geq 1$ with $\frac{x_N}{N} < M - 1$. For any $n \in \mathbb{N}$ we have

$$\frac{x_n}{n} = \frac{x_{kN+r}}{kN+r} \leq \frac{kx_N + x_r}{kN+r} = \frac{kx_N}{kN}\frac{kN}{kN+r} + \frac{x_r}{kN+r} < M + \frac{\max\{x_1, x_2, \ldots, x_{N-1}\}}{n}.$$

We can find an $K \geq 1$ such that for $n \geq K$ we have

$$\frac{\max\{x_1, x_2, \ldots, x_{N-1}\}}{n} < 1.$$

We completes the proof of (ii). □

Theorem 14. ***Density of the sequence*** $n \cdot \alpha$, $\alpha \notin \mathbb{Q}$ *Let* $\alpha \notin \mathbb{Q}$. *Then, the sequence* $x_n = \mathrm{Fr}(n\cdot\alpha)$, $n = 1, 2, 3, \ldots$ *is dense in the interval* $[0, 1]$. $\mathrm{Fr}(t) = t - \mathrm{Int}(t)$ *is the fractional part of number* t.

Proof. Let $\alpha \notin \mathbb{Q}$ and $x_n = \mathrm{Fr}(n \cdot \alpha)$, $n = 1, 2, 3, \ldots$. Then, $x_n \neq x_m$ for $n \neq m$. Assume that we have $n \neq m$ such that $x_n = x_m$. Then,

$$n\alpha + k = m\alpha + w\ , k, w \in \mathbb{N} \implies \alpha = \frac{w-k}{n-m},$$

which contradicts $\alpha \notin \mathbb{Q}$.

Let us fix an $\varepsilon > 0$ and find an $n \in \mathbb{N}$ with $\varepsilon > \frac{1}{n}$. Let us divide the interval $[0, 1]$ into n subintervals I_k, $k = 1, 2, \ldots, n$ of equal length $\frac{1}{n} < \varepsilon$. Consider $n+1$ first elements of the sequence: $\{x_1, x_2, \ldots, x_n, x_{n+1}\}$. By the Pigeon Hole principle two of these elements are in the same subinterval I_{k_0} so there exist n_0 and m_0 such that

$$|x_{n_0} - x_{m_0}| < \varepsilon.$$

We have $\mathrm{Fr}[(n_0 - m_0) \cdot \alpha)] = \mathrm{Fr}[(n_0 \cdot \alpha) - (m_0 \cdot \alpha)] = \mathrm{Fr}[x_{n_0} + k_1 - x_{m_0} - k_2]$ $= \mathrm{Fr}[x_{n_0} - x_{m_0} + k_1 - k_2]$. Thus,

$$\mathrm{Fr}[(n_0 - m_0) \cdot \alpha)] = 1 - \mathrm{Fr}[(m_0 - n_0) \cdot \alpha)].$$

One of them, say $\mathrm{Fr}[(n_0 - m_0) \cdot \alpha)]$ is smaller than ε. Then the sequence $y_k = \mathrm{Fr}[k \cdot (n_0 - m_0) \cdot \alpha)] = x_{k(n_0 - m_0)}$ makes "turtle" steps of length smaller than ε and visits all intervals I_k, $k = 1, 2, \ldots, n$.

We proved that the sequence $(x_n)_{n \geq 1}$ is ε-dense in the interval $[0, 1]$. Since $\varepsilon > 0$ was arbitrary this proves our claim. $\square$

7. Functions and Limits

7.1. Limit of a function.

Definition 7.1. ***Limit of a function: Cauchy definition***

A function $f : A \to \mathbb{R}$ has a limit equal to L at point x_0, $\lim_{x\to x_0} f(x) = L \iff$

$$\forall_{\varepsilon>0}\ \exists_{\delta>0}\ 0 < |x - x_0| < \delta \implies |f(x) - L| < \varepsilon.$$

Note that the value $f(x_0)$ is irrelevant to the existence or value of the limit $\lim_{x\to x_0} f(x)$, and only the values of f in a neighbourhood of x_0 are taken into account.

In "neighbourhood" language this definition is formulated as follows.

Definition 7.2. ***Limit of a function: Cauchy definition*** *A function $f : A \to \mathbb{R}$ has a limit equal to L at point x_0, $\lim_{x\to x_0} f(x) = L \iff$ for any neighbourhood $\mathcal{U}$ of L, we can find a neighbourhood $\mathcal{V}$ of x_0, with the point x_0 removed, which goes under f into $\mathcal{U}$, i.e.,*

$$f(\mathcal{V} \setminus \{x_0\}) \subset \mathcal{U}.$$

Example 7.1.

Show $\lim_{x\to 0} x \sin\frac{1}{x} = 0$. Fix an $\varepsilon > 0$. We need to find $\delta > 0$ such that $|x| < \delta \implies |x \sin\frac{1}{x}| < \varepsilon$. Since $|\sin\frac{1}{x}| < 1$ it is enough to have $|x| < \varepsilon$. Thus, $\delta = \varepsilon$ is sufficient.

Example 7.2.

Show $\lim_{x\to 2} x^2 = 4$. Fix an $\varepsilon > 0$. We need to find $\delta > 0$ such that $|x - 2| < \delta \implies |x^2 - 4| < \varepsilon$. We have $|x^2 - 4 = (x - 2)(x + 2)|$. It is enough to consider x's close to 2 so we can assume $|x - 2| < 1$ or $1 < x < 3$. Then, we have $|x^2 - 4 = (x - 2)(x + 2)| < |x - 2| \cdot 5$ and it is enough to have $|x - 2| < \varepsilon/5$. To have both conditions satisfied we set $\delta = \min\{1, \varepsilon/5\}$.

Example 7.3.

Show

$$\lim_{x\to 2} \frac{x^2}{x^2+4x+5} = \frac{4}{17}.$$

Fix an $\varepsilon > 0$. We need to find $\delta > 0$ such that $|x-2| < \delta \implies \left|\frac{x^2}{x^2+4x+5} - \frac{4}{17}\right| < \varepsilon$. We have

$$\left|\frac{x^2}{x^2+4x+5} - \frac{4}{17}\right| = \left|\frac{17x^2 - 4x^2 - 16x - 20}{x^2+4x+5}\right| = \left|\frac{13x^2-16x-20}{17(x^2+4x+5)}\right|.$$

The polynomial $p(x) = 13x^2 - 16x - 20$ has root $x = 2$ so it is divisible by $x - 2$. We have

$$\begin{array}{l}
13x+10 \\
\hline
13x^2 - 16x - 20 \ : \ x-2 \\
-(13x^2 - 26x) \\
\hline
10x - 20 \\
-(10x-20) \\
\hline
0 \qquad .
\end{array}$$

Thus,

$$\left|\frac{13x^2-16x-20}{17(x^2+4x+5)}\right| = \left|\frac{(x-2)(13x+10)}{17(x^2+4x+5)}\right|.$$

It is enough to consider x's close to 2 so we can assume $|x-2| < 1$ or $1 < x < 3$. Then, we have $\left|\frac{(x-2)(13x+10)}{17(x^2+4x+5)}\right| = |x-2|\left|\frac{(13x+10)}{17(x^2+4x+5)}\right| < |x-2| \cdot \frac{49}{170}$ and it is enough to have $|x-2| \cdot \frac{49}{170} < \varepsilon$, or $|x-2| < \frac{170\varepsilon}{49}$. To have both conditions satisfied we set $\delta = \min\{1, \frac{170\varepsilon}{49}\}$.

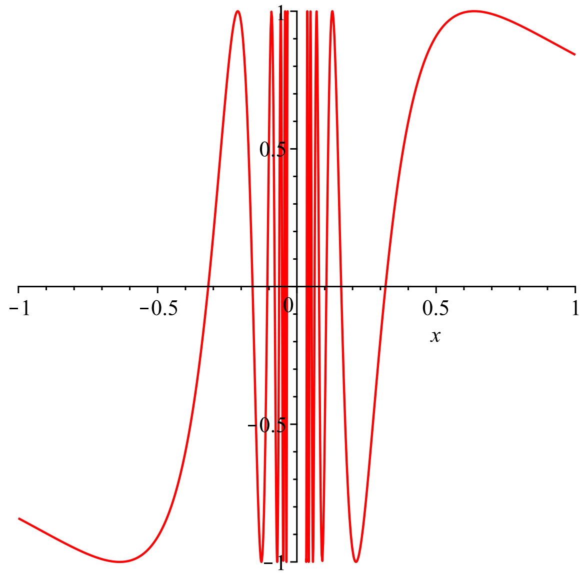

FIGURE 12. Graph of $\sin\frac{1}{x}$ around 0 .

Example 7.4.

Consider the function

$$f(x) = \begin{cases} \sin\frac{1}{x} & \text{if } x \neq 0; \\ 0 & \text{if } x = 0. \end{cases}$$

We will show that the $\lim_{x\to 0} f(x)$ does not exists. We need to show that for any potential limit L we have

$$\exists_{\,\varepsilon>0}\ \forall_{\,\delta>0}\ 0<|x-x_0|<\delta \text{ and } |f(x)-L|\geq\varepsilon.$$

Note that the $\sin\frac{1}{x}=1$ when $\frac{1}{x}=\frac{\pi}{2}+2k\pi$, $k\in\mathbb{Z}$, i.e., for $x_k^{(1)}=1/(\frac{\pi}{2}+2k\pi)$, $k\in\mathbb{Z}$. Similarly, $\sin\frac{1}{x}=-1$ for $x_k^{(-1)}=1/(\frac{-\pi}{2}+2k\pi)$, $k\in\mathbb{Z}$. Both sequences, $(x_k^{(1)})$ and $(x_k^{(-1)})$, converge to 0 as $k\to\infty$.

Let us fix $\varepsilon=1/4$. For arbitrary $\delta>0$ we can find points $x_k^{(1)}$ and $x_k^{(-1)}$ in the interval $(0,\delta)$. We have $|f(x_k^{(1)})-f(x_k^{(-1)})|=2$ so they cannot be both ε-close to an L, whatever it is.

Example 7.5.

Let $f(x)=x+Int(x)$. Show that f does not have a limit at any integer point. We will show this at $x_0=1$. Let $\varepsilon=1/4$.

For arbitrary $0<\delta<1$ the values of $f(x)<1$ for $x\in(1-\delta,1)$ and the values of $f(x)\geq 2$ for $x\in[1,1+\delta)$. Thus, we can find x_1,x_2 with $|x_1-x_2|<\delta$ and $|f(x_1)-f(x_2)|\geq 1\geq\varepsilon$. Again, no limit L is possible.

Definition 7.3. ***Limit of a function: Heine (sequential) definition***

A function $f:A\to\mathbb{R}$ has a limit equal to L at point x_0, $\lim_{x\to x_0} f(x)=L \iff$ for any sequence $(x_n)_{n\geq 1}$ such that $x_n \underset{n\to\infty}{\longrightarrow} x_0$ and $x_n\neq x_0$, $n=1,2,\dots$ we have $f(x_n)\underset{n\to\infty}{\longrightarrow} L$.

Example 7.6.

Consider

$$f(x)=\begin{cases}0 & \text{if } x\notin\mathbb{Q};\\ \frac{1}{q} & \text{if } x=\frac{p}{g} \text{ (no common factors).}\end{cases}$$

Show that $\lim_{x\to x_0} f(x)=0$, for any $x_0\in\mathbb{R}$. We will use Heine definition. Let us fix an $x_0\mathbb{R}$. If $x_n\to x_0$ and all $x_n\in\mathbb{R}\setminus\mathbb{Q}$, then $f(x_n)=0\underset{n\to\infty}{\longrightarrow}0$. Assume that

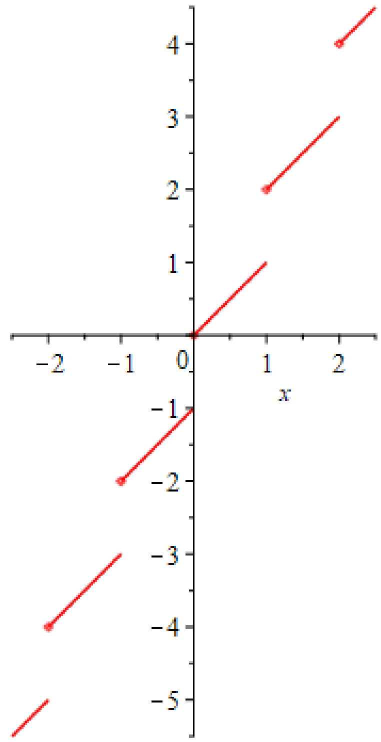

FIGURE 13. Graph of $f(x) = x + \mathrm{Int}(x)$.

$x_n \to x_0$ and all $x_n \in \mathbb{Q}$, $x_n = \frac{p_n}{q_n}$. The fractions $\frac{p_n}{q_n}$ with bounded denominators $q_n \leq M$ form a discreet subset of $\mathbb{R}$. Thus, if $\frac{p_n}{q_n} \underset{n\to\infty}{\longrightarrow} x_0$ we have $q_n \underset{n\to\infty}{\longrightarrow} +\infty$ and $f(x_n) = \frac{1}{q_n} \underset{n\to\infty}{\longrightarrow} 0$. The general sequence $x_n \underset{n\to\infty}{\longrightarrow} x_0$ we can break into a subsequence of rational numbers and a subsequence of irrational numbers.

Theorem 15. ***Equivalence of Cauchy and Heine definitions of a limit:*** *Both definitions of a limit, Cauchy and Heine, are equivalent.*

Proof. **Cauchy $\implies$ Heine**

We assume

$$(*) \quad \forall_{\varepsilon>0} \, \exists_{\delta>0} \; 0 < |x - x_0| < \delta \implies |f(x) - L| < \varepsilon.$$

Take a sequence $x_n \to x_0$, with $x_n \neq x_0$ for all $n \geq 1$. We want to show

$$\forall_{\varepsilon>0} \, \exists_{N\geq 1} \, \forall_{n\geq N} \; |f(x_n) - L| < \varepsilon.$$

Let us fix an $\varepsilon > 0$. From $(*)$ we know that we can find a $\delta > 0$ such that if $0 < |x - x_0| < \delta$, then $|f(x) - L| < \varepsilon$. Since $x_n \to x_0$, with $x_n \neq x_0$ for all $n \geq 1$, for such a δ we can find an $N \geq 1$ such that $0 < |x_n - x_0| < \delta$ for all $n \geq N$. Thus, for $n \geq N$ we have $|f(x_n) - L| < \varepsilon$ and we proved that $\lim_{x\to x_0} f(x) = L$ in the sense of Heine.

Heine $\implies$ Cauchy

We will show the contrapositive: if $\lim_{x\to x_0} f(x) \neq L$ in the sense of Cauchy, then we can find a sequence $x_n \to x_0$, with $x_n \neq x_0$ for all $n \geq 1$, such that $f(x_n) \not\to L$, as $n \to \infty$. We assume

$$(**) \quad \exists_{\,\varepsilon>0} \, \forall_{\,\delta>0} \, \exists_{\,x} \; 0 < |x - x_0| < \delta \text{ and } |f(x) - L| \geq \varepsilon.$$

Take $\delta_n = 1/n$ and choose points x_n, $n = 1, 2, \ldots$ satisfying $(**)$ for δ_n. Then, $x_n \to x_0$, with $x_n \neq x_0$ for all $n \geq 1$, and for any $n \geq 1$ we have $|f(x_n) - L| \geq \varepsilon$. This means $f(x_n) \not\to L$, as $n \to \infty$. □

Remark 7.1. *In the second part of the proof of Theorem 15 we used an additional axiom not mentioned before, Axiom of Choice. It says that given any collection of non-empty sets, it is possible to construct a new set by choosing one element from each set. We used it to choose the points* $\{x_n\}_{n\geq 1}$. *Some mathematicians accept this axiom, some do not. We, in this course, accept and use the Axiom of Choice.*

The Axiom of Choice has some surprising consequences, for example, so called Banach-Tarski paradox. Wikipedia summarizes it as follows.

Given a solid ball in three-dimensional space, there exists a decomposition of the ball into a finite number of disjoint subsets that can be put back together in a different way to yield two identical copies of the original ball. Indeed, the reassembly process involves only moving the pieces around and rotating them, without changing their original shape. The pieces themselves are not "solids" in the traditional sense, but infinite scatterings of points. The reconstruction can work with as few as five pieces.

Example 7.7.

We will redo here Example 7.3 using Heine definition. Show

$$\lim_{x\to 2}\frac{x^2}{x^2+4x+5}=\frac{4}{17}.$$

We take any sequence $(x_n)_{n\geq 1})$ such that $x_n \to 2$ and $x_n \neq 2$ for $n = 1, 2, \ldots$. Then, using Proposition 6.4 we obtain

$$\frac{x_n^2}{x_n^2+4x_n+5}\underset{n\to\infty}{\longrightarrow}\frac{2^2}{2^2+4\cdot 2+5}=\frac{4}{17}.$$

We can see that the proof using Heine definition is much easier.

Now, we will prove a number of propositions about the limits of functions. The proofs are very similar to those of the sequence propositions so they are usually skipped.

Proposition 7.1. ***Uniqueness of the limit of a function:*** *If* $\lim_{x\to x_0} f(x) = A$ *and* $\lim_{x\to x_0} f(x) = B$*, then* $A = B$*.*

Proof. The proof is similar to the proof of Proposition 6.1. □

Proposition 7.2. *If* $\lim_{x\to x_0} f(x) = A$*, then there exist a* $\delta > 0$ *such that the function* f *is bounded on* $(x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$*.*

Proof. The proof is similar to the proof of Proposition 6.2. □

Proposition 7.3. *If* $\lim_{x\to x_0} f(x) = A$ *and* $A < B$ *then there exist a* $\delta > 0$ *such that we have* $f(x) < B$ *on* $(x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$*.*

Proof. Again the proof is similar to the proof of Proposition 6.2. □

Proposition 7.4. *If* $f(x) \leq g(x)$ *on some neighbourhood of* x_0*, then*

$$\lim_{x\to x_0} f(x) \leq \lim_{x\to x_0} g(x).$$

Proof. Let $\lim_{x\to x_0} f(x) = A$ and $\lim_{x\to x_0} g(x) = B$. Assume that $B < A$ and let $\varepsilon = (A - B)/3$. We can find a common $\delta > 0$ such that for $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$ we have

$$g(x) < B + \varepsilon < A - \varepsilon < f(x),$$

which contradicts the assumption $f(x) \leq g(x)$. □

Proposition 7.5. ***Three Functions Theorem*** *If $f(x) \leq h(x) \leq g(x)$ on some neighbourhood of x_0 and $\lim_{x\to x_0} f(x) = A = \lim_{x\to x_0} g(x)$, then*

$$\lim_{x\to x_0} h(x) = A.$$

Proof. This is the corollary of Proposition 7.4. □

Proposition 7.6. ***Arithmetic of Limits of Functions*** *Let us assume that $\lim_{x\to x_0} f(x) = A$ and $\lim_{x\to x_0} g(x) = B$. Then*

(a) $\lim_{x\to x_0}(f + g)(x) = A + B$;

(b) $\lim_{x\to x_0}(f \cdot g)(x) = A \cdot B$;

(c) *if, additionally, $B \neq 0$, then $\lim_{x\to x_0} \left(\frac{f}{g}\right)(x) = \frac{A}{B}$;*

(d) $\lim_{x\to x_0} \max\{f, g\}(x) = \max\{A, B\}$;

(e) $\lim_{x\to x_0} \min\{f, g\}(x) = \min\{A, B\}$.

Proof. For an effortless proof one can use Heine definition of the limit and Proposition 6.4. The following formulas are useful for the proof of parts (d) and (e).

$$\max\{x, y\} = \frac{1}{2}(x + y + |x - y|);$$
$$\min\{x, y\} = \frac{1}{2}(x + y - |x - y|).$$

□

Definition 7.4. ***Infinite Limits:***

$$\lim_{x\to x_0} f(x) = +\infty \iff \forall_{M\in\mathbb{R}}\ \exists_{\delta>0}\ 0<|x-x_0|<\delta \implies f(x)>M.$$
$$\lim_{x\to x_0} f(x) = -\infty \iff \forall_{M\in\mathbb{R}}\ \exists_{\delta>0}\ 0<|x-x_0|<\delta \implies f(x)<M.$$

Definition 7.5. ***Limits at*** $\pm\infty$***:***

$$\lim_{x\to+\infty} f(x) = L \iff \forall_{\varepsilon>0}\ \exists_{M\in\mathbb{R}}\ x>M \implies |f(x)-L|<\varepsilon.$$
$$\lim_{x\to-\infty} f(x) = L \iff \forall_{\varepsilon>0}\ \exists_{M\in\mathbb{R}}\ x<M \implies |f(x)-L|<\varepsilon.$$

All definitions of limits can be unified using the "neighbourhood" language.

7.2. Increasing and Decreasing Functions:

Definition 7.6. *A function* $f : A \to \mathbb{R}$ *is called increasing* $\iff$ $x<y \implies f(x)\le f(y)$, *for all* $x, y \in A$.

A function $f : A \to \mathbb{R}$ *is called decreasing* $\iff$ $x<y \implies f(x)\ge f(y)$, *for all* $x, y \in A$.

If the inequalities between the values of function above are strict, then the function is called strictly increasing or strictly decreasing, correspondingly.

Increasing and decreasing functions are called together monotonic functions.

7.3. One-sided Limits of a Function:

Definition 7.7. *We say that function* $f : A \to \mathbb{R}$ *has left one-sided limit at* x_0 *equal to* L, $\lim_{x\to x_0^-} f(x) = L$, $\iff$

$$\forall_{\varepsilon>0}\ \exists_{\delta>0}\ x_0-\delta<x<x_0 \implies |f(x)-L|<\varepsilon.$$

We say that function $f : A \to \mathbb{R}$ *has left one-sided limit at* x_0 *equal to* L, $\lim_{x\to x_0^-} f(x) = L$, $\iff$

$$\forall_{\varepsilon>0}\ \exists_{\delta>0}\ x_0<x<x_0+\delta \implies |f(x)-L|<\varepsilon.$$

The above definitions are Cauchy definitions. They can be formulated in the language of sequences as Heine definitions of one-sided limits.

Proposition 7.7.

$$\lim_{x \to x_0} f(x) = L \iff \lim_{x \to x_0^-} f(x) = L \text{ and } \lim_{x \to x_0^+} f(x) = L.$$

Proof. We skip the proof as it follows immediately from the definitions. □

Proposition 7.8. *A monotonic function has one-sided limits at any point.*

Proof. We prove this for an increasing function $f : A \to \mathbb{R}$ and a left one-sided limit at a point x_0. Consider an interval $I = (x_1, x_0) \subset A$. Since f is increasing the set $V = \{f(x) : x \in I\}$ is bounded above by $f(x_0)$ and has a supremum, say $\alpha = \sup V$. By the definition of supremum, for any $\varepsilon > 0$ we can find an element $f(x_2) \in V$ with $\alpha - \varepsilon < f(x_2) \leq \alpha$. Let $\delta = x_0 - x_2$. Then, for any $x_0 - \delta < x < x_0$ we have $f(x_2) \leq f(x) \leq \alpha$ so $\alpha - \varepsilon < f(x) \leq \alpha$. We proved that

$$\lim_{x \to x_0^-} f(x) = \alpha = \sup\{f(x) : x \in (x_1, x_0)\}.$$

□

Proposition 7.9. *A monotonic function has at most countably many points where the limit does not exist.*

Proof. Let f be a monotonic, say increasing, function and let NL be the set of points where the limit of f does not exist. This means that

$$NL = \{x_0 : L(x_0)^- = \lim_{x \to x_0^-} f(x) < \lim_{x \to x_0^+} f(x) = L(x_0)^+\}.$$

The intervals $\{(L(x_0)^-, L(x_0)^+)\}_{x_0 \in NL}$ are pairwise disjoint since f is an increasing function. For any $x_0 \in NL$ we can find a rational number $r(x_0) \in (L(x_0)^-, L(x_0)^+)$. The function $h : NL \to \mathbb{Q}$ defined as $h(x_0) = r(x_0)$ is an injection because of the disjointness of intervals. Thus, Card(NL) $\leq \aleph_0$. □

Corollary 7.1. *The set of increasing functions $F(\nearrow)$ is of cardinality continuum.*

Proof. In the notation of Proposition 7.9 an increasing function is uniquely determined by the collection $\{L(x_0)^-, f(x_0), L(x_0)^+\}_{x_0 \in NL}$. This means that

$$\mathrm{Card}(\mathrm{F}(\nearrow)) = (\mathfrak{c}^3)^{\aleph_0} = \mathfrak{c}^{\aleph_0} = \mathfrak{c}.$$

□

8. Continuous Functions

In short, a function f is continuous at a point $x_0 \iff f(x_0) = \lim_{x \to x_0} f(x)$. $f : A \to \mathbb{R}$ is continuous on $A \iff f$ is continuous at any $x \in A$. Since we have two equivalent definitions of the limit, there are two equivalent definitions of continuity of a function $f : \mathbb{R} \to \mathbb{R}$:

Definition 8.1. *(1) Cauchy (or $\varepsilon - \delta$) definition: f is continuous at point $x_0 \in \mathbb{R}$ $\iff$*

$$\forall_{\varepsilon > 0} \ \exists_{\delta > 0} \ \forall_{x \in \mathbb{R}} \ |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon \,.$$

(2) Heine (or sequential) definition: f is continuous at point $x_0 \in \mathbb{R} \iff$ for any sequence $\{x_n\}$ such that $x_n \underset{n\to\infty}{\longrightarrow} x_0$ we have $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x_0)$.

Theorem 16. *Definitions (1) and (2) are equivalent.*

Proof. The proof is almost the same as that of Theorem 15.

(1) $\implies$ (2) : Let $\{x_n\}$ be such that $x_n \underset{n\to\infty}{\longrightarrow} x_0$. We need to prove that $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x_0)$, i.e.,

$$(*) \ \forall_{\varepsilon > 0} \ \exists_{N \geq 1} \ \forall_{n \geq N} \ |f(x_n) - f(x_0)| < \varepsilon \,.$$

Let us fix an $\varepsilon > 0$. By (1), we can find a $\delta > 0$ such that $|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon$. Since $x_n \underset{n\to\infty}{\longrightarrow} x_0$, we can find an $N \geq 1$ such that for $n \geq N$ we have $|x_n - x_0| < \delta$ and then $|f(x_n) - f(x_0)| < \varepsilon$. $(*)$ has been proved.

$(2) \Longrightarrow (1)$: We will prove contrapositive statement $\neg(1) \Longrightarrow \neg(2)$. Let us assume that (1) does not hold, i.e.,

$$\exists_{\,\varepsilon>0}\ \forall_{\,\delta>0}\ \exists_{\,x\in\mathbb{R}}\ |x-x_0|<\delta \text{ and } |f(x)-f(x_0)|\geq\varepsilon\,.$$

Let $\varepsilon_0 > 0$ be the ε whose existence is claimed above. It says "for any δ" so we will use a sequence od δ's. Let $\delta_n = 1/n > 0$, $n = 1, 2, \ldots$. For each δ_n we can find an x_n such that $|x_n - x_0| < \delta_n$ and $|f(x_n) - f(x_0)| \geq \varepsilon_0$. Thus, the sequence $x_n \underset{n\to\infty}{\longrightarrow} x_0$, but $f(x_n) \underset{n\to\infty}{\not\longrightarrow} f(x_0)$. We proved $\neg(2)$. □

Example: We will prove that $f(x) = x^2 + 3$ is continuous at $x_0 = 3$. Note that $f(3) = 12$.

Using definition (1): Let us fix an $\varepsilon > 0$. We have to find $\delta > 0$ such that $|x-3| < \delta \implies |f(x) - 12| < \varepsilon$. This means $|x^2 + 3 - 12| < \varepsilon$ or $|x^2 - 9| < \varepsilon$ or

$$(**)\ |x-3||x+3| < \varepsilon\,.$$

We have $|x - 3| < \delta$. To estimate $|x + 3|$ (which is unbounded on real line) we make first assumption on δ: Let $\delta < 1$. Then, $|x - 3| < \delta$ is $|x - 3| < 1$ which implies $2 < x < 4$. This, in turn implies $|x + 3| < 7$. Inequality (**) becomes $\delta \cdot 7 < \varepsilon$. We will satisfy it making second assumption on δ: Let $\delta < \varepsilon/7$. We define

$$\delta = \frac{1}{2}\min\{1, \varepsilon/7\}\,.$$

This δ satisfies both assumptions. Above we proved that if these assumptions are satisfied and $|x - 3| < \delta$, then $|f(x) - 12| < \varepsilon$. This proves that f is continuous at $x_0 = 3$.

Using definition (2): Let $\{x_n\}$ be any sequence such that $x_n \to 3$ as $n \to \infty$. Using the theorems about sums and products of limits we obtain:

$$f(x_n) = x_n^2 + 3 \to 3^2 + 3 = 12 = f(3)\,.$$

We proved that f is continuous at $x_0 = 3$.

8.1. **Properties of Continuous Functions:**

Proposition 8.1. *If f is continuous and $A < f(x_0) < B$, then for any numbers $A < D < f(x_0) < E < B$ we can find $\delta > 0$ such that $f((x_0 - \delta, x_0 + \delta)) \subset (D, E)$.*

Proof. Let $\varepsilon = \frac{1}{2}\min\{E - f(x_0), f(x_0) - D\} > 0$. For this ε we can find a $\delta > 0$ such that for $x \in (x_0 - \delta, x_0 + \delta)$ we have $|f(x) - f(x_0)| < \varepsilon$ or in other words $f((x_0 - \delta, x_0 + \delta)) \subset (f(x_0) - \varepsilon, f(x_0) + \varepsilon) \subset (D, E)$. □

Proposition 8.2. ***Arithmetic on Continuous Functions:*** *Let $f, g : A \to \mathbb{R}$ be two continuous functions. Then,*

(a) *the function $f + g$ is continuous ;*

(b) *the function $f \cdot g$ is continuous ;*

(c) *if, additionally, $g(x) \neq 0$, $x \in A$, then the function $\frac{f}{g}$ is continuous ;*

(d) *the function $\max\{f, g\}(x)$ is continuous ;*

(e) *the function $\min\{f, g\}(x)$ is continuous.*

Proof. This is a corollary of Proposition 7.6. □

Theorem 17. ***Continuity of the inverse function:*** *Let $f : [a, b] \to [c, d]$ be invertible. If f is continuous, then the inverse function f^{-1} is also continuous.*

Proof. We will present two proofs. The first proof uses the fact that f is monotonic (an invertible function on $[a, b] \subset \mathbb{R}$ is monotonic), the second does not use this fact.

(1) Let us assume that f is increasing and that f^{-1} is not continuous at $y_0 = f(x_0)$. f^{-1} is also increasing and by Proposition 7.8 we know that the one sided limits of f^{-1} at y_0 exist. See Figure 14.

If f^{-1} is not continuous at y_0, these one sided limits are different and the function f is not defined on the interval between them. This is a contradiction as f was supposed to be a well defined continuous function.

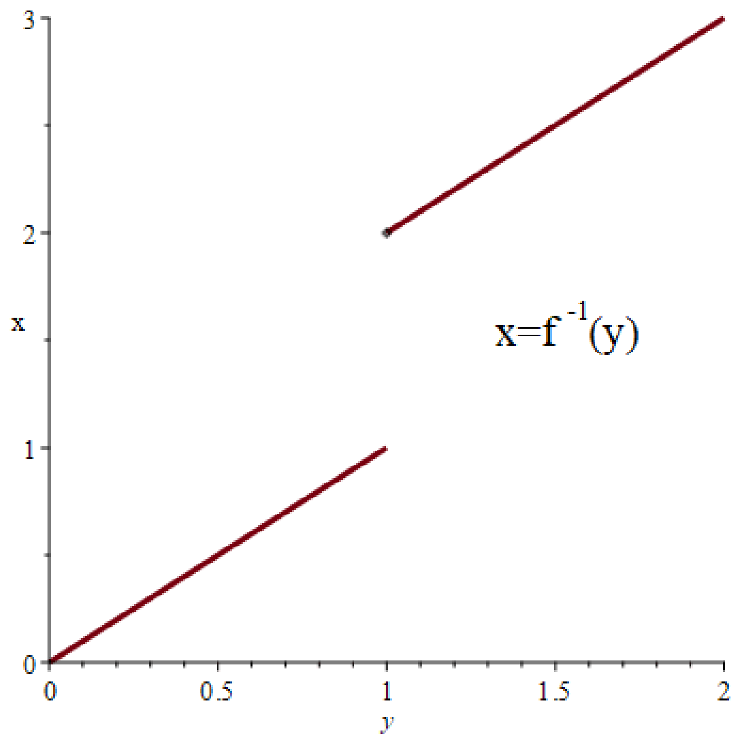

FIGURE 14. Discontinuous inverse function.

(2) Now, we do not use monotonicity of f. Let f be continuous on $[a, b]$. Let f^{-1} be not continuous at $y_0 = f(x_0) \in [c, d]$, i.e., we have a sequence $y_n = f(x_n) \to y_0$ such that $f^{-1}(y_n) = x_n \not\to x_0 = f^{-1}(y_0)$. In particular, $(x_n)_{n\geq 1}$ contains a subsequence $(x_{n_k})_{k\geq 1}$, which stays away from x_0 (Proposition 6.19). The sequence $(x_{n_k})_{k\geq 1}$ is bounded (it is in $[a, b]$) so by Bolzano-Weierstrass Theorem 6 it contains a further convergent subsequence $(x_{n_{k_j}})_{j\geq 1}$, say $x_{n_{k_j}} \underset{j\to\infty}{\longrightarrow} x_s \neq x_0$. Since f is continuous we have $y_{n_{k_j}} = f(x_{n_{k_j}}) \underset{j\to\infty}{\longrightarrow} f(x_s) \neq f(x_0)$. At the same time, $y_{n_{k_j}} \underset{j\to\infty}{\longrightarrow} y_0 = f(x_0)$, as a subsequence of a sequence with this limit. Thus, the subsequence $(y_{n_{k_j}})_{j\geq 1}$ converges to two different limits. A contradiction. $\square$

Proposition 8.3. ***Composition of Continuous Functions is Continuous:*** *Let $f : A \to B$ and $g : B \to C$. If f is continuous at $x_0 \in A$ and g is continuous at $y_0 = f(x_0) \in B$, then the composition $g \circ f$ is continuous at x_0.*

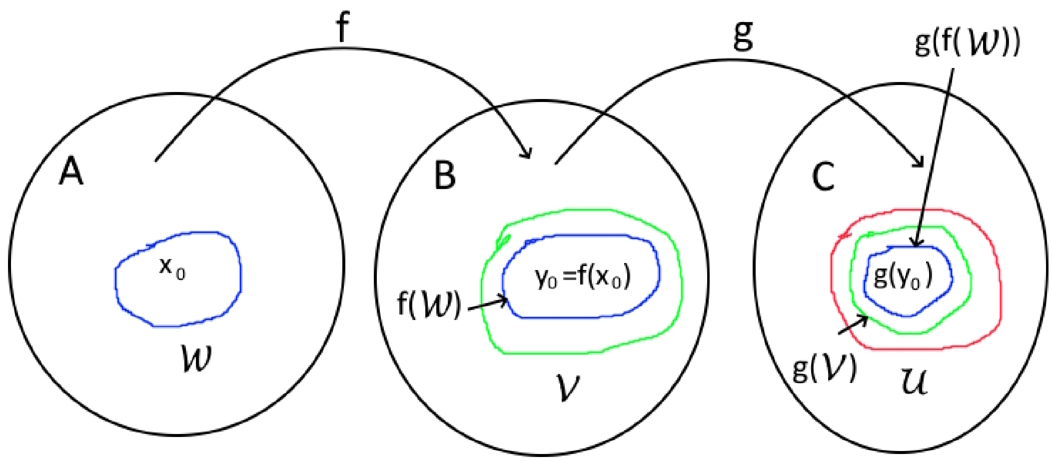

FIGURE 15. Composition of continuous functions.

Proof. The easiest proof uses the neighbourhood language. We want to show that for any neighbourhood $\mathcal{U}$ of $g \circ f(x_0)$ we can find a neighbourhood $\mathcal{W}$ of x_0 with $g \circ f(\mathcal{W}) = g(f(\mathcal{W})) \subset \mathcal{U}$.

g is continuous at y_0 so for any neighbourhood $\mathcal{U}$ of $g(y_0)$ we can find a neighbourhood $\mathcal{V}$ of y_0 with $g(\mathcal{V}) \subset \mathcal{U}$.

f is continuous at x_0 so for the neighbourhood $\mathcal{V}$ of $y_0 = f(x_0)$ we can find a neighbourhood $\mathcal{W}$ of x_0 with $f(\mathcal{W}) \subset \mathcal{V}$.

Then, $g(f(\mathcal{W})) \subset g(\mathcal{V}) \subset \mathcal{U}$, and the proof is completed.

□

8.2. **Examples of Continuous Functions:**

(1) It is easy to prove that the function $f(x) = x$ is continuous on $\mathbb{R}$.
(2) By Proposition 8.2 it implies that all polynomials are continuous functions.
(3) By the same Proposition 8.2 it implies that all rational functions, i.e., fractions of polynomial over polynomial, are continuous as long as the denominator is not 0.
(4) We will prove that $f(x) = x^\alpha$, $0 < \alpha < 1$, are continuous on the interval $[0, +\infty)$.

Note that if $0 < a < 1$, then $a < a^\alpha$. Let $0 < t < 1$. Then, $0 < 1 - t < 1$ as well and we have

$$0 < 1 - t^\alpha < 1 - t < (1 - t)^\alpha \ .$$

Let $0 < x < y$. Then $0 < t = \frac{x}{y} < 1$ and we obtain

$$1 - \left(\frac{x}{y}\right)^\alpha < \left(1 - \frac{x}{y}\right)^\alpha,$$

or

$$y^\alpha - x^\alpha < (y - x)^\alpha.$$

This inequality shows that $f(x) = x^\alpha$, $0 < \alpha < 1$ is continuous, as we can use $\delta = \varepsilon^{1/\alpha}$ in Cauchy definition of continuity.

Remark 8.1. *This is an example of more general situation, when a function f satisfies* ***Hölder's inequality****:*

(9) $$|f(x) - f(y)| \leq H(x - y)^\alpha \ , 0 < \alpha \leq 1, \ x, y \in A.$$

Hölder's inequality implies continuity of f.

If $\alpha = 1$, the inequality (9) is called ***Lipschitz's inequality.***

(5) Exponential function $f(x) = e^x$ is continuous on $\mathbb{R}$. Since $e^x - e^y = e^x(1 - e^{y-x})$ it is enough to show that $\lim_{x\to 0} e^x = 1$. We will use Heine definition of a limit. We know that $\lim_{n\to\infty} e^{1/n} = 1$. Let $(x_n)_{n\geq 1}$ will be an arbitrary sequence such that $0 < x_k$, $k = 1, 2, \ldots$ and $x_k \underset{k\to\infty}{\longrightarrow} 0$. For any x_k we can find a natural number n_k such that $\frac{1}{n_k+1} \leq x_k < \frac{1}{n_k}$. Then,

$$e^{\frac{1}{n_k+1}} \leq e^{x_k} < e^{\frac{1}{n_k}}$$

and since sequences on both sides converge to 1, we have $e^{x_k} \to 1$ as well.

For sequences with negative elements we use

$$e^{x_k} = \frac{1}{e^{-x_k}} \underset{k\to\infty}{\longrightarrow} \frac{1}{1} = 1.$$

(6) In view of Theorem 17 and the above example, the function

$$f(x) = \ln x$$

is continuous on $(0, +\infty)$.

(7) The functions $\sin x$, $\cos x$ are continuous on $\mathbb{R}$. We use the trigonometric formulas:

$$\sin x - \sin y = 2\sin\frac{x-y}{2}\cdot\cos\frac{x+y}{2} \text{ and } \cos x - \cos y = -2\sin\frac{x-y}{2}\cdot\sin\frac{x+y}{2},$$

to obtain

$$|\sin x - \sin y| \leq |x-y| \text{ and } |\cos x - \cos y| \leq |x-y|.$$

Thus, both sin and cos functions satisfy Lipschitz condition which implies continuity.

Proposition 8.4. ***Monotonic function is continuous except on a countable set.***

Let $f : A \to \mathbb{R}$ be a monotonic function. Then, f is continuous except on a countable set, possibly empty.

Proof. We prove this for an increasing function f. We know that f has one sided limits at any point. Let D be the set of points of discontinuity of f. If $x_1 < x_2$ are two points where f is not continuous then

$$f(x_1^-) < f(x_1^+) \leq f(x_2^-) < f(x_2^+),$$

so the intervals $(f(x_1^-), f(x_1^+))$ and $(f(x_2^-), f(x_2^+))$ are disjoint. If rational numbers r_1, r_2 are such that $r_1 \in (f(x_1^-), f(x_1^+))$ and $r_2 \in (f(x_2^-), f(x_2^+))$, then $r_1 \neq r_2$. This means that the function $\phi : D \to \mathbb{Q}$ defined by $\phi(x) = r$, where $r \in (f(x^-), f(x^+))$ is an injection and Card(D) $\leq \aleph_0$. $\square$

8.3. Three "Hard" Theorems about Continuous Functions.

Theorem 18. ***Continuous Function on a Compact Interval is Bounded:*** *Let $f : [a,b] \to \mathbb{R}$ be continuous on a bounded (compact) interval $[a,b]$. Then, f is bounded, i.e., there exists an $M > 0$ such that $|f(x)| \le M$ for all $x \in [a,b]$.*

Proof. Let us assume that function f is not bounded above, i.e., for any $n = 1, 2, 3, \ldots$ we can find a point $x_n \in [a,b]$ such that $f(x_n) \ge n$. The sequence $\{x_n\}$ is bounded so by Bolzano-Weierstrass theorem it contains a convergent subsequence $x_{n_k} \to x_0$, as $k \to \infty$. Then, $x_0 \in [a,b]$. Since f is continuous we have (Heine definition)

$$f(x_{n_k}) \to f(x_0) \ , \ k \to \infty \, .$$

On the other hand, we have

$$f(x_{n_k}) \ge n_k \ , \text{ so } f(x_{n_k}) \to +\infty \ , \ k \to \infty \, .$$

A contradiction. □

Theorem 19. ***Continuous Function on a Compact Interval attains its Extremal Values, the Maximum and the Minimum:*** *Let $f : [a,b] \to \mathbb{R}$ be continuous on a bounded (compact) interval $[a,b]$. Then, there exists a point $x_1 \in [a,b]$ such that*

$$f(x_1) = m = \inf_{x \in [a,b]} f(x) \, ,$$

and, there exists a point $x_2 \in [a,b]$ such that

$$f(x_2) = M = \sup_{x \in [a,b]} f(x) \, .$$

This also means that $m = \min_{x \in [a,b]} f(x)$ and $M = \max_{x \in [a,b]} f(x)$.

Proof. We will prove the existence of x_1. Since $m = \inf_{x \in [a,b]} f(x)$ for any $n = 1, 2, 3, \ldots$ we can find a point $x_n \in [a,b]$ such that $m \le f(x_n) \le m + 1/n$. The sequence $\{x_n\}$ is bounded so by Bolzano-Weierstrass theorem it contains a convergent subsequence $x_{n_k} \to x_1$, as $k \to \infty$. Then, $x_1 \in [a,b]$. Since f is continuous we have (Heine definition)

$$f(x_{n_k}) \to f(x_1)\ ,\ k \to \infty\,.$$

We also have

$$m \leq f(x_{n_k}) \leq m + 1/n_k\ ,\ \text{so}\ f(x_{n_k}) \to m\ ,\ k \to \infty\,.$$

Thus,

$$f(x_1) = m\,,$$

and f attains its infimum on $[a,b]$. □

Application of attaining extremal values

Example 8.1. ***The closest point:***

Find the point of the circle $x^2 + (y-1)^2 = 1$ which is the closest to the point $(2,3)$.

Let $P = (x,y)$ be a point on the circle. Then $P = (x, 1 \pm \sqrt{1-x^2})$, $x \in [-1,1]$. The square of the distance between P and $(2,3)$ is

$$d_{\pm}^2(x) = (x-2)^2 + (1 \pm \sqrt{1-x^2} - 3)^2.$$

We need to consider two functions d_+^2 for upper half of the circle, and d_-^2 for lower half of the circle. Both are continuous. By Theorem 19, both functions attain the infimum. The point where the smaller of the infima is attained is the point of the circle closest to $(2,3)$. Theorem 19 gives only the existence of the closest point, not a method to actually find it. This we will do later using derivatives.

Theorem 20. ***Intermediate Value Theorem:*** *Let $f : [a,b] \to \mathbb{R}$ be continuous on a bounded (compact) interval $[a,b]$. If $f(a) < 0$ and $f(b) > 0$, then there exist a point $c \in (a,b)$ such that $f(c) = 0$.*

Proof. Let $d = b - a$. We will construct approximations of a point c by induction:

1st step: Consider the point $t = (a + b)/2$ (middle point between a and b).

If $f(t) < 0$, then define $a_1 = t$, $b_1 = b$. Note that $b_1 - a_1 = d/2$.

If $f(t) > 0$, then define $a_1 = a$, $b_1 = t$. Note that also in this case $b_1 - a_1 = d/2$.

2nd step: Consider the point $t = (a_1 + b_1)/2$ (middle point between a_1 and b_1).

If $f(t) < 0$, then define $a_2 = t$, $b_2 = b_1$. Note that $b_2 - a_2 = d/4$.

If $f(t) > 0$, then define $a_2 = a_1$, $b_2 = t$. Note that also in this case $b_2 - a_2 = d/4$.

Assume that we have points $a_n < b_n$ with $f(a_n) < 0 < f(b_n)$ and $b_n - a_n = d/2^n$. (If at any time $f(t) = 0$, then we set $c = t$ and stop the procedure.)

(n+1)st step: Consider the point $t = (a_n + b_n)/2$ (middle point between a_n and b_n).

If $f(t) < 0$, then define $a_{n+1} = t$, $b_{n+1} = b_n$. Note that $b_{n+1} - a_{n+1} = d/2^{n+1}$.

If $f(t) > 0$, then define $a_{n+1} = a_n$, $b_{n+1} = t$. Note that also in this case $b_{n+1} - a_{n+1} = d/2^{n+1}$.

This way we constructed two sequences: increasing $\{a_n\}$ and decreasing $\{b_n\}$ with $b_n - a_n = d/2^n$. Thus, they converge to the same limit, say c:

$$a_n \to c\ ,\ b_n \to c\ ,\ n \to \infty\,.$$

Moreover, we have $f(a_n) < 0$ and $f(b_n) > 0$ for all n. Since, f is continuous we have

$$f(a_n) \to f(c)\ ,\ f(b_n) \to f(c)\ ,\ n \to \infty\,.$$

Thus, $f(c) \le 0$ and $f(c) \ge 0$, which means that $f(c) = 0$. □

Applications of Intermediate Value Theorem:

Existence of $\sqrt{2}$:

The function $f(x) = x^2 - 2$ is continuous on $\mathbb{R}$. We have $f(1) = -1 < 0$ and $f(2) = 4 - 2 > 0$. Thus, there exists an $x \in (1, 2)$ such that $x^2 - 2 = 0$, or $x^2 = 2$ which defines the $\sqrt{2}$.

Polynomial $P(x) = 18x^3 - 27x^2 + 13x - 2$ has 3 zeros in $(0, 1)$:

$P(x)$ is continuous on $\mathbb{R}$ and we have $P(0) = -2 < 0$, $P(4/10) = 4/125$, $P(6/10) = -4/125$, $P(1) = 2$. Thus, P changes sign three times and by IVTh it has at least three zeros in $(0, 1)$. As a polynomial of third order it has at most three real zeros, so P has exactly three zeros in $(0, 1)$.

A continuous function $f : [a, b] \to [a, b]$ has a fixed point:

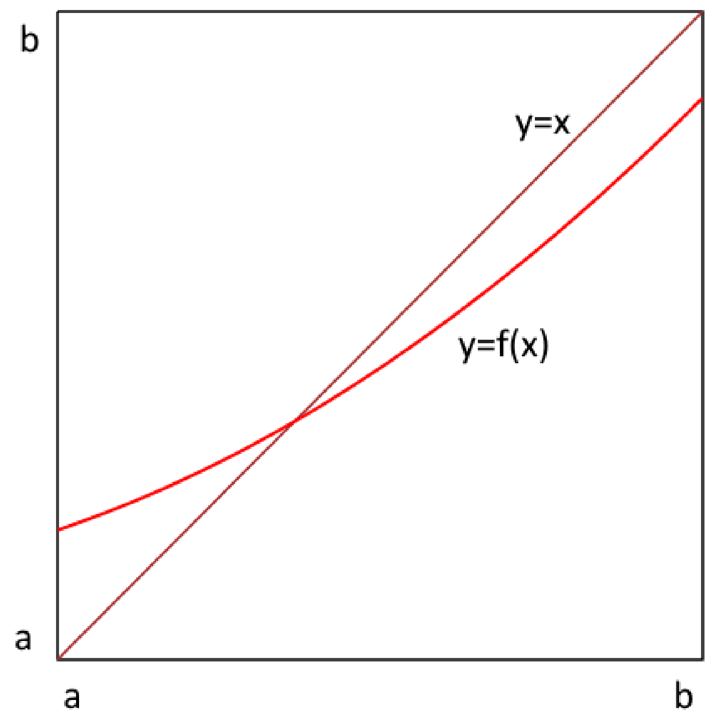

FIGURE 16. Graph of a continuous function intersects the diagonal.

Let $f : [a, b] \to [a, b]$ be a continuous function. Then, there exists $x_0 \in [a, b]$ with $f(x_0) = x_0$. Consider $g(x) = f(x) - x$. Then, $g(a) = f(a) - a \geq 0$ and $g(b) = f(b) - b \leq 0$. Since g is also continuous on $[a, b]$, by IVTh there exists $x_0 \in [a, b]$ with $g(x_0) = 0$ or $f(x_0) = x_0$.

We see this in the Figure 16. The graph of a continuous function has to intersect the diagonal.

A more difficult to prove is the following

Proposition 8.5. *Let $f : [a,b] \to [a,b]$ be an increasing function. Then, there exists $x_0 \in [a,b]$ with $f(x_0) = x_0$.*

Example 8.2. ***Cutting a piece of bread in halves:***

Let A be a bounded convex region of the plain $\mathbb{R}^2$. See Figure 17. For any fixed line L_0 there exists a line L parallel to L_0 cutting A into two halves of equal areas.

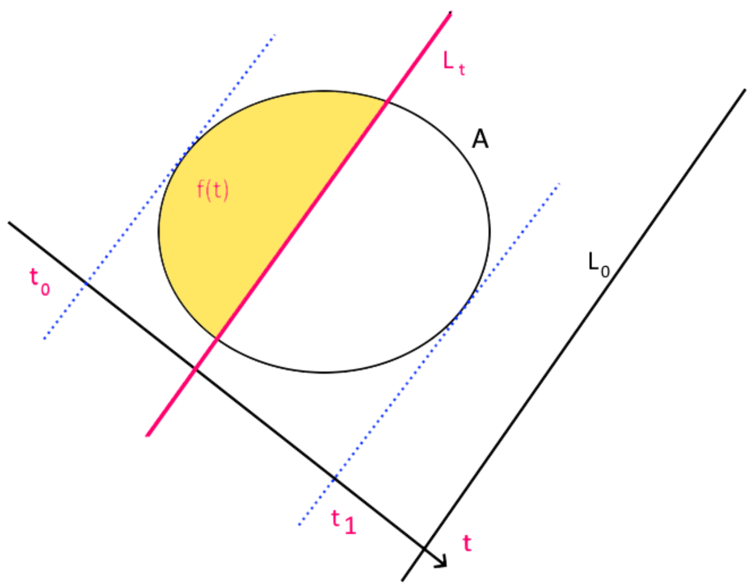

FIGURE 17. Cutting A into halves of equal areas.

We draw a t-coordinate axis perpendicular to L_0. Let the projection of A onto the axis be the interval $[t_0, t_1]$. We draw a line L_t, $t_0 \leq t \leq t_1$, parallel to the line L_0 and cutting the t-axis at point t. Let $f(t)$ be the **relative** area of the part of A to the left and above the line L_t (yellow). By geometric considerations, we can assume that f is a continuous function of t. We have $f(t_0) = 0$ and $f(t_1) = 1$, so by the IVTh, there exists a $t_0 < t_s < t_1$ with $f(t_s) = 1/2$. The line $L = L_{t_s}$ cuts A into halves of equal areas.

Remark: In a similar way we can prove that for a given point x_0 we can find a line L going through the point x_0 and cutting A into halves of equal areas.

Example 8.3. ***Cutting a ham sandwich in halves:***

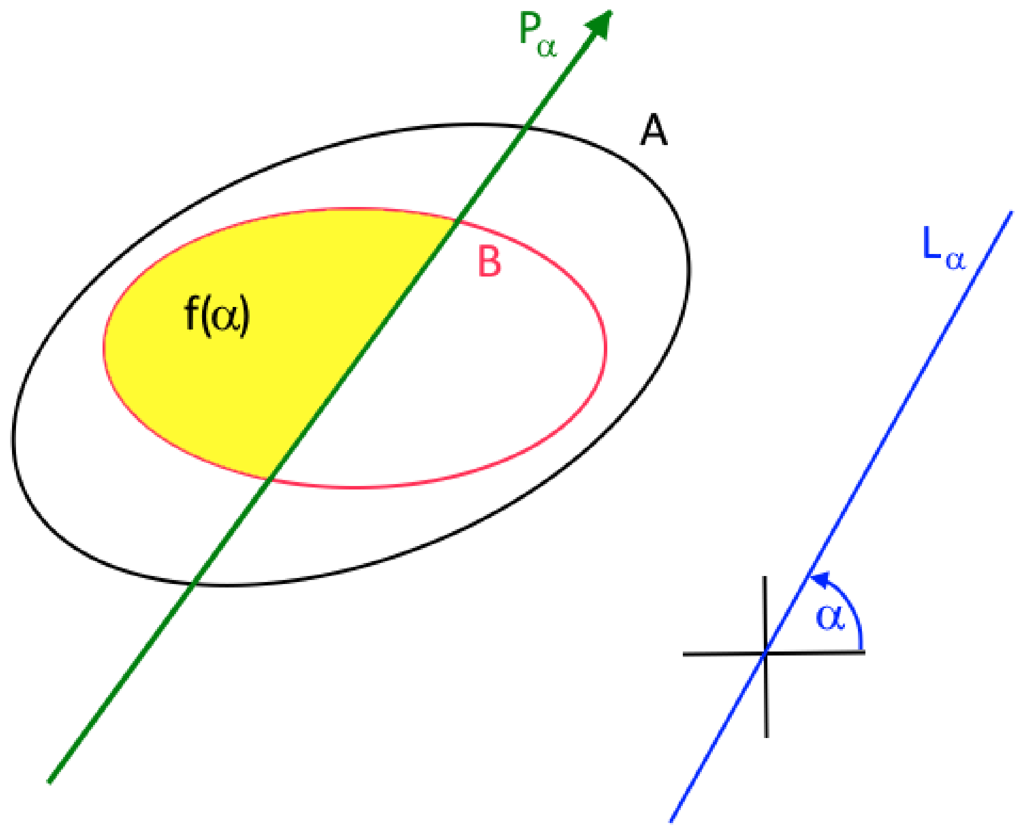

FIGURE 18. Cutting the ham sandwich in halves.

We want to cut the ham sandwich with one line in such a way that each half has the same amount of bread and the same amount of ham. More mathematically: Let A and B be bounded convex regions of the plain $\mathbb{R}^2$. For analogy with the ham sandwich, we will assume that $B \subset A$. See Figure 18. We will show that there exists a line P which at the same time cuts A into two halves of equal areas and cuts B into two halves of equal areas.

Let us consider a family of lines L_α making an angle α with some fixed axis, $0 \leq \alpha \leq \pi$. From Example 8.2 we know that for any L_α we can find line P_α dividing the region A into halves of equal areas and parallel to L_α. We can set an orientation on P_α so we can say what is on the RHS or LHS of P_α. Let $f(\alpha)$ denote the **relative** area of the part of region B on the LHS of P_α (yellow). By geometric considerations we can assume that $f(\alpha)$ is a continuous function of α.

If $f(0) = 1/2$ we found the required line P (note that each P_α cuts region A in halves). Let us assume that $f(0) < 1/2$. Then, $f(\pi) > 1/2$. By IVTh there is an $0 < \alpha_0 < \pi$ such that $f(\alpha_0) = 1/2$, i.e., the line $P = P_{\alpha_0}$ cuts B into halves of equal areas, at the same time cutting A into halves of equal areas. The case $f(0) > 1/2$ is done the same way.

8.4. **Two Definitions of Uniform Continuity.** There are two equivalent definitions of the uniform continuity of a function $f : A \to \mathbb{R}$:

Definition 8.2. *(1U) Cauchy (or $\varepsilon - \delta$) definition: f is uniformly continuous on set $A \subset \mathbb{R} \iff$*

$$\forall_{\varepsilon>0} \exists_{\delta>0} \forall_{x,y\in A} \ |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon \,.$$

(2U) Heine (or sequential) definition: f is uniformly continuous on set $A \subset \mathbb{R} \iff$ for any sequences $\{x_n\}, \{y_n\}$ (contained in A) such that $|x_n - y_n| \underset{n\to\infty}{\longrightarrow} 0$ we have $|f(x_n) - f(y_n)| \underset{n\to\infty}{\longrightarrow} 0$.

(We do not make any other assumptions on these sequences, in particular they do not have to be convergent.)

Example 8.4.

We will consider the function $f(x) = x^2$ on $\mathbb{R}$ and show how δ in the definition of continuity at x_0 depends on x_0. This will indicate that $f(x) = x^2$ is not uniformly continuous on $\mathbb{R}$. Let us fix an $\varepsilon > 0$. We have $|f(x) - f(x_0)| = |x^2 - x_0^2| = |x - x_0||x + x_0|$. For x close to x_0 we have $|x + x_0| \sim 2|x_0|$. Assuming $|x - x_0| < \delta$ we want $|x - x_0||x + x_0| < \varepsilon$ or $\delta \cdot 2|x_0| < \varepsilon$. Thus, for fixed $\varepsilon > 0$ we need $\delta < \varepsilon/(2|x_0|)$. We see that for $f(x) = x^2$, δ depends on the point x_0 at which which consider continuity.

Theorem 21. *Definitions (1U) and (2U) are equivalent.*

Proof. The proof is quite similar to the proof of Theorem 16.

(1U) $\Longrightarrow$ (2U) : Let $\{x_n\}$ and $\{y_n\}$ be such that $|x_n - y_n| \underset{n\to\infty}{\longrightarrow} 0$. We need to prove that $|f(x_n) - f(y_n)| \underset{n\to\infty}{\longrightarrow} 0$, i.e.,

$$(*U)\ \forall_{\varepsilon>0}\ \exists_{N\geq 1}\ \forall_{n\geq N}\ |f(x_n) - f(y_n)| < \varepsilon\,.$$

Let us fix an $\varepsilon > 0$. By (1U), we can find a $\delta > 0$ such that $|x_n - y_n| < \delta \implies |f(x_n) - f(y_n)| < \varepsilon$. Since $|x_n - y_n| \underset{n\to\infty}{\longrightarrow} 0$, we can find an $N \geq 1$ such that for $n \geq N$ we have $|x_n - y_n| < \delta$ and then $|f(x_n) - f(y_n)| < \varepsilon$. $(*U)$ has been proved.

(2U) $\Longrightarrow$ (1U) : We will prove contrapositive statement $\neg(1U) \Longrightarrow \neg(2U)$. Let us assume that (1U) does not hold, i.e.,

$$\exists_{\varepsilon>0}\ \forall_{\delta>0}\ \exists_{x,y\in A}\ |x - y| < \delta \text{ and } |f(x) - f(y)| \geq \varepsilon\,.$$

Let $\varepsilon_0 > 0$ be the ε whose existence is claimed above. It says "for any δ" so we will use a sequence od δ's. Let $\delta_n = 1/n > 0$, $n = 1, 2, \ldots$. For each δ_n we can find an $x_n, y_n \in A$ such that $|x_n - y_n| < \delta_n = 1/n$ and $|f(x_n) - f(y_n)| \geq \varepsilon_0$. Thus, we have $|x_n - y_n| \underset{n\to\infty}{\longrightarrow} 0$, but $|f(x_n) - f(y_n)| \underset{n\to\infty}{\not\longrightarrow} 0$. We proved $\neg(2U)$. □

8.5. **Main Theorems about Uniformly Continuous Functions.**

Theorem 22. *If f satisfies Lipschitz or Hölder condition on A,*

$$|f(x) - f(y)| \leq L|x - y|^\alpha\ ,\ \ x, y \in A\,,\ \ 0 < \alpha \leq 1,$$

then f is uniformly continuous on A.

Proof. We use Cauchy definition (1U). Let us fix an $\varepsilon > 0$. Set $\delta = (\varepsilon/L)^{1/\alpha}$. If $|x - y| < \delta$, then

$$|f(x) - f(y)| \leq L|x - y|^\alpha < L \cdot \delta^\alpha = \varepsilon\,.$$

□

Example 8.5.

Let $f(x) = 1/x^{2014}$, $x \in [1, +\infty)$. We will show that f is uniformly continuous on $[1, +\infty)$. We will show Lipschitz inequality and invoke Theorem 22. By Mean Value Theorem we have

$$|f(x) - f(y)| = |f'(c)||x - y| = \left| \frac{-2014}{c^{2015}} \right| |x - y| \leq 2014|x - y|,$$

since $c \geq 1$. So f satisfies Lipschitz inequality with $L = 2014$.

Theorem 23. *If f is continuous on a closed bounded interval $[a, b]$, then f is uniformly continuous on $[a, b]$.*

Proof. The proof is immediate if we use Heine definition (2U): Assume that f is not uniformly continuous on $[a, b]$, i.e, There exist sequences $\{x_n\}, \{y_n\}$ (contained in $[a, b]$) such that $|x_n - y_n| \underset{n\to\infty}{\longrightarrow} 0$ and $|f(x_n) - f(y_n)| \underset{n\to\infty}{\not\longrightarrow} 0$.

Since $|f(x_n) - f(y_n)| \underset{n\to\infty}{\not\longrightarrow} 0$ there is a subsequence of natural numbers $\{n_k\}$ such that $|f(x_{n_k}) - f(y_{n_k})| \geq \eta$ for some $\eta > 0$. Both sequences $\{x_{n_k}\}$ and $\{y_{n_k}\}$ are bounded (contained in $[a, b]$) so we can find a common convergent subsequences $\{x_{n_{k_j}}\}$ and $\{y_{n_{k_j}}\}$. Since $|x_n - y_n| \underset{n\to\infty}{\longrightarrow} 0$, they both converge to the same point, say $c \in [a, b]$, i.e., $x_{n_{k_j}} \underset{j\to\infty}{\longrightarrow} c$ and $y_{n_{k_j}} \underset{j\to\infty}{\longrightarrow} c$. Since f is continuous we have $f(x_{n_{k_j}}) \underset{j\to\infty}{\longrightarrow} f(c)$ and $f(y_{n_{k_j}}) \underset{j\to\infty}{\longrightarrow} f(c)$ but this contradicts $|f(x_{n_k}) - f(y_{n_k})| \geq \eta$. □

Example 8.6.

Consider $f(x) = \sqrt{x}$ for $x \in [0, 1]$. We know that f is continuous on $[0, 1]$ (it follows by inequality $\sqrt{x} - \sqrt{y} \leq \sqrt{x - y}$ for $x \geq y$). By Theorem 23 f is uniformly continuous on $[0, 1]$. Note that f DOES NOT satisfy Lipschitz inequality on $[0, 1]$. For the proof: Assume that it does. Then, in particular, there exist a constant L such that $\sqrt{x} - 0 \leq L(x - 0)$ or $\frac{1}{\sqrt{x}} \leq L$ for all $x \in [0, 1]$ which is impossible.

Example 8.7.

Consider a function $f(x) = \frac{x \cos(x^6)}{x^2+1}$, $x \in \mathbb{R}$. We will show that f is uniformly continuous on $\mathbb{R}$. We will use Cauchy definition (1U) and Theorem 23. Let us fix

$\varepsilon > 0$. Since cos is bounded we have $\lim_{x\to\pm\infty} f(x) = 0$, i.e., there exists an $M > 0$ such that for any $x \in (M, +\infty)$ we have $|f(x)| < \varepsilon/2$ and also for any $x \in (-\infty, -M)$ we have $|f(x)| < \varepsilon/2$. This means that

$$(\heartsuit) \quad \text{for any } x, y \in (M, +\infty) \quad \text{we have } |f(x) - f(y)| < \varepsilon\,,$$

and also

$$(\clubsuit) \quad \text{for any } x, y \in (-\infty, -M) \quad \text{we have } |f(x) - f(y)| < \varepsilon\,.$$

Consider f on the interval $I = [-M - 3, M + 3]$. f is continuous on I (as a combination of continuous functions) so by Theorem 23 f is uniformly continuous on I. This means that for our ε we can find a $\delta > 0$ such that

$$(\diamondsuit) \text{ for any } x, y \in I, |x - y| < \delta \quad \text{implies } |f(x) - f(y)| < \varepsilon\,.$$

We can assume that $\delta < 3$. If it is not, then we make it smaller and it will also work.

We will show that this δ works on the whole $\mathbb{R}$. Let $|x - y| < \delta$. Assuming $x \leq y$, there are five possibilities :

(a) $x, y \in (-\infty, -M)$. Then, $|f(x) - f(y)| < \varepsilon$ by $(\clubsuit)$;

(b) $x \in (-\infty, -M)$, y outside. Then, since $\delta < 3$ both $x, y \in [-M - 3, M + 3]$ and $|f(x) - f(y)| < \varepsilon$ by $(\diamondsuit)$;

(c) $x, y \in [-M, M]$. Then, both $x, y \in [-M - 3, M + 3]$ and $|f(x) - f(y)| < \varepsilon$ by $(\diamondsuit)$;

(d) $y \in (M, +\infty)$, x outside. Then, since $\delta < 3$ both $x, y \in [-M - 3, M + 3]$ and $|f(x) - f(y)| < \varepsilon$ by $(\diamondsuit)$;

(e) $x, y \in (M, +\infty)$. Then, $|f(x) - f(y)| < \varepsilon$ by $(\heartsuit)$.

We proved that f is uniformly continuous on $\mathbb{R}$. Note that

$$f'(x) = \frac{(\cos x^6 - 6x^6 \sin x^6)(x^2 + 1) - 2x^2 \cos x^6}{(x^2 + 1)^2}\,,$$

is unbounded on $\mathbb{R}$ so f does not satisfy Lipschitz condition, i. e., Mean Value theorem method would not work to prove uniform continuity of this function.

Example 8.8.

Consider the function

$$f(x) = \begin{cases} 1/\ln x\ , & \text{for } 0 < x \le 1/2; \\ 0\ , & \text{for } x = 0. \end{cases}$$

It is easy to see that it is continuous on $[0, 1/2]$ so by Theorem 23 f is uniformly continuous on $[0, 1/2]$. We will show that it does not satisfy Hölder condition for any $0 < \alpha \le 1$. Let us assume that it does. Then, for $0 < x \le 1/2$ we have

$$|\frac{1}{\ln x}| \le H x^{\alpha},$$

for some $H > 0$. Then,

$$|\frac{1}{x^{\alpha} \ln x}| \le H,\ 0 < x \le 1/2,$$

which is impossible since $\lim_{x\to 0^+} x^{\alpha} \ln x = 0$, see Corollary 6.1.

Theorem 24. *If f is uniformly continuous on A and $\{x_n\}$ is a Cauchy sequence contained in A, then the sequence $\{f(x_n)\}$ is also Cauchy.*

Proof. We want to prove

$$\forall_{\ \varepsilon>0}\ \exists_{\ N\ge 1}\ \forall_{\ n,m\ge N}\ |f(x_n) - f(x_m)| < \varepsilon\,.$$

Let us fix an $\varepsilon > 0$. Since f is uniformly continuous, for this ε we can find a $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$. Since $\{x_n\}$ is Cauchy, there exists an $N \ge 1$ such that for any $n, m \ge N$ we have $|x_n - x_m| < \delta$. We see that this N works also for the sequence $\{f(x_n)\}$ and ε: if $n, m \ge N$, then $|x_n - x_m| < \delta$ and then $|f(x_n) - f(x_m)| < \varepsilon$. □

Theorem 24 can be used to show that a function IS NOT uniformly continuous.

Example 8.9.

Show that $f(x) = 1/x^{2014}$ is not uniformly continuous on $(0, 1]$. We take a sequence $x_n = 1/n$ contained in $(0, 1]$. It is Cauchy since it converges to 0. The sequence $f(x_n) = n^{2014}$ diverges to $+\infty$ so it is not Cauchy. By Theorem 24 f is not uniformly continuous on $(0, 1]$.

Another method to prove that a function is not uniformly continuous is just to use directly the definition. It often requires some ingenuity.

Example 8.10.

Show that $f(x) = \sin x^2$ is not uniformly continuous on $[0, +\infty)$. This is an example of a continuous bounded function, which is not uniformly continuous. Looking for a hint we calculate $f'(x) = 2x \cos x^2$ and see that the slope of f will be arbitrary large close to points where $\cos x^2 = 1$ or $x^2 = 2n\pi$ or $x = \sqrt{2n\pi}$. Let us define two sequences $x_n = \sqrt{2n\pi}$ and $y_n = \sqrt{2n\pi + a}$ for some small $a > 0$. We have

$$|x_n - y_n| = |\sqrt{2n\pi} - \sqrt{2n\pi + a}| = \left| \frac{2n\pi - 2n\pi - a}{\sqrt{2n\pi} + \sqrt{2n\pi + a}} \right| = \left| \frac{a}{\sqrt{2n\pi} + \sqrt{2n\pi + a}} \right| \to 0\,,$$

as $n \to +\infty$. On the other hand, we have

$$|f(x_n) - f(y_n)| = |\sin(2n\pi) - \sin(2n\pi + a)| = \sin a > 0\,.$$

By Heine definition (2U), f is not uniformly continuous on $[0, +\infty)$.

Example 8.11.

Show that $f(x) = x^{16}$ is not uniformly continuous on $[0, +\infty)$. We will use Cauchy definition (1U). Its negation is:

$$\exists_{\,\varepsilon>0}\ \forall_{\,\delta>0}\ \exists_{\,x,y\in A}\ |x - y| < \delta \text{ and } |f(x) - f(y)| \geq \varepsilon\,.$$

Let us consider points $x = z$ and $y = z + \delta/2$. Then, always $|x - y| < \delta$. We have

$$|f(y) - f(x)| = (z + \delta/2)^{16} - z^{16} > (z + \delta/2)z^{15} - z^{16} = z^{15}\delta/2\,.$$

Set $\varepsilon = 1$. For arbitrary $\delta > 0$ we can find z such that $|f(y) - f(x)| > 1$. We proved that f is not uniformly continuous on $[0, +\infty)$.

The following theorem complements Theorem 24.

Theorem 25. *If for any Cauchy sequence $(x_n)_{n\geq 1}$ the sequence $(f(x_n))_{n\geq 1}$ is also Cauchy, then f is uniformly continuous.*

Proof. We will do a proof by contrapositive. We assume that f is not uniformly continuous and we will show a Cauchy sequence $(x_n)_{n\geq 1}$ for which the sequence $(f(x_n))_{n\geq 1}$ is not Cauchy. If $f : [a,b] \to \mathbb{R}$ is not uniformly continuous, then

$$\exists_{\,\varepsilon>0}\ \forall_{\,\delta>0}\ \exists_{\,x,y\in[a,b]}\ |x-y|<\delta\ \wedge\ |f(x)-f(y)|\geq\varepsilon. \tag{10}$$

Fix an $\varepsilon_0 > 0$ given by formula (10). For $\delta_n = 1/n$, we can find $x_n, y_n \in [a,b]$ with $|x_n - y_n| < 1/n$ and $|f(x_n) - f(y_n)| \geq \varepsilon_0$. We can find a common convergent subsequences $(x_{n_k})_{k\geq 1}$ and $(y_{n_k})_{k\geq 1}$. Since $|x_{n_k} - y_{n_k}| < 1/n_k$, they have a common limit, say $\lim_{k\to\infty} x_{n_k} = c = \lim_{k\to\infty} y_{n_k}$. The sequence $(x_{n_1}, y_{n_1}, x_{n_2}, y_{n_2}, x_{n_3}, y_{n_3}, x_{n_4}, y_{n_4}, \dots)$ also converges to c so it is Cauchy. At the same time we have $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon_0$, for all $k \geq 1$, so the sequence $(f(x_{n_1}), f(y_{n_1}), f(x_{n_2}), f(y_{n_2}), f(x_{n_3}), f(y_{n_3}), f(x_{n_4}), f(y_{n_4}), \dots)$ is not Cauchy.

□

9. Derivatives

9.1. Definition and Basic Properties of Derivatives.

Definition 9.1. *Let $f : [a, b] \to \mathbb{R}$ and $x_0 \in (a, b)$. The derivative $f'(x_0)$ of f at the point x_0 is*

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

if the limit exists.

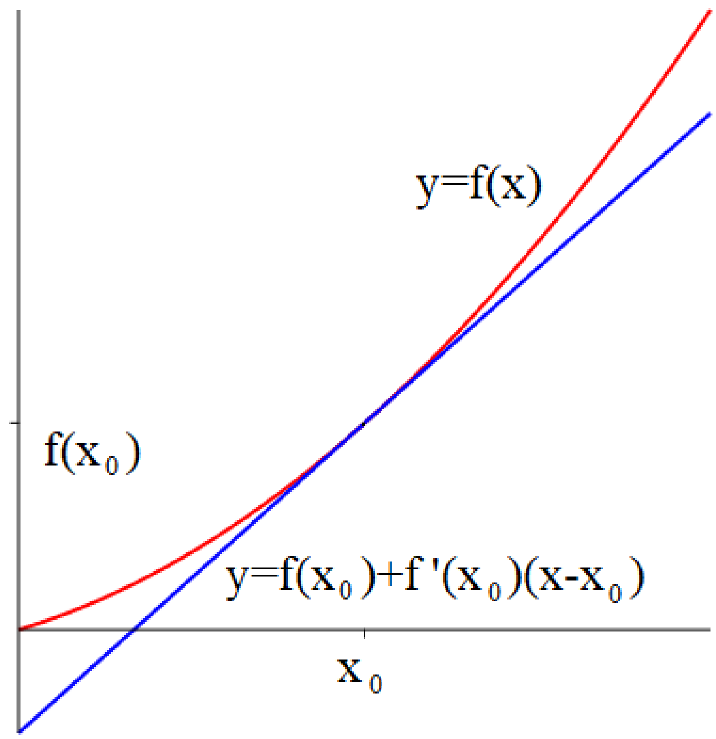

Figure 19. $f'(x_0)$ is the slope of the tangent line at x_0

Geometrically, the derivative $f'(x_0)$ is the slope of the line tangent to the graph of f at the point x_0. See Figure 19. If the derivative at x_0 exists, we say that f is differentiable at x_0. The function f is differentiable on the interval (a, b) when it is differentiable at any point of the interval.

In the similar way we define left hand sided and right hand sided derivatives.

Definition 9.2. *Let* $f : [a,b] \to \mathbb{R}$ *and* $x_0 \in (a,b)$*. The left hand sided derivative* $f'(x_0^-)$ *of* f *at the point* x_0 *is*

$$f'(x_0^-) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}.$$

if the limit exists.

The right hand sided derivative $f'(x_0^+)$ *of* f *at the point* x_0 *is*

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

if the limit exists.

Proposition 9.1. *Function* f *is differentiable at* $x_0 \iff$ *both one sided derivatives at* x_0 *exist and are equal,* $f'(x_0^-) = f'(x_0^+)$*.*

Proof. It is a corollary of Proposition 7.7. □

Often the definitions of derivative are formulated in the following form:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \; ; f'(x_0^{\pm}) = \lim_{h \to 0^{\pm}} \frac{f(x_0 + h) - f(x_0)}{h} \; .$$

Proposition 9.2. *If a function* f *is differentiable at* x_0*, then* f *is continuous at* x_0*.*

Proof. We have

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

□

Example 9.1.

Let

$$g(x) = \begin{cases} x^2 & , \text{ for } x \geq 0; \\ 0 & , \text{ for } x \geq 0. \end{cases}$$

We have $g'(0^-) = 0$, and

$$g'(0^+) = \lim_{h \to 0+} \frac{g(h) - g(0)}{h} = \lim_{h \to 0+} \frac{h^2}{h} = 0,$$

so $g'(0) = 0$.

Example 9.2.

Let

$$f(x) = \begin{cases} x \sin \frac{1}{x} & , \text{ for } x \neq 0; \\ 0 & , \text{ for } x = 0. \end{cases}$$

We have

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} \frac{h \sin \frac{1}{h}}{h},$$

which does not exist. Thus, f is not differentiable at 0, although it is continuous at 0.

Example 9.3.

Let

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

We have

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} \frac{h^2 \sin \frac{1}{h}}{h} = \lim_{h \to 0} h \sin \frac{1}{h} = 0.$$

Thus, f is differentiable at 0. For $x \neq 0$ we have

$$f'(x) = 2x \sin \frac{1}{x} + x^2 \cos \frac{1}{x} \left(\frac{-1}{x^2} \right) = 2x \sin \frac{1}{x} - \cos \frac{1}{x},$$

which does not have a limit at $x = 0$. Thus, f' is not continuous at 0.

Theorem 26. ***Rules of Differentiation:***

(a) $f(x) \equiv c \implies f'(x) = 0$;

(b) $f(x) = x \implies f'(x) = 1$;

Now, we assume that functions f and g are differentiable. Then,

(c) $(f \pm g)'(x) = f'(x) \pm g'(x)$;

(d) $(f \cdot g)'(x) = f(x) \cdot g'(x) + f'(x) \cdot g(x)$;

(e) $\left(\frac{1}{g}\right)'(x) = \frac{-g'(x)}{g^2(x)}$;

(f) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$;

Chain Rule : $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$.

Proof. We will prove (d), (e) and the Chain Rule.

(d) We have

$$\begin{aligned}(f \cdot g)'(x) &= \lim_{h\to 0} \frac{f(x+h)\cdot g(x+h) - f(x)\cdot g(x)}{h} \\ &= \lim_{h\to 0} \frac{f(x+h)\cdot g(x+h) - f(x+h)\cdot g(x) + f(x+h)\cdot g(x) - f(x)\cdot g(x)}{h} \\ &= \lim_{h\to 0} \frac{f(x+h)[g(x+h)-g(x)]}{h} + \lim_{h\to 0}\frac{[f(x+h)-f(x)]g(x)}{h} \\ &= f(x)g'(x) + f'(x)g(x).\end{aligned}$$

(e) We have

$$\left(\frac{1}{g}\right)'(x) = \lim_{h\to 0}\frac{1}{h}\left(\frac{1}{g(x+h)} - \frac{1}{g(x)}\right) = \lim_{h\to 0}\frac{1}{h}\left(\frac{g(x)-g(x+h)}{g(x+h)g(x)}\right) = \frac{-g'(x)}{g^2(x)}.$$

Chain Rule: We use Heine definition of a limit. Let $(x_n)_{n\geq 1}$ be any sequence such that $x_n \underset{n\to\infty}{\longrightarrow} x_0$ and $x_n \neq x_0$ for all $n \geq 1$. We need to show that

$$\frac{f(g(x_n)) - f(g(x_0))}{(x_n - x_0)} \underset{n\to\infty}{\longrightarrow} f'(g(x_0))\cdot g'(x_0).$$

We consider two cases: (i) for all $n \geq 1$ we have $g(x_n) = g(x_0)$. Since both f and g are differentiable, then $g'(x_0) = 0$ and $f'(g(x_0)) = 0$ and the formula holds.

(ii) We have $g(x_n) \neq g(x_0)$ for all $n \geq 1$. Then,

$$\frac{f(g(x_n)) - f(g(x_0))}{(x_n - x_0)} = \frac{f(g(x_n)) - f(g(x_0))}{g(x_n) - g(x_0)}\frac{g(x_n)-g(x_0)}{(x_n - x_0)} \underset{n\to\infty}{\longrightarrow} f'(g(x_0))g'(x_0).$$

We also used that fact that g is continuous at x_0. □

Example 9.4. $(x^n)' = nx^{n-1}$**,** $n = 1, 2, \ldots$

We will prove it by induction. We have $x' = 1$ so the formula works for $n = 1$. Assume that $(x^n)' = nx^{n-1}$, for some $n \geq 1$. We have

$$(x^{n+1})' = (x \cdot x^n)' = 1 \cdot x^n + x \cdot (nx^{n-1}) = (n+1)x^n,$$

which proves it for $n+1$. By the Principle of Induction the formula holds for all $n \geq 1$.

Example 9.5. $(e^x)' = e^x$

We need to show that for any $x \in \mathbb{R}$ we have $\lim_{h\to 0} \frac{e^{x+h}-e^x}{h} = e^x$. Since $e^{x+h} - e^x = e^x(e^h - 1)$ it is enough to show that

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1$$

. This follows from Example 6.2 using the method of 5.

We will now show that an inverse function of a differentiable function is also differentiable, under n additional assumption.

Theorem 27. *Let f be an invertible differentiable function and $f'(x_0) \neq 0$. The the inverse function f^{-1} is differentiable at $y_0 = f(x_0)$ and*

$$(f^{-1})'(y_0) = \frac{1}{f(x_0)}.$$

Proof. Since f is invertible we have $f(x) \neq f(x_0)$ in some neighbourhood of x_0. (This also follows by the fact that $f'(x_0) \neq 0$.) Let $y_0 = f(x_0)$ and $y = f(x)$. We can write

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} \underset{y\to y_0}{\longrightarrow} \frac{1}{f'(x_0)}.$$

□

Example 9.6. *The logarythm function* $\ln : (0, +\infty) \to \mathbb{R}$ *is differentiable and* $(\ln x)' = \frac{1}{x}$, $x \in \mathbb{R}$. *We have* $\ln x = y$ *if and only if* $x = e^y$, *i.e., logarythm and the exponential function are mutually inverse. By Theorem 27 and Example 9.5 we have*

$$(\ln x)' = \frac{1}{e^y} = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

Example 9.7. *The* arctan *function* $\arctan : (-\infty, +\infty) \to (-\pi/2, \pi/2)$ *is differentiable and*

$$(\arctan x)' = \frac{1}{1 + x^2}\ , \ x \in \mathbb{R}.$$

We have $\arctan x = y$ *if and only if* $x = \tan(y)$, *i.e., the* tan *and* arctan *functions are mutually inverse. We have*

$$(\tan y)' = \left(\frac{\sin y}{\cos y}\right)' = \frac{\cos y \cdot \cos y - \sin y \cdot (-\sin y)}{\cos^2 y} = 1 + \tan^2 y.$$

By Theorem 27 we have

$$(\arctan x)' = \frac{1}{1+\tan^2 y} = \frac{1}{1+x^2}.$$

Example 9.8. $(\sin x)' = \cos x$, $(\cos x)' = -\sin x$.

This follows by the trigonometric formulas of 7 and Example 4.1.

Example 9.9. $(x^\alpha)' = \alpha x^{\alpha-1}$, $\alpha \in \mathbb{R}$, $x \in (0, +\infty)$.

We will use the Chain Rule 26 and the just developed formulas. We have

$$(x^\alpha)' = (e^{\alpha \ln x})' = e^{\alpha \ln x} \cdot \alpha \frac{1}{x} = \alpha x^{\alpha-1}.$$

Example 9.10. $f(x) = \sin\left(\frac{x^2}{\cos x^3}\right)$.

We will use the Chain Rule and some of the formulas above. We have

$$f'(x) = \cos\left(\frac{x^2}{\cos x^3}\right) \cdot \left(\frac{2x \cdot \cos x^3 - x^2 \cdot (-(\sin x^3)3x^2)}{\cos^2 x^3}\right).$$

9.2. **Applications of Derivatives.**

Definition 9.3. ***Maximum Point and Minimum Point:*** *Let $f : D \to \mathbb{R}$ and $A \subset D$. We say that $x_0 \in A$ is a maximum point of f on $A \iff f(x_0) \geq f(x)$ for all $x \in A$.*

Similarly, we say that $x_0 \in A$ is a minimum point of f on $A \iff f(x_0) \leq f(x)$ for all $x \in A$.

Definition 9.4. ***Local Maximum Point and Local Minimum Point:*** *Let $f : D \to \mathbb{R}$. We say that $x_0 \in D$ is a local maximum point of f on $D \iff f(x_0) \geq f(x)$ for all $x \in (x_0 - \delta, x_0 + \delta)$, for some $\delta > 0$.*

Similarly, we say that $x_0 \in D$ is a local minimum point of f on $D \iff f(x_0) \leq f(x)$ for all $x \in (x_0 - \delta, x_0 + \delta)$, for some $\delta > 0$.

Definition 9.5. ***Critical point:*** *Let $f : D \to \mathbb{R}$. A point $x_0 \in D$ is called a critical point of $f \iff f'(x_0) = 0$.*

Theorem 28. ***Local maximum or local minimum implies critical point.***

If x_0 is a local maximum or local minimum point of f, and f is differentiable at x_0, then x_0 is a critical point of f, i.e., $f'(x_0) = 0$.

Proof. Let x_0 be a local maximum point, i.e., $f(x_0) \geq f(x)$ for $x \in (x_0 - \delta, x_0 + \delta)$, for some $\delta > 0$. When we calculate a limit at x_0 it is enough to consider only points $x \in (x_0 - \delta, x_0 + \delta)$. Then, we have

$$f'(x_0^-) = \lim_{x \to x_0 \ , \ x < x_0} \frac{f(x_0) - f(x)}{x_0 - x} \geq 0,$$

and

$$f'(x_0^+) = \lim_{x \to x_0 \ , \ x > x_0} \frac{f(x_0) - f(x)}{x_0 - x} \leq 0.$$

Since f is differentiable at x_0, both limits are equal, so we have $f'(x_0) = 0$.

The proof for a minimum point is similar. □

Remark 9.1. *The statement reverse Theorem 28 is not true. The function $f(x) = x^3$ is strictly increasing on $\mathbb{R}$ and $f'(0) = 0$.*

Theorem 29. ***Rolle's Theorem.*** *If f is continuous on $[a, b]$ and differentiable on (a, b) and $f(a) = f(b)$, then there exists a point $x_0 \in (a, b)$ such that $f'(x_0) = 0$.*

Proof. Since f is continuous on $[a, b]$ it attains its maximum value and its minimum value, by Theorem 19.

If f is constant on $[a, b]$, then $f'(x) = 0$ on (a, b).

If for some $x \in (a, b)$ we have $f(x) > f(a)$, then f attains its maximum at some point $x_0 \in (a, b)$ and $f'(x_0) = 0$, by Theorem 28. See Figure 20.

If for some $x \in (a, b)$ we have $f(x) < f(a)$, then f attains its minimum at some point $x_0 \in (a, b)$ and $f'(x_0) = 0$, again by Theorem 28. □

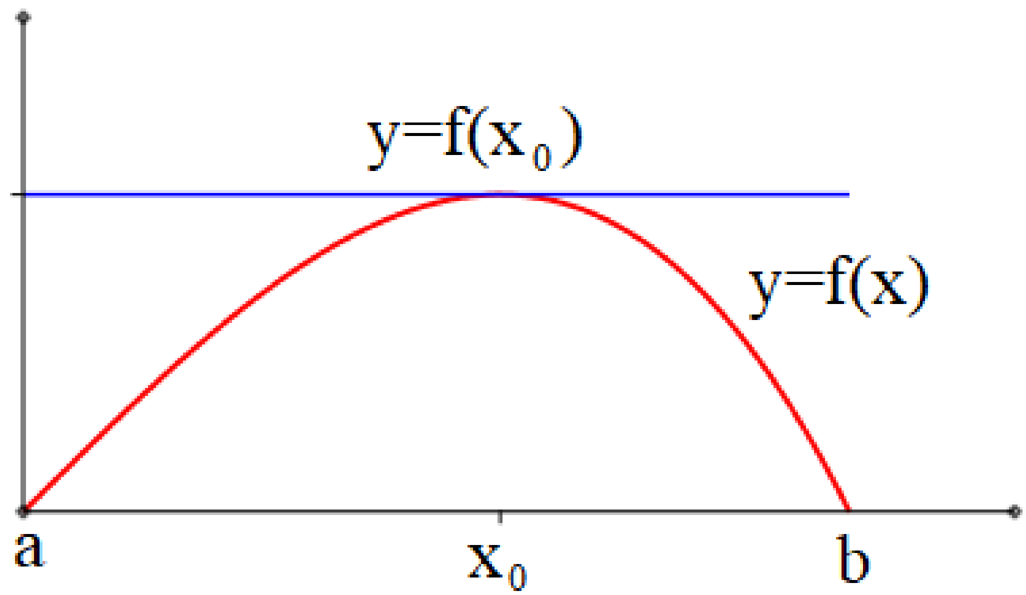

FIGURE 20. Rolle's theorem.

Example 9.11.

Show that the function $f(x) = \ln x + 2x - x^2$ has exactly two zeros in the interval $(0, 3)$. We see that $f(1/e) = -1 + 2/e - 4/e^2 < 0$ since $2 < e$. We also have $f(1) = 0 + 2 - 1 > 0$ and $f(3) = \ln 3 + 6 - 9 < 0$ since $\ln 3 \sim 1$. By the Intermediate Value Theorem f has at least two zeros in $(0, 3)$. We will use Rolle's theorem to show that f does not have more zeros in $(0, 3)$. Assume that f has three zeros in $(0, 3)$. Then, by Rolle's Theorem its derivative $f'(x) = 1/x + 2 - 2x$ has two zeros in $(0, 3)$ and its second derivative $f''(x) = -1/x^2 - 2$ has one zero in $(0, 3)$. This is impossible, since $-1/x^2 - 2 < 0$ for all $x \in \mathbb{R}$. Thus, f has exactly two zeros in the interval $(0, 3)$.

Theorem 30. ***Lagrange Theorem or Mean Value Theorem***

If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists a point $x_0 \in (a, b)$ such that

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

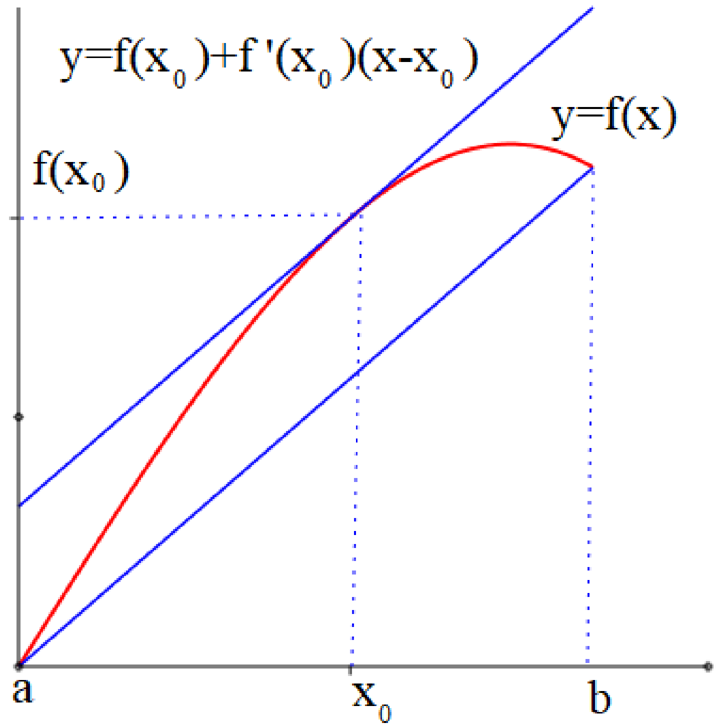

FIGURE 21. Mean Value Theorem

Proof. Geometrically the Lagrange theorem is Rolle's theorem from the point of view of the secant line going through points $(a, f(a))$ and $(b, f(b))$, or "rotated" Rolle's theorem. We use this in the proof.

Consider

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

Then, g is continuous on $[a, b]$, differentiable on (a, b) and $g(a) = 0 = g(b)$. By Rolle's Theorem 29, at some point $x_0 \in (a, b)$ we have $g'(x_0) = 0$ or $f'(x_0) = \frac{f(b)-f(a)}{b-a}$. □

Applications of Mean Value Theorem:

If f is continuous on $[a,b]$ and differentiable on (a,b), then for any $x \in [a,b]$ we can write

$$f(x) = f(a) + f'(c)(x-a),$$

for some $c \in (a,x)$.

This implies

$$|f(x) \leq |f(a)| + \sup_{(a,x)} |f'| \cdot |x-a|.$$

Corollary 9.1.

If f and g are continuous on $[a,b]$ and differentiable on (a,b), then

(a) If $f' \equiv 0$, then f is constant on $[a,b]$.

(b) If $f' = g'$ on (a,b), then $f = g + \text{const}$ on $[a,b]$.

(c) If $f' \geq 0$ on (a,b), then f is increasing on $[a,b]$.

Let $a \leq x < y \leq b$. We have $f(y) - f(x) = f'(c) \cdot (y-x) \geq 0$, $x < c < y$.

(d) If $f' \leq 0$ on (a,b), then f is decreasing on $[a,b]$.

The proof is similar as for (c).

(e) If $f(a) = g(a)$ and $f'(x) \leq g'(x)$ on (a,b) then $f(b) \leq g(b)$.

This follows by (c) applied to $g - f$.

Remark 9.2. *The statements reverse to (c) and (d) follow by the definition of the derivative.*

Example 9.12. *Show that*

$$\frac{b-a}{b} < \ln\frac{b}{a} < \frac{b-a}{a}\ ,\ \ 0 < a < b. \tag{11}$$

Since $(\ln x)' = \frac{1}{x}$, *by Mean Value theorem, we have*

$$\ln\frac{b}{a} = \ln b - \ln a = \frac{1}{c}(b-a),$$

for some $a < c < b$. *This implies (11).*

Example 9.13. *Show that*

$$|\arctan b - \arctan a| < \frac{1}{1+a^2}(b-a)\ ,\ 0 < a < b. \tag{12}$$

Since $(\arctan x)' = \frac{1}{1+x^2}$ *this follows by Mean Value theorem.*

Example 9.14.

$$e^x > 1 + x\ , x \neq 0.$$

Let $f(x) = e^x$ *and* $g(x) = 1 + x$. *We have* $f(0) = 1 = g(0)$ *and* $f'(x) = e^x$, $g'(x) = 1$.

For $x > 0$ *we have* $f'(x) > g'(x)$ *and the claim follows by part (e), slightly modified.*

For $x < 0$ *we have* $f'(x) < g'(x)$ *and the claim again follows by part (e), used "backwards".*

Example 9.15.

$$x - \frac{x^2}{2} < \ln(1+x) < x\ ,\ x > 0.$$

Let $f(x) = x - \frac{x^2}{2}$, $g(x) = \ln(1+x)$ *and* $h(x) = x$. *We have* $f(0) = h(0) = g(0) = 0$ *and* $f'(x) = 1 - x$, $g'(x) = \frac{1}{1+x}$ *and* $h'(x) = 1$.

For $x > 0$ *we have* $f'(x) < g'(x) < h'(x)$ *and the claim again follows by part (e).*

Example 9.16.

$$x - \frac{x^3}{6} < \sin x < x\ ,\ x > 0.$$

First, we will show $\sin x < x\ ,\ x > 0$. *We know this inequality from Example 4.1, but we will reprove it.*

Let $g(x) = \sin x$ *and* $h(x) = x$. *We have* $h(0) = g(0) = 0$ *and* $g'(x) = \cos x$ *and* $h'(x) = 1$. *Using part (e) of Corollary 9.2 we obtain that* $\sin x < x$ *for* $0 < x < 2\pi$. *For larger values of* x *it holds as* $\sin x \leq 1$. *Now, we prove*

$$x - \frac{x^3}{6} < \sin x\ ,\ x > 0.$$

We define $f(x) = x - \frac{x^3}{6}$, $g(x) = \sin x$. *We have* $f(0) = g(0) = 0$ *and* $f'(x) = 1 - \frac{x^2}{2}$, $g'(x) = \cos x$. *To show that* $1 - \frac{x^2}{2} < \cos x$, *for* $x > 0$ *we introduce* $f_1(x) = 1 - \frac{x^2}{2}$, $f_2(x) = \cos x$. *We have* $f_1(0) = 1 = f_2(0)$ *and* $f_1'(x) = -x$, $f_2'(x) = -\sin x$. *We have* *just proved* $\sin x < x$ *so* $f_1'(x) < f_2'(x)$, $x > 0$. *Thus, by part (e) we have* $1-\frac{x^2}{2} < \cos x$ *and again by part (e) we obtain* $f(x) < g(x)$ *or* $x - \frac{x^3}{6} < \sin x$ *for* $x > 0$.

Example 9.17.

Show that there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that

$$(f(x))^7 + (f(x))^5 = x, \ x \in \mathbb{R}.$$

Let us consider $g(x) = x^7+x^5$. Since $g'(x) = 7x^6+5x^4 \geq 0$, the function g is increasing and hence invertible. Let $f = g^{-1}$. Then, $g(f(x) = x$ or $(f(x))^7+(f(x))^5 = x, \ x \in \mathbb{R}$.

Example 9.18.

Show: Let $f : [0,1] \to \mathbb{R}$ be twice differentiable. If $f(0) = 0, \ f(1) = 1, \ f'(0) = f'(1) = 0$, then $|f''(x)| \geq 4$, for some $x \in [0,1]$.

Let us assume that $|f''(x)| < 4$ for $x \in [0,1]$. Then,

$$\frac{f'(x) - f'(0)}{x - 0} = f''(\bar{x}) < 4.$$

Thus, $f'(x) < 4x$, for $x \in [0,1]$. Then, using $f(0) = 0$ we obtain $f(x) < 2x^2$ and $f(1/2) < 1/2$.

Let $g(x) = 1 - f(1-x)$. Then, $g(0) = 0$, $g'(0) = -f'(1-0)(-1) = 0$. As we did with the function f, also for g we obtain $g(1/2) < 1/2$. This implies $1 - f(1/2) < 1/2$ or $f(1/2) > 1/2$ and we have a contradiction.

Definition 9.6. ***Higher order derivatives:*** *Higher order derivatives of a function* f *are defined by*

$$f^{(n)} = (f^{n-1})' \ , n = 1, 2, 3, \ldots$$

In particular, $f^{(0)} = f$, $f^{(1)} = f'$, $f^{(2)} = f''$, *etc.*

Theorem 31. ***Second derivative test for maxima, minima:***

Let $f : (a, b) \to \mathbb{R}$ be differentiable twice and $f'(x_0) = 0$ for some $x_0 \in (a, b)$.

(a) *If $f''(x_0) > 0$, then f has a local minimum at x_0.*

(b) *If $f''(x_0) < 0$, then f has a local maximum at x_0.*

Proof. We prove (a). We have

$$f''(x_0) = \lim_{h \to 0} \frac{f'(x_0 + h) - f'(x_0)}{h} = \lim_{h \to 0} \frac{f'(x_0 + h)}{h} > 0,$$

which means that

for $h < 0$, $f'(x_0 + h) < 0$, i.e, the function f is decreasing to the left of x_0,

for $h > 0$, $f'(x_0 + h) > 0$, i.e, the function f is increasing to the right of x_0.

Together, these facts show that x_0 is a local minimum point of f. □

Theorem 32. *Let $f : (a, b) \to \mathbb{R}$ be differentiable twice and $x_0 \in (a, b)$.*

(a) *If f has a local minimum at x_0, then $f''(x_0) \geq 0$.*

(b) *If f has a local maximum at x_0, then $f''(x_0) \leq 0$.*

Proof. This follows by the definition of the second derivative. □

Theorem 33. ***Generalized (or Cauchy) Mean Value Theorem:*** *Assume that f and g are continuous on $[a, b]$ and differentiable on (a, b). Then, there exists $x_0 \in (a, b)$ such that*

$$(f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0),$$

or, if we can divide,

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

Proof. Consider

$$h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)).$$

We have

$$h(a) = f(a)g(b) - f(a)g(a) - g(a)f(b) + g(a)f(a) = f(a)g(b) - g(a)f(b),$$

and

$$h(b) = f(b)g(b) - f(b)g(a) - g(b)f(b) + g(b)f(a) = -f(b)g(a) + g(b)f(a) = h(a).$$

By Rolle's theorem 29 $h'(x_0) = 0$, for some $x_0 \in (a, b)$. This implies the claim. □

Theorem 34. ***L'Hôspital Rule:*** *Assume that f and g are continuous on $[a, b]$ and differentiable on (a, b). Let $x_0 \in (a, b)$, $\lim_{x\to x_0} f(x) = 0 = \lim_{x\to x_0} g(x)$ and $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ exists. If $g'(x) \neq 0$ in some neighbourhood of x_0, then*

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}.$$

Proof. Since $g'(x) \neq 0$ in some neighbourhood of x_0, $g(x) - g(x_0)$ cannot be 0 in this neighbourhood (Mean Value theorem). We have

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(\bar{x})}{g'(\bar{x})},$$

where $\bar{x}$ is between x and x_0, by Generalized Mean Value theorem. Note that if $x \to x_0$, then $\bar{x} \to x_0$, as well. This proves the claim. □

Example 9.19.

We have $\lim_{x\to 0}(1 - \cos x) = 0 = \lim_{x\to 0} x^2$, and using L'Hôspital Rule we obtain

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2} = \lim_{x\to 0} \frac{\sin x}{2x} = \frac{1}{2}.$$

Corollary 9.2. ***L'Hôspital Rule at infinity:*** *Assume that f and g are continuous and differentiable on $(0, +\infty)$. Let $\lim_{x\to +\infty} f(x) = 0 = \lim_{x\to +\infty} g(x)$ and $\lim_{x\to +\infty} \frac{f'(x)}{g'(x)}$ exists. If $g'(x) \neq 0$ in some neighbourhood of $+\infty$, then*

$$\lim_{x\to +\infty} \frac{f(x)}{g(x)} = \lim_{x\to +\infty} \frac{f'(x)}{g'(x)}.$$

Proof. Introducing $x = \frac{1}{t}$ we have $\lim_{\ t\to 0^+} f(\frac{1}{t}) = 0 = \lim_{\ t\to 0^+} g(\frac{1}{t})$ and by L'Hôspital Rule

$$\lim_{t\to 0^+} \frac{f(\frac{1}{t})}{g(\frac{1}{t})} = \lim_{t\to 0^+} \frac{(f(\frac{1}{t}))'}{(g(\frac{1}{t}))'} = \lim_{t\to 0^+} \frac{f'(\frac{1}{t})\frac{-1}{t^2}}{g'(\frac{1}{t})\frac{-1}{t^2}} = \lim_{t\to 0^+} \frac{f'(\frac{1}{t})}{g'(\frac{1}{t})},$$

which is equivalent to the claim. □

Theorem 35. ***More difficult version of L'Hôspital Rule:*** *Assume that f and g are continuous on $[a,b]$ and differentiable on (a,b). Let $x_0 \in (a,b)$, $\lim_{x\to x_0} g(x) = +\infty$ and $\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = A$ exists. If $g'(x) \neq 0$ in some neighbourhood of x_0, then*

$$\lim_{x\to x_0} \frac{f(x)}{g\,(x)} = A.$$

Note, that x_0 can be $\pm\infty$.

Proof. We will consider only left hand side limits. Let $\varepsilon > 0$ be fixed.

There exists $x_1 < x_0$ such that for $x_1 < x < x_0$ we have

$$A - \varepsilon < \frac{f'(x)}{g'(x)} < A + \varepsilon.$$

Since $\lim_{x\to x_0^-} g(x) = +\infty$ there exists $x_1 < x_2 < x_0$ such that for $x_2 < x < x_0$ we have $g(x) > g(x_1)$. We can also assume $g(x_1) > 0$. By Generalized Mean Value theorem we have

$$\frac{f(x_1) - f(x)}{g(x_1) - g(x)} = \frac{f'(\zeta)}{g'(\zeta)},\ x_2 < x < x_0,\ x_1 < \zeta < x.$$

Thus, for $x_2 < x < x_0$, we have

$$A - \varepsilon < \frac{f(x_1) - f(x)}{g(x_1) - g(x)} < A + \varepsilon.$$

We have

$$\frac{f(x_1) - f(x)}{g(x_1) - g(x)} = \frac{f(x_1)}{g(x_1) - g(x)} - \frac{f(x)}{g(x_1) - g(x)}\frac{g(x)}{g(x)},$$

so

$$-\frac{f(x_1)}{g(x_1) - g(x)} + A - \varepsilon < -\frac{f(x)}{g(x_1) - g(x)}\frac{g(x)}{g(x)} < A + \varepsilon - \frac{f(x_1)}{g(x_1) - g(x)}.$$

Multiplying by $\frac{g(x)-g(x_1)}{g(x)}$ we obtain

$$\frac{g(x)-g(x_1)}{g(x)}\left(-\frac{f(x_1)}{g(x_1)-g(x)}+A-\varepsilon\right)<\frac{f(x)}{g(x)}<\frac{g(x)-g(x_1)}{g(x)}\left(-\frac{f(x_1)}{g(x_1)-g(x)}+A+\varepsilon\right).$$

This means that for x close to x_0 we have

$$(1+\varepsilon)(A-2\varepsilon)<\frac{f(x)}{g(x)}<(1+\varepsilon)(A+2\varepsilon),$$

which implies

$$\lim_{x\to x_0^-}\frac{f(x)}{g(x)}=A.$$

□

Example 9.20.

Find $\lim_{x\to+\infty} xe^{-x}$. We have $\lim_{x\to+\infty} x=+\infty$ and $\lim_{x\to+\infty} e^{-x}=0$. We can write

$$\lim_{x\to+\infty} xe^{-x}=\lim_{x\to+\infty}\frac{x}{e^x}=\lim_{x\to+\infty}\frac{1}{e^x}=0,$$

using the more difficult version of L'Hôspital Rule.

Example 9.21.

We have $\lim_{x\to 0} x=0$ and $\lim_{x\to 0^+}\ln x=-\infty$. Using L'Hôspital Rule we obtain

$$\lim_{x\to 0^+} x\ln x=\lim_{x\to 0^+}\frac{\ln x}{\frac{1}{x}}=\lim_{x\to 0^+}\frac{\frac{1}{x}}{\frac{-1}{x^2}}=\lim_{x\to 0^+}(-x)=0.$$

Example 9.22.

In this example $\lim_{x\to+\infty} f(x)=+\infty$, $\lim_{x\to+\infty} g(x)=+\infty$, $\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}=0$, but $\lim_{x\to+\infty}\frac{f(x)}{g(x)}$ does not exists. This is not a counterexample to L'Hôspital Rule, as $g'(x)$ has zeros in any neighbourhood of $+\infty$.

Let $f(x)=1+x+\sin x\cos x$ and $g(x)=(x+\sin x\cos x)e^{\sin x}$. It is easy to see that $\lim_{x\to+\infty} f(x)=+\infty=\lim_{x\to+\infty} g(x)$. Let us consider

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}\frac{1+x+\sin x\cos x}{(x+\sin x\cos x)e^{\sin x}}.$$

For $x_n=\pi n$ we have

$$\frac{f(x_n)}{g(x_n)}=\frac{1+\pi n}{\pi n}\underset{n\to\infty}{\longrightarrow}1.$$

For $x_n = 2\pi n + \pi/2$ we have

$$\frac{f(x_n)}{g(x_n)} = \frac{1 + 2\pi n + \pi/2}{(2\pi n + \pi/2) \cdot e} \underset{n\to\infty}{\longrightarrow} \frac{1}{e}.$$

Thus, the limit $\lim_{x\to+\infty} \frac{f(x)}{g(x)}$ does not exists. We have

$$f'(x) = 1 + \cos^2 x - \sin^2 x = 2\cos^2 x,$$

and

$$g'(x) = 2\cos^2 x \cdot e^{\sin x} + (x + \sin x \cos x) \cos x \cdot e^{\sin x}.$$

Thus,

$$\begin{aligned} \frac{f'(x)}{g'(x)} &= \frac{2\cos^2 x}{2\cos^2 x \cdot e^{\sin x} + (x + \sin x \cos x)\cos x \cdot e^{\sin x}} \\ &= \frac{2\cos x}{2\cos x \cdot e^{\sin x} + (x + \sin x \cos x) \cdot e^{\sin x}} \underset{x\to+\infty}{\longrightarrow} 0, \end{aligned}$$

as the limit is of the form

$$\frac{\text{bounded}}{\text{bounded} + (+\infty + \text{bounded}) \cdot \text{bounded}}$$

and $e^{-1} \leq e^{\sin x} \leq e$.

Graphing functions

The examples below show the standard techniques for graphing functions. First we check the domain of the function. Then, we find the zeros and the signs of the derivative, to decide where function is increasing and where it is decreasing.

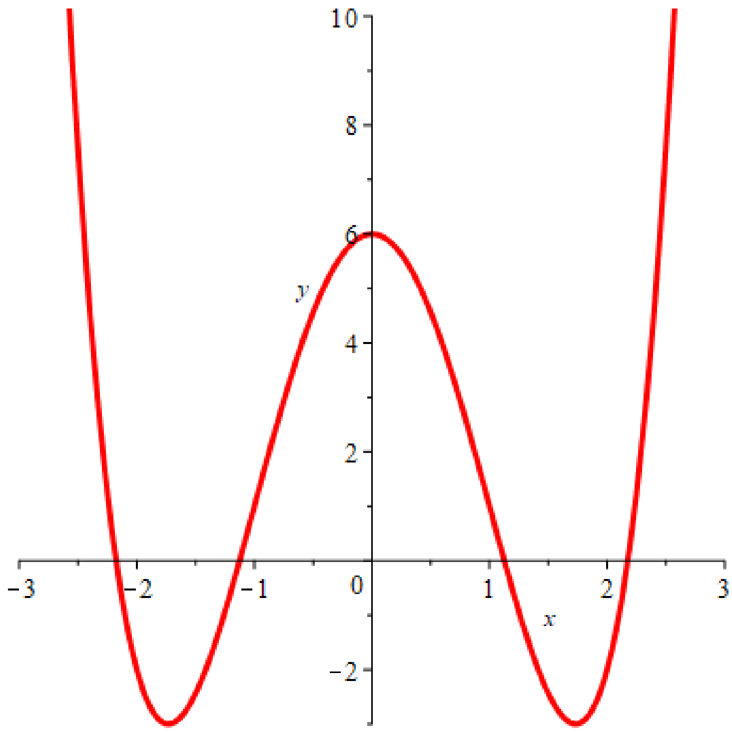

FIGURE 22. Graph of the function $f(x) = x^4 - 6x^2 + 6$

Example 9.23.

Graph $f(x) = x^4 - 6x^2 + 6$. The function is differentiable on the whole $\mathbb{R}$. We have $f'(x) = 4x^3 - 12x$. Zeros od the derivative: $f'(x) = 0$ for $x_1 = 0$, $x_2 = -\sqrt{3}$ and $x_3 = \sqrt{3}$. The sign of the derivative changes at these points and is, from left to right, $-$, $+$, $-$, $+$. Thus, $x_1 = 0$ is a local maximum, $f(0) = 6$, while $x_2 = -\sqrt{3}$ and $x_3 = \sqrt{3}$ are local minima, $f(\pm\sqrt{3}) = -3$. The graph of f is shown in Figure 22.

Example 9.24.

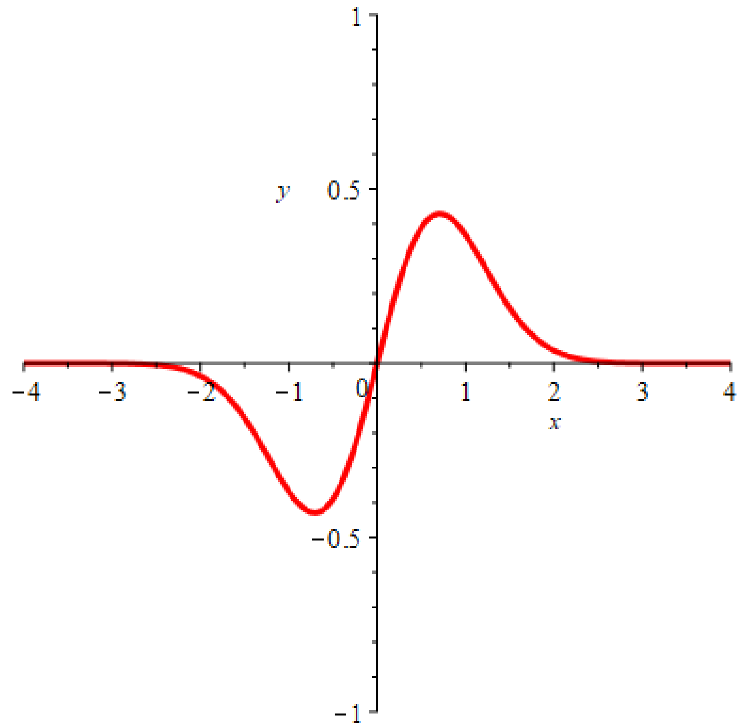

FIGURE 23. Graph of the function $f(x) = xe^{-x^2}$

Graph $f(x) = xe^{-x^2}$. The function is differentiable on the whole $\mathbb{R}$. We have $f'(x) = e^{-x^2} - 2x^2e^{-x^2} = (1-2x^2)e^{-x^2}$. Zeros od the derivative: $f'(x) = 0$ for $x_1 = -\sqrt{1/2}$ and $x_2 = \sqrt{1/2}$. The sign of the derivative changes at these points and is, from left to right, $-$, $+$, $-$. Thus, $x_1 = -\sqrt{1/2}$ is a local minimum , while $x_2 = \sqrt{1/2}$ is a local maximum. We have $f(\pm\sqrt{1/2}) = \mp e^{-1/2}/\sqrt{2}$.

We also see that $f''(x) = x(4x^2-6)e^{-x^2}$ has a 0 at $x = 0$. The first derivative changes sign there. This is an **inflection** point, where the curvature of the graph changes (from concave to convex in this case).

Another feature of this graph are the horizontal asymptotes, both in this cases equal to the x-axis. We $\lim_{x\to\pm\infty}(f(x) - 0) = 0$ and $\lim_{x\to\pm\infty}(f'(x) - 0) = 0$. The graph of f is shown in Figure 23.

Example 9.25.

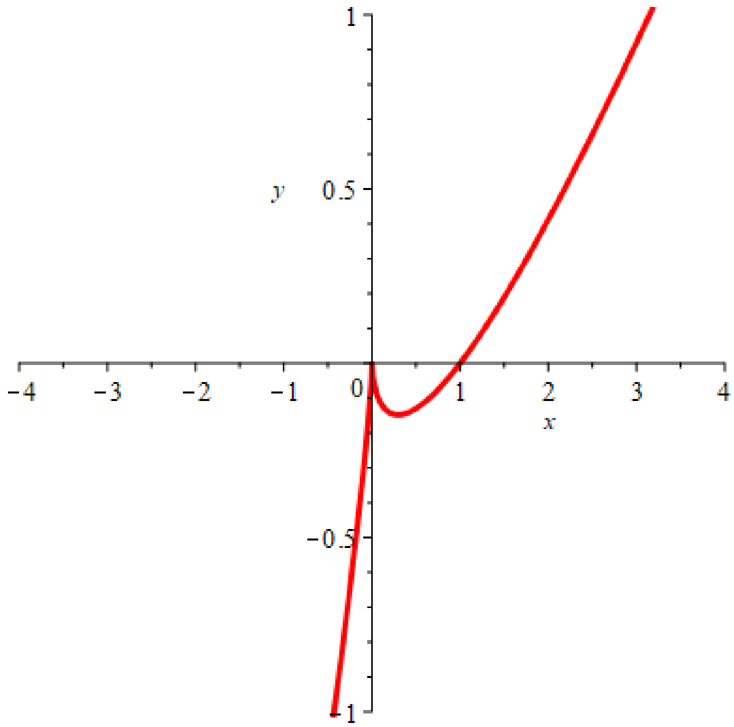

FIGURE 24. Graph of the function $f(x) = x - x^{2/3}$

Graph $f(x) = x - x^{2/3}$. The function is differentiable on $\mathbb{R} \setminus \{0\}$. Outside 0 we have $f'(x) = 1 - \frac{2}{3}x^{-1/3}$. We have $\lim_{x\to\pm 0} f'(x) = \mp\infty$. Zero od the derivative: $f'(x) = 0$ for $x_1 = \frac{8}{27}$. The sign of the derivative changes between points $-\infty$, 0 and $\frac{8}{27}$, from left to right, $+$, $-$, $+$. Thus, 0 is a local maximum , while $x_1 = \frac{8}{27}$ is a local minimum. We have $f(0) = 0$ and $f(\frac{8}{27}) \sim -0.1481$. Note, that this function has a local maximum at a point where it is not differentiable. The graph of f is shown in Figure 24.

Example 9.26.

Graph $f(x) = x + \frac{1}{x}$. The function is defined and differentiable on $\mathbb{R} \setminus \{0\}$. Outside 0 we have $f'(x) = 1 - x^{-2}$. We have $\lim_{x\to\pm 0} f'(x) = \mp\infty$. Zeros od the derivative: $f'(x) = 0$ for $x_1 = -1$ and $x_2 = 1$. The sign of the derivative changes between points

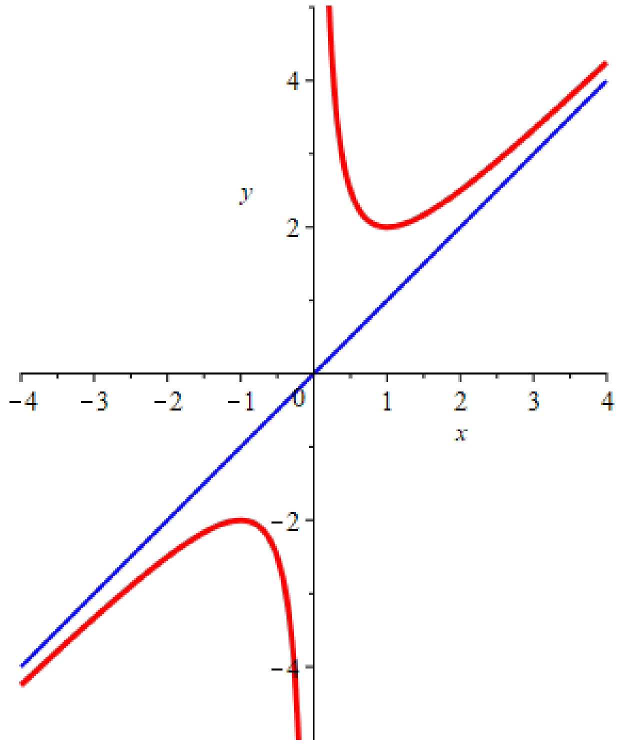

FIGURE 25. Graph of the function $f(x) = x + \frac{1}{x}$

$-\infty$,-1, 0, 1 and $+\infty$, from left to right, $+$, $-$, $-$, $+$. Thus, $x_1 = -1$ is a local maximum , while $x_2 = 1$ is a local minimum. We have $f(-1) = -2$ and $f(1) = 2$.

We see that $\lim_{x\to\pm\infty}(f(x) - x) = 0$ and $\lim_{x\to\pm\infty}(f'(x) - 1) = 0$. Thus, the diagonal $y = x$ is the **oblique** asymptote to the graph of f both in $+\infty$ and in $-\infty$.

Since $\lim_{x\to\pm 0} f'(x) = \mp\infty$ and $\lim_{x\to\pm 0} f(x) = \pm\infty$, the y-axis is the **vertical** asymptote to the graph of f at 0 (from both sides). The graph of f is shown in Figure 25.

Theorem 36. *If f is differentiable at x_0, $a_n < x_0 < b_n$, $n = 1, 2, \ldots$ and $\lim_{n\to\infty} a_n = x_0 = \lim_{n\to\infty} b_n$, then*

$$\lim_{n\to\infty} \frac{f(a_n) - f(b_n)}{a_n - b_n} = f'(x_0).$$

Proof. We write

$$\frac{f(a_n)-f(b_n)}{a_n-b_n}-f'(x_0)$$
$$=\frac{a_n-x_0}{a_n-b_n}\left(\frac{f(a_n)-f(x_o)}{a_n-x_0}-f'(x_0)\right)+\frac{x_0-b_n}{a_n-b_n}\left(\frac{f(x_0)-f(b_n)}{x_0-b_n}-f'(x_0)\right)\underset{n\to\infty}{\longrightarrow}0$$

Note that both $\frac{a_n-x_0}{a_n-b_n}$ and $\frac{x_0-b_n}{a_n-b_n}$ are positive and sum up to 1. □

Example 9.27.

Find a rectangle with perimeter 40 and the largest area.

Let x denote one side of the rectangle. Then, the other side is $20-x$ and the area is $f(x)=x(20-x)=20x-x^2$, $0<x<20$. To find the maximum we calculate the derivative $f'(x)=20-2x$ which has the only 0 at $x=10$. This point is the maximum by the geometric reasons. The maximal area is $f(10)=100$.

Theorem 37. *If f is continuous on $[a,b]$ and differentiable on (a,b), then f' has the Intermediate Value Property, i.e., if $f'(a)<r<f'(b)$, then there exists $c\in(a,b)$ with $f'(c)=r$.*

Proof. Consider the function $F(x)=f(x)-rx$. If it has a maximum or minimum at $c\in(a,b)$, then $F'(c)=f'(c)-r=0$ and the proof is completed. If not, then either

(i) $F(a)<F(x)$ for $x\in(a,b]$ and $F'(a)\geq 0$ or $f'(a)-r\geq 0$ which implies $f'(x)\geq r$, a contradiction; or

(ii) $F(b)<F(x)$ for $x\in[a,b)$ and $F'(b)\leq 0$ or $f'(a)-r\leq 0$ which implies $f'(x)\leq r$, again a contradiction.

□

Theorem 38. ***Taylor formula:*** *If f is a function such that $f,f',f'',\dots,f^{(n)}$ are continuous on $[a,b]$ and $f^{(n+1)}$ exists on (a,b), then there is a $c\in(a,b)$ such that*

$$f(b)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(b-a)^k+\frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$

Proof. **Proof from Larsen's book:** Let us define the constant α in such a way that

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{\alpha}{(n+1)!}(b-a)^{n+1}. \tag{13}$$

and define

$$F(x) = f(b) - \left(\sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(b-x)^k + \frac{\alpha}{(n+1)!}(b-x)^{n+1} \right).$$

By (13) we have $F(a) = 0$. We also see that $F(b) = 0$. By the assumptions of the theorem the function F is continuous on $[a,b]$ and differentiable on (a,b). By Rolle's theorem 29 there exists $c \in (a,b)$ such that $F'(c) = 0$. We have We will calculate $F'(x)$ using induction. Let

$$G_j(x) = f(b) - \left(\sum_{k=0}^{j} \frac{f^{(k)}(x)}{k!}(b-x)^k \right).$$

We have

$$G_1'(x) = -\left(f(x) + \frac{f'(x)}{1}(b-x) \right)' = -f'(x) - f''(x)(b-x) + f'(x) = -f''(x)(b-x).$$

$$\begin{aligned} G_2'(x) &= -f''(x)(b-x) - \left(\frac{f^{(2)}(x)}{2!}(b-x)^2 \right)' \\ &= -f''(x)(b-x) - \frac{f^{(3)}(x)}{2!}(b-x)^2 + \frac{f^{(2)}(x)}{2!} \cdot 2(b-x) = -\frac{f^{(3)}(x)}{2!}(b-x)^2. \end{aligned}$$

Now, assume that

$$G_j'(x) = -\frac{f^{(j+1)}(x)}{j!}(b-x)^j.$$

Then, we have

$$\begin{aligned} G_{j+1}'(x) &= -\frac{f^{(j+1)}(x)}{j!}(b-x)^j - \left(\frac{f^{(j+1)}(x)}{(j+1)!}(b-x)^{j+1} \right)' \\ &= -\frac{f^{(j+1)}(x)}{j!}(b-x)^j - \frac{f^{(j+2)}(x)}{(j+1)!}(b-x)^{j+1} + \frac{f^{(j+1)}(x)}{(j+1)!} \cdot (j+1)(b-x)^j \\ &= -\frac{f^{(j+2)}(x)}{(j+1)!}(b-x)^{j+1}. \end{aligned}$$

By induction we obtain

$$G_n'(x) = -\frac{f^{(n+1)}(x)}{n!}(b-x)^n.$$

Thus,

$$F'(c) = -\frac{f^{(n+1)}(c)}{n!}(b-c)^n + \frac{\alpha}{(n+1)!}(n+1)(b-c)^n.$$

Since $F'(c) = 0$ we obtain

$$-\frac{f^{(n+1)}(c)}{n!}(b-c)^n + \frac{\alpha}{(n+1)!}(n+1)(b-c)^n = 0,$$

or $\alpha = f^{(n+1)}(c)$, which completes the proof. □

Remark 9.3. *We can write the Taylor formula in the form*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}, \ c \text{ between } x_0 \text{ and } x.$$

The polynomial

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k,$$

is called Taylor's polynomial of the n-th order for the function f at the point x_0 and

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

is called the remainder of order n. This is Cauchy form of the remainder. There are other equivalent but sometimes more useful forms of the remainder.

Example 9.28.

Let $f(x) = e^x$, $x \in \mathbb{R}$. For any $n \geq 1$ we have $f^{(n)}(x) = e^x$ so for $x_0 = 0$ we have $f^{(n)}(0) = 1$ and

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + e^c \frac{x^{n+1}}{(n+1)!}, \tag{14}$$

for some c between 0 and x. We see that $R_n(x) = e^c \frac{x^{n+1}}{(n+1)!}$ converges to 0 as $n \to +\infty$ for any $x \in \mathbb{R}$.

Example 9.29.

We will show that the base of natural logarithms e is not rational.

Let us assume that $e = \frac{m}{n}$. By formula (14) we have

$$\frac{m}{n} = e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + e^c \frac{1}{(n+1)!}, \quad 0 < c < 1.$$

Multiplying both sides by $n!$ we obtain

$$m(n-1)! = n! + n! + \frac{n!}{2!} + \frac{n!}{3!} + \cdots + \frac{n!}{n!} + e^c \frac{1}{(n+1)}.$$

Since $e^c < 3$ this is impossible for $n \geq 2$ as the difference of two integers cannot be less than 1.

Example 9.30.

Let $f(x) = \sin x$, $x \in \mathbb{R}$. We have:

$$f(x) = \sin x, \quad f(0) = 0;$$
$$f'(x) = \cos x, \quad f'(0) = 1;$$
$$f''(x) = -\sin x, \quad f''(0) = 0;$$
$$f^{(3)}(x) = -\cos x, \quad f^{(3)}(0) = -1;$$
$$f^{(4)}(x) = \sin x, \quad f^{(4)}(0) = 0;$$

and the higher derivatives repeat. Thus, we obtain

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} + \cdots \pm (\sin c \text{ or } \cos c) \frac{x^{n+1}}{(n+1)!},$$

for some c between 0 and x. We see that $R_n(x) = \pm(\sin c \text{ or } \cos c)\frac{x^{n+1}}{(n+1)!}$ converges to 0 as $n \to +\infty$ for any $x \in \mathbb{R}$.

Example 9.31.

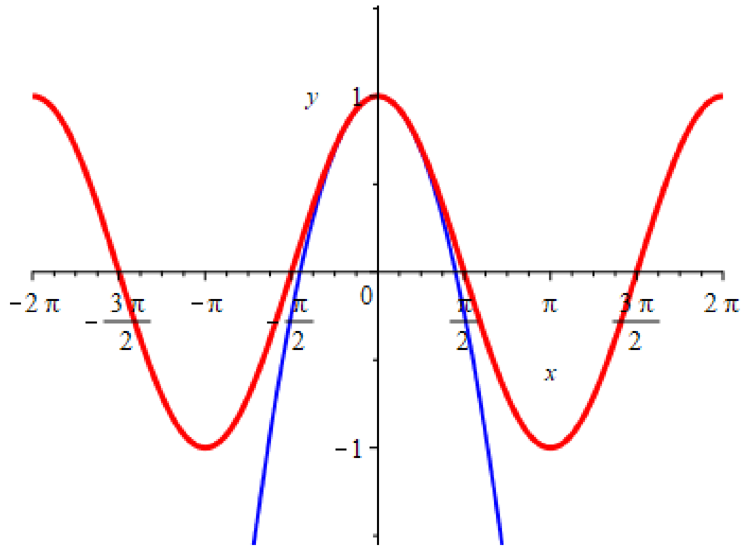

FIGURE 26. $P_2(x) = 1 - \frac{x^2}{2!}$ approximates $\cos x$

Let $f(x) = \cos x$, $x \in \mathbb{R}$. We can use the above calculations with a shift:

$$f(x) = \cos x, \quad f(0) = 1;$$

$$f'(x) = -\sin x, \quad f'(0) = 0;$$

$$f''(x) = -\cos x, \quad f''(0) = -1;$$

$$f^{(3)}(x) = \sin x, \quad f^{(3)}(0) = 0;$$

$$f^{(4)}(x) = \cos x, \quad f^{(4)}(0) = 1;$$

and the higher derivatives repeat. Thus, we obtain

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} + \cdots \pm (\sin c \text{ or} \cos c)\frac{x^{n+1}}{(n+1)!},$$

for some c between 0 and x. Again, we see that $R_n(x) = \pm(\sin c \text{ or} \cos c)\frac{x^{n+1}}{(n+1)!}$ converges to 0 as $n \to +\infty$ for any $x \in \mathbb{R}$.

The approximations $P_2(x)$, $P_4(x)$ and $P_8(x)$ of $\cos x$ are shown in Figures 26, 27 and 28, respectively.

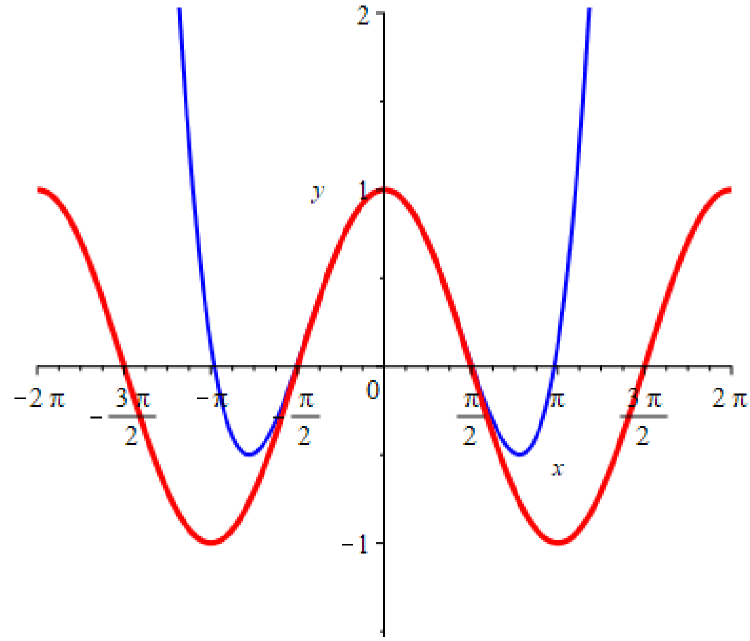

FIGURE 27. $P_4(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$ approximates $\cos x$

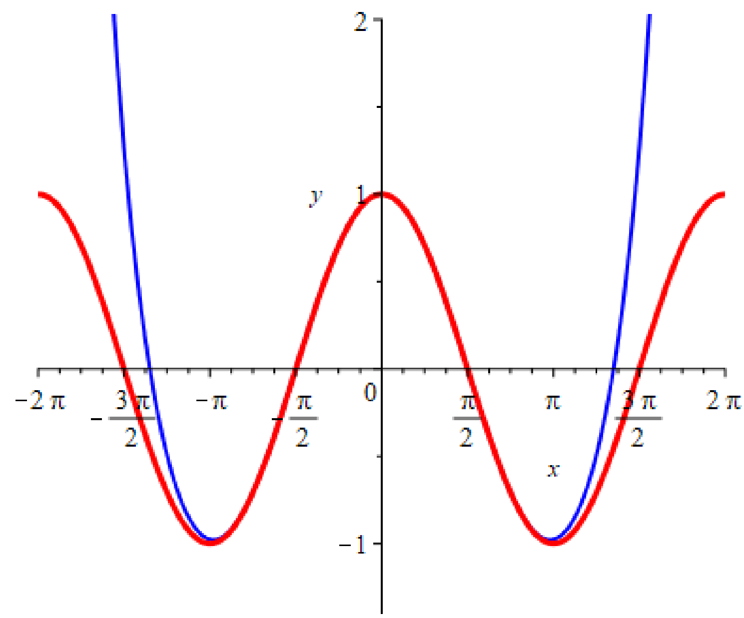

FIGURE 28. $P_8(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!}$ approximates $\cos x$

Example 9.32.

Let $f(x) = \ln(1+x)$. We have:

$$\begin{aligned} f(x) &= \ln(1+x), \quad f(0) = 0; \\ f'(x) &= \frac{1}{1+x}, \quad f'(0) = 1; \\ f''(x) &= -\frac{1}{(1+x)^2}, \quad f''(0) = -1; \\ f^{(3)}(x) &= \frac{2}{(1+x)^3}, \quad f^{(3)}(0) = 2!; \\ f^{(4)}(x) &= -2\cdot 3\frac{1}{(1+x)^4}, \quad f^{(4)}(0) = -3!; \\ f^{(5)}(x) &= 2\cdot 3\cdot 4\frac{1}{(1+x)^5}, \quad f^{(5)}(0) = 4!. \end{aligned}$$

By induction we can prove $f^{(n)}(x) = (-1)^{n-1}(n-1)!\frac{1}{(1+x)^n}$ and $f^{(n)}(0) = (-1)^{n-1}(n-1)!$. Thus, we obtain

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} + \cdots \pm \frac{1}{(1+c)^{n+1}}\frac{x^{n+1}}{n+1},$$

for some c between 0 and x. We see that $R_n(x) = \pm\frac{1}{(1+c)^{n+1}}\frac{x^{n+1}}{n+1}$ converges to 0 as $n \to +\infty$, as long as $|x| \le 1$, or for $x \in [-1,1]$.

10. Convex Functions

We base our presentation on [3] and [1].

Definition 10.1. *A function defined on an interval I, $f : I \to \mathbb{R}$ is called convex* $\Longleftrightarrow$

(15) $$f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y),$$

for all $x, y \in I$ and $\lambda \in [0,1]$.

A function is called **strictly convex** if the inequality in (15) is strict and **concave** or **strictly concave** if the inequalities are reversed.

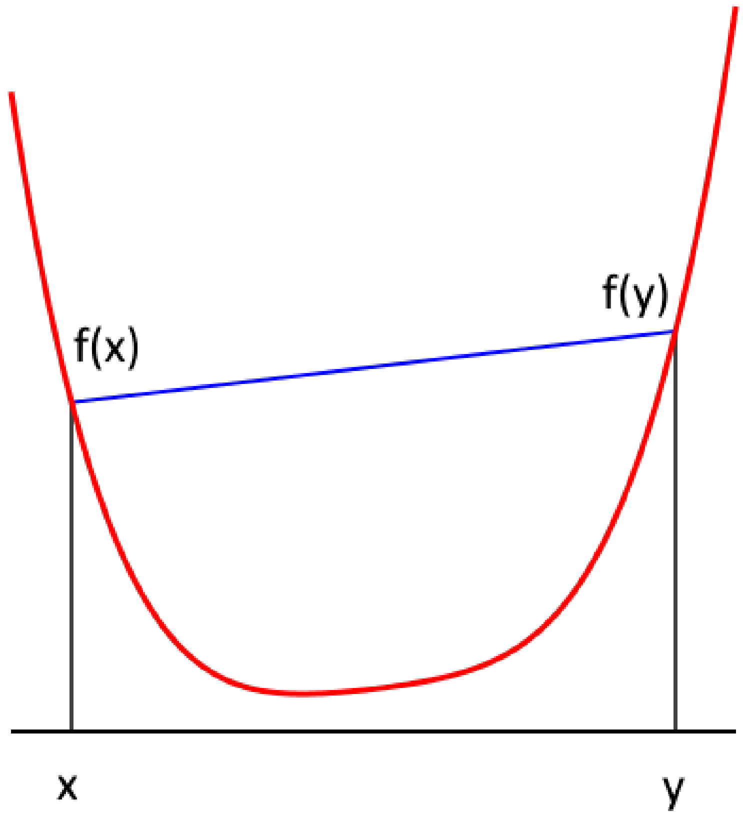

FIGURE 29. Graph of a convex function

Geometrically, this means that the chord through points $(x, f(x))$ and $(y, f(y)$ is always above the graph of f on the interval (x, y). See Figure 29. This means

$$f(x) \leq f(a) + \frac{f(b) - f(a)}{b - a}(x - a), \tag{16}$$

for any $a < x < b$ in the interval I.

Theorem 39. *A function convex on an open interval (a, b) is continuous.*

Proof. Let $x_0 \in (a, b)$. The function $g(x) = f(x + x_0) - f(x_0)$ is also convex and $g(0) = 0$. This shows that we can assume that $0 \in (a, b)$ and prove that f is continuous at 0.

For sufficiently small $\delta > 0$, $I = [-\delta, \delta] \subset (a, b)$. For $x \in I$, we have

$$\begin{aligned} f(x) = f(|x|\text{sgn(x)}) &= \text{f}\left(\frac{|\text{x}|}{\delta}\text{sgn(x)}\delta\right) \\ &= f\left(\frac{|x|}{\delta}\text{sgn(x)}\delta + \left(1 - \frac{|\text{x}|}{\delta}\right)\cdot 0\right) \leq \frac{|x|}{\delta}f(\text{sgn(x)}\delta). \end{aligned}$$

If $M = \max\{f(-\delta), f(\delta)$, then we obtain $f(x) \leq \frac{M}{\delta}|x|$, for $x \in I$.

We also have

$$0 = f(0) = f(\frac{1}{2}(-x) + \frac{1}{2}x) \leq \frac{1}{2}f(-x) + \frac{1}{2}f(x),$$

which implies $-f(-x) \leq f(x)$. Thus, we showed

$$|f(x)| \leq \frac{M}{\delta}|x|,$$

implying $\lim_{x\to 0} f(x) = 0$. □

Example 10.1.

A function convex on a closed interval $[a, b]$ does not need to be continuous. This is shown by example

$$f(x) = \begin{cases} x^2, \text{ for } x \in [-1, 1); \\ 2, \text{ for } x = 1. \end{cases}$$

Definition 10.2. *A function defined on an interval I, $f : I \to \mathbb{R}$ is called mid-point convex* $\iff$

(17) $$f\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y),$$

for all $x, y \in I$.

A convex function is obviously mid-point convex but the reverse statement is not always true.

Example 10.2.

The real numbers $\mathbb{R}$ form a linear space over the rational numbers $\mathbb{Q}$. Let the $B = \{v_\alpha\}_{\alpha\in\Lambda}$ be the basis of $\mathbb{R}$ over $\mathbb{Q}$. Then, the conjugate basis $B^* = \{v^*_\alpha\}_{\alpha\in\Lambda}$

consists of functions on $\mathbb{R}$ which are non-constant and linear over $\mathbb{Q}$, i.e., for any $r_1, r_2 \in \mathbb{Q}$ and any $x, y \in \mathbb{R}$ we have

$$v_\alpha^*(r_1 x + r_2 y) = r_1 v_\alpha^*(x) + r_2 v_\alpha^*(y).$$

This imply that v_α^* is a mid-point convex function on $\mathbb{R}$. Since $v_\alpha^*(\mathbb{R}) = \mathbb{Q}$ the function v_α^* cannot be continuous. Thus, it is not convex.

Theorem 40. *If the function f is mid-point convex and continuous, then it is convex.*

Proof. Let f be mid-point convex on the interval (a, b). We will show f satisfies inequality (15) for all $x, y \in (a, b)$ and λ of the form $p/2^n$, $p, n \in \mathbb{N}$. Assume that inequality holds for n and we will show it for $n + 1$. Let $p + q = 2^{n+1}$ and

$$z = \frac{p}{2^{n+1}} x + \frac{q}{2^{n+1}} y = \frac{1}{2} \left(\frac{p}{2^n} x + \frac{r}{2^n} y + y \right).$$

We may assume that $p < q$, hence $p < 2^n < q = 2^n + r$ and $p + r = 2^n$. Then,

$$\begin{aligned} f(z) &\leq \frac{1}{2} \left[f\left(\frac{p}{2^n} x + \frac{r}{2^n} y \right) + f(y) \right] \\ &\leq \frac{1}{2} \left[\frac{p}{2^n} f(x) + \frac{r}{2^n} f(y) + \frac{2^n}{2^n} f(y) \right] \leq \frac{p}{2^{n+1}} f(x) + \frac{q}{2^{n+1}} f(y). \end{aligned}$$

We proved by induction that inequality (15) holds for all λ of the form $p/2^n$, $p, n \in \mathbb{N}$. Since these numbers are dense in $[0, 1]$ continuity of f implies inequality (15) for all $\lambda \in [0, 1]$ and f is convex. □

Remark 10.1. *For a mid-point convex function to be convex a much weaker conditions are sufficient, for example Lebesgue measurability.*

Theorem 41. *If the function f is mid-point convex on interval (a, b) and discontinuous at one point it unbounded on any subinterval.*

Proof. We can assume that f is defined on $(-a,a)$, $f(0) = 0$ and there exists a sequence $x_n \to 0$ with $\lim_{n\to\infty} f(x_n) = m > 0$. If $m < 0$, then we have $f(x_n) < 0$ for large n and

$$-f(-x_n) \leq f(x_n) \quad \text{or} \quad f(-x_n) \geq -f(x_n) > 0,$$

and we can consider the sequence $(-x_n)_{n\geq 1}$ instead.

The sequence $2x_n$ also converges to 0 and we have

$$2f(x_n) \leq f(0) + f(2x_n) = f(2x_n),$$

and $\liminf_{x\to\infty} f(2x_n) \geq 2m$. By induction we can prove that

$$\liminf_{x\to\infty} f(2^k x_n) \geq 2^k m.$$

Thus, f is unbounded above in any neighbourhood of 0. There exists a sequence y_n convergent to 0 with $f(y_n) \to +\infty$.

Let z be any point in the interval $(-a,a)$. The sequence $z + 2y_n$ converges to z and

$$f(y_n) = f\left(\frac{z + 2y_n - z}{2}\right) \leq \frac{1}{2}(f(z+2y_n) + f(-z)).$$

Since $f(y_n) \to +\infty$ we also have $f(z + y_n) \to +\infty$. The function f is unbounded in any neighbourhood of z. Since z was arbitrary the function f on any neighbourhood of any point in $(-a,a)$ and hence it is unbounded on any subinterval of $(-a,a)$. □

Remark 10.2. *Since convex functions are bounded on subintervals by inequality (16), Theorem 41 proves again that a convex function is continuous.*

10.1. **Differentiability of convex functions.** Below we study the differentiability of convex functions. Let f be a convex function o $\mathbb{R}$. We consider the difference quotient

$$\frac{f(x+h) - f(x)}{h}\ ,\ h > 0.$$

Inequality (16) show that the difference quotient is increasing as a function of h, i.e.,

$$\frac{f(x+h_1) - f(x)}{h_1} \leq \frac{f(x+h_2) - f(x)}{h_2}\ ,\ 0 < h_1 < h_2.$$

It is also and increasing function of x, i.e.,

$$\frac{f(x_1+h)-f(x_1)}{h} \leq \frac{f(x_2+h)-f(x_2)}{h}\ ,\ x_1 < x_2,$$

or

$$f(x_1+h)-f(x_1) \leq f(x_2+h)-f(x_2)\ ,\ x_1 < x_2,$$

or

$$f(x_1+h)+f(x_2) \leq f(x_2+h)+f(x_1)\ ,\ x_1 < x_2. \tag{18}$$

To justify the last inequality we set $d = x_2 - x_1 + h$ and write

$$x_2 = \frac{h}{d}x_1 + \frac{x_2-x_1}{d}(x_2+h)\quad ,\quad x_1+h = \frac{x_2-x_1}{d}x_1 + \frac{h}{d}(x_2+h).$$

By the convexity of f we obtain

$$f(x_2) \leq \frac{h}{d}f(x_1) + \frac{x_2-x_1}{d}f(x_2+h),$$

and

$$f(x_1+h) = \frac{x_2-x_1}{d}f(x_1) + \frac{h}{d}f(x_2+h).$$

Adding up we obtain equality (18).

Since the difference quotient is increasing in h, $h>0$, the right hand side derivative $f'(x_+)$ exists at any point x. Since the difference quotient is increasing in x, the right hand side derivative $f'(x_+)$ is an increasing function of x.

Now, we will in a similar way consider the difference quotient from the left hand side

$$\frac{f(x-h)-f(x)}{-h}\ ,\ h>0.$$

First, we note that this difference is decreasing in h, i.e.,

$$\frac{f(x-h_1)-f(x)}{-h_1} \geq \frac{f(x-h_2)-f(x)}{-h_2}\ ,\ 0<h_1<h_2.$$

This follows from the other form of inequality (16), which says

$$f(x) \leq f(b) + \frac{f(b)-f(a)}{b-a}(x-b)\ ,$$

or

$$\frac{f(x)-f(b)}{x-b} \leq \frac{f(a)-f(b)}{a-b}\quad ,\ a<x<b.$$

The monotonicity of this difference quotient in h implies that the left hand side derivative $f'(x_-)$ exists at any point x.

The left hand side difference quotient is also an increasing function of x, i.e.,

$$\frac{f(x_1-h)-f(x_1)}{-h} \le \frac{f(x_2-h)-f(x_2)}{-h}\ ,\ x_1 < x_2,$$

or

$$f(x_1-h)-f(x_1) \ge f(x_2-h)-f(x_2)\ ,\ x_1 < x_2,$$

or

$$f(x_1-h)+f(x_2) \ge f(x_2-h)+f(x_1)\ ,\ x_1 < x_2. \tag{19}$$

For small h, we have $x_1-h < x_1 < x_2-h < x_2$ while for the equation (18) we had $x_1 < x_1+h < x_2 < x_2+h$. Thus, the equality (19) is equivalent to (18) with the renamed points.

Since the left hand side difference quotient is an increasing function of x, the left hand side derivative $f'(x_-)$ is an increasing function of x.

The inequality of the form of (18) or (19) implies also that

$$\frac{f(x-h)-f(x)}{-h} \le \frac{f(x+h)-f(x)}{h}\ ,\ h>0\ ,$$

and

$$\frac{f(x+h)-f(x)}{h} \le \frac{f(x+s-h)-f(x+s)}{h}\ ,\ s>0, h>0\ ,$$

for h sufficiently small. This implies

$$f'(x_-) \le f'(x_+) \le f'((x+s)_-),$$

for all x and $s>0$. Thus, $f'(x_+) = f'(x_-)$ at any point of continuity of $f'(x_-)$.

We proved the following:

Theorem 42. ***Differentiability of a convex function:*** *Let f be a convex function on $\mathbb{R}$. Then, both left and right hand side derivatives exists at every point and are both increasing. They are equal except at at most countably many points, i.e., the function f is differentiable except at at most countably many points.*

Proposition 10.1. *(i) If the function $f : I \to \mathbb{R}$ is differentiable and the derivative is increasing, then f is convex.*

(ii) If the function $f : I \to \mathbb{R}$ is twice differentiable and the second derivative is nonnegative, then f is convex.

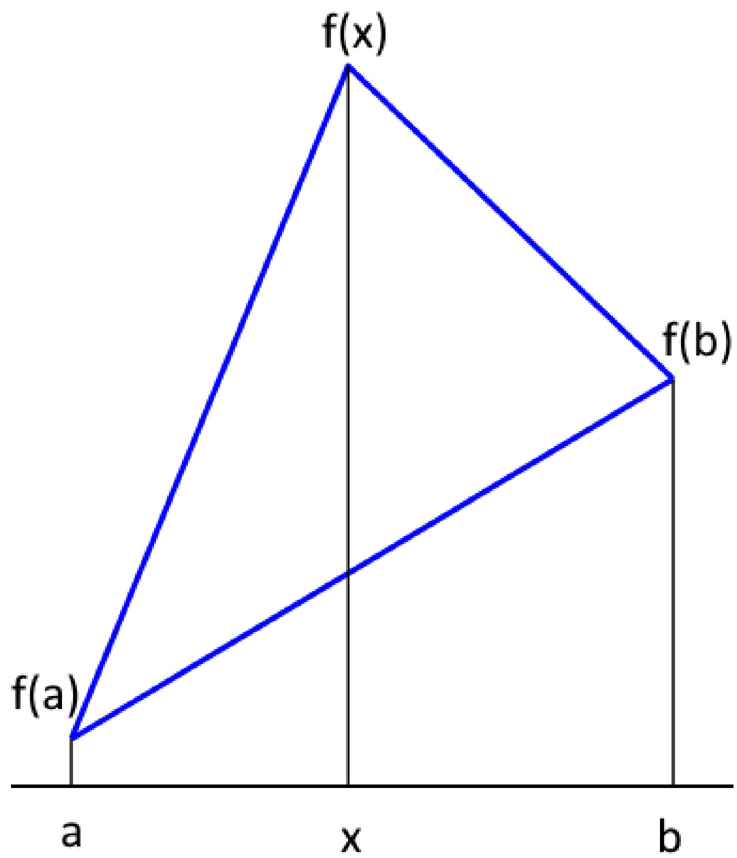

FIGURE 30. Non-convexity implies derivative not increasing.

Proof. It is enough to prove (i). We will prove the contrapositive statement, if f is not convex then f' is not increasing.

Assume, that we can find points $a < x < b$ with $f(x)$ strictly above the chord joining $(a, f(a))$ and $(b, f(b))$, i.e.,

$$f(x) > f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \implies \frac{f(x) - f(a)}{x - a} > \frac{f(b) - f(a)}{b - a},$$

and

$$f(x) > f(b) + \frac{f(b)-f(a)}{b-a}(x-b) \implies \frac{f(x)-f(b)}{x-b} < \frac{f(b)-f(a)}{b-a}.$$

Thus,

$$\frac{f(x)-f(a)}{x-a} > \frac{f(x)-f(b)}{x-b}.$$

By Mean Value theorem it implies that for some points $x_1 \in (a,x)$ and $x_2 \in (x,b)$ we have $f'(x_1) > f'(x_2)$, so f' is not increasing. $\square$

Example 10.3.

$f(x) = e^x$ is convex on $\mathbb{R}$. It follows by Proposition 10.1.

Example 10.4.

Young's inequality Let $a, b > 0$ and $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{p} = 1$. Then,

$$a\,b \leq \frac{1}{p}\,a^{\,p} + \frac{1}{q}\,b^{\,q}.$$

Since e^x is convex we have

$$e^{\frac{1}{p}\,x + \frac{1}{q}\,y} \leq \frac{1}{p}\,e^x + \frac{1}{q}\,e^y.$$

Substituting $e^x = a^p$ and $e^y = b^q$ we obtain the Young's inequality.

11. Problems

The problems are divided into 10 assignments.

Assignment # 1

Problem 1.: Express the sentence $\alpha \Rightarrow (\beta \Rightarrow \gamma)$ using symbols $\wedge$, $\vee$ and $\neg$.

Problem 2.: Write negation of the sentence

$$\alpha \wedge (\beta \Rightarrow \gamma).$$

Problem 3.: Prove that the following sentences are theorems, i.e., are true for all choices of sentences α, β, γ.

$$a) \qquad [(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \gamma)] \Rightarrow (\alpha \Rightarrow \gamma),$$

$$b) \qquad [(\alpha \wedge \neg\beta) \Rightarrow (\gamma \wedge \neg\gamma)] \Rightarrow (\alpha \Rightarrow \beta).$$

Problem 4.: Use the theorem 3b) to prove by contradiction:

Theorem: (n is an integer and n^2 is odd) $\Rightarrow$ (n is an integer and n is odd).

Explicitly write sentences α, β and γ.

Problem 5.: Prove by contraposition ($(\alpha \Rightarrow \beta) \Longleftrightarrow (\neg\beta \Rightarrow \neg\alpha)$):

Theorem: (n is an integer and n^2 is odd) $\Rightarrow$ (n is an integer and n is odd).

Explicitely write sentences α, β.

Problem 6.: Let $A, B, C, D \subset \mathbb{R}$. Prove that

$$(A \setminus B) \cap (C \setminus D) = (A \cap C) \setminus (B \cup D)\,.$$

Hint: For arbitrary $K, L \subset \mathbb{R}$ we have $K \setminus L = K \cap L^c$, where L^c is the complement of L.

Problem 7.: Write the negation of the sentence

$$\forall_{\,x} \, \exists_{\,y} \, (x < z) \wedge (z < y)\,.$$

Assignment # 2

Problem 1:

(a) Represent $987654321_{(10)}$ in base 16. For digits $(10),(11),(12),(13),(14)$ and (15) you can use A,B,C,D,E,F, correspondingly. This is the accepted standard.

(b) Represent the fraction $\frac{3}{7}$ in base 4. Digits for base 4 are $\{0,1,2,3\}$.

Problem 2: What decimal number is $0.1414141414\ldots_{(5)}$ (base 5) ?.

Problem 3: Use axioms listed in class to prove that $\forall_a\ (-1)\cdot a = -a$. Write which axiom you are using on each step.

Problem 4: Write the negation of the sentence

$$\forall_{\varepsilon>0}\ \exists_{N\geq 1}\ \forall_{n\geq N}\ |x_n - L| < \varepsilon\ .$$

Problem 5: If $5\cdot 12 = 104$, how much is $10\cdot 11$?

Problem 6: Prove by induction:

$$1^3+2^3+3^3+\cdots+n^3 = \frac{1}{4}n^2(n+1)^2 = (1+2+3+\cdots+n)^2\ .$$

Problem 7: Prove by induction or any other method:

$$\binom{k}{l} + \binom{k}{l+1} = \binom{k+1}{l+1},$$

$0\leq l,\ l+1\leq k$.

Problem 8: Prove by induction: For any $n\geq 1$ the number $2^{n+2}\cdot 3^n + 5n - 4$ is divisible by 25.

Problem 9: Prove that for any $n\geq 1$ we have

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

Assignment # 3

Problem 1. Prove that the following sets have cardinality $\aleph_0$ by giving explicitly an bijective function $f: A\to\mathbb{N}$.

(i) $A=\{n\in\mathbb{N} : n \text{ has remainder 2 when divided by 3}\}$;

(ii) $A=\{x\in(0,1) : \sin\frac{1}{x}=0\}$.

Problem 2. Show that the following pairs of sets $A, B \subset \mathbb{R}$ have the same cardinality by giving explicitly a bijection $f : A \to B$ or by using the Cantor-Bernstein-Schröder Theorem.

(i) $A = (0, 1]$, $B = [1, \infty)$;

(ii) $A = (0, 1]$, $B = (-\infty, \infty)$.

Problem 3: Let $\mathcal{C}$ be a family of all circles in the plane $\mathbb{R}^2$ which have rational centers and rational radiuses, i.e., for each $C \in \mathcal{C}$ the center is a point with both coordinates rational and the radius is a rational number as well. Prove that the family $\mathcal{C}$ is countable.

Problem 4: Let $I = [0, 1)$ and $SQ = [0, 1) \times [0, 1) = \{(x, y) \in \mathbb{R}^2 : 0 \leq x, y < 1\}$. Show that

$$\mathrm{Card(I)} = \mathrm{Card(SQ)}.$$

Hint: Represent real numbers by their decimal expansions and construct a bijection mapping a pair of expansions into one.

Problem 5: Prove that the set of sequences

$$\Sigma_2^+ = \{(a_0, a_1, a_2, \dots) : a_i \in \{0, 1\}\} \ ,$$

is uncountable.

Problem 6: Prove that the set $A = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is countable.

Problem 7: Recall that $|x| = \begin{cases} x & \text{for } x \geq 0; \\ -x & \text{for } x < 0. \end{cases}$. Prove:

$$(a) \quad ||x_1 - y_1| - |x_2 - y_2|| \leq |x_1 - x_2| + |y_1 - y_2| \, ;$$

$$(b) \quad |x - y| \geq ||x| - |y|| \ , \ x, y \in \mathbb{R} \, .$$

Hint: To show that $|a| \leq b$ it is enough to show that $a \leq b$ and $-a \leq b$.

Problem 8: Prove that

$$\sqrt{n} < \sqrt[n]{n!} \ , \quad n \geq 3.$$

Hint: To prove $a_1 < b_1 \implies a_2 < b_2$ it is enough to show

$$\frac{a_2}{a_1} \leq \frac{b_2}{b_1}.$$

You may also need the information that $(1+1/n)^n < e$ for all $n \geq 1$.

Problem 9: The unit circle in the plane is the set of point $\mathbb{T} = \{(x,y) : x^2 + y^2 = 1\}$. Show that $\mathbb{T}$ has the same cardinality as the real line $\mathbb{R}$.

Problem 10: Consider a set A such that there exists a surjective map $f : \mathbb{N} \to A$. Construct an injection $g : A \to \mathbb{N}$.

Hint: suppose $f : \mathbb{N} \to A$ is surjective; for each $a \in A$, consider $f^{-1}(a) = \{n \in \mathbb{N} : f(n) = a\}$ and apply the well-ordering of $\mathbb{N}$. Well-ordering says that every nonempty subset of $\mathbb{N}$ has the least element.

Assignment # 4

Problem 1: Let $S \subset \mathbb{R}$. If there exists $x \in S$ such that $y \leq x$ for every $y \in S$, we say that x is the largest element of S, or write $x = \max S$.

(i) Prove that if S has a largest element then it is unique.

(ii) Prove that if S has a largest element then S has a supremum and $\sup S = \max S$.

(iii) Prove that if S is finite then S has a largest element.

(iv) Give two examples of a set S which does NOT have a largest element: one which has a supremum (i.e. $\sup S$ exists in $\mathbb{R}$) and one which does not.

Problem 2: (i) Prove that if $r \in \mathbb{Q}$ and $r \neq 0$ then $r\sqrt{2}$ is irrational.

(ii) Prove that for every $a, b \in \mathbb{R}$ with $a < b$, there exists $r \in \mathbb{Q}$ with $a < r\sqrt{2} < b$. What property of the irrationals in $\mathbb{R}$ does this prove? Explain.

Problem 3: Find the limits and use the definition to prove convergence:

$$(a) \quad a_n = \frac{\sin(2^{2021 \cdot n})}{n^2 - 3}, \quad n = 1, 2, 3, \ldots$$

$$(b) \quad b_n = \frac{n^2 + 3n + 2021}{n^3 + n^2 + 1}, \quad n = 1, 2, 3, \ldots$$

Problem 4:

(a) Negate the definition of the limit to get the mathematical statement of $\lim a_n \neq L$

(b) Show that for $a_n = (-1)^n$, $\lim a_n \neq 0$

Hint: use part (a) with $\varepsilon = 1/2$.

(c) Use the definition of the limit to show that for $a_n = (-1)^n$,

$$\lim \frac{a_1 + \ldots + a_n}{n} = 0.$$

Problem 5: Assume that $a_n \to 2$ and $b_n \to 3$ as $n \to \infty$. Use the definition to prove

$$(3a_n - b_n) \to 3.$$

Problem 6: Assume that $a_n \to 2$ as $n \to \infty$. Prove that for large enough indices n we have

$$a_n > 1.$$

Problem 7: Use "squeeze" theorem to find the limit

$$\lim_{n\to\infty} \sqrt[n]{7^n + e^{2n} + \cos^2(n! + n^n)}.$$

Problem 8: Assume that $a_n \to A$ as $n \to \infty$ and that $a_n > 0$ for all $n \geq 1$. Prove that

$$\sqrt{a_n} \to \sqrt{A}, \quad n \to \infty.$$

Hint: Prove first that

$$\sqrt{a} - \sqrt{b} \leq \sqrt{a - b},$$

for any $a \geq b \geq 0$.

Problem 9: Prove that

$$\left(\frac{n}{e}\right)^n < n! < e\left(\frac{n}{2}\right)^n, \quad n \geq 1.$$

Hint: To prove $a_1 < b_1 \implies a_2 < b_2$ it is enough to show

$$\frac{a_2}{a_1} \leq \frac{b_2}{b_1}.$$

You may also need the information that $(1 + 1/n)^n < e < (1 + 1/n)^{n+1}$ for all $n \geq 1$.

Problem 10: Find the limit

$$\lim_{n\to\infty}\left(\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{n(n+1)}\right)\ .$$

Problem 11: Consider the sequence $x_1 = 4$, $x_{n+1} = \frac{x_n^2+11}{2x_n}$, $n = 1, 2, \ldots$. Prove that it is convergent and find the limit.

Hint 1: Prove that $\{x_n\}$ is decreasing and bounded below.

Hint 2: You need to prove that $x_n^2 \geq 11$. Induction and inequality $a^2 + b^2 \geq 2ab$, for $a, b \geq 0$ can be useful.

Problem 12: (difficult) Is $\sqrt[3]{2+\sqrt{5}}+\sqrt[3]{2-\sqrt{5}}$ a rational number? Prove or disprove.

Assignment # 5

Problem 1:

The sequence $(x_n)_{n=1}^{\infty}$ is decreasing iff

$$x_n \geq x_{n+1} \quad , \ n = 1, 2, \ldots$$

Prove the following theorem: A decreasing sequence bounded below is convergent.

Problem 2: Let $x_1 = 3$ and $x_{n+1} = \sqrt{2 + x_n}$ for $n = 1, 2, \ldots$ Prove that $(x_n)_{n=1}^{\infty}$ is convergent and find the limit.

Hint: Show that $(x_n)_{n=1}^{\infty}$ is decreasing, bounded below by 2.

Problem 3: Use squeeze theorem to find the following limits

$$(a) \quad \lim_{n\to\infty} \sqrt[n]{2^n + e^n + \cos^2(10^n + 10!)}\ ;$$

$$(b) \quad \lim_{n\to\infty} \sqrt[n]{2^n + \frac{1}{1000^{1000}}e^n - \cos^2(1000^n + 1000!)}\ ;$$

$$(c) \quad \lim_{n\to\infty} \frac{1}{n}\log(3^n - e^n)\ .$$

Problem 4: Prove that the sequence

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

is decreasing.

Problem 5: Let A be a nonempty bounded subset of $\mathbb{R}$ with $\alpha = \sup A$ and $\beta = \inf A$. Show that A contains a monotone increasing sequence with limit α and a monotone decreasing sequence with limit β. Hint: consider cases $\alpha \in A$ and $\alpha \notin A$ and then $\beta \in A$ and $\beta \notin A$.

Problem 6: Show that a bounded sequence in $\mathbb{R}$ that does not converge has more than one subsequential limit. That is, show that a non-convergent bounded sequence has two subsequences each with a different limit.

Problem 7: Which of the sequences below are Cauchy sequences?

(a) $a_n = 1 + \frac{1}{4} + \frac{2^2}{4^2} + \cdots + \frac{n^2}{4^n}$,

(b) $b_n = \alpha_1 q^1 + \alpha_2 q^2 \cdots + \alpha_n q^n$, for $|\alpha_i| \leq M$ and $|q| < 1$,

(c) $c_n = 1 + \frac{1}{2^2} + \frac{2}{3^2} + \cdots + \frac{n}{(n+1)^2}$.

Problem 8: Find the limit of

$$x_n = \left(\sqrt[n]{n} - 1\right)^{\frac{1}{n}} .$$

Hint: You can use the fact that $(1 + 1/n)^n \to e$ as $n \to \infty$.

Assignment # 6

Problem 1: Consider a sequence $\{a_n\}_{n=1,2,\ldots}$ such that the subsequences $\{a_{2k}\}_{k=1,2,\ldots}$, $\{a_{2k-1}\}_{k=1,2,\ldots}$ and $\{a_{3k}\}_{k=1,2,\ldots}$ are convergent. Prove that the sequence $\{a_n\}_{n=1,2,\ldots}$ converges.

Problem 2:

Find $\limsup x_n$ and $\liminf x_n$ for:

(a) $x_n = \pi^{-n(-1)^n}$, (b) $x_n = 2^n + (-2)^n$, (c) $x_n = \sin(n\pi/2) + \cos(n\pi/2)$,

Problem 3: Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be **positive** sequences in $\mathbb{R}$. Show that

$$\liminf a_n \cdot \liminf b_n \leq \liminf(a_n \cdot b_n) .$$

Give example with the strict inequality.

Remark: You can use the theorem proved in class: For a sequence (x_n) there exist subsequences (x_{n_k}) and (x_{n_ℓ}) such that

$$\lim_{k\to\infty} x_{n_k} = \limsup_{n\to\infty} x_n \quad , \quad \lim_{\ell\to\infty} x_{n_\ell} = \liminf_{n\to\infty} x_n .$$

Problem 4: Let (x_n) be a bounded sequence and let (y_n) be a convergent sequence. Show

$$\liminf x_n + \liminf y_n = \liminf(x_n + y_n)\ .$$

Problem 5: Let the sequence $\{a_n\}$ satisfy

$$\left|\frac{a_{n+1} - a_n}{a_n - a_{n-1}}\right| \le r < 1,$$

for $a = 2, 3, 4, \ldots$. Prove that $\{a_n\}$ is Cauchy.

Problem 6: Let $\{a_n\}$ be the sequence defined by the recursive formula

$$a_1 = 1, \quad a_{n+1} = \frac{2 + a_n}{1 + a_n}, \quad n = 1, 2, 3, \ldots$$

Show that the sequence is Cauchy and find its limit.

Hint: Show that $a_n \ge 1$ for $n \in \mathbb{N}$. Use Problem 5.

Problem 7: Let $f : \mathbb{N} \to \mathbb{Q}$ be a bijection. Does the sequence $\{f(n)\}$ contain a convergent subsequence?

Problem 8: Find $\limsup a_n$ and $\liminf a_n$ for

$$a_n = |\sin n| + |\cos n|\ , \quad n = 1, 2, \ldots$$

Hints: (i) Prove $1 \le |\sin n| + |\cos n| \le \sqrt{2}$.

(ii) Accept as true the Theorem: If $\alpha \notin \mathbb{Q}$, then the sequence $n\alpha \pmod 1$, $n = 1, 2, 3, \ldots$, is dense in the interval $[0, 1]$.

(iii) Use the Theorem to show that the sequence $n \pmod{2\pi}$, $n = 1, 2, 3, \ldots$, is dense in the interval $[0, 2\pi]$. You can use the fact that $\pi \notin \mathbb{Q}$.

Problem 9: Find $\limsup x_n$ and $\liminf x_n$ for

$$x_n = \sin\left(\pi\sqrt{N^2 + n}\right), \quad n = 1, 2, \ldots$$

Assignment # 7

Problem 1: Use ε-δ definition to prove:

(a) $\lim_{x\to 3} x^2 + 1 = 10$;

(b) $\lim_{x\to 1} \frac{x}{x^2+2014} = \frac{1}{2015}$;

(c) $\lim_{x\to 0} \frac{\sin x}{x^2+1} = 0$.

Problem 2: Let $f(5) = \lim_{x\to 5} f(x) = 7$. Prove that there exists a small open interval I containing 5, such that $f(x) > 6$ for all $x \in I$.

Problem 3: Let $\lim_{x\to 6} f(x) = 15$. Use Cauchy definition of limit to prove the following:

(a) Prove that there exists a sequence (x_n) such that $x_n \to 6$ and $f(x_n) \to 15$ as $n \to \infty$.

(b) Prove, that if (x_n) is any sequence such that $x_n \neq 6$ for all n and $x_n \to 6$, then

$$f(x_n) \to 15 \ , \ \text{as } n \to \infty \, .$$

Problem 4:

Use the ε-δ definition to prove that $f(x) = x^3 + 3$ is continuous at $x_0 = 3$.

Problem 5: The function f satisfies $|f(x)| \le x^2$, for all $x \in \mathbb{R}$. Prove that $\lim_{n\to 0} f(x) = 0$.

Problem 6: Use the ε-δ definition to prove that $f(x) = \dfrac{1}{x^2+1}$ is continuous at $x_0 = 1$.

Problem 7: Find the limit

$$\lim_{n\to 0} x^3 \cos\left(\frac{x^3+x+1}{x^6+x^4+64}\right) .$$

Problem 8: Assume $f : (-a, a) \setminus \{0\} \to (0, +\infty)$ and

$$\lim_{x\to 0}\left(f(x) + \frac{1}{f(x)}\right) = 2.$$

Prove that $\lim_{x\to 0} f(x) = 1$.

Hint: First show that $f(x) + \frac{1}{f(x)} \ge 2$.

Problem 9: Assume that n is a prime number. Show that $n^2 - 1$ is divisible by 12.

Problem 10: Find the limits

$$(a) \quad \lim_{n\to\infty} \frac{3^n + (-2)^n}{3^{n+1} + (-2)^{n+1}} \ .$$

$$(b) \quad \lim_{n\to\infty} \frac{1 + a + a^2 + a^3 + \cdots + a^n}{1 + b + b^2 + b^3 + \cdots + b^n} \quad , \ |a| < 1, \ |b| < 1 \ .$$

Assignment # 8

Problem 1: Show that the polynomial $P(x) = x^{2018} + 3x - 1$ has a zero in the interval $(0, 1)$. In which of the intervals $(0, 1/4), [1/4, 1/2), [1/2, /3/4), [3/4, 1)$ is the zero?

Problem 2:

We say that $f : \mathbb{R} \to \mathbb{R}$ satisfies Hölder condition if

$$|f(x) - f(y)| \leq H|x - y|^{\alpha}, \quad x, y \in \mathbb{R}, \quad H, \alpha > 0.$$

If $\alpha = 1$ then, the above is called Lipschitz condition.

(a) Prove: If f satisfies Hölder condition, then f is uniformly continuous.

(b) Prove: If f satisfies Hölder condition with $\alpha > 1$, then f is constant.

Problem 3: Prove: If f is continuous on $I = (0, 1]$ and $\lim_{x \to 0^+} f(x) = L$ is finite then f is bounded on I.

Hint: As always there are many ways to do this problem. One of them can use the fact that a continuous function on any closed finite interval is bounded.

Problem 4: (a) Use Cauchy definition of continuity and necessary results on continuous functions to prove that if f and g are continuous at point x_0, then their product fg is also continuous at x_0.

(b) Use Heine definition of continuity and necessary results on arithmetics of limits to prove that if f and g are continuous at point x_0, then their product fg is also continuous at point x_0.

Problem 5: (a) Prove that

$$\lim_{n \to \infty} \frac{1}{2^n} \ln \left(\frac{1}{2^n} \right) = 0.$$

Hint: Use the properties of the logarithms.

(b) Use result in (a) to show that

$$\lim_{x \to 0^+} x \ln x = 0.$$

Hint: Use Heine definition of the limit of function and note that for any number t close to 0 we can find a natural n such that $1/2^{n+1} < t \leq 1/2^n$.

(c) Use result in (b) to prove that for any $\alpha > 0$ we have

$$\lim_{x\to 0^+} x^\alpha \ln x = 0.$$

Problem 6: **True or false?** Justify Your answer.

a) If M is the maximum value of f, then $|M|$ is the maximum value of $|f|$.

b) If M_1 is the maximum value of f and M_2 is the maximum value of g , then $M_1 + M_2$ is the maximum value of $f + g$.

c) If M_1 is the minimum value of f and M_2 is the minimum value of g , then $M_1 \cdot M_2$ is the minimum value of $f \cdot g$.

d) If f and g are bounded on $[a, b]$ and $g(x) \neq 0$ for all $x \in [a, b]$, then f/g is bounded on $[a, b]$.

e) If f is bounded on $[a, b]$, g is continuous on $[a, b]$, and $g(x) \neq 0$ for all $x \in [a, b]$, then f/g is bounded on $[a, b]$.

f) If f and g are bounded on (a, b), g is continuous on (a, b), and $g(x) \neq 0$ for all $x \in (a, b)$, then f/g is bounded on (a, b).

g) If f is bounded on every closed subinterval of (a, b), then f is bounded on (a, b).

h) There exists a bounded function on $[0, 1]$ which achieves neither an infimum nor a supremum.

Problem 7: Let $f : [0, 1] \to [0, 1]$ be an increasing function (not necessarily continuous). Prove that it has a fixed point (i.e., there exists $x \in [0, 1]$ such that $f(x) = x$).

Problem 7: Prove that the number $\log_3 5$ is irrational.

Assignment # 9

Problem 1. Let $f : D \to \mathbb{R}$ be a continuous function, where $D \subset \mathbb{R}$.

(i) If $D = \mathbb{R}$ and $xf(x) \geq 0$ for all x, prove that there is some x_0 in $\mathbb{R}$ with $f(x_0) = 0$.

(ii) What is false in the following argument? We are assuming $f : \mathbb{R} \to \mathbb{R}$ is continuous, and $f(x) \neq 0$ for all $x \in \mathbb{R}$. Then the function $g = \frac{1}{f}$ is continuous on $\mathbb{R}$.

Suppose there are points $a, b \in \mathbb{R}$ with $f(a) < 0 < f(b)$. Then $g(a) < 0 < g(b)$, so we can use the Intermediate Value Theorem on g to conclude that there exists an x between a and b with $g(x) = 0$, meaning

$$\frac{1}{f(x)} = 0.$$

Problem 2.

(i) Write the definition and the **negation** of the definition of uniform continuity for a function f on a domain D in $\mathbb{R}$.

(ii) Prove that the function $x \to \frac{1}{x}$ is not uniformly continuous on $D = \mathbb{R} \setminus \{0\}$.

Hint: You can either use the negation of the definition or a theorem about uniform continuity and Cauchy sequences.

(iii) If f is continuous on $D = [a, b]$, and $f \neq 0$ on D, prove that the function $\frac{1}{f}$ is uniformly continuous on D.

Hint: You can either use the definition (Write $\left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| = \frac{|f(y)-f(x)|}{|f(x)f(y)|}$.) or a theorem about continuous function on $[a, b]$ and uniform continuity.

Problem 3. Prove that if $f : (a, b) \to \mathbb{R}$ is continuous and f is NOT bounded above or below, then f is surjective, i.e. for every $y \in \mathbb{R}$, there exists $x \in \mathbb{R}$ with $f(x) = y$.

Hint: Show that the assumptions mean that for every m and M there have to be points a', b' in (a, b) with $f(a') < m$ and $f(b') > M$.

Problem 4 A tourist walked on Monday from village A to village B. He started at 9 a.m. and finished at 9 p.m. (He was a dedicated tourist.) Next day he went back from B to A along the same path but in the opposite direction. Again, he started at 9 a.m. and finished at 9 p.m. Prove, that there exists a point on the path at which he was at the same time on both days. Are the starting and finishing times important in this problem?

Problem 5 Show that

$$f(x) = \begin{cases} 0, & \text{for } x = 0; \\ x^7 \sin(1/x^6), & \text{for } x \neq 0; \end{cases}$$

is differentiable, but the derivative is not continuous at 0.

Problem 6 Consider the function

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

Show that f is differentiable at 0 and $f'(0) = -\frac{1}{2} < 0$. Prove that arbitrarily close to 0 there are intervals on which f is increasing and intervals on which f is decreasing.

Comment: This shows that value of the derivative at one point does not determine the increasing/decreasing in a neighbourhood of this point.

Interesting Limits

(1) $\lim_{n\to\infty} \frac{a^x}{x^\alpha}$, $a > 1$, $\alpha \in \mathbb{R}$.

(2) $\lim_{n\to\infty} \frac{\ln x}{x^\alpha}$, $\alpha > 0$.

(3) $\lim_{n\to 0} a^x$, $a > 0$.

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

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

(6) $\lim_{n\to 0} (1 + x)^{\frac{1}{x}}$.

(7) $\lim_{n\to 0} \frac{\ln(1+x)}{x}$.

(8) $\lim_{n\to 0} \frac{(1+x)^\alpha - 1}{x}$, $\alpha \in \mathbb{R}$.

(9) $\lim_{n\to 0} \frac{a^x - 1}{x}$, $a > 0$.

(9a) $\lim_{n\to\infty} \ln\left(e^x - 2^x \ln x + \cos^2(x^{x^x})\right)$.

(10) $\lim_{n\to 0} (\ln x)^{\frac{1}{x}}$.

(11) $\lim\limits_{n\to 0} (\cos x)^{\frac{1}{\sin^2 x}}$.

(12) $\lim_{n\to 0^+} (\sin x)^{\frac{1}{\ln x}}$.

(13) $\lim_{n\to 0^+} x^{\sin x}$.

(14) $\lim_{n\to\infty} (e^x - 1)^{\frac{1}{x}}$.

(15) $\lim_{n\to\infty} \left(\sqrt{x + 2014} - \sqrt{x}\right)$.

(16) $\lim_{n\to 0} \frac{\sin 2x + 2 \arctan 3x + 3x^2}{\ln(1 + 3x + \sin^2 x) + xe^x}$.

(17) $\lim_{n\to 0} \frac{\ln\cos x}{\tan x^2}$.

(18) $\lim_{n\to 0^+} \frac{\sqrt{1-e^{-x}}-\sqrt{1-\cos x}}{\sqrt{\sin x}}$.

(19) $\lim_{n\to 0}(1+x^2)^{\frac{1}{\tan x}}$.

(20) $\lim_{n\to\infty} \left(\tan\frac{\pi x}{2x+1}\right)^{\frac{1}{x}}$.

(21) $\lim_{n\to\infty} x\left(\ln\left(1+\frac{x}{2}\right)-\ln\frac{x}{2}\right)$.

(22) $\lim_{n\to 0^+} \left(2\sin\sqrt{x}+\sqrt{x}\sin\frac{1}{x}\right)^{x}$.

(23) $\lim_{n\to 0} \left(1+xe^{-\frac{1}{x^2}}\sin\frac{1}{x^4}\right)^{e^{\left(\frac{1}{x^2}\right)}}$.

(24) $\lim_{n\to 0} \left(1+e^{-\frac{1}{x^2}}\arctan\frac{1}{x^2}+xe^{-\frac{1}{x^2}}\sin\frac{1}{x^4}\right)^{e^{\left(\frac{1}{x^2}\right)}}$.

References

[1] Donoghue, William F., Jr., Distributions and Fourier transforms, Pure Appl. Math., 32, Academic Press, New York, 1969, viii+315 pp., MR3363413.

[2] Lee Larson, Introduction to Real Analysis, University of Louisville, available at https://www.kufunda.net/publicdocs/IntroRealAnal.pdf.

[3] Niculescu, Constantin P.; Persson, Lars-Erik, Convex functions and their applications – a contemporary approach, CMS/CAIMS Books Math., 14, Springer, Cham, 2025, xxii+494 pp., MR4888856.

Index