HIGHER MATHEMATICS

Lectures

Part One

HIGHER MATHEMATICS

Lectures

Part One

Stefan Wurm

A·T·I·C·E

ATICE LLC, Albany NY

First English language ATICE e-Book Edition | ISBN 978-1-951894-13-9.

Contents

Preface

At German universities, lectures on higher mathematics are an integral part of the curriculum in natural and engineering sciences. These lectures aim to provide students with the mathematical foundations for their respective subject areas, typically in the first four semesters. This was also the case for me when a good forty years ago at the beginning of my physics studies I first entered the lecture hall of the Technische Universität München (TUM), the place where Prof. Dr. Armin Leutbecher taught Higher Mathematics. I realize of course that not everyone can or wants to share the same enthusiasm for mathematics. However, I hope that those who are reading these lines will understand what I mean when I say that to me as a student those mathematics courses have been a real source of happiness. Happiness in the sense that back then I always looked forward to each and every one of these lectures. This certainly did not only have to do with the content of the lectures, but at least as much with the way they were delivered by Prof. Leutbecher. Of course, one always expects clarity from a mathematician. But the clarity with which professional mathematicians generally conduct their discussions does not necessarily carry over to how a mathematician might then impart his knowledge to students. Prof. Leutbecher's clarity and style of delivery made his Higher Mathematics lectures an intellectual delight. In addition, I also had the good fortune that the exercises for Prof. Leutbecher lectures were given by Dr. Peter Vachenauer. At the beginning of the 1990s the first edition of a two-volume textbook on Higher Mathematics co-authored by Dr. Vachenauer was published. The exemplary methodology and care with which the lecture materials were studied during my time at TUM in Dr. Vachenauer's tutorial exercises is reflected in this textbook.

A little over a year ago, while tidying up, I stumbled across my transcripts of the Higher Mathematics lectures from the years 1981-1983 and the corresponding exercises. At first I was surprised that these forty-year-old documents were not lost during various moves over four decades, some of them between continents. When I then curiously began to leaf through my rediscovered lecture notes I all of a sudden experienced the same kind of joy which I once felt when I was sitting in the lecture hall, listening spellbound to Prof. Leutbecher's lectures some forty years ago. Although these notes, my transcript of Prof. Leutbecher's lectures, cannot replace a textbook, they do convey the essential content of Higher Mathematics with a vividness that I believe should make them a reading pleasure for students or anyone else seriously interested in mathematics. All too often such lecture notes are riddled with errors, and this was no different here. After reviewing and correcting my notes several times, hopefully the vast majority of them have been corrected. Preserving the clarity and style of Prof. Leutbecher's lectures, as I captured them in my notes more than forty years ago, was something I attached great importance to when revising my notes. Translating those notes from their original German into English added of course another challenge. Quite likely some of the elegance of the German language lectures may have been lost in translation. However, I do hope that the English language version still conveys the essence of the lectures original style and clarity. This volume, **HIGHER MATHEMATICS - Lectures Part One**, contains the material of the Higher Mathematics I lectures as given by Prof. Leutbecher in the winter semester 1981/82 at the TUM.

Stefan Wurm

Albany, New York
April 2022

1. Complete Mathematical Induction

1.1 The Principle of Complete Induction

Examples:

(1) Summation of the first n numbers

Sum the first n numbers			n
$1 + 2$	$=$	3	2
$1 + 2 + 3$	$=$	6	3
$1 + 2 + 3 + 4$	$=$	10	4
$1 + 2 + 3 + 4 + 5$	$=$	15	5

From the table it can be seen that for the sums of the first n numbers s_n with $n = 2, 3$ and 4:

$$2s_2 = 2 \cdot 3 \quad ; \quad 2s_3 = 3 \cdot 4 \quad ; \quad 2s_4 = 4 \cdot 5 \qquad \curvearrowright$$

presumably: $\quad 2s_n = n\,(n+1)$

Writing the terms of the sum in s_n one above the other from 1 to n and from n to 1 and adding:

$$\begin{array}{ccccccccccccc} 1 & + & 2 & + & 3 & + & ... & + & (n-1) & + & n & = & s_n \\ n & + & (n-1) & + & (n-2) & + & ... & + & 2 & + & 1 & = & s_n \\ \hline (n+1) & + & (n+1) & + & (n+1) & + & ... & + & (n+1) & + & (n+1) & = & 2\,s_n \end{array}$$

The number of summands in the last line is n, therefore: $2s_n = n\,(n+1)$

(2) Summation of the first n squares

Sum the first n squares					n
$1+4$	$=$	5	$=$	$1\cdot 5$	2
$1+4+9$	$=$	14	$=$	$2\cdot 7$	3
$1+4+9+16$	$=$	30	$=$	$\frac{10}{3}\cdot 9$	4
$1+4+9+16+25$	$=$	55	$=$	$5\cdot 11$	5

Discovery: Multiplying the first factor after the second equals sign by 3 gives the sequence in Example (1).

$$\text{Bet that:}\qquad s_n \;=\; \frac{1}{6}\,n\,(n+1)\,(2n+1)$$

At least that is correct for $n = 1, 2, 3, 4$, and 5. Assuming the bet assertion holds for n, then

$$s_{n+1} \;=\; s_n \;+\; (n+1)^2 \;=\; \frac{1}{6}\,(n+1)\,(n+2)\,(2n+3)$$

This is the assertion for $n+1$ instead of n.

(3) Recursive definition of a sequence

- let $x \;\geq\; 0$ be given
- $w_0 \;=\; 1+x$ assume that n is defined
- $w_{n+1} \;=\; \dfrac{1}{2}\left(w_n + \dfrac{x}{w_n}\right)$

this is possible because w_n is always greater than 0. According to the principle of complete induction, w_n is thus explained for any natural number $n \geq 0$.

The principle of complete induction expresses a property which belongs to the totality of natural numbers $n \geq 1$.

Definition

If M is a set of natural numbers, which contains the number 1 and also contains $n+1$ if it contains n, then M is the set $\mathbb{N}$ of all natural numbers.

Example:

(4) The factorials $n!$

$$0! \;=\; 1 \quad ; \quad (n+1)! \;=\; n!\,(n+1)$$

n	0	1	2	3	4	5
$n!$	1	1	2	6	24	120

<u>The combinatorial significance of factorials:</u>

$n!$ = number of possibilities for arranging n items $a_1, a_2, \cdots, a_n$ in a row.

n = 2:	1	2	n = 3:	1	2	3	n = 4:	1	2	3	4
	2	1		1	3	2		1	2	4	3
				2	1	3		1	$\vdots$	$\vdots$	$\vdots$
				2	1	3		$\vdots$	$\vdots$	$\vdots$	$\vdots$
				2	3	1					
				3	1	2					
				3	3	1					

Proof of the claim about the number $n!$ = number A_n of possible permutations of n objects by complete induction.

$n = 1$: $A_1 = 1$

$n \to n+1$: There are exactly $n+1$ possibilities to fill the first position from $n+1$ objects, after the proven part there are $n!$ possibilities to fill the remaining positions.

Combined: $(n+1)\,n! \;=\; A_{n+1} \;=\; (n+1)!$

$\square$

1.2 Bernoulli's Inequality

Let $k > -1$. Then the following inequality holds for all natural numbers n:

$$\boxed{(1+k)^n \;\geq\; 1 + n\,k}$$

Proof by complete induction:

$n = 1$: Left as right we have $1 + k$

$n \to n+1$: $(1+k)^{n+1} \;=\; (1+k)\,(1+k)^n$ according to the definition of powers

$\geq\; (1+k)\,(1+n\,k)$ according to the induction hypothesis and because inequalities can be multiplied by positive factors

$$\text{Combined:} \quad (1+k)^{n+1} \;\geq\; 1 + (n+1)\,k + n\,k^2$$
$$\geq\; 1 + (n+1)\,k$$

because $n\,k^2 \geq 0$ this is the assertion for $n+1$ instead of n

□

Remarks:

(1) The importance of Bernoulli's inequality lies in the estimate given for the powers. For k close to 0 this estimate is the best possible.

(2) The powers a^n for arbitrary real base numbers a are also explained recursively:

$$a^n \;:=\; 1 \quad ; \quad a^{n+1} \;:=\; a^n\,a$$

1.3 Binomial coefficients

$\binom{n}{m}$ are defined for all pairs n, m of natural numbers $\leq\ 0$

$$\binom{n}{0} \;:=\; 1 \quad ; \quad \binom{n}{m+1} \;:=\; \binom{n}{m}\frac{n-m}{m+1}$$

$\binom{n}{m}$ is always a natural number $\geq\ 0$ $\left[\binom{n}{m} \in \mathbb{N}\right]$ and in the case $0 \leq m \leq n$

$$\binom{n}{m} \;=\; \frac{n!}{m!\,(n-m)!} \;=\; \binom{n}{n-m}$$

while $\binom{n}{m} = 0$ if $m > n$. Furthermore, the identity of Pascal's triangle applies

$$\binom{n}{m} + \binom{n}{m+1} \;=\; \binom{n+1}{m+1}$$

Proof of the first equation by induction on m

$$[m = 0: \qquad \binom{n}{m} \;=\; 1 \quad ; \quad \frac{n!}{m!\,(n-m)!} \;=\; 1$$

$$m \to m+1: \qquad \binom{n}{m+1} \;=\; \binom{n}{m}\frac{n-m}{m+1}$$

Right side of the equation:

$$\frac{n!}{m!\,(m+1)\,(n-m+1)!} \;=\; \frac{n!}{m!\,(n-m)!}\,\frac{n-m}{m+1}$$

So the assertion is true for ür $m+1$ instead of m. In the case of $m = n$ consider the recursion formula

$$\binom{n}{n+1} = * \cdot \frac{0}{*}$$

Hence for $m > n$ always $\binom{n}{m} = 0$

□

The formula of Pascal's triangle is correct if $m \geq n$ (according to the proven = upper side of Pascal's triangle)

$$\binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}$$

It remains to show the validity for $0 \leq m < n$. Formula for the summands:

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$$

$$\binom{n}{m+1} = \frac{n!}{(m+1)!\,(n-m+1)!}$$

$$\binom{n}{m} + \binom{n}{m+1} = \frac{n!}{(m+1)!\,(n-m)!}\,\underbrace{[(m+1)+(n-m)]}_{(n+1)}$$

$$= \frac{(n+1)!}{(m+1)!\,(n+1-m-1)!} = \binom{n+1}{m+1}$$

In tabular form the formula of Pascal's triangle gives for the possible binomial coefficients $\binom{n}{m}$ with $0 \leq m < n$ and n, m each in the value range of 0 to 5 and $\binom{n}{m} = 0$ for $m > n$:

		m					
		0	1	2	3	4	5
	0	1	0	0	0	0	0
	1	1	1	0	0	0	0
n	2	1	2	1	0	0	0
	3	1	3	3	1	0	0
	4	1	4	6	4	1	0
	5	1	5	10	10	5	1

1.4 Combinatorial Significance of $\binom{n}{m}$

It is the number of possible ways to select $a_1, a_2, \cdots, a_n$ subsets of m elements from a set of n elements.

Proof:

This number is abbreviated with $C^{\mathrm{n}}_{\mathrm{m}}$. So the assertion is that:

$$C^{\mathrm{n}}_{\mathrm{m}} = \binom{n}{m}$$

Proof preparation:

$$C^{\mathrm{n}}_{0} = \binom{n}{0}$$

This statement is true because nothing is to be selected (the empty set $\emptyset$ is to be selected).

Complete induction by n:

$n = 0$: $C^{0}_{\mathrm{k}} = \binom{0}{k}$ for $k \geq 0$ is correct for a similar reason, since $\emptyset$ has no k-element subset

$n \to n+1$: Being sought after is the number of ways to select $(m+1)$-element subsets from $a_1, a_2, \cdots, a_{n+1}$!

Divide these sets T into two classes:

$$A = \text{sets } T \text{ with } a_{n+1} \in T$$

$$B = \text{sets } T \text{ with } a_{n+1} \notin T$$

A has $C^{\mathrm{n}}_{\mathrm{m}}$ elements T, because there are still n elements from $a_1, a_2, \cdots, a_n$ to be added.

B has $C^{\mathrm{n}}_{\mathrm{m+1}}$ elements T, because all $m+1$ elements of T have to be chosen from the smaller set $a_1, a_2, \cdots, a_n$!

$$\begin{aligned} C^{\mathrm{n+1}}_{\mathrm{m+1}} &= C^{\mathrm{n}}_{\mathrm{m}} + C^{\mathrm{n}}_{\mathrm{m+1}} \\ &= \binom{n}{m} + \binom{n}{m+1} \qquad \text{according to the induction premise} \end{aligned}$$

$$= \binom{n+1}{m+1} \qquad \text{according to the identity of Pascal's triangle}$$

□

1.5 The Sum Sign

$$\sum_{m=n}^{N} a_m$$

is declared in the special case that $n > N$ by 0 and otherwise recursively by:

$$\sum_{m=n}^{N+1} a_m = \sum_{m=n}^{N} a_m + a_{N+1}$$

Important is the possibility of reindexing, for example:

$$\sum_{m=n}^{N} a_m = \sum_{m=n+1}^{N+1} a_{m-1}$$

The sum formula of the finite geometric series! For every real (and also every complex number) $x \neq 1$ the following applies to all $n \in \mathbb{N}_0$:

$$\sum_{m=0}^{n} x^m = \frac{x^{n+1}-1}{x-1}$$

Proof by complete induction

$n = 0$: Left side $= 1$,

right side $= \dfrac{x-1}{x-1} = 1$,

hence the value 1 is on both sides.

$n \to n+1$: Consider

$$\sum_{m=0}^{n+1} x^m = \sum_{m=0}^{n} x^m + x^{n+1}$$

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

□

1.6 Binomial Theorem

Let a, b be real (or complex). Then for all $n \in \mathbb{N}_0$ holds

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

Proof by induction on n

$n = 0$: Left side $= (a+b)^0$, right side $= \binom{0}{0} a^0\, b^0$,

hence on both sides there is the value 1 .

$n \to n+1$: $(a+b)^{n+1}$

$$\begin{aligned}
&= (a+b)\,(a+b)^n \\
&= (a+b) \sum_{k=0}^{n} \binom{n}{k} a^{n-k}\, b^k \qquad \text{induction premise} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{n+1-k}\, b^k + \sum_{k=0}^{n} \binom{n}{k} a^{n-k}\, b^{k+1} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{n+1-k}\, b^k + \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n+1-k}\, b^k \\
&= a^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1}\right] a^{n+1-k}\, b^k + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k}\, b^k
\end{aligned}$$

This proves the formula for $n+1$ instead of n

$\square$

The cases $n = 2, 3, 4$

$$(a+b)^2 = a^2 + 2\,a\,b + b^2$$

$$(a+b)^3 = a^3 + 3\,a^2 b + 3\,a\,b^2 + b^3$$

$$(a+b)^4 = a^4 + 4\,a^3 b + 6\,a^2 b^2 + 4\,a\,b^3 + b^4$$

Two cases of special a, b and general n:

$$a = b = 1: \qquad \sum_{k=0}^{n} \binom{n}{k} \cdot 1 \cdot 1 = 2^n$$

Because $\binom{n}{k}$ is the number of k-element subsets of a set with n elements, the formula shows: A set with n elements has 2^n subsets.

$$a = 1, \; b = -1 \qquad \sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0 \qquad \text{if} \quad n > 0$$

The sum of the alternating sign elements of the n-th row of Pascal's triangle is 0.

2. Inequalities and Absolute Value

Inequalities are mathematical sentences which express size comparisons like $a \leq b$ or $a < b$ between real numbers. All inequalities (in this sense) can be traced back to the concept of positivity: $a \leq b$ means $b - a$ is positive or 0, while $a < b$ means $b - a$ is positive. The first kind of inequality is called ordinary inequality, while the second kind is called strict inequality. The significance of positivity: The positive numbers (the set of all positive real numbers $]0, \infty[$) contain with any two elements their sum and their product.

From this follow 3 basic rules:

U-1 One can add the same number to both sides of an inequality.

U-2 One may multiply left and right sides of an inequality by the very same positive factor.

U-3 When multiplying by (-1), each inequality is reversed (for example: $a \leq b \Rightarrow -b \leq -a$).

All 3 rules apply to ordinary and strict inequalities. The comparability of any two real numbers can be discerned from the positions of real numbers as displayed on the number line.

U-4 For each pair of real numbers a, b exactly one of the following 3 relationships applies:

$$a = b, \quad a < b, \quad b < a$$

Application to the comparison of the reciprocals of positive numbers.

$$0 < a < b, \quad \Rightarrow \quad 0 < \frac{1}{b} < \frac{1}{a}$$

Proof: Asserted is also that if b is positive

$$\Rightarrow \quad \frac{1}{b} \quad \text{positive} \quad (0 < b \Rightarrow 0 < \frac{1}{b})$$

Indirect: certainly $\frac{1}{b} \neq 0$. Assumption: $\frac{1}{b} < 0$. Application of U-2 gives:

$$1 = b\frac{1}{b} < b \cdot 0 = 0$$

$1 < 0$ is wrong, hence the assumption $\frac{1}{b} < 0$ was wrong. Therefore, after U-4 $\frac{1}{b} > 0$.
The full assertion follows from consideration of

$$\frac{1}{a} - \frac{1}{b} = \frac{1}{a\,b}(b - a)$$

On the right is the product of two positive numbers, hence a positive number, hence

$$\frac{1}{b} < \frac{1}{a}$$

□

2.1 The Absolute Value

The absolute value of real numbers x is defined as:

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

Elementary conclusions!

- Always $|x| \geq 0$, "=" stands for $x = 0$
- Always $|-x| = |x|$
- Always $|x\,y| = |x|\,|y|$
- If $y \neq 0$, then $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$

The triangle inequality

For all real x, y the following applies:

$$\boxed{|x + y| \leq |x| + |y|}$$

Proof: It always holds true that

$$x \;\leq\; |x| \quad ; \quad y \;\leq\; |y|$$

likewise

$$-x \;\leq\; |x| \quad ; \quad -y \;\leq\; |y|$$

Applying U-1 twice gives

$$x \,+\, y \;\leq\; |x| \,+\, |y|$$

likewise

$$-x \,-\, y \;\leq\; |x| \,+\, |y|$$

That means:
$|x| + |y|$ is an upper bound for the two numbers $x + y$ and $-(x + y)$. Because one of these two numbers is equal to $|x + y|$, the triangle inequality holds.

□

Remarks:

(1) The background of the name becomes clear with the transition from real to complex numbers.

(2) The inverse triangle inequality

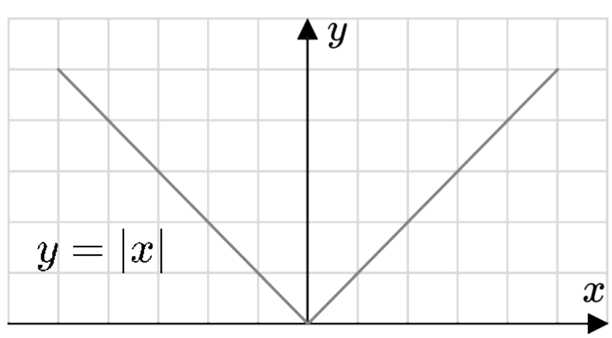
$$|x + y| \;\geq\; |x| \,-\, |y|$$

follows easily from the triangle inequality. It is also often used in estimates.

(3) The instruction

$$x \;\rightarrow\; |x|$$

declares a function of the real numbers, the absolute value function.



Because the graph does not change when it is reflected on the y-axis, the absolute value function is an “even function”.

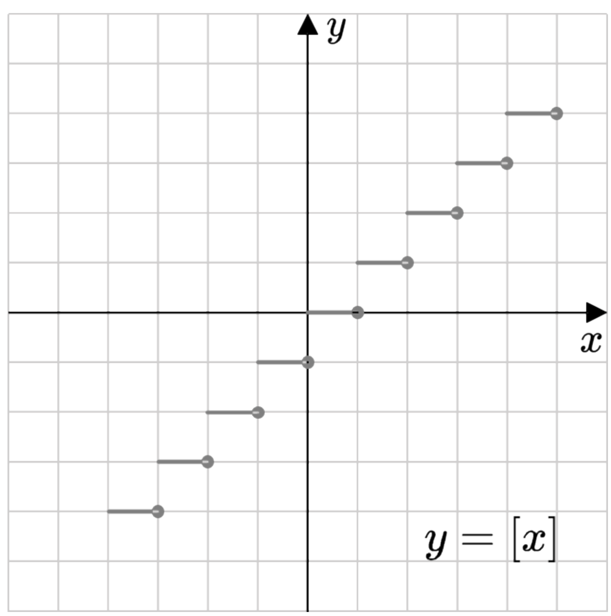
The Archimedean property of the arrangement of $\mathbb{R}$, the set of all real numbers, means that:

> For every real number x there is a natural number n such that $x < n$.

For every real number x this property returns an integer number $m = [x]$ with the property

$$\boxed{[x] \;\leq\; x \;\leq\; [x] + 1}$$

The Entire function is declared by the prescription $x \to [x]$.



Conclusions from the Archimedean property

(1) For every number ϵ, no matter how small, there is a $n \in \mathbb{N}$ with $\dfrac{1}{n} < \epsilon$

Proof: $x := \dfrac{1}{\epsilon}$

According to Archimedes, there exists $\dfrac{1}{n} < \epsilon$ with $x < n$, hence $\dfrac{1}{n} < \epsilon$

$\square$

(2) For given $b > 1$ and arbitrary bound $M > 0$, there exists a $n \in \mathbb{N}$ with $b^n < M$

Proof:

We have $\quad b^n \;\geq\; 1 + n\,k \qquad (b \;=\; 1 + k)$

With Archimedes, for $x := \dfrac{M}{k}$ there exists a $n \in \mathbb{N}$ with $x < n$, i.e. $M < n\,k$

Therefore $\quad b^n \;\geq\; 1 + n\,k \;>\; M$

□

(3) For given q with $0 < q < 1$ and for every $\epsilon > 0$ there exists a $n \in \mathbb{N}$ with $q^n < \epsilon$

Proof: Insert in conclusion (2) $\;b^n := \dfrac{1}{\epsilon} \;;\; M := \dfrac{1}{\epsilon}$

After conclusion (2) there is a $n \in \mathbb{N}$ with $b^n > M$, hence $q^n < \epsilon\,.$

□

(4) If $a < b$ are real numbers, then a rational number exists

$$c \;=\; \frac{m}{n} \qquad (m, n \text{ integer}, \; n > 0)$$

for which the following applies: $\quad a \;<\; c \;<\; b$

Proof: $\quad \epsilon := \dfrac{b-a}{2}\,;\quad$ Then a $n \in \mathbb{N}$ exists with $\dfrac{1}{n} \;<\; \epsilon$

this means $\qquad 2 \;<\; n\,b \,-\, n\,a$

$$m \;:=\; [n\,a] \,+\, 1$$

$$n\,a \;<\; m \;<\; n\,b \quad / \quad \frac{1}{n}$$

$$a \;<\; \frac{m}{n} \;<\; b$$

□

With the arrangement of $\mathbb{R}$ by "$<$" the subsets called intervals can be described. Let $a < b$ be real:

$$[a, b] \;:=\; \{x \in \mathbb{R} / a \leq x \leq b\}$$

is called the closed interval between a and b.

$$]a, b[\;:=\; \{x \in \mathbb{R} / a < x < b\}$$

is called the open interval between a and b.

The two half-open intervals $I =]a, b]$ and $I = [a, b[$ both satisfy the inclusions $]a, b[\subset I \subset [a, b]$.

To the intervals one also counts the sets $]-\infty, a]$, $]-\infty, a[$, $]b, \infty[$, $[b, \infty[$ and $]-\infty, \infty[$. The second, third and fifth of these sets are open!

2.2 The Concept of Monotony

is used for real sequences and functions f whose domain D is a subset of $\mathbb{R}$. In reality, such sequences can also be understood as functions with the domain being $D = \mathbb{N}$ or $D = \mathbb{N}_0$.

A function $f : D \to \mathbb{R}$ is called monotonically increasing (rising) if

$$x_1, x_2 \in D \quad x_1 \leq x_2 \;\Rightarrow\; f(x_1) \leq f(x_2)$$

if more than that, in addition

$$x_1, x_2 \in D \quad x_1 < x_2 \;\Rightarrow\; f(x_1) < f(x_2)$$

then f is called strictly monotonically increasing!

Examples:

(1) $f(x) = [x]$, the Entier function is monotonically increasing, but not strictly monotonically increasing

(2) A sequence is considered which includes a parameter x.

$$e_n(x) = \left(1 + \frac{x}{n}\right)^n, \qquad x \in \mathbb{R}$$

defines a monotonically increasing sequence for $n > -x$

Proof with Bernoulli's inequality:

If $n > -x$ the bracket is > 0, hence also $e_n(x) > 0$. Consider $\dfrac{e_{n+1}}{e_n}$

$$\frac{e_{n+1}}{e_n} = \frac{n^n}{(n+1)^{n+1}} \frac{(n+1+x)^{n+1}}{(n+x)^n} \qquad / \quad \times \; \frac{n+x}{n+x} \frac{n^{n+1}}{n^{n+1}}$$

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

$$= \frac{n+x}{n}\left(1-\underbrace{\frac{x}{(n+1)\,(n+x)}}_{-k \text{ in Bernoulli inequality}}\right)^{n+1}$$

$$\geq \frac{n+x}{n}\left(1-\frac{x}{n+x}\right) \geq 1$$

Result: $\big(e_n(x)\big)_{n>-x}$ increases monotonically!

□

A function $f : D \to \mathbb{R}$ is called (strictly) monotonically decreasing if for all $x_1, x_2 \in D$ the following holds:

$$x_1 \leq x_2 \Rightarrow f(x_1) \geq f(x_2)$$

$$x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$$

Example:

(3) Example (3) from chapter 1

$D = \mathbb{N}$, a parameter t appears, $t \geq 0$

$$w_0 = 1 + t \qquad ; \qquad w_{n+1} = \frac{1}{2}\left(w_n + \frac{t}{w_n}\right)$$

defines a strictly monotonically decreasing sequence.

Proof: The terms of the sequence are positive

Consider: $\displaystyle \frac{w_{n+1}}{w_n} = \frac{1}{2}\left(1+\frac{t}{w_n^2}\right)$

It remains to show: $w_n^2 > t$

Induction start: $n = 0$:

$$w_0^2 - t = t^2 + t + 1 > 0$$

$n \to n+1$: We have

$$w_{n+1}^2 - t = \frac{1}{4}\left(w_n^2 + 2\,t + \frac{t^2}{w_n^2}\right) - t$$

$$= \frac{1}{4}\left(w_n^2 - 2\,t + \frac{t^2}{w_n^2}\right) = \frac{1}{4}\left(w_n - \frac{t}{w_n}\right) > 0$$

because according to the induction hypothesis the numerator $w_n^2 - t$ is not 0.

□

2.3 The Completeness of $\mathbb{R}$

is based on the fact that every upper-bound and non-empty subset A of $\mathbb{R}$ has a smallest upper bound, called supremum of A, abbreviated $\sup(A)$. Here upper bound of A refers to every real number s with $a \in A \Rightarrow a \leq s$. With s, every number greater than s is also an upper bound of A.

The gist of it: If A is not empty and has upper bounds in the first place, then the set of all upper bounds of A has a smallest element.

Three simple examples:

$A := \mathbb{N}$ has no upper bound (according to Archimedes).

$A =]-\infty, 1]$ has the supremum 1.

$A := \{x \in \mathbb{Q} \,|\, x^2 \leq 2\}$ is bounded above by $\frac{3}{2}$ and below by $-\frac{3}{2}$. The $\sup(A)$, namely $\sqrt{2}$ does not belong to A.

Remarks:

(1) A set $A \subset \mathbb{R}$ is labeled as being bounded below if it has a lower bound s, i.e. $a \in A \Rightarrow s \leq a$.

(2) From the basic rule U-3 follows: Every set $A \subset \mathbb{R}$ that is both non-empty and bounded below has a greatest lower bound, the infimum of A, abbreviated $\inf(A)$.

2.4 The Monotone Convergence Criterion for Sequences

Let $(a_n)_{n \geq 1}$ be a real sequence which increases monotonically and has an upper bound. Then this sequence converges to the supremum of its value set $\{a_n \,|\, n \in \mathbb{N}\}$.

Remark:

(1) From U-3 follows: Every monotonically decreasing sequence that is bounded below $(b_n)_{n \in \mathbb{N}}$ converges to the infimum of its value set.

Proof of the monotone convergence criterion

let $\alpha = \sup \{a_n \,|\, n \in \mathbb{N}\}$

α is the upper bound, hence $a_n \leq \alpha$ for each n

α is the smallest upper bound, so if $\epsilon > 0$ (but arbitrarily small) then $\alpha - \epsilon$ itself is not an upper bound. This implies the existence of an index n_ϵ such that:

$$\alpha - \epsilon \;<\; a_{n_\epsilon}$$

Now, since the sequence increases monotonically, $a_{n_\epsilon} \leq a_n$ for all $n \geq n_\epsilon$. In particular, it was shown that for every $\epsilon > 0$ there exists an index n_ϵ with

$$\boxed{n \;\geq\; n_\epsilon \;\Rightarrow\; |a_n - \alpha| \;<\; \epsilon}$$

The inequality after the "$\Rightarrow$" is expressed in such a form that the sequence members a_n fall within the ϵ neighborhood of α:

ϵ-neighborhood: $\;]\alpha - \epsilon, \alpha + \epsilon[$

This shows:

$$\lim_{n \to \infty} a_n \;=\; \alpha$$

Returning to examples (2) and (3):

$$e_n(x) \;=\; \left(1 + \frac{x}{n}\right)^n , \qquad n \;>\; -x$$

The existence of an upper bound for this sequence is clear in the case $x \leq 0$ because then $e_n(x) \leq 1$.

On the other hand, if $x > 0$, then the sequence

$$\frac{1}{\left(1 - \frac{x}{n}\right)^n} , \qquad n \;>\; x$$

decreases monotonically as the reciprocal of a monotonically increasing sequence of positive numbers.

$$0 \;<\; \left(1 - \frac{x^2}{n^2}\right)^n \;\leq\; 1 \quad \text{applies analogously to the above}$$

Multiplication by the positive number $\dfrac{1}{\left(1 - \frac{x}{n}\right)^n}$ yields

$$0 \quad < \quad \left(1+\frac{x}{n}\right)^n \quad \leq \quad \frac{1}{\left(1-\frac{x}{n}\right)^n}$$

Because on the right side stands a monotonically decreasing sequence, it is bounded (upwards) by the term with the smallest index, which is $n = [x] + 1$.

The sequence $w_n(x)$ declared for $x \in [0, \infty[$

$$\begin{aligned} w_0(x) \quad &= \quad 1 + x \\ w_{n+1}(x) \quad &= \quad \frac{1}{2}\left(w_n + \frac{x}{w_n}\right) \end{aligned}$$

has only terms ≥ 0, hence it is bounded below by 0.

Remark:

(2) The term upper or lower bound is not only defined for the $A \subseteq \mathbb{R}$ but also for functions $f : D \to \mathbb{R}$. Thus f is called bounded above if for all $x \in D$ $f(x) \leq t$. In this case $D \subset \mathbb{R}$ is not even necessary, e.g. D could also be a circle. For $A \subset \mathbb{R}$ with no upper bound one uses $sup(A) = +\infty$, correspondingly $\inf(A) = -\infty$ is used for $A \subset \mathbb{R}$ with no lower bound.

3. Limit Values

3.1 Definition of a Limit

Let $D \subset \mathbb{R}$ with $\sup D = \infty$, also $f : D \to \mathbb{R}$. One says f converges to $\alpha \in \mathbb{R}$, if x goes to ∞, if for every $\epsilon \in \mathbb{R}$, $\epsilon > 0$ a bound t_ϵ exists with $x \in D$ and $x > t_\epsilon \Rightarrow |f(x) - \alpha| < \epsilon$.

This state of affairs is being expressed by the “equation”:

$$\boxed{\lim_{\substack{x \to \infty \\ x \in D}} f(x) = \alpha}$$

Remark:

(1) If f is a sequence $(a_n)_{n \in \mathbb{N}}$ then $D = \mathbb{N}$, one writes $\lim\limits_{n \to \infty} a_n = \alpha$

$f : D \to \mathbb{R}$; D with no upper bound; $\sup D = +\infty$

$$\lim_{\substack{x \to \infty \\ x \in D}} f(x) = \alpha \quad \text{means that:}$$

For each tolerance limit $\epsilon > 0$ there is a bound t_ϵ for the argument x

$$x \geq t_\epsilon \ \wedge \ x \in D \quad \Rightarrow \quad f(x) \in \,]\alpha - \epsilon, \alpha + \epsilon[$$

The domain of the considered function is also frequently omitted in a number of other limit terms.

Examples:

(1) Let $a_n = {}^1\!/_n$. The a_n form a monotonically decreasing sequence of positive numbers. Conclusion (1) from the Archimedean property states

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

(2) $\lim\limits_{n\to\infty} x^n \;=\; 0 \;, \quad \text{if} \quad |x| < 1$

Proof:

Let $q = |x|$. Then (in the case $q > 0$) it follows from conclusion (3) drawn from the Archimedean property of $\mathbb{R}$ that:

$$\lim_{n\to\infty} q^n \;=\; 0 \qquad \text{therefore} \qquad \lim_{x\to\infty} x^n \;=\; 0$$

$\square$

Expressed through the geometric series:

Its terms form a null sequence. Implication for the sequence

$$\boxed{s_n \;:=\; \sum_{k=0}^{n} x^k}$$

From the sum formula for s_n

$$s_n \;=\; \frac{1}{1-x} - \frac{x^{n+1}}{1-x}$$

Let x be given with $|x| < 1$, furthermore an accuracy bound

$$\epsilon' > 0 \quad \text{and} \quad \frac{1}{1-x} \;>\; 0\,.$$

The convergence statement about the sequence $\left(x^k\right)_{k\geq 0}$ results in an index n_ϵ for $\epsilon := \epsilon'(1-x) > 0$ with $|x^n| < \epsilon$ for all $n > n_\epsilon$. Therefore

$$\frac{x^n}{1-x} \;<\; \frac{\epsilon}{1-x} \qquad \text{if only} \quad n \geq n_\epsilon \;, \text{hence}$$

$$\left| s_n - \frac{1}{1-x} \right| \;=\; \left| \frac{x^{n+1}}{1-x} \right| \;<\; \epsilon' \;, \quad \text{if} \quad n \geq N_{\epsilon'}\,!$$

Observation: The convergence goodness depends on the position of the point x:

$$\boxed{\sum_{k=0}^{n} x^k \;=\; \frac{1}{1-x} \;, \quad \text{if} \quad |x| < 1}$$

This is the sum formula of the geometric series.

Example:

(3) In every infinite decimal fraction there is (if the sign is ignored) a monotonically increasing sequence!

$$a_n = \sum_{k=0}^{n} \frac{z_k}{10^k}$$

where $z_1, z_2, \ldots$ is the sequence of digits of the decimal fraction after the comma. The digits z_k with $k > 0$ satisfy $z_k \in \mathbb{N}_0 \cap [0, 10[$.

The sequence $(a_n)_{n \geq 0}$ is bounded from above

$$0 \leq z_k \leq 9 \quad \Rightarrow$$

$$a_n \leq z_0 + \sum_{k=1}^{n} \frac{9}{10^k} = z_0 + 9 \cdot \frac{1}{1 - \frac{1}{10}}$$

As a monotonically increasing sequence with an upper bound, $(a_n)_{n \geq 0}$ converges.

Conversely: Every real number $x \geq 0$ can also be written as an infinite decimal fraction.

$$x_0 := x \qquad z_0 := [x]$$

$$x_{n+1} = 10\,(x_n - z_n) \qquad z_{n+1} := [x_{n+1}]$$

Remark:

(2) Concrete calculations almost always use finite decimal numbers. The occurring inaccuracies are to be checked. To be determined: Has the fraction used been obtained by truncating or by rounding?

$$\pi = 3.1415 \text{ (truncated)} \quad ; \quad \pi = 3.1416 \text{ (rounded)}$$

Example:

(4) $x \;\to\; \dfrac{x}{x-1}$

does not return a real value for all real x. The denominator becomes 0 for $x = -1$. Largest possible definition range:

$$D = \left]-\infty, +\infty\right[\;\setminus\; \{-1\}$$

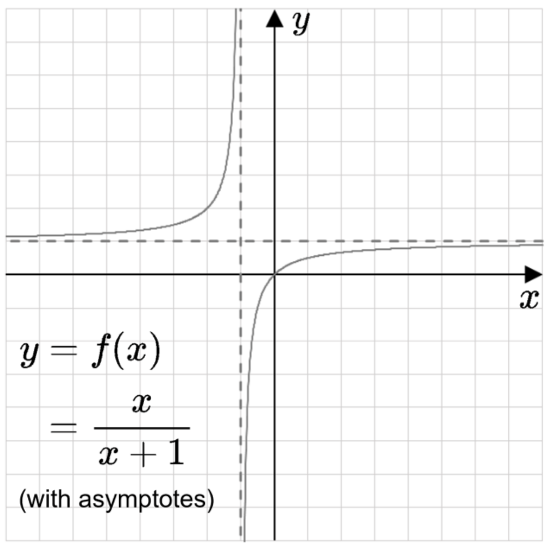
$$f(x) = \frac{x}{x+1} = \sum_{k=0}^{\infty} (-1)^k \cdot x^{k+1}$$

returns a so-called power series representation from the geometric series, only useful for $|x| < 1$.

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

Shows: For $\epsilon > 0$ there is a t_ϵ with $|f(x) - 1| < \epsilon$ for all $x \geq t_\epsilon$; put differently $\lim\limits_{x \to \infty} f(x) = 1$.

The graph of $f(x)$ reveals $\lim\limits_{x \to -\infty} f(x) = 1$. Furthermore, it also indicates the necessity to consider other limit scenarios, such as $x \searrow -1$ or respectively $x \nearrow -1$ or $x \to 0$.



Limits are compatible with ordinary inequalities. Let $D \subset \mathbb{R}$ be a set with no upper bound on which f_1 and f_2 are both defined. Further, $\lim\limits_{\substack{x \to \infty \\ x \in D}} f_k(x) = \alpha_k$ $(k = 1, 2)$ shall hold. Then:

a) $$\boxed{\lim_{\substack{x \to \infty \\ x \in D}} (f_1 + f_2)(x) = \alpha_1 + \alpha_2}$$

b) $$\boxed{\lim_{\substack{x \to \infty \\ x \in D}} (f_1\, f_2)(x) = \alpha_1\, \alpha_2}$$

With $\alpha_2 \neq 0$, $\dfrac{f_1}{f_2}(x)$ is declared for sufficiently large $x \in D$ and it applies that:

c) $$\boxed{\lim_{\substack{x \to \infty \\ x \in D}} \frac{f_1}{f_2}(x) = \frac{\alpha_1}{\alpha_2}}$$

d) If $f_1(x) \leq f_2(x)$ applies to all $x \in D$, then $\alpha_1 \leq \alpha_2$ also applies.

Proof: let $\epsilon > 0$ be given

1) According to the assumption, there exists a bound t_1 for $\dfrac{\epsilon}{2}$ with

$$x \in D \;\wedge\; x \geq t_1 \quad \Rightarrow \quad |f_1(x) - \alpha_1| \;<\; \frac{\epsilon}{2}$$

Similarly, $t_2 \in \mathbb{R}$ exists with

$$x \in D \;\wedge\; x \geq t_2 \quad \Rightarrow \quad |f_2(x) - \alpha_2| \;<\; \frac{\epsilon}{2}$$

To be estimated $\quad |f_1(x) + f_2(x) - \alpha_1 - \alpha_2|\,.$

According to the triangle inequality:

$$|f_1(x) - \alpha_1 + f_2(x) - \alpha_2| \;\leq\; |f_1(x) - \alpha_1| \,+\, |f_2(x) - \alpha_2|$$

with $\;t_\epsilon := \max(t_1, t_2)\;$ follows

$$|f_1(x) + f_2(x) - \alpha_1 - \alpha_2| \;\leq\; |f_1(x) - \alpha_1| \,+\, |f_2(x) - \alpha_2| \;<\; \frac{\epsilon}{2} + \frac{\epsilon}{2} \;=\; \epsilon$$

if $\;x \geq t_\epsilon\,,\; x \in D$

2) Sei $M \;:=\; 1 + |\alpha_1| + |\alpha_2| \;;\;$ to be estimated:

$$\begin{aligned} &|f_1(x)\, f_2(x) - \alpha_1\, \alpha_2| \\ &\quad = \; |f_1(x)\, f_2(x) - \alpha_1\, f_2(x) + \alpha_1\, f_2(x) - \alpha_1\, \alpha_2| \\ &\qquad\qquad \text{(according to the triangle inequality)} \\ &\quad \leq \; |f_1(x) - \alpha_1|\, |f_2(x)| \,+\, |\alpha_1|\, |f_2(x) - \alpha_2| \end{aligned}$$

According to the precondition, t_ϵ exists with

$$x \in D \;\wedge\; x \geq t_\epsilon \quad \Rightarrow \quad |f_2(x) - \alpha_2| \;<\; 1$$

$$|f_2(x) - \alpha_2| \;<\; \frac{\epsilon}{M} \quad , \quad |f_1(x) - \alpha_1| \;<\; \frac{\epsilon}{M}$$

Estimate first

$$|f_2(x)| \;=\; |f_2(x) - \alpha_2 + \alpha_2| \;\leq\; |f_2(x) - \alpha_2| \,+\, |\alpha_2| \;=\; 1 \,+\, |\alpha_2|$$

Therefore together

$$\begin{aligned} |f_1(x)\, f_2(x) - \alpha_1\, \alpha_2| \;&\leq\; |f_1(x) - \alpha_1|\, |f_2(x)| \,+\, |\alpha_1|\, |f_2(x) - \alpha_2| \\ &<\; \frac{\epsilon}{M}\,(1 + |\alpha_2|) \,+\, |\alpha_1|\, \frac{\epsilon}{M} \;=\; \epsilon \end{aligned}$$

for all $\;x \;\geq\; t_\epsilon$

3) Because of **2)** it is sufficient to discuss the special case of $f_1 = 1$ (constant). f_2 can have roots, but because of $\alpha_2 \neq 0$, $\epsilon_1 = |\alpha_2|/2$ can be used to define the limit conditions.

If $t_0 \in \mathbb{R}$ exists with $x \in D,\ x > t_0 \quad \Rightarrow \quad |f_2(x) - \alpha_2| \ < \ \frac{1}{2}|\alpha_2|$

With the so-called inverse triangle inequality

$$\begin{aligned} |f_2(x)| &= |\alpha_2 + f_2(x) - \alpha_2| \\ &\geq |\alpha_2| - |f_2(x) - \alpha_2| \ > \ |\alpha_2| - \left|\frac{\alpha_2}{2}\right| \end{aligned}$$

for $x \in D \ \wedge \ [t_0, \infty[\ , \ \frac{1}{f_2(x)}$ is declared

$$\left|\frac{1}{f_2(x)} - \frac{1}{\alpha_2}\right| = \left|\frac{\alpha_2 - f_2(x)}{\alpha_2\, f_2(x)}\right| \leq \frac{|f_2(x) - \alpha_2|}{|\alpha_2|^2/2} \ < \ \epsilon$$

$$\text{if} \quad |f_2(x) - \alpha_2| \ < \ \frac{\epsilon\,|\alpha_2|^2}{2} \ = \ \epsilon\,'$$

According to the presupposition about $f_2(x)$ there is a bound T for $\epsilon\,'$ with

$$x \in D, \quad x \geq T \quad \Rightarrow \quad |f_2(x) - \alpha_2| \ < \ \epsilon\,'$$

This inserted $\quad\Rightarrow\quad$ the assertion.

4) Indirect: assumption $f_1(x) \leq f_2(x) \quad \forall\ x \in D$ and $\alpha_1 > \alpha_2$

$$\text{set} \quad \epsilon \ := \ \frac{\alpha_1 - \alpha_2}{2}$$

A bound t exists with $x \in D, \quad x \geq t \quad \Rightarrow$

$|f_1(x) - \alpha_1| \ < \ \epsilon \quad ; \quad |f_2(x) - \alpha_2| \ < \ \epsilon$

$$\begin{aligned} \alpha_1 - \alpha_2 &\leq \alpha_1 - \alpha_2 + f_2(x) - f_1(x) \\ &\leq \alpha_1 - f_1(x) + f_2(x) - \alpha_2 \\ &< 2\,\epsilon \ = \ \alpha_1 - \alpha_2 \qquad \text{contradiction} \quad \Rightarrow \quad \text{wrong!} \end{aligned}$$

therefore $\alpha_1 > \alpha_2$ wrong, hence $\alpha_1 \leq \alpha_2$

□

Application to the sequence

$$w_0(t) \ = \ 1 + t \quad ; \qquad w_{n+1}(t) \ = \ \frac{1}{2}\left(w_n(t) + \frac{t}{w_n(t)}\right)$$

According to the monotonicity criterion

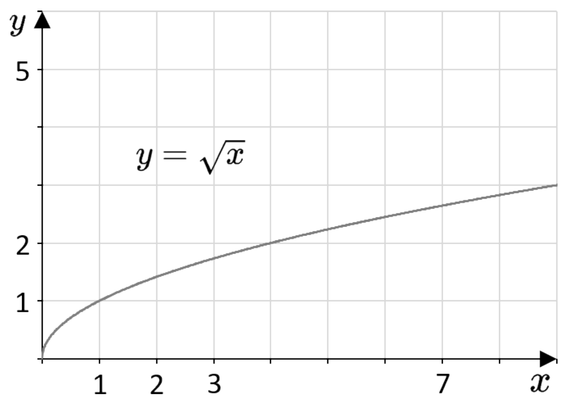
$(w_n)_{n\geq 0}$ converges, like $w(t) := \lim\limits_{n\to\infty} w_n(t)$

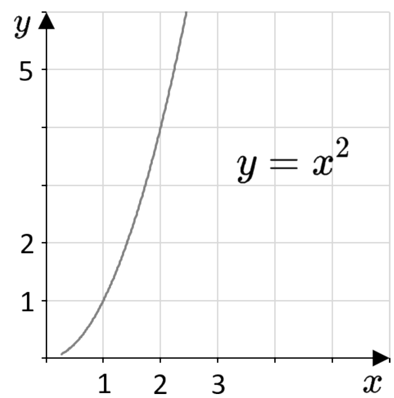
In the case $t = 0$ we have $\left(\frac{1}{2}\right)^n_{n\geq 0}$, therefore $w_0 = 0$

If $t > 0$, then because of $w_n^2(t) > t \quad w_n^2(t) \geq t > 0$

Transition to the limit value in the recursion formula: $w(t) = \frac{1}{2}\left(w(t) + \frac{t}{w(t)}\right)$

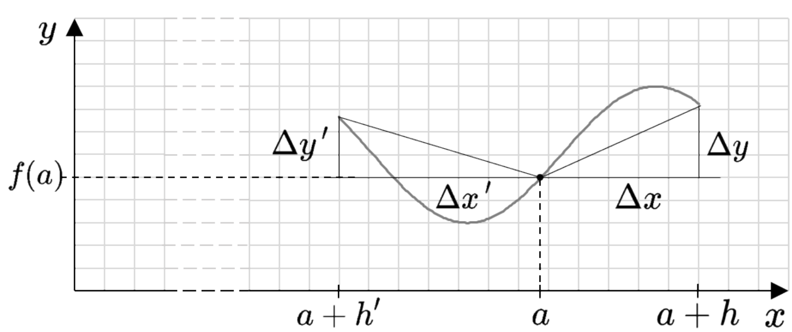
Hence $w^2(t) = t$; in this way the function $t \to \sqrt{t}$ was constructed.





The most graphic and (for now) most important limit concept in physics and technology is differentiation.

3.2 Differentiation



Considered will be a real function f, which is declared in a neighborhood of the point a (perhaps also somewhere else). The fixed point $\begin{pmatrix} a \\ f(a) \end{pmatrix}$ lies on the graph of f, as does $\begin{pmatrix} a+h \\ f(a+h) \end{pmatrix}$, another variable point.

Examined will be the change in the straight line connecting these points when the second point is varied.

The slope of the line is

$$\frac{\Delta y}{\Delta x} = \frac{f(a+h) - f(a)}{h}$$

The question is whether the corresponding limit value exists

$$\boxed{\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h}}$$

This is the slope of the tangent to the graph of f at the point $\begin{pmatrix} a \\ f(a) \end{pmatrix}$

Usual nomenclature: $\left.\frac{df}{dx}\right|_{x=a}$; $f'(a)$

Often in applications: $x = t$, the time and $f'(a)$ is written as $\dot{f}(a)$. Interpretation: the change in current speed! The quest to find the derivative can also be seen as an approximation problem. The given function f shall be approximated in the best possible way by using a linear function (= straight line) in the vicinity of a. Differentiation = linear approximation.

Examples:

(1) Any constant function

$$f(x) = c \quad (x \in \mathbb{R})$$

is differentiable everywhere with the derivative 0: $f'(x) = 0 \quad \forall\ x \in \mathbb{R}$

(2) The function $f(x) = x \ \forall\ x \in \mathbb{R}$, called identity function, is also differentiable everywhere with the derivative $f'(x) = 1 \quad forall \ x \in \mathbb{R}$

(3) The absolute value function $x \to |x|$ is differentiable everywhere outside the zero point with the derivative $+1$ for $a > 0$ and derivative -1 for $a < 0$. The derivative does not exist at the zero point.

$$\lim_{h \searrow 0} \frac{f(h) - f(0)}{h} \quad , \quad \lim_{h \nearrow 0} \frac{f(h) - f(0)}{h}$$

(4) The root function $x \to \sqrt{x} = w(x)$ is differentiable at all points a with the derivative $w'(a) = \dfrac{1}{2\sqrt{a}}$

Proof: The difference quotient

$$\frac{\sqrt{a+h}-\sqrt{a}}{h} = \frac{1}{\sqrt{a+h}+\sqrt{a}} \qquad \text{goes against} \quad \frac{1}{2\sqrt{a}} \quad \text{if} \quad h \to 0$$

$$\lim_{h\to 0} \sqrt{a+h} = a \qquad \text{(continuity of the root function in } a\text{)}$$

For example: If $\epsilon > 0$ then $d_\epsilon > 0$ exists with $|h| < d_\epsilon$

$$\Rightarrow \quad \left|\sqrt{a+h}-\sqrt{a}\right| < \epsilon$$

So that the root function $\sqrt{a+h}$ is declared at all: first restriction $d_\epsilon \leq a$. If $|h| < a \quad \Rightarrow$

$$\left|\sqrt{a+h}-\sqrt{a}\right| = \frac{|h|}{\sqrt{a+h}+\sqrt{a}} < \frac{|h|}{\sqrt{a}}$$

The additional requirement $d_\epsilon \leq \epsilon\sqrt{a}$ is therefore sufficient.

□

3.3 Basic Rules of Differentiation

Let D be a set of real numbers which contains a neighborhood $]a-\delta, a+\delta[$ $(\delta > 0)$ of the point a, further let g, f be real functions defined on D, which are also differentiable in a. Then the following applies:

a) $f+g$ is differentiable in a with the derivative

$$\boxed{(f+g)'(a) = f'(a) + g'(a)} \qquad \text{Sum rule}$$

b) The product function $f\,g$ is differentiable in a with the derivative

$$\boxed{(f\,g)'(a) = f'(a)\,g(a) + f(a)\,g'(a)} \qquad \text{Product rule}$$

c) If g has no roots, then f/g is defined on D and differentiable at point a with the derivative

$$\boxed{\left(\frac{f}{g}\right)'(a) = \frac{f'(a)\,g(a) - f(a)\,g'(a)}{g^2(a)}} \qquad \text{Quotient rule}$$

Proof:

The first statement **a)** is a special case of the general limit rule for sums. (In the cases **b)** and **c)** the corresponding is not correct)

For **b)**: Difference quotient

$$\frac{f(a+h)\,g(a+h) - f(a)\,g(a)}{h}$$

$$= \frac{f(a+h) - f(a)}{h}\,g(a+h) \;+\; f(a)\,\frac{g(a+h) - g(a)}{h}$$

reveals the correctness of the product rule, because:

$$\lim_{h \to 0} g(a+h) \;=\; g(a).$$

Because of **b)**, the special case where $f = 1$ is the constant 1 suffices to prove the quotient rule **c)**. Then the difference quotient for $1/g$

$$\frac{1}{h}\left(\frac{1}{g(a+h)} - \frac{1}{a}\right) \;=\; -\frac{g(a+h) - g(a)}{g(a+h)\cdot g(a)\,h}$$

converges to $\dfrac{g\,'(a)}{g^2(a)}$ as h approaches 0.

□

> Every differentiable function is continuous.

Let the real function f be differentiable at point a. Then it holds true that

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

Proof: According to the premise

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \;=\; f\,'(0)$$

Besides, it is obvious that $\lim\limits_{h \to 0} h \;=\; 0$

From the compatibility of limits with the product it follows

$$\lim_{h \to 0}(f(a+h) \;-\; f(a)) \;=\; 0$$

which is another version of the claim.

□

The chain rule deals with the behavior of complicated functions (chain application)

$$f(g(x)) \;=\; f \circ g(x) \quad \text{say } f \text{ after } g$$

The chain rule

Let g be a real function differentiable in a. Let f be a real function differentiable at the point $g(a)$. Then $f \circ g$ (the compound) is differentiable in a with the derivative

$$(f \circ g)'(a) \;=\; f'(g(a)) \cdot g'(a)$$

Proof sketch: The difference quotient

$$\frac{f(g(a+h)) - f(g(a))}{h} \;=\; \frac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)} \, \frac{g(a+h) - g(a)}{h}$$

lets one suspect what the formula looks like! A measure is needed to supplement this if g is not strictly monotonic.

4. Important Functions

4.1 The Exponential Function in the Real

There is exactly one function with value 1 at the origin that is differentiable on all $\mathbb{R}$ and which is the solution of the differential equation $y' = y$, the exponential function

$$x \;\to\; \exp(x)$$

It is positive everywhere and has the functional equation

$$\exp(x_1 + x_2) \;=\; \exp(x_1) \cdot \exp(x_2) \qquad x_1, x_2 \in \mathbb{R}$$

Finally, it is strictly monotonically increasing.

Proof:

1) The functional equation

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

is a positive function everywhere according to the examples in the last paragraphs. Given x_1, x_2: for sufficiently large n

$$\frac{\left(1 + \frac{x_1}{n}\right)^n \left(1 + \frac{x_2}{n}\right)^n}{\left(1 + \frac{x_1 + x_2}{n}\right)^n} \;=\; \underbrace{\left[1 + \frac{x_1 x_2 / n^2}{1 + \frac{x_1 + x_2}{n}}\right]^n}_{\text{(Bernoulli)}}$$

$$\geq\; 1 + \frac{x_1 + x_2}{n + x_1 + x_2}$$

$$\frac{\left(1+\frac{x_1+x_2}{n}\right)^n}{\left(1+\frac{x_1}{n}\right)^n\left(1+\frac{x_2}{n}\right)^n} \;=\; \underbrace{\left[1-\frac{x_1\,x_2/n^2}{\left(1+\frac{x_1}{n}\right)\left(1+\frac{x_2}{n}\right)}\right]^n}_{\text{(Bernoulli)}}$$

$$\geq\; 1-\frac{n\,x_1x_2}{(n+x_1)(n+x_2)}$$

The two inequalities in the limit case $n\to\infty$:

$$\frac{f(x_1)\,f(x_2)}{f(x_1+x_2)} \;\geq\; 1 \quad ; \quad \frac{f(x_1+x_2)}{f(x_1)\,f(x_2)} \;\geq\; 1$$

Therefore:

$$f(x_1)\,f(x_2) \;=\; f(x_1+x_2)$$

2) If $|h|<1$ then

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

On the left is Bernoulli's inequality. Using the binomial theorem:

$$\left(1+\frac{h}{n}\right)^n \;=\; \sum_{j=0}^{n}\binom{n}{j}\left(\frac{h}{n}\right)^j$$

$$=\; 1+h+\sum_{j=2}^{n}\binom{n}{j}\frac{1}{n^j}\,h^j$$

Therefore

$$\left(1+\frac{h}{n}\right)^n \;\leq\; 1+h+\sum_{j=2}^{n}|h|^j \;\leq\; 1+h+\frac{h^2}{1-|h|}$$

Transition to the limit case$n\to\infty$

$$1+h \;\leq\; f(h) \;\leq\; 1+h+\frac{h^2}{1-|h|}$$

3) Differentiability:

Subtraction of $1+h$ on both sides and division by h

$$0 \;\leq\; \left| \frac{f(h)-1}{h} - 1 \right| \;\leq\; \frac{|h|}{1-|h|}$$

$f(0) = 1$; consequence for $h \to 0$

$$f'(0) \;=\; \lim_{h\to 0} \frac{f(h)-f(0)}{h} \;=\; 1$$

Given $x \in \mathbb{R}$. Difference quotient:

$$\begin{aligned} \frac{f(x+h)-f(x)}{h} &= \frac{f(x)\,f(h)-f(x)\,f(0)}{h} \\ &= f(x)\,\frac{f(h)-f(0)}{h} \end{aligned}$$

Consequence:

$$f'(x) \;=\; f(x)\cdot 1 \;=\; f(x)$$

4) Monotony:

Inequality at the end of **2)** yielded $f(h)-1 \geq h$. Now let h be positive:

$$\begin{aligned} f(x+h) \;-\; f(x) &= f(x)\,(f(h)-1) \\ &\geq f(x)\,h \;>\; 0 \end{aligned}$$

5) Uniqueness:

Suppose also g solves the differential equation: $y' = y$

Consider the quotient $\quad q(x) \;=\; \dfrac{g(x)}{f(x)}$

$q(x)$ is differentiable with the derivative:

$$\begin{aligned} q'(x) &= \frac{g'(x)\,f(x)-g(x)\,f'(x)}{f^2(x)} \qquad \text{(differential equation!)} \\ &= \frac{g(x)\,f(x)-g(x)\,f(x)}{f^2(x)} \;=\; 0 \end{aligned}$$

According to the mean value theorem (proven in the next chapter), this is only possible if q on $\mathbb{R}$ is a constant function.

□

4.2 The Natural Logarithm

There is exactly one function l declared on $]0,\infty[$ with the properties

a) $\boxed{l(x) \;\leq\; x - 1}$

b) $\boxed{l(x_1\,x_2) \;=\; l(x_1) + l(x_2)}$

the natural logarithm ln. ln is a strictly monotonically increasing function and in its domain $]0,\infty[$ differentiable with derivative: $\ln'(x) = \dfrac{1}{x}$

Proof: for $x > 0$ declare 3 sequences

$$x_0 := x \quad ; \quad x_{n+1} = \sqrt{x_n} \quad ; \quad g_n := 2^n\left(1 - \frac{1}{x_n}\right) \quad ;$$

and let $\quad h_n := 2^n(x_n - 1) = x_n\,g_n$

1) $\lim\limits_{n\to\infty} x_n = 1$: let $x \geq 1 \;\Rightarrow\; 1 \leq \underset{\underset{x_{n+1}}{\|}}{\sqrt{x}} \leq x_n$

$(x_n)_n$ is as a monotonically decreasing sequence bounded below and convergent.

$\alpha := \lim\limits_{n\to\infty} x_n$, Recursion formula and $x \geq 1$ result in:

$\alpha^2 = \alpha \geq 1$, i.e.: $\alpha = 1$; analogously for $0 < x < 1$

2) $$g_{n+1} - g_n = 2^n\left(2 - \frac{2}{\sqrt{x_n}} - 1 + \frac{1}{x_n}\right) = 2^n\left(1 - \frac{1}{\sqrt{x_n}}\right)^2 \geq 0$$

$$h_n - g_n = 2^n\left(x_n - 1 - 1 + \frac{1}{x_n}\right) = 2^n\left(\sqrt{x_n} - \frac{1}{\sqrt{x_n}}\right)^2 \geq 0$$

$$h_n - h_{n+1} = 2^n\left(x_n - 1 - 2\sqrt{x_n} + 2\right) = 2^n\left(\sqrt{x_n} - 1\right)^2 \geq 0$$

$$g_n \leq g_{n+1} \leq h_{n+1} \leq h_n$$

Hence both limits exist and because of $h_n = x_n\,g_n$ and according to **1)**

$$\lim_{n\to\infty} g_n = \lim_{n\to\infty} h_n =: l(x)$$

Because h_n falls monotonically, the following applies in particular:

$$l(x) \;\leq\; h_0(x) \;=\; x - 1$$

3) Functional equation:

Apparently $\quad \sqrt{x\,y} \;=\; \sqrt{x}\,\sqrt{y} \qquad x, y \geq 0$

Therefore $\quad (x\,y)_n \;=\; x_n\,y_n \quad$ and thus:

$$h_n(x\,y) \;=\; 2^n(x_n\,y_n - 1)$$

$$h_n(x\,y) \;-\; h_n(x)\,h_n(y) \;=\; 2^n(x_n\,y_n - 1 - x_n + 1 - y_n + 1)$$

$$=\; 2^n(x_n - 1)\,(y_n - 1) \;=\; (x_n - 1)\,h_n(y)$$

Transition to the limit case:

$$l(x\,y) \;-\; l(x) \;-\; l(y) \;=\; \lim_{n\to\infty}(x_n - 1)\,h_n(y) \;=\; 0$$

because according to **1)** $\quad \lim_{n\to\infty} x_n \;=\; 1$

4) Uniqueness:

Assume $\mathscr{L}\colon\,]0,\infty[\; \to \mathbb{R}$ satisfies the properties **a)** and **b)**. Functional equation for $x = y = 1$

$$\mathscr{L}(1) \;=\; \mathscr{L}(1\cdot 1) \;=\; \mathscr{L}(1) + \mathscr{L}(1) \;\Rightarrow\; \mathscr{L}(1) \;=\; 0$$

Functional equation with $y = \dfrac{1}{x}$

$$0 \;=\; \mathscr{L}\left(x\cdot\frac{1}{x}\right) \;=\; \mathscr{L}(x) + \mathscr{L}\left(\frac{1}{x}\right)$$

that means:

$$\mathscr{L}\left(\frac{1}{x}\right) \;=\; -\mathscr{L}(x)$$

use property **a)** of function 1 for $\dfrac{1}{x}$ instead of x

$$-\mathscr{L}(x) \;=\; \mathscr{L}\left(\frac{1}{x}\right) \;\leq\; \frac{1}{x} - 1 \quad\Rightarrow$$

$$1 - \frac{1}{x} \;\leq\; \mathscr{L}(x) \;\leq\; x - 1 \qquad (*) \quad \text{more than statement } \mathbf{a)}$$

Because of $x_n^{2^n} = x$, **b)** results in

$$\mathscr{L}(x) \;=\; \mathscr{L}\left(x_n^{2^n}\right) \;=\; 2^n \mathscr{L}(x_n)$$

$$g_n \;=\; 2^n\left(1 - \frac{1}{x_n}\right) \underset{(*)}{\leq} 2^n \mathscr{L}(x_n) \;=\; \mathscr{L}(x) \underset{(*)}{\leq} 2^n(x_n - 1)$$

$$g_n \;\leq\; \mathscr{L}(x) \;\leq\; h_n(x) \qquad \forall\, n$$

Transition to the limit case $n \to \infty$

$$l(x) \;\leq\; \mathscr{L}(x) \;\leq\; l(x)$$

therefore

$$l(x) \;=\; -\mathscr{L}(x)$$

5) Monotony:

Let $0 < q < 1$. In this case, according to **a)**

$$l(q) \;\leq\; q - 1 \;<\; 0$$

If $0 < x_1 < x_2$ then set $q = \dfrac{x_1}{x_2}$.

With that

$$l(x_1) \;=\; l(x_2\, q) \;=\; l(x_2) + l(q)$$

Because $l(q) < 0$ it follows that $l(x_1) < l(x_2)$

6) The derivative:

The estimate $(*)$ also applies to $l(x)$. Moreover $l(1) = 0$. For $|h| < 1$ insert $x = 1 + h$ into $(*)$.

$$1 - \frac{1}{1+h} \;\leq\; l(1+h) - l(1) \;\leq\; h$$

$$\frac{h}{1+h} \;\leq\; l(1+h) - l(1) \;\leq\; h$$

if $h > 0$ then

$$\frac{1}{1+h} \;\leq\; \frac{l(1+h)-l(1)}{h} - 1 \;\leq\; 0$$

that means:

$$\lim_{h \searrow 0} \frac{l(1+h)-l(1)}{h} \;=\; 1 \qquad , \qquad \lim_{h \nearrow 0} \frac{l(1+h)-l(1)}{h} \;=\; 1$$

With that: $\quad l\,'(1) \;=\; 1$

Let a be arbitrary and greater than 0. Difference quotient:

$$\frac{l(a+h)-l(a)}{h} \;=\; \frac{l\left(1+\frac{h}{a}\right)-l(1)}{h/a}\,\frac{1}{a}$$

with the proven: $\quad l\,'(a) \;=\; \dfrac{1}{a}$

□

The exponential function and the natural logarithm form a pair of inverse functions, that means:

$$\boxed{\begin{array}{lcll} \ln(\exp(x)) & = & x & \forall\ x \in \mathbb{R} \\ \exp(\ln(y)) & = & y & \forall\ y \in \mathbb{R} \end{array}}$$

Consider:

$$f(x) \;:=\; x \;-\; \ln(\exp(x))$$

has the derivative

$$f\,'(x) \;=\; 1 - \frac{1}{\exp(x)} \cdot \exp(x) \;=\; 0$$

according to the mean value theorem it follows: $f \;=$ constant. Since $f(0) \;=\; 0 - \ln(\exp(0)) \;=\; 0$, the constant is 0. Now put $\ln(y) \;=\; x$ in the first equation

$$\ln(\exp(\ln(y))) \;=\; \ln(y)$$

We have the same value for two arguments under the strictly monotonic function ln, the arguments must be the same

$$\exp(\ln(y)) \;=\; y$$

4.3 The General Power Function in the Real

The general power function is declared for base $a > 0$ by

$$\boxed{a^x \;=\; \exp(x\ln(a)) \quad ; \quad x \in \mathbb{R}}$$

This generalizes the ordinary power function a^n, because $a^n = \exp(n\ln(a))$.

This is easy to show by complete induction:

$$n = 1; \quad a^0 \;=\; 1$$

$n \to n+1$ just proven: $\quad a \;=\; \exp(\ln(a))$

$$\begin{aligned} a^{n+1} \;=\; a^n\, a \;&=\; \exp(n\;\ln(a))\,\exp(\ln(a)) \\ &=\; \exp((n+1)\;\ln(a)) \end{aligned}$$

The ordinary power rules remain valid:

$$a^{x_1}\,a^{x_1} \;=\; a^{x_1+x_2} \quad ; \quad (a^{x_1})^{x_2} \;=\; a^{x_1\,x_2} \qquad \forall\; x_1, x_2 \in \mathbb{R}$$

$$a^x\,b^x \;=\; (a\,b)^x \qquad \forall\; x \in \mathbb{R}$$

According to the chain rule:

$$\boxed{a^x \;=\; \frac{da^x}{dx} \;=\; a^x\ln(a)}$$

There is a positive number e (Euler) with $\ln(\mathrm{e}) = 1$, namely

$$\mathrm{e} \;=\; \exp(1) \;=\; \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

The usual notation $\exp(and) = \mathrm{e}^x$ is consistent with the definition of the general power function. The general power function $x \to a^x$ has an inverse function, namely

$${}^a\log(x) \;=\; \frac{1}{\ln(a)}\cdot\ln(x) \qquad \text{the logarithm to the base } a$$

The general power function $x \to x^v$ (for $x \geq 0$) is differentiable with derivative

$$\frac{dx^v}{dx} \;=\; v\,x^{v-1}$$

According to the definition $x^v = \exp(v\ \ln(x))$. Chain rule:

$$\begin{aligned}\frac{dx^v}{dx} &= \exp(v\ \ln(x))\,\frac{v}{x}\\ &= \exp(v\ \ln(x))\,v\ \exp(-\ln(x))\\ &= v\ \exp((v-1)\ln(x)) \;=\; v\,x^{v-1}\end{aligned}$$

Among the functions x^v we find the usual power functions but also various square root functions $\sqrt[k]{x} = x^{1/k}$, however, only for $x > 0$.

Growth comparisons:

$$\begin{array}{llll} \lim\limits_{x\to\infty} \dfrac{e^x}{x^k} = \infty & \lim\limits_{x\to\infty} \dfrac{x^k}{e^x} = 0 & \forall\ k \in \mathbb{N}_0 \\ \lim\limits_{x\to\infty} \ln(x) = \infty & \lim\limits_{x\searrow 0} \ln(x) = -\infty & \\ \lim\limits_{x\to\infty} \dfrac{\ln(x)}{x^\alpha} = \infty & \forall\ \alpha > 0 & \\ \lim\limits_{x\searrow 0} x^\alpha = 0 & \lim\limits_{x\searrow 0} x^{-\alpha} = \infty & \end{array}$$

Proof sketch:

By definition

$$\left(1 + \frac{x}{x+1}\right)^{k+1} \le e^x \quad \text{therefore}$$

$$\frac{x^{k+1}}{(k+1)^{k+1}} \le e^x \qquad \text{hence} \qquad \frac{x}{(k+1)^{k+1}} \le \frac{e^x}{x^k}$$

$$\ln(e) = 1 > 0 \qquad ; \qquad \ln(e^n) = n\ \ln(e)$$

The first limit statement follows from the monotony of ln. Second statement of the first line: Replace x by $\alpha\ln(x)$, consider $k = 1$. This results in:

$$\lim_{x\to\infty} \frac{\alpha\ \ln(x)}{x^\alpha} = 0 \quad ; \quad x^\alpha \text{ is by definition } = \exp(\alpha\ln(x))$$

4.4 Polynomials

The polynomials arise from the constant function and from the identity function ($x \to x$) through possibly repeated multiplications or additions.

Typical example:

$$P(x) \;=\; a_0 x^n \,+\, a_1 x^{n-1} \,+\, \ldots \,+\, a_{n-1} x \,+\, a_n$$

The a_k, real at the moment, later also complex, are called "coefficients" of P, in particular a_0 is the "leading coefficient". If $a_0 \neq 0$, then n denotes the "degree" of P ($\deg P$). If $a_0 = 1$ then P is called "normalized".

With regard to differentiability, the polynomials are manageable. The totality of polynomials is abbreviated as $\mathbb{R}[x]$ or $\mathbb{C}[x]$, respectively, depending on the restriction for the coefficients.

One can calculate with polynomials just like with whole numbers. Concurrent indexing is preferable for describing addition and multiplication of polynomials!

$$P \;=\; \sum_{k=0}^{n} a_k x^k \quad ; \quad Q \;=\; \sum_{j=0}^{r} b_j x^r$$

Formation of $P \circ Q$ coefficient-wise. The product is generated through ordering by fixed x-power

$$P \circ Q \;=\; \sum_{l=0}^{n+r} \Big(\sum_{k=0}^{l} a_k b_{l-k} \Big) x^l$$

where the above introduced coefficients $a_k = 0$, $b_j = 0$ if $k > n$ or if respectively $j > r$. For $l = n+r$ the sum reduces to $a_n b_r$. From this follows $\deg P \circ Q = \deg P + \deg Q$.

Division with remainder is also possible:

Let f be a polynomial of degree n with leading coefficient a, g be a polynomial with leading coefficient b of degree m. If the degree n of f is less than $m = \deg g$, f itself becomes the remainder of the division. Otherwise one forms

$$f_1(x) \;=\; f(x) \,-\, \frac{a}{b} x^{n-m} g(x)$$

f_1 has a smaller degree than f. In any case, possible repetition leads to a remainder with degree $< m = \deg g$

$$f(x) \;=\; q(x)\,g(x) \,+\, h$$

where g, h are polynomials with $\deg h < \deg g$. Here $h = 0$ could also be the zero polynomial, to which the degree $-\infty$ is assigned.

A case of particular importance is

$$g(x) \;=\; x - \alpha$$

It leads to the so-called Horner method and an algorithm for evaluating polynomials.

$$f(x) \;=\; a_0 x^n \,+\, a_1 x^{n-1} \,+\, \ldots \,+\, a_n$$

Polynomial evaluation of $f(x)$ with Horner's method at the point $x = \alpha$ using

$$(\ldots \quad (((a_0\alpha + a_1)\alpha + a_2)\alpha + a_3)\ldots \quad + a_{n-1})\,\alpha$$

Multiply from the inside out. Number of multiplications $\leq n$

$$\begin{array}{c|cccccc}
 & a_0 & a_1 & a_2 & a_3 & \ldots & a_n \\
+ & & (a_0^{(1)}\cdot\alpha) & (a_1^{(1)}\cdot\alpha) & (a_2^{(1)}\cdot\alpha) & \ldots & (a_{n-1}^{(1)}\cdot\alpha) \\
\hline
 & a_0^{(1)} & a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & \ldots & a_n^{(1)} \;=\; f(\alpha) \;=\; h
\end{array}$$

$a_n^1 \;=\; f(\alpha)$ is also the remainder.

Example:

$$f(x) \;=\; x^3 + x^2 - 2x - 1$$

Evaluate at positions $\alpha = 2, -2, -3$

$\alpha = 2:$

$$\begin{array}{c|cccc|}
 & 1 & 1 & -2 & -1 \\
+ & & 1\cdot 2 & 3\cdot 2 & 4\cdot 2 \\
\hline
 & 1 & 3 & 4 & 7
\end{array}$$

hence: $a_0 = 1$, $a_1 = 3$, $a_2 = 4$ and remainder $= 7$

$$\Rightarrow \quad f(x) = (x-2)\,(x^2+3x+4) + 7$$

$\underline{\alpha = -2:}$

$$\begin{array}{rcccc|} & 1 & 1 & -2 & -1 \\ + & & 1\cdot(-2) & (-1)\cdot(-2) & 0\cdot(-2) \\ \hline & 1 & -1 & 0 & -1 \end{array}$$

hence: $a_0 = 1$, $a_1 = -1$, $a_2 = 0$ and remainder $= -1$

$$\Rightarrow \quad f(x) = (x+2)\,(x^2-x) - 1$$

$\underline{\alpha = -3:}$

$$\begin{array}{rcccc|} & 1 & 1 & -2 & -1 \\ + & & 1\cdot(-3) & (-2)\cdot(-3) & 4\cdot(-3) \\ \hline & 1 & -2 & 4 & -13 \end{array}$$

hence: $a_0 = 1$, $a_1 = -2$, $a_2 = 4$ and remainder $= -13$

$$\Rightarrow \quad f(x) = (x+3)\,(x^2-2x+4) - 13$$

$$a_n^{(1)} = f(\alpha) = h \quad ; \quad g = \alpha - x$$

$$q = \sum_{k=0}^{n-1} a_k x^{n-1-k}$$

It could be that α is a root of f; then $h = a_n^{(1)} = 0$, hence

$$f(x) = q(x)(x-\alpha)$$

The root can be split off, the degree of q is:

$$\deg q = \deg f - 1$$

In particular, every non-constant polynomial f has at most $n = \deg f$ roots.

Another formulation of this observation is the identity theorem for polynomials.

4.4.1 Identity Theorem for Polynomials

Two polynomials f_1, f_2 of degree $\leq n$, which agree at $n+1$ places, are identical! Because $f = f_1 - f_2$ has (more) $n+1$ zeros and $n \geq \deg f$ is the zero polynomial.

Example:

$$
\begin{array}{lrcl}
(1) & (x+1)^{2n} & = & \displaystyle\sum_{k=0}^{2n} \binom{2n}{k} x^k \\
& \| & & \\
& (x+1)^n (x+1)^n & = & \displaystyle\left(\sum_{k=0}^{n} \binom{n}{k} x^k\right)\left(\sum_{k=0}^{n} \binom{n}{k} x^k\right) \\
& & = & \displaystyle\sum_{r=0}^{2n} \left(\sum_{k=0}^{r} \binom{n}{k}\binom{n}{r-k}\right) x^r
\end{array}
$$

The two outer polynomials have the same value at more than $2n$ points, i.e. they are identical. In particular, the coefficients at x^n are equal:

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2$$

Let there be $n+1$ "node positions":

$$x_0,\, x_1,\, x_2,\, \ldots\ldots\, x_n$$

and $n+1$ "node values":

$$y_0,\, y_1,\, y_2,\, \ldots\ldots\, y_n$$

To this exists an "interpolation polynomial" P with $\deg P \leq n$

$$P(x_i) = y_i \quad ; \quad 0 \leq i \leq n$$

Newton's recursive construction!

$$P_0 = y_0 \quad ; \quad P_{k+1} = P_k + c_{k+1}(x - x_0)(x - x_1)\ldots(x - x_k)$$

In that c_{k+1} is to be chosen as the solution to the equation

$$y_{k+1} = P_k(x_k) + c_{k+1}(x_{k+1} - x_0)\ldots(x_{k+1} - x_k)$$

Example:

(2) Nodes positions: $-2, -1, \quad 0, \quad 1, \quad 2$

Node values: $-1, \quad 1, -1, -1, \quad 7$

$$P_0 \;=\; -1; \quad P_1 \;=\; P_0 + c_1(x - x_0)$$

$$P_1 \;=\; -1 + c_1(x+2)$$

$$P_1(-1) \;=\; 1 \;=\; -1 + c_1 \cdot 1 \;\Rightarrow\; c_1 \;=\; 2$$

$$P_1 \;=\; -1 + 2(x+2) \;=\; 2x + 3$$

$$P_2 \;=\; P_1 + c_2(x+2)(x+1)$$

$$P_2(0) \;=\; -1 \;=\; P_1(0) + c_2 \cdot 2 \;=\; 3 + c_2 \cdot 2 \;\Rightarrow\; c_2 \;=\; -2$$

$$P_2 \;=\; P_1 - 2(x+2)(x+1) \;=\; -2x^2 - 4x - 1$$

$$P_3 \;=\; P_2 + c_3(x+2)(x+1)(x-0)$$

$$P_3(1) \;=\; -1 \;=\; P_2(1) + c_3 \cdot 6 \;=\; -7 + c_3 \cdot 6 \;\Rightarrow\; c_3 \;=\; 1$$

$$\begin{aligned} P_3 &= -1 + 2(x+2) - 2((x+2)(x+1) + (x+2)(x+1) \cdot x \\ &= \; x^3 + x^2 - 2x - 1 \end{aligned}$$

$$P_4 \;=\; P_3 + c_4(x+2)(x+1)(x-0)(x-1)$$

$$P_4(2) \;= 7 \;=\; P_3(2) + c_4(*) \;=\; 7 + c_4(*) \;\Rightarrow\; c_4 \;=\; 0$$

Combined:

$$P_0 \;=\; -1$$

$$P_1 \;=\; 2x + 3$$

$$P_2 \;=\; -2x^2 - 4x - 1$$

$$P_3 \;=\; x^3 + x^2 - 2x - 1$$

$$P_4 \;=\; P_3$$

5. Existence Theorems

5.1 The Nested Interval Theorem

Let $I_n = [a_n, b_n]$, $n \in \mathbb{N}$

be a monotonically decreasing sequence of closed intervals with respect to inclusion $\subset$. Then there exists a point α that lies in all intervals.

Proof:

$$I_{n+1} \subset I_n \ ; \quad [a_{n+1}, b_{n+1}] \subset [a_n, b_n]$$

means $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$. In particular, $(a_n)_{n \geq 1}$ is monotonically increasing and bounded. Therefore:

$$\lim_{n \to \infty} a_n \;=\; \alpha \;=\; \sup_{n \in \mathbb{N}} a_n$$

Assertion: $a_k \leq \alpha \leq b_k \qquad \forall\ k$

If $k \leq n$ then $a_k \leq a_n \leq b_k$; Transition to the limit case $n \to \infty$: $a_k \leq \alpha \leq b_k$

$\square$

First application: The Bolzano-Weierstraß theorem.

5.2 The Bolzano-Weierstraß Theorem

Every bounded sequence of real numbers c_n , $n \geq 1$ has a convergent subsequence $(c_{n_k})_{k \ \ geq\,1}$.

Proof:

Let $a_1 < b_1$ be the lower and upper bounds of the sequence

$$I_1 \;:=\; [a_1, b_1]$$

Halve I_1 into I_k, $k = 1, 2$. If $I_k = [a_k, b_k]$ is such that for infinitely many $n \in \mathbb{N}$ $c_n\ inI_k$, then bisect and choose as I_{k+1} the first partial interval I on the left for which $c_n \in I$ for infinitely many n. The sequence of intervals I_k, $k \geq 1$ is monotonically decreasing with respect to inclusion. Therefore an $\alpha \in I_k\ \forall\ k$ exists.

Recursive definition of $n_k:$ $n_1 = 1$

If n_k is chosen such that $c_{nk} \in I_k$, then we choose $n_{k+1} > n_k$ such and minimal with the property

$$c_{n_{k+1}} \in I_{k+1}$$

Because the sequence of interval lengths converges to 0 it follows that α is the only point in all I_k therefore

$$\lim_{k \to \infty} c_{n_k} = \alpha$$

$\square$

Definition of Cauchy sequences (in the real).
A real sequence $(a_n)_{n \geq 0}$ is called a "Cauchy sequence" if for each $\epsilon > 0$ there exists an index n_ϵ with

$$m, n \geq n_\epsilon \Rightarrow |a_m - a_n| \leq \epsilon$$

An application of the nested interval theorem is the <u>Cauchy criterion for $\mathbb{R}$</u>: Every real Cauchy sequence is convergent.

Proof idea:

Let $(a_n)_{n \geq 1}$ be the given Cauchy sequence. For $\epsilon = \dfrac{1}{2 \cdot 2^k}$ one finds a n_k with

$$m, n > n_k \Rightarrow |a_m - a_n| < \frac{1}{2} \cdot 2^{-k}$$

Moreover, $n_1 < n_2 < n_3 \ldots$ can be achieved. Interval sequence:

$$I_k := [a_{n_k} - 2^{-k}, a_{n_k} + 2^{-k}]$$

I_k is monotonically decreasing with respect to inclusion. Hence there is an $\alpha \in I_k\ \forall\ k$. Because the interval lengths converge to 0 it follows

$$\lim_{n \to \infty} a_n = \alpha$$

Definition of continuity)
Let a be a point in the domain D of the real function f. f is called "continuous" in a if

$$\lim_{\substack{x \to a \\ x \in D}} f(x) = f(a)$$

If this is the case for all $a \in D$, then f is called continuous!

Remarks:

(1) Expressed by the concept of neighborhood, continuity of f in a means: For each neighborhood $U =]f(a) - \epsilon, f(a) + \epsilon[$ of the pixel $f(a)$ there is a neighborhood $V =]a - \delta, a + \delta[$ of the point a with $f(V \cap D) \subset U$.

(2) Use continuity of f at the point a. For every sequence of points $a_n \in D$

$$\text{with} \quad \lim_{n \to \infty} a_n = a \quad \text{also} \quad \lim_{n \to \infty} f(a_n) = f(a)$$

(see also in chapter 3 before the chain rule.). Every differentiable function is continuous.

Example:

The root function $w(x) = \sqrt{x}$ is also continuous in $x = 0$. $D = [0, \infty[\,,\ w(0) = 0$

$$\lim_{x \searrow 0} w(x) = \lim_{x \searrow 0} x^{1/2} = 0 = w(0)$$

Remark:

(3) Analogously, the general power function $x \to x^{\alpha}$ $(\alpha > 0)$ can be continuously extended to the zero point by $0^{\alpha} := 0$.

5.3 The Extreme Value Theorem

Let f be a continuous real function on $[a, b]$. Then the value set $f([a, b]) := \{f(x),\, x \in [a, b]\}$ is limited and therefore has a sup M and a inf m. Further, numbers c, $d \in [a, b]$ exist and $f(c) = m$, $f(d) = M$.

Proof sketch:

1) Boundedness: (indirect)

Assumption: An $a_n \in [a, b]$ exists with $\lim_{n \to \infty} |f(a_n)| = \infty$

According to Bolzano-Weierstraß a convergent subsequence exists. It is denoted by $(a_n)_{n\geq 1}$. For this subsequence $\lim\limits_{n\to\infty} |f(a_n)| = \infty$

$$\alpha \;:=\; \lim_{n\to\infty} a_n \in [a,b] \qquad a \leq a_n \leq b$$

However, since f is continuous at α, $\lim\limits_{n\to\infty} f(a_n) = \alpha$ should hold.

Contradiction: Hence $f([a,b])$ was bounded after all.

2) Maximum and minimum:

There are $c_n \in [a,b]$ with $\lim\limits_{n\to\infty} f(c_n) = M$. According to Bolzano-Weierstraß, $(c_n)_{n\geq 1}$ has a convergent subsequence. Again, it is denoted c_n.

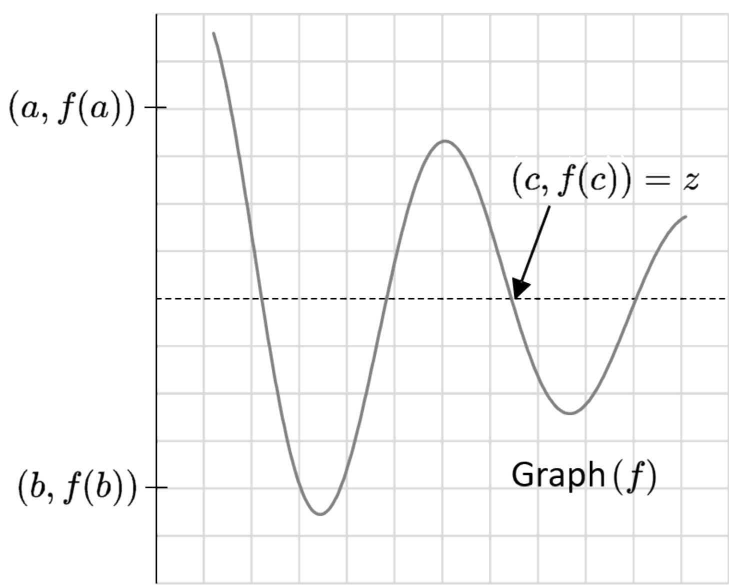
$$\lim_{n\to\infty} c_n \;=:\; d$$

Therefore, because of the continuity of f at d

$$\lim_{n\to\infty} f(c_n) \;=\; M \;=\; f(d)$$

Similarly, one finds $f(c) = m$ for a matching $c \in [a,b]$.

5.4 The Intermediate Value Theorem



Let f be continuous on the interval $[a,b]$, z be a number between the values $f(a)$, $f(b)$ of f in the end points of the interval. Then there is at least one $c \in [a,b]$ with $f(c) = z$.

Proof:

Let (without restriction)

$$f(b) \; < \; z \; < \; f(a) \qquad A : \{x \in [a,b], \;\; f(x) > z\}$$

According to the assumption $A \neq \emptyset$ is bounded. Therefore, A has a supremum $(\sup(A))$.

Assertion: $f(c) = z$

1) If $x_1 \in [a,b]$ with $f(x_1) > z$ then $\epsilon_1 \; = \; f(x_1) - z \; > \; 0$.

Because f is continuous in x_1 a $\delta_1 > 0$ exists and

$$|f(x) - f(x_1)| \;\; < \;\; \epsilon_1 \qquad \forall \;\, x \in [x_1, x_1 + \delta_1[$$

Then it follows $f(x) > z$ here, hence $x_1 \neq c$.

2) Let $x_2 \in [a,b]$ with $f(x_2) < z$, choose $\epsilon_2 \; = \; z - f(x_2) \; > \; 0$.

Because f is continuous in x_2 a $\delta_2 > 0$ exists with

$$|f(x) - f(x_2)| \;\; < \;\; \epsilon_2 \qquad \forall \;\, x \in \,]x_2 - \delta_2, x_2]$$

In particular $f(x) < z$, hence $x_2 \neq c$.

3) $f(x) > z$ is wrong, $f(x) < z$ is wrong, hence $f(x) = z$,.

$\square$

Example:

$$\left.\begin{array}{l} f(x) \;\; = \;\; x^3 \; - \; x \; - \; 1 \\[2ex] a \; = \; 1, \; b \; = \; 2, \; z \; = \; 0 \end{array}\right\} \quad f(c) = 0 \quad \text{for a} \quad c \in [1,2]$$

5.5 The Inverse Function

Let $f \; : \; [a,b] \; \to \; \mathbb{R}$ be continuous and strictly monotonically increasing. Let $A :=$ $f(a)$, $B := f(b)$ be the values in the end points of the interval. For each $y \in [A,B]$ there is exactly one point $x = g(y)$, $g(y) \in [a,b]$, with the property $f(x) = y$. The thus from f declared function, the "inverse function of f", is strictly monotonically increasing and continuous on $[A,B]$. Furthermore:

$$\boxed{\begin{array}{ccc} f(g(y)) & = & y \\ \forall\ y \in [A,B] & & \end{array} \qquad \begin{array}{ccc} g(f(x)) & = & x \\ \forall\ x \in [a,b] & & \end{array}}$$

Proof:

1) The existence of $x \in [a,b]$ with $f(x) = y$ follows from the intermediate value theorem. The uniqueness follows from the strict monotony of f. If $y_1 < y_2$ from $[A,B]$, then $x_1 = g(y_1) \geq x_2 = g(y_2)$ is not possible because otherwise f would not be strictly monotonically increasing.

2) The continuity of g:
First let $y \in\ ]A,B[$; then $g(y) \in\ ]a,b[\ =: c$. Consider $\epsilon > 0$ and so small that $]c-\epsilon, c+\epsilon[\ \subset [a,b]$. The existence of a $\delta > 0$ with $g(\,]y-\delta, y+\delta[\,)\ subset]c-\epsilon, c+\epsilon[$ then follows from the monotony of f and g. For continuity at the points A and B the proof can easily be modified.

3) The equation $f(g(y)) = y$ holds per definition of g. In particular, the first equation applies to $y = f(x)$:

$$f(\underbrace{g(f(x))}_{x_1}) \ = \ f(x) \qquad \forall\ x \in [a,b]$$

That $f(x_1) \geq f(x)$ for the strictly monotone function f is only possible if $x_1 = x$, i.e.

$$g(f(x)) \ = \ x \qquad \forall\ x \in [a,b]$$

$\square$

Remarks:

(1) A corresponding theorem also applies to strictly monotonically decreasing functions. The inverse functions then also become strictly monotonically decreasing!

(2) A similar theorem also holds if one has strictly monotonic, continuous functions on open or half-open intervals. (compare exp and ln).

(3) Even if an explicit formula (e.g. $f(x) = x^5 - x - 1$) is given for f, no such formula needs to exist for the inverse function. Occasionally, the derivative of the inverse function provides a formula for the inverse function.

Derivative of the inverse function:

Let $f:\ [a,b] \to \mathbb{R}$ be continuous and strictly monotonically increasing and $g:\ [A,B] \to \mathbb{R}$ be its inverse function. If f is differentiable in the inner points $x \in\ ]a,b[$ and $f'(x) \neq 0$ then g in $y = f(x)$ is differentiable with derivative

$$g'(y) \quad = \quad \frac{1}{f'(g(y))}$$

Proof:

Consider the quotient of g

$$\frac{g(y+h)-g(h)}{y+h-y} \quad = \quad \frac{1}{\dfrac{f(g(y+h))-f(g(y))}{g(y+y)-g(y)}}$$

Because g is continuous especially at the point y, the limit value exists for $h \to 0$, i.e

$$g'(y) \quad = \quad \frac{1}{f'(g(y))}$$

$\square$

Example:

$$f(x) \quad = \quad \sinh(x) \quad = \quad \frac{1}{2}\left(\mathrm{e}^{x}-\mathrm{e}^{-x}\right)$$

Monotony: $x_1 < x_2\,,$ with that

$$(f(x_2)-f(x_1))\cdot 2 \quad = \quad \underbrace{(\mathrm{e}^{x_2}-\mathrm{e}^{x_1})}_{>0} \quad + \quad \underbrace{(\mathrm{e}^{-x_1}-\mathrm{e}^{-x_2})}_{>0} \quad > \quad 0$$

f is continuous because it is differentiable everywhere. Let g be the inverse function

$$f'(x) \quad = \quad \sinh'(x) \quad = \quad \cosh(x)$$

$\operatorname{arsinh}(y)$ is the abbreviation of *area sinus hyperbolicus* of (y) and hence the inverse function of $\sinh(x) = \frac{1}{2}\left(\mathrm{e}^{x}-\mathrm{e}^{-x}\right)$ follows with

$$\sinh'(x) \quad = \quad \cosh(x) = \sqrt{\sinh^2(x)+1}$$

and the derivative of the inverse function $g = \operatorname{arsinh}$

$$g'(y) \quad = \quad \frac{1}{f'(g(y))} \quad = \quad \frac{1}{\sqrt{(\sinh(g(y)))^2+1}} \quad = \quad \frac{1}{\sqrt{y^2+1}}$$

Also consider: $\quad h(y) \quad = \quad \ln\left(y+\sqrt{y^2+1}\right)$

Everywhere on $\mathbb{R}$ differentiable with derivative

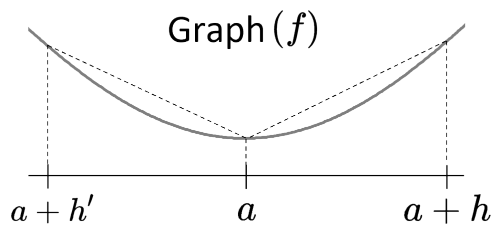
$$h'(y) \;=\; \frac{1}{y+\sqrt{y^2+1}}\left(1+\frac{2y}{2\sqrt{y^2+1}}\right) \;=\; \frac{1}{\sqrt{y^2+1}} \;=\; g'(y)$$

The difference $g-h$ is a function defined on $\mathbb{R}$ with derivative 0, therefore constant. Calculation of the constants by evaluation at any point:

$$y = 0 \quad ; \quad g(0) = 0 \quad ; \quad h(0) = 0 \quad \text{therefore} \quad g = h: \;\Rightarrow$$

$$\operatorname{arsinh}(y) \;=\; \ln\left(y+\sqrt{y^2+1}\right)$$

5.6 Local Extrema



Local extrema consitute a generic term for local maxima and local minima. If f is a real function declared in a neighborhood U of the point a, then a is called the local maximum of f if a possibly smaller neighborhood $]a-\delta, a+\delta[$ of a exists with $f(x) \le f(a)$, if $|x-a| \le \delta$; Correspondingly, a is called a local minimum of f if $f(x) \ge f(a)$ for a neighborhood of a. Every differentiable real function f has a root of the derivative $f'(0) = 0$ in each local extremum a.

Proof:

A proof suffices in the case of a local minimum a of f. For positive h with $a+h$ in the relevant neighborhood the difference quotient is

$$\frac{f(a+h)-f(a)}{h} \;\ge\; 0$$

correspondingly for $h' < 0$ with $a+h'$ in the relevant neighborhood the difference quotient is

$$\frac{f(a+h')-f(a)}{h'} \;\le\; 0$$

Then, because f is differentiable in a by assumption

$$\left.\begin{array}{lcl} \lim\limits_{h \searrow 0} \dfrac{f(a+h)-f(a)}{h} & = & f'(a) \\ \lim\limits_{h' \nearrow 0} \dfrac{f(a+h')-f(a)}{h'} & = & f'(a) \end{array}\right\} \Rightarrow f'(a) = 0$$

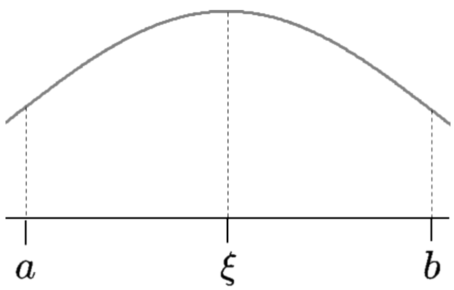
□

Remark:

$f'(a) = 0$ is possible without a being a local extremum. For example consider

$$f(x) = x^{2n+1} \quad n \in \mathbb{N} \,, \quad \text{at} \quad a = 0$$

5.7 Rolle's Theorem



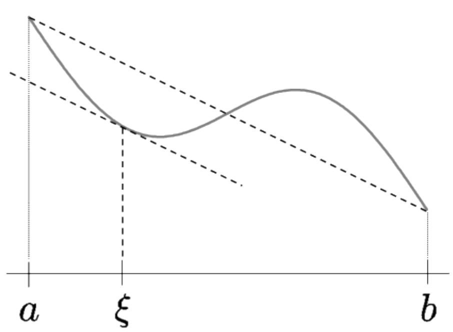
Let f be a real function differentiable on the open interval I. Then for points $a < b$ in I with $f(a) = f(b)$ the following holds:

A ξ exists with $a < \xi < b$ and $f'(\xi) = 0$.

Proof:

If $f(x) = f(a)$ for all $x \in \,]a, b[$ then f is constant there and $f'(\xi) = 0 \;\forall\; \xi \in \,]a, b[$. Otherwise there is a $c \in \,]a, b[$ with $f(c) \neq f(a)$ and without restriction $f(c) > f(a)$. Then the maximum of the function on $[a, b]$ is not in the end points. Let ξ be a point where f has a maximimum. Then $a < \xi < b$ and ξ is also a local maximum of f, hence the theorem about local extrema yields the assertion $f'(\xi) = 0$.

5.8 The Mean Value Theorem



Let f be a real function differentiable on the open interval I, for points $a < b$ on I there is a $\xi \in \,]a, b[$ with

$$\boxed{\frac{f(b) - f(a)}{b - a} = f'(\xi)}$$

Proof:

Based on Rolle's theorem with auxiliary function

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}\,(x - a)$$

F is differentiable with f:

$$F(a) \;=\; f(a) \quad ; \quad F(b) \;=\; f(b) - \frac{f(b)-f(a)}{b-a}\,(b-a) \;=\; f(a)$$

Rolle $\Rightarrow$ a $\xi \in]a,b[$ exists with

$$F(\xi) \;=\; 0 \;=\; f'(\xi) \;=\; \frac{f(b)-f(a)}{b-a}$$

$\square$

First application:

If f satisfies the precondition of the theorem and also $f'(c) = 0 \;\forall\; x \in I$ then f is constant on I. Because according to the mean value theorem, for points $a < b$ in I

$$\frac{f(b)-f(a)}{b-a} \;=\; 0$$

Therefore:

$$f(b) \;=\; f(a) \qquad \forall\; a \in I \quad \forall\; b \in I$$

6. Applications of the Mean Value Theorem

6.1 Monotony, Local Extrema, Convexity

6.1.1 The Monotony Criterion of Differential Calculus

Let f be a real function differentiable on an open interval I. f is then and only then monotonically increasing on I if therein everywhere $f'(x) \geq 0$.

Supplement: Under the condition of monotonicity f is not already strictly increasing if $f'(x) = 0$ on a whole subinterval.

Proof:

Assumption:

$$f'(x) \geq 0 \qquad \forall\ x \in I$$

Let $a < b$ both be in I, then by the mean value theorem

$$\frac{f(b) - f(a)}{b - a} = f'(\xi) \geq 0 \quad (a < \xi < b)$$

Therefore $\quad f(b) \geq f(a)$

$\square$

<u>Remark:</u>

Of course, the theorem also provides a criterion for monotone decay! In each local extremum a of the differentiable function f we necessarily have $f'(a) = 0$. Whether it is then a question of a local maximum or minimum or neither of the two can often be decided directly!

6.1.2 Criterion for Strict Local Extrema

Let x_0 be a zero of the derivative $f\,'$ of a twice differentiable function f. Then x_0 is a local minimum of f in the case $f\,''(x_0) > 0$ and a local minimum in the case $f\,''(x_0) < 0$ x_0.

Proof: In the event of $f\,''(x) < 0$

Let $f\,''(x) = \lim_{x \to x_0} \dfrac{f\,'(x) - f\,'(x_0)}{x - x_0} < 0$

By assumption $f\,'(x_0) = 0$. Hence a neighborhood of x_0 exists in which $f\,'(x)$ has opposite sign as the difference $x - x_0$. Therefore, if (in the neighborhood $\coloneqq |x - x_0| < \delta$) $x_0 - \delta < a < x_0 < b < x_0 + \delta$, then according to the mean value theorem:

$$\frac{f(a) - f(x_0)}{a - x_0} = f\,'(\xi) \qquad ; \qquad \frac{f(b) - f(x_0)}{b - x_0} = f\,'(\eta)$$

with matching ξ and η

$$\xi \in \,]a, x_0[\; > 0 \qquad \text{i.e.} \qquad f(a) < f(x_0)$$

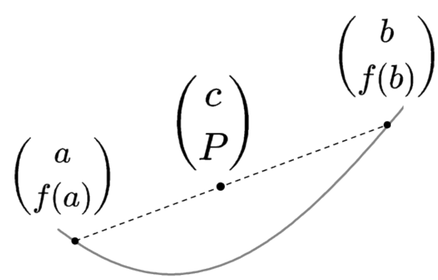
$$\eta \in \,]x_0, b[\; > 0 \qquad \text{i.e.} \qquad f(b) < f(x_0)$$

$\square$

A real function f defined on an interval is called convex (convex from below) if for every three points $a < c < b$ of the interval, the function value for $f(c)$ does not produce a point above the connecting line from $\begin{pmatrix} a \\ f(a) \end{pmatrix}$ to $\begin{pmatrix} b \\ f(b) \end{pmatrix}$. This means:

$$f(c) \leq \frac{b-c}{b-a} f(a) + \frac{c-a}{b-a} f(b)$$

$$c = \frac{b-c}{b-a} a + \frac{c-a}{b-a} b$$



Expressed differently: for all $\lambda \in \,]0, 1[$

$$f\left(\lambda\, a + (1 - \lambda)\, b\right) \leq \lambda\, f(a) + (1 - \lambda)\, f(b)$$

on the connecting line the following applies:

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

that means:

$$
\begin{aligned}
(b-a)\,P &= (b-c)\,f(a) + (c-a)\,f(b) \\
P &= \frac{b-c}{b-a}\,f(a) + \frac{c-a}{b-a}\,f(b)
\end{aligned}
$$

6.1.3 Convexity Criterion

A function f that is twice differentiable on an open interval I is convex if and only if everywhere on I $f''(x) \geq 0$

Proof:

1) $f''(x) \geq 0 \quad \forall\ x \in I$, means that f' increases monotonically on I. Let $a < c < b$ in I. According to the mean value theorem

$$
\begin{aligned}
\frac{f(c)-f(a)}{c-a} &= f'(\xi) \qquad \text{for a} \qquad \xi \in\,]a,c[\\
\frac{f(b)-f(c)}{b-c} &= f'(\eta) \qquad \text{for a} \qquad \eta \in\,]c,b[
\end{aligned}
$$

Combined, because of $\quad f'(\xi) \leq f'(\eta)$

$$
\begin{aligned}
\frac{f(c)-f(a)}{c-a} &\leq \frac{f(b)-f(c)}{b-c} \\
(b-a)\,f(c) &\leq (b-c)\,f(a) + (c-a)\,f(b) \\
f(c) &\leq \frac{b-c}{b-a}\,f(a) + \frac{c-a}{b-a}\,f(b)
\end{aligned}
$$

2) Assume $f''(x) < 0$ for a $x_0 \in I$ and $c := f'(x_0)$. Then consider $F(x) = f(x) - c\,(x-x_0)$. With f, F is twice differentiable: specifically $F'(x_0) = f'(x_0) - c = 0$. x_0 is a stationary point of F; $F''(x_0) = f''(x_0) < 0$. According to the criterion about strict local extrema x_0 is a strict local maximum of F.

In particular, for small $\delta > 0$

$$
F(x_0-\delta) < F(x_0) \quad ; \quad F(x_0+\delta) < F(x_0)
$$

that means:

$$
\begin{aligned}
&f(x_0-\delta) + c\,\delta < f(x_0) \quad ; \quad f(x_0+\delta) - c\,\delta < f(x_0) \\
&f(x_0) > \frac{1}{2}f(x_0-\delta) + \frac{1}{2}f(x_0+\delta)
\end{aligned}
$$

Hence f is not convex $\quad (a = x_0 - \delta,\ b = x_0 + \delta,\ c = x_0\,)$.

$\square$

Remark:

There is also a criterion for convexity from above, the associated functions f are called concave!

6.2 Hölder's Inequality

Let p, $q \geq 1$ and let $1/p + 1/q = 1$. Then for every two systems $(x_i)_{1 \leq i \leq n}$ or respectively $(y_i)_{1 \leq i \leq n}$ of non-negative real numbers, the following estimate holds:

$$\boxed{\sum_{i=1}^{n} x_i \cdot y_i \;\leq\; \left(\sum_{i=1}^{n} x_i^p\right)^{1/p} \cdot \left(\sum_{i=1}^{n} y_i^q\right)^{1/q}}$$

Remark:

(1) The most commonly used special case is $p = q = 2$. In this case the inequality is also called the Cauchy-Schwarz inequality.

Proof of Hölder's inequality:

1) Lemma: For all pairs $a \geq 0$, $b \geq 0$ the following holds

$$a\,b \;\leq\; \frac{a^p}{p} + \frac{b^q}{q}$$

In the case of $a = 0$ or respectively $b = 0$ this is correct. Now let $a > 0$, $b > 0$

$$(\ln(x))'' \;=\; -\frac{1}{x^2} \quad \text{therefore } \ln(x) \text{ is concave on } \;]0, \infty[$$

Use this for the convex combination $\quad \dfrac{1}{p}\,a^p + \dfrac{1}{q}\,b^q$

$$\ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \;\geq\; \frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q \;=\; \ln(a\,b)$$

Application of the monotonically increasing function exp

$$\frac{a^p}{p} + \frac{b^q}{q} \;\geq\; a\,b$$

2) If with $a_i \geq 0\,, \quad b_i \geq 0$

$$\sum_{i=1}^{n} a_i^p \;=\; 1 \;=\; \sum_{i=1}^{n} b_i^q$$

then according to **1)**

$$a_i\, b_i \;\leq\; \frac{1}{p}\, a_i^p + \frac{1}{q}\, b_i^q$$

Summation over i results in

$$\sum_{i=1}^{n} a_i\, b_i \;\leq\; \frac{1}{p} \sum_{i=1}^{n} a_i^p + \frac{1}{q} \sum_{i=1}^{n} b_i^q \;=\; \frac{1}{p} + \frac{1}{q} \;=\; 1$$

3) Conclusion of proof

Abbreviations $\quad s = \sum_{i=1}^{n} x_i^p \quad , \quad t = \sum_{i=1}^{n} y_i^p$

$$a_i \;=\; \frac{x_i}{s^{1/p}} \quad , \quad b_i \;=\; \frac{y_i}{t^{1/q}} \quad \text{satisfy} \quad \mathbf{2)}$$

The assertion: $\quad \sum_{i=1}^{n} x_i\, y_i \;\leq\; s^{1/p}\, t^{1/q}$

is true if $s = 0$ or if $t = 0$; in case $s,\, t > 0$ holds, then

$$a_i := \frac{1}{s^{1/p}}\, x_i \quad ; \quad b_i := \frac{1}{t^{1/q}}\, y_i$$

meet the requirement of **2)**, therefore

$$1 \;\geq\; \sum_{i=1}^{n} a_i\, b_i \;=\; \frac{1}{s^{1/p}\, t^{1/q}} \cdot \sum_{i=1}^{n} x_i\, y_i$$

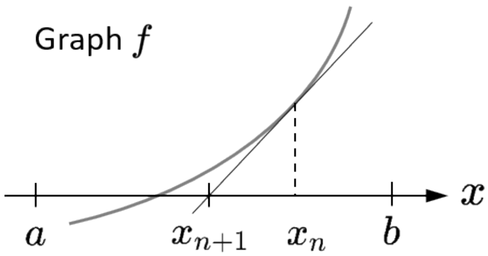
$\square$

<u>Remark:</u>

(2) Another consequence is Minkowski's inequality. For $p \geq 1$ and every pair of systems $(z_i)_{1 \leq i \leq n}$; $(w_i)_{1 \leq i \leq n}$ the following estimation holds:

$$\left(\sum_{k=1}^{n} |z_k + w_k|^p \right)^{1/p} \;=\; \left(\sum_{k=1}^{n} |z_k|^p \right)^{1/p} + \left(\sum_{k=1}^{n} |w_k|^p \right)^{1/p}$$

6.3 Newton's Method



Newton's method for approximating a zero α of a differentiable function f obtains a second approximation x_{n+1} from an initial value x_n in the vicinity of α from the intersection of the x axis with the tangent to the graph of f at the point $\begin{pmatrix} x_n \\ f(x_n) \end{pmatrix}$, in formulas

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

Does not work in every case, does not produce in every case a sequence that converges to α!

A theorem about Newton's method

Let the at least twice differentiable function f be real and convex on the interval I. Moreover, with $a < b$ in I, $f(a) < 0 < f(b)$

Then the following holds:

a) To the right of a in I lies exactly one zero α of f. There

$$f'(\alpha) \;\geq\; -\frac{f(a)}{\alpha - a} \;>\; 0$$

b) For each start value x_0 to the right of α ($x_0 \geq \alpha$), Newton's method returns a monotonically decreasing sequence $(x_n)_{n\geq 0}$ with limit value α

c) If $f''(x) \leq K \quad \forall\ x \in [\alpha, x_n]$ then the following estimate holds

$$0 \;\leq\; x_{n+1} - \alpha \;\leq\; \frac{K}{f'(x)}\,(x_n - \alpha)^2$$

Proof:

1) Intermediate value theorem $\quad\Rightarrow$

There exists a root $\alpha \in\ ]a, b[$ of f. According to the monotony criterion, the derivative f' increases monotonically

$$\frac{-f(a)}{\alpha - a} \;=\; \frac{f(\alpha) - f(a)}{\alpha - a} \;\leq\; f'(\alpha)$$

Therefore f to the right of α is strictly increasing, i.e. positive. In particular, α is unique!

2) In $I \cap [a, \infty[$ the function f is negative to the left of α and positive to the right of α. There $\alpha \leq x$ is equivalent to $0 \leq f(x)$. Now if $x_n = \alpha$, then $x_{n+1} = \alpha$, and then all further assertions are obvious. If $x_n > \alpha$

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

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

since α is the root of $f(x)$. Application of the mean value theorem results in

$$x_{n+1} - \alpha = (x_n - \alpha)\left[1 - \frac{f'(\xi)}{f'(x_n)}\right] \qquad (*)$$

with positive $\xi \in \,]\alpha, x_n[$.

Because of the monotonicity of f', it follows in particular that $0 \leq x_{n+1} \leq x_n - \alpha$

$(x_n)_{n \geq 0}$ increases monotonically with limit value $x_* \geq \alpha$, , $n \to \infty$ in the recursion formula

$$x_* = \frac{f(x_*)}{f'(x_*)} \qquad \text{hence} \qquad \begin{cases} f(x_*) = 0 \\ x_* = 0 \end{cases}$$

3) The mean value theorem can be applied again to the equation $(*)$

$$\begin{aligned} x_{n+1} - \alpha &= \frac{x_n - \alpha}{f'(x_n)}\left[f'(x_n) - f'(\xi)\right] \\ &= \frac{(x_n - \alpha)(x_n - \xi)}{f'(x_n)} f''(\eta) \end{aligned}$$

with matching $\eta \in \,]\xi, x_n[$, and

$$0 \leq x_{n+1} - \alpha \leq \frac{(x_n - \alpha)^2 K}{f'(\alpha)}$$

Remark:

The estimation can be improved by the following right-hand side:

$$\frac{K}{2 f'(\alpha)} (x_n - x_{n+1})^2$$

□

6.4 The Extended Mean Value Theorem

Given are two differentiable real functions f, g on $]a,b[$, which are also continuous in the end points of the interval a, b. Furthermore, let $g'(x) \neq 0 \;\forall\; x \in \,]a,b[\,$. Then it follows that $g(b) \neq g(a)$ and also

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)} \qquad , \qquad \xi \in \,]a,b[$$

Proof:

$g(b) = g(a)$ is certainly wrong, because according to a generalization of our version of Rolle's theorem, $g'(\eta) = 0$ for $\eta \in \,]a,b[$ as opposed to the precondition.

The auxiliary function:

$$h(x) \;=\; (g(b)-g(a))\,f(x) \;-\; (f(b)-f(a))\,g(x)$$

is differentiable on $]a,b[$ and also continuous in the end points a, b

$$h(b) \;=\; f(a)\,g(b) \;-\; g(a)\,f(b) \;=\; h(a)$$

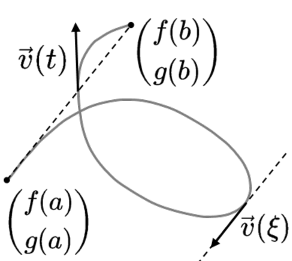
Hence according to Rolle $h'(\xi) = 0$ for a suitable $\xi \in \,]a,b[$. Substituting this into the derivative of the auxiliary function gives the assertion directly.

□

Remark:

For $g(x) = x$ the ordinary mean value theorem results.

Geometric interpretation of the result, let $x = t$ be the time:



$t \to \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}$ shall describe the trajectory of a point in the plane during the course of time in the interval $[a,b]$. The "velocity vector" $\vec{v} := \begin{pmatrix} \dot{f}(t) \\ \dot{g}(t) \end{pmatrix}$ is to be thought of as originating at the path point $\begin{pmatrix} f(t) \\ g(t) \end{pmatrix}$, its direction = direction of the tangent.

Hence the statement: On the trajectory there is a point where the velocity has the same direction as the connecting line through the end points of the trajectory.

6.5 The Rule of de l'Hospital

Let f, g be differentiable functions on $]a,b[$ with real values. Furthermore let

$$g'(x) \;\neq\; 0 \quad \begin{cases} \text{and} & \lim\limits_{x\searrow a} f(x) \;=\; \lim\limits_{x\searrow a} g(x) \;=\; 0 \\ \text{or} & \lim\limits_{x\searrow a} f(x) \;=\; \lim\limits_{x\searrow a} g(x) \;=\; \infty \end{cases}$$

Then the following applies

$$\lim_{x\searrow a} \frac{f(x)}{g(x)} \;=\; \lim_{x\searrow a} \frac{f'(x)}{g'(x)} \qquad \text{if the right-side limit value exists.}$$

Proof in the first case:

If one sets $f(a) = g(a) = 0$, then f and g also become continuous in a. The generalized mean value theorem is then being applied to the interval $[a,x]$ with any $x \in\,]a,b[$. With this we get

$$\frac{f(x)}{g(x)} \;=\; \frac{f(x)-f(a)}{g(x)-g(a)} \;=\; \frac{f'(\xi(x))}{g'(\xi(x))} \qquad \text{with} \qquad \xi \;=\; \xi(x) \in\,]a,x[$$

The assertion follows from this formula.

$\square$

Remark:

An analogous sentence holds for left-sided and thus also for two-sided limit values.

Example:

(1) $\mu,\, \nu \neq 0 \quad , \quad a > 0$

$$\lim_{x\searrow a} \frac{x^{\mu}-a^{\nu}}{x^{\nu}-a^{\mu}} \;=\; \lim_{x\searrow a} \frac{\mu\, x^{\mu-1}}{\nu\, x^{\nu-1}} \;=\; \lim_{x\searrow a} \frac{\mu}{\nu}\, x^{\mu-\nu} \;=\; \frac{\mu}{\nu}\, a^{\mu-\nu}$$

Supplement: If f and g are differentiable in $]a,\infty[$ so that

$$g'(x) \;\neq\; 0 \quad \begin{cases} \text{and} & \lim\limits_{x\to\infty} f(x) \;=\; \lim\limits_{x\to\infty} g(x) \;=\; 0 \\ \text{or} & \lim\limits_{x\to\infty} f(x) \;=\; \lim\limits_{x\to\infty} g(x) \;=\; \infty \end{cases}$$

Then the following holds

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} \;=\; \lim_{x\to\infty} \frac{f'(x)}{g'(x)} \qquad \text{if the right-side limit value exists!}$$

Example:

$$(2)\ \lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = \lim_{x\to\infty}\exp\left(x\,\ln\left(1+\frac{a}{x}\right)\right)$$

$$= \exp\left(\lim_{x\to\infty}\frac{\ln(1+{}^{a}/_{x})}{{}^{1}/_{x}}\right)$$

$$= \exp\left(\lim_{x\to\infty}\frac{-\,{}^{a}/_{x^2}}{-\,{}^{1}/_{x^2}\cdot(1+{}^{a}/_{x})}\right)$$

From $\lim_{x\to\infty}\left(1+\frac{a}{x}\right) = 1$ follows $\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = \mathrm{e}^a$

6.6 Taylor's Formula

Let f be a real function that is at least $N+1$ times ($N \in \mathbb{N}_0$) differentiable in a neighborhood of $a \in \mathbb{R}$. Then the follwing holds there

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \dots$$
$$\dots + \frac{1}{N!}f^{(N)}(a)(x-a)^N + R_{N+1}(x)$$

$$\boxed{f(x) = \sum_{n=0}^{N}\frac{1}{n!}f^{(n)}(a)(x-a)^n + R_{N+1}(x)}$$

with a remainder $R_{N+1}(x)$ which according to Lagrange has the following form:

$$\boxed{R_{N+1}(x) = \frac{1}{(N+1)!}f^{(N+1)}(a+\vartheta(x-a))(x-a)^{N+1}}$$

where $\vartheta \in\,]0,1[$; and according to Cauchy, with $\vartheta' \in\,]0,1[$

$$\boxed{R_{N+1}(x) = \frac{1}{N!}f^{(N+1)}(a+\vartheta'(x-a))(1-\vartheta')^N(x-a)^{N+1}}$$

Remarks:

(1) Taylor's formula gives an approximation for f in a neighborhood of a freely chosen reference point a by a polynomial T_N of degree $\leq N$. This polynomial is distinguished by $T_N^{(n)}(a) = f_N^{(n)}(a)\,, n = 0, 1, 2\ldots, N$; the expression for the remainder $R_{N+1}(x) = f(x) - T_N(x)$ reflects the goodness of the approximation.

(2) The case $N = 0$

$$f(x) \;=\; f(a) \,+\, f\,'(a + \vartheta(x-a))(x-a)$$

is a version of the mean value theorem.

(3) Other forms exist for the remainder R_{N+1}.

Proof for the Lagrangian case:

Trick: The reference point $a = t$ is made variable.

$$F(t) \;=\; f(x) \,-\, \sum_{n=0}^{N} \frac{1}{n!} f^{\,(n)}(t)(x-t)^n$$

$$G(t) \;=\; (x-t)^{N+1}$$

are considered (for the interesting case $x \neq a$) on the interval with the end points a and x. Generalized mean value theorem

$$\frac{F(a) - F(x)}{G(a) - G(x)} \;=\; \frac{F\,'(\xi)}{G\,'(\xi)} \qquad \text{for a } \xi \text{ in between } a \text{ and } x$$

hence $\quad \xi \;=\; a + \vartheta(x-a) \quad$ with $\quad 0 < \vartheta < 1$

$$G\,'(t) \;=\; -(N+1)(x-t)^N$$

$$F\,'(t) \;=\; -\sum_{n=0}^{N} \frac{1}{n!} f^{\,(n+1)}(t)(x-t)^n \,+\, \sum_{n=1}^{N} \frac{1}{(n-1)!} f^{\,(n)}(t)(x-t)^{n-1}$$

$$F\,'(t) \;=\; -\frac{1}{N!} f^{N+1}(t)(a-t)^N$$

Therefore

$$\frac{F\,'(\xi)}{G\,'(\xi)} \;=\; \frac{1}{(N+1)!} f^{\,(n+1)}(\xi)$$

$$\frac{F(a) - F(x)}{G(a) - G(x)} \;=\; \frac{R_{N+1}(x)}{(x-a)^{N+1}} \quad ; \quad F(x) = 0 \quad , \quad F(a) = R_{N+1}(x)$$

$\square$

Remark:

(4) In the case of a continuous $f^{(N+1)}$ with $f^{(N+1)}(a) \neq 0$ it is possible to make a statement in a neighborhood of a regarding the sign of $R_{N+1}(x)$: for x that are sufficiently close to a, $f^{(N+1)}(a+\vartheta(x-a))$ has the same sign as $f^{(N+1)}(a)$. If N is odd and hence $N+1$ even, then $R_{N+1}(x)$ has the same sign on both sides of a. On the other hand, if N is even and $N+1$ therefore odd, then the remainder changes sign when x changes from one side to the other.

Example:

Let $f^{(n)}(a) = 0$ $(1 \leq n \leq N)$ and $f^{(N+1)}(a) \neq 0$. The Taylor polynomial is then constant $= f(a)$.

If N is odd and hence $N+1$ even, then f has a local extremum in a. If N is even, then f has no local extremum in a.

Remark:

(5) If f is arbitrarily often differentiable then there is an obvious question: Is $\lim\limits_{N \to \infty} R_{N+1}(x) = 0$? If this applies to all x of an environment, one has the "Taylor expansion" of f!

6.6.1 The Exponential Series

$$f(x) \quad = \quad \exp(x) \quad , \quad a = 0$$

From the characterization of the function exp, solution of $y' = y$ with value 1 at 0, follows $\exp^{(n)}(x) = \exp(x)$, the n-th term of the Taylor formula is:

$$\frac{1}{n!}\, x^n$$

We use the Lagrangian remainder

$$R_{N+1}(x) \quad = \quad \frac{\exp(\vartheta_N\, x)}{(N+1)!}\, x^{N+1}$$

Because of the monotony of exp it follows for $x \in [-T, T]$

$$|R_{N+1}(x)| \quad \leq \quad \frac{\mathrm{e}^T\, T^{N+1}}{(N+1)!}$$

and from that in case $N \geq 2T$

$$|R_{N+k}(x)| \quad \leq \quad \frac{\mathrm{e}^T\, T^N}{N!} \left(\frac{1}{2}\right)^k$$

It follows

$$\lim_{k\to\infty} R_{N+k}(x) \;=\; 0 \qquad \forall\ x \in [-T,T]$$

$$\exp(x) \;=\; \mathrm{e}^x \;=\; 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \ldots$$

$$\boxed{\mathrm{e}^x \;=\; \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad x \in \mathbb{R}}$$

Remarks:

(1) Analogously one gets $\mathrm{e}^{-x} \;=\; 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 \pm \ldots$

$$\mathrm{e}^{-x} \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\,x^n \qquad x \in \mathbb{R}$$

(2) Euler's number e is irrational

Proof: Suppose $p, q \in \mathbb{N}$. The equation

$$1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{N!} + \frac{\mathrm{e}^{\vartheta_N}}{(N+1)!} \;=\; \frac{p}{q} \qquad (*)$$

is impossible for $N > q$ and any $\vartheta_N \in \,]0,1[$! Otherwise multiplication with $N!$ with the result

$$\frac{N!\,p}{q} - \sum_{n=0}^{n} \frac{N!}{n!} \;=\; \frac{\mathrm{e}^{\vartheta_N}}{N+1}$$

Now there is an integer on the left because it is a sum of integers, but there is a positive number on the right $< {}^{\mathrm{e}}/_{3} < 1 \;\Rightarrow\;$ equation $(*)$ impossible.

$\square$

6.6.2 The Logarithm Series

Taylor expansion of the natural logarithm around the reference point $a = 1$

$$x \;\to\; \ln(1+x)$$

Derivatives

$$\ln^{(n)}(1+x) \;=\; \frac{(-1)^{n-1}\,(n-1)!}{(1+x)^n} \qquad n \geq 1 \quad \text{(by induction)}$$

The term to the 0-th power of x in the Taylor expansion is $\ln(1+0) \;=\; 0$, whereas the term to the n-th power

$$\frac{(-1)^{n-1}}{n} \cdot x^n$$

What about $\quad R_{N+1}(x) \;\to\; 0$?

For $\;0 \le x \le 1\;$ Lagrangian remainder:

$$R_{N+1}(x) \;=\; \frac{(-1)^N}{N+1}\,\frac{1 \cdot x^{N+1}}{(1+\vartheta\, x)^{N+1}}$$

For $0 \le x \le 1$ the second factor becomes ≤ 1, the first factor forms a null sequence, hence

$$\lim_{N\to\infty} R_{N+1}(x) \;=\; 0 \qquad \text{if} \quad 0 \le x \le 1$$

For $\;-1 < x \le 0\;$ Cauchy remainder:

$$R_{N+1}(x) \;=\; \frac{(-1)^N\,(1-\vartheta\,')^N}{(1+\vartheta\,'x)^{N+1}}\,x^{N+1}$$

Because of

$$1 \,-\, \vartheta'\,|x| \;\ge\; 1 \,-\, \vartheta\,' \qquad \text{follows} \qquad \left(\frac{1-\vartheta\,'}{1+\vartheta\,'x}\right)^N \;\le\; 1$$

Therefore

$$|R_{N+1}(x)| \;\le\; \frac{|x|^{N+1}}{1-|x|}$$

and therefore

$$\lim_{N\to\infty} R_{N+1}(x) \;=\; 0 \qquad \text{if} \qquad -1 < x \le 0$$

$$\ln(1+x) \;=\; \frac{x}{1} - \frac{x^2}{2} + \frac{x^3}{3} \pm \ldots$$

$$\boxed{\ln(1+x) \;=\; \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\,x^n}{n} \qquad -1 < x \le 1}$$

In particular, the value of the alternating harmonic series becomes

$$1 \,-\, \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \pm \ldots \;=\; \ln 2$$

However, this series is useless for the calculation of $\ln 2$.

The binomial formula

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$

possesses a generalization for arbitrary exponents α instead of n.

6.6.3 The Binomial Series

For $\alpha \in \mathbb{N}$ the generalized binomial coefficient $\binom{\alpha}{n}$ for natural $n \geq 0$ is recursively declared by

$$\binom{\alpha}{0} = 1 \quad , \qquad \binom{\alpha}{n+1} = \frac{\alpha - n}{n+1} \cdot \binom{\alpha}{n}$$

written out

$$\binom{\alpha}{n} = \frac{\alpha}{n} \cdot \frac{\alpha-1}{n-1} \cdot \frac{\alpha-2}{n-2} \cdot \ \ldots \ \cdot \frac{\alpha-n+1}{1}$$

For $-1 < x < 1$ the following holds

$$(1+x)^\alpha = 1 + \alpha\, x + \frac{\alpha\,(\alpha-1)}{1\cdot 2}\, x^2 + \frac{\alpha\,(\alpha-1)\,(\alpha-2)}{1\cdot 2\cdot 3}\, x^3 + \ldots$$

$$\boxed{(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n}$$

For $\quad f(x) = (1+x)^\alpha \quad$ in $\quad -1 < x < \infty \quad$ the following applies

$$f^{(n)}(x) = n! \cdot \binom{\alpha}{n} \cdot (1+x)^{\alpha-n}$$

Proof by induction!

for $n = 0$ correct: $\qquad f^{(n+1)} = (n+1)! \binom{\alpha}{n+1} (1+x)^{\alpha-(n+1)}$

The remainder after Lagrange

$$R_{N+1}(x, a=0) = \binom{\alpha}{N+1} (1+\vartheta\, x)^{\alpha-N-1} x^{N+1}$$

From this one can conclude

$$\lim_{N\to\infty} R_{N+1}(x) = 0 \ , \qquad \text{if} \quad 0 \leq x < 1$$

For $-1 < x \leq 0$ use the Cauchy remainder with some added consideration.

<u>Special case $N = 1$:</u> $\quad (1+x)^\alpha \;=\; 1 + \alpha\, x + R_2$

According to Lagrange

$$R_2(x, a=0) \;=\; \frac{\alpha\,(\alpha-1)}{2}\,(1-\vartheta\, x)^{\alpha-2}\,x^2$$

According to Cauchy

$$R_2(x, a=0) \;=\; \alpha\,(\alpha-1)\,(1+\vartheta' x)^{\alpha-2}\,(1-\vartheta')\,x^2$$

Specifically: remainders for $\alpha = \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}$

<u>$\alpha = {}^1\!/_2$:</u> $\displaystyle \binom{\alpha}{2} = \binom{{}^1\!/_2}{2} = \frac{{}^1\!/_2 \cdot (-{}^1\!/_2)}{2} = -\frac{1}{8}$

$$\sqrt{1+x} \;=\; 1 + \frac{x}{2} + R_2\,,$$

with $R_2 < 0$ and $|R_2| \leq \dfrac{x^2}{8}$, if $x > 0$

<u>$\alpha = -{}^1\!/_2$:</u> $\displaystyle \binom{\alpha}{2} = \binom{-{}^1\!/_2}{2} = \frac{-{}^1\!/_2 \cdot (-{}^3\!/_2)}{2} = \frac{3}{8}$

$$\frac{1}{\sqrt{1+x}} \;=\; 1 - \frac{x}{2} + R_2\,,$$

with $R_2 > 0$ and $|R_2| \leq \dfrac{3 \cdot x^2}{8}$, if $x > 0$

<u>$\alpha = {}^1\!/_3$:</u> $\displaystyle \binom{\alpha}{2} = \binom{{}^1\!/_3}{2} = \frac{{}^1\!/_3 \cdot (-{}^2\!/_3)}{2} = -\frac{1}{9}$

$$\sqrt[3]{1+x} \;=\; 1 + \frac{x}{3} + R_2\,,$$

with $R_2 < 0$ and $|R_2| \leq \dfrac{x^2}{9}$, if $x > 0$

7. Oscillation Equation and Trigonometric Functions

$$\boxed{y'' + y \;=\; 0 \qquad \text{(L)}}$$

We are looking for all twice differentiable functions that satisfy (L)!

Properties of the solutions of (L):

(1) If f is a solution of (L), then $f'' = -f$; hence f is even four times differentiable, $-f$ is again a solution, i.e. $f^{(IV)} = f$. In particular, f is arbitrarily often differentiable.

(2) For any two solutions f, g of (L) and any constants a, $b \in \mathbb{R}$,

$$af + bg$$

is again another solution.

(3) For any two solutions f, g of (L), the "Wronski determinant" is declared as

$$w(x) \;=\; w_{f,g}(x) \;=\; (fg' - f'g)(x)$$

w is differentiable with the derivative

$$\begin{aligned} w' &= fg'' - f''g \\ &= -fg + fg \;=\; 0 \end{aligned}$$

Therefore $w_{f,g} = w$ is a constant.

(4) For each pair of initial conditions $f(0) = a\,,\; f'(0) = b \;\; (a, b \in \mathbb{R})$ exists at most one solution f of (L).

Proof:

Let us assume that f and g are two solutions of (L) with the same initial conditions. Then it follows that $h = f - g$ is also a solution of (L) with the initial conditions $h(0) = h'(0) = 0$.

Consider:

$$h^2(x) + h'^2(x) \quad \text{is differentiable with derivative}$$

$$2\,h'(x)\,(h(x) + h''(x)) \equiv 0$$

Therefore (for every solution h) $h^2(x) + h'^2(x)$ is constant. Evaluation of our solution h in the point 0 results in 0 for the constant.

$$h^2(x) + h'^2(x) = 0 \qquad \forall\ x$$

In particular $h(x) = 0\ \forall\ x$. Therefore $f = g$.

□

(5) If f is any solution of (L) with initial conditions $f(0) = a$, $f'(0) = b$, then, according to (1), one has the Taylor expansion of f around the zero point.

$$f(x) = a + \frac{b}{1!}x - \frac{a}{2!}x^2 - \frac{b}{3!}x^3 + \frac{a}{4!}x^4 + \frac{b}{5!}x^5 -, -, +, +, \cdots$$

$\cos x$ or respectively $\sin x$ is declared as the solution of the oscillation equation (L) with the initial conditions

$$\begin{array}{rcl} \cos(0) &=& 1 \\ \cos'(0) &=& 0 \end{array} \quad \text{or respectively} \quad \begin{array}{rcl} \sin(0) &=& 0 \\ \sin'(0) &=& 1 \end{array}$$

(6) The Wronski determinant for $f = \cos$, $g = \sin$ is calculated from the knowledge of the functions $\cos' x$ and $\sin' x$ which are the solutions of (L) with the initial conditions

$$\begin{array}{rcl} \cos'(0) &=& 0 \\ \cos''(0) &=& -1 \end{array} \quad \text{or respectively} \quad \begin{array}{rcl} \sin'(0) &=& 1 \\ \sin''(0) &=& 0 \end{array}$$

that means: $\cos' x = -\sin x$ and $\sin' x = \cos x$ and hence the constant Wronski determinant is:

$$\cos^2 x + \sin^2 x = 1$$

(7) Hence a solution of (L) with $f(0) = a$, $f'(0) = b$ is therefore

$$f(x) = a\cos x + b\sin x$$

It is uniquely determined according to (4) by the initial conditions. In short: The solution space of the oscillation equation is

$$\{f \;=\; a\cos + b\sin\,;\;\; a,b \in \mathbb{R}\}$$

(8) With $f(x)$, $g(x) := f(x_1 + x)$ is also a solution of (L). In particular for $f = \cos$ or respectively $f = \sin$ one searches for the representation of

$$f(x_1 + x) \;=\; a\cos(x) \,+\, b\sin(x)$$

The constants are obtained by evaluating at the position $x = 0$ for the function and its derivative

$$f(x_1) \;=\; a \quad ; \quad f'(x_1) \;=\; b$$

Specifically:

Addition theorem or functional equation

$$\cos(x_1 + x_2) \;=\; \cos x_1 \cos x_2 \,-\, \sin x_1 \sin x_2$$

$$sin(x_1 + x_2) \;=\; \sin x_1 \cos x_2 \,+\, \cos x_1 \sin x_2$$

(9) Commonly used corollaries

$$\cos(-x) \;=\; \cos x \quad , \quad \sin(-x) \;=\; -\sin x$$

$$\cos 2x \;=\; \cos^2 x \,-\, \sin^2 x$$

$$\sin 2x \;=\; 2\,\sin x \cos x$$

$$\sin x \,-\, \sin y \;=\; 2\,\cos\frac{x+y}{2}\,\sin\frac{x-y}{2} \qquad (IV)$$

$$\cos x \,-\, \cos y \;=\; -2\,\sin\frac{x+y}{2}\,\sin\frac{x-y}{2} \qquad (V)$$

Because $\cos(-x)$, $\sin(-x)$ are solutions of (L) with the initial conditions

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or respectively} \quad \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

The doubling formulas follow through $x_1 = x_2 = x$ from the addition theorems!

With $\quad x_1 = \dfrac{x+y}{2} \quad , \quad x_2 = \dfrac{x-y}{2}$ we have $\begin{cases} x_1 + x_2 = x \\ x_1 - x_2 = y \end{cases}$

Conclusion (IV) left side:

$$\sin(x_1 + x_2) - \sin(x_1 - x_2) =$$

$$= \sin x_1 \cos x_2 + \cos x_1 \sin x_2 - \sin x_1 \cos x_2 + \cos x_1 \sin x_2$$

$$= 2 \cos x_1 \sin x_2$$

Conclusion (V) left side:

$$\cos(x_1 + x_2) - \cos(x_1 - x_2) = -2 \sin x_1 \sin x_2$$

7.1 Intermediate Section on Power Series

The label "(infinite) series" denotes an expression $\displaystyle\sum_{n=0}^{\infty} a_n$

It has two meanings:

(1) It denotes the sequence $s_n = \displaystyle\sum_{k=0}^{n} a_k$, $n \in \mathbb{N}$ of the so-called "partial sums" s_n .

(2) If this sequence has a limit value $s = \lim\limits_{n \to \infty} s_n$, then $s = \displaystyle\sum_{n\ ,=0}^{\infty} a_n$ is called an infinite series.

The arithmetic laws about convergent sequences can be transferred to the convergent infinite series!

Because of $a_n = s_n - s_{n-1}$ one has as a necessary condition for the convergence of the sequence of partial sums: The sequence of "terms" a_n of a convergent series $\displaystyle\sum_{n=0}^{\infty} a_n$ is a null sequence.

This condition is not sufficient! Example (harmonic series):

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots = \sum_{m=1}^{\infty} \frac{1}{m} \qquad \text{is divergent}$$

$$\begin{aligned}
s_1 &= 1 \\
s_2 &= 1 + \frac{1}{2} \\
s_4 &= s_2 + \frac{1}{3} + \frac{1}{4} \geq s_2 + \frac{1}{4} + \frac{1}{4} = s_2 + \frac{1}{2} \\
s_8 &= s_4 + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \geq s_4 + 4\frac{1}{8} = s_4 + \frac{1}{2}
\end{aligned}$$

In general

$$s_{2^{k+1}} = s_{2^k} + \sum_{m=2^k+1}^{2\cdot 2^k} \frac{1}{m} \geq s_{2^k} + 2^k \sum \frac{1}{2^{k+1}} = s_{2^k} + \frac{1}{2}$$

The monotonically increasing sequence of s_n is therefore not bounded, i.e. not actually convergent!

7.1.1 Leibniz' Convergence Criterion

For every monotonically decreasing zero sequence $(a_n)_{n\geq 0}$ the "alternating" series

$$\sum_{n=0}^{\infty} (-1)^n a_n$$

is convergent and for the partial sums a_n and the limit value s the following holds

$$s_{2n-1} \leq s_{2n+1} \leq s \geq s_{2n} \geq s_{2n+2}$$

and

$$|s - s_n| \leq a_{n+1} \quad \forall\ n \in \mathbb{N}$$

Proof:

$$\begin{aligned}
s_n &= s_{n-1} + (-1)^n a_n \\
s_{n+1} &= s_{n-1} + (-1)^n \underbrace{(a_n - a_{n+1})}_{\geq 0}
\end{aligned}$$

The odd indexed partial sums form an increasing sequence, the even indexed ones a monotonically decreasing sequence, always $s_{2n+1} \leq s_0$, $s_{2n} \geq s_1$. According to the monotony criterion of sequences both sequences converge and $|s_{2n+1} - s_{2n}| = a_{2n+1}$ shows that the two limits values agree!

□

Examples:

(1) Alternating harmonic series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} = \ln 2$$

(2) Leibniz series

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \frac{\pi}{4}$$

Series of the form

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots$$

are known as "power series" centered at the point x_0. The numbers a_n (real or complex) are called "coefficients" of the power series. The shift $x \mapsto x + x_0$ turns this into a power series

$$S = \sum_{n=0}^{\infty} a_n x^n$$

with the center point 0. Each such series is assigned an element $\varrho \in [0, \infty]$ (N.H. Abel):

$$\varrho \coloneqq \sup\{r \geq 0 \,/\, \text{ the sequence } (a_n r^n) \text{ bounded}\}$$

is the radius of convergence of S.

7.1.2 Theorem on the Radius of Convergence

If ϱ is the convergence radius of $S = \sum_{n=0}^{\infty} a_n x^n$, then S converges for $|x| < \varrho$ and diverges for $|x| > \varrho$.

Proof:

1) If $|x| > \varrho$, then $r = |x|$ is not allowed under the supremum bracket for ϱ. Hence the sequence $(a_n x^n)_{n \geq 0}$ of the series members is not bounded. Because the sequence of the terms of any convergent series is necessarily bounded as a zero sequence, the given series does not converge!

2) Now let $|x| < \varrho$. By the definition of ϱ there exists a r with $|x| < r \leq \varrho$ with $(a_n x^n)_{n \geq 0}$ bounded, somewhat like $|a_n| r^n \leq M \;\; \forall \;\, n \in \mathbb{N}_0$.

Handle the sequence of partial sums according to the Cauchy criterion:

$$|s_{n+p}(x) - s_n(x)| \;=\; \left| \sum_{m=n+1}^{n+p} a_m x^m \right|$$

(with triangle inequality)

$$\leq \sum_{m=n+1}^{n+p} |a_m||x|^m \;\leq\; \sum_{m=n+1}^{n+p} |a_m|\, r^m \left(\frac{|x|}{r}\right)^m$$

(according to the premise)

$$\leq \left(\frac{|x|}{r}\right)^{n+1} \cdot M \cdot \sum_{k=0}^{\infty} \left|\frac{x}{r}\right|^k \;=\; \frac{M\left(\frac{|x|}{r}\right)^{n+1}}{1 - \frac{|x|}{r}} \;<\; \epsilon$$

in the case $n \geq n_\epsilon$, because $q := \dfrac{|x|}{r} < 1$

$\square$

Remark:

The proof even allows the conclusion that $\sum_{n=0}^{\infty} |a_n||x|^n$ converges if $|x| < \varrho$; in words, inside their region of convergence, power series converge absolutely.

7.1.3 A Criterion for Determining the Radius of Convergence

If $\lim_{n\to\infty} \left|\dfrac{a_n}{a_{n+1}}\right| = R$, then $R = \varrho$ is the radius of convergence of $S = \sum_{n=0}^{\infty} a_n x^n$

Proof: the premise implies in particular $a_n \neq 0$ if $n \geq 0$.

1) Let $0 \leq r < R$. Then

$$|a_{n+1}|\, r \;\leq\; |a_n| \quad \text{if} \quad n \geq n_1 \geq n_0$$

Therefore $|a_n|\, r^n \;\geq\; |a_{n+1}|\, r^{n+1}$ and the sequence $(|a_n|\, r^n)$ is a monotonically decreasing and bounded sequence! Hence $R \leq \varrho$.

2) Let $R < r$, then $(|a_n|\, r^n)$ is not bounded. Choose r_1 with $R < r_1 < r$

$$|a_n| \;\leq\; |a_{n+1}|\, r_1 \quad \text{if} \quad n \geq n_2 \geq n_0$$

Therefore

$$0 \;\le\; M \;:=\; |a_{n_2}| r_1^{n_2} \;\le\; |a_n| r_1^n \quad ; \quad n \ge n_2 \quad \text{(inductive)}$$

$$|a_n| r^n \;=\; |a_n| r_1^n \left(\frac{r}{r_1}\right)^n \;\ge\; M \left(\frac{r}{r_1}\right)^n \;\to\; \infty$$

In particular, this sequence $(|a_n| r^n)_{n \ge 0}$ is not bounded!

□

Examples:

(1) The geometric series: $\sum_{n=0}^{\infty} x^n$

has nothing but coefficients $a_n = 1$, therefore $\left|\frac{a_n}{a_{n+1}}\right| = 1$, i.e. $\varrho = 1$.

(2) The exponential series: $\sum_{n=0}^{\infty} \frac{x^n}{n!}$

has the coefficients $a_n = \frac{1}{n!}$, hence $\left|\frac{a_n}{a_{n+1}}\right| = n+1$, i.e. $\varrho = \infty$; therefore every sequence $\left(\frac{r^n}{n!}\right)_{n\;,\ge 0}$ $(r \ge 0)$ is bounded!

(3) The series:

$$C(x) \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \quad ; \quad S(x) \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

both have the radius of convergence $\varrho = \infty$ because the sequences

$$\left(\frac{r^{2n}}{(2n)!}\right)_{n \ge 0} \quad , \quad \left(\frac{r^{2n+1}}{(2n+1)!}\right)_{n \ge 0} \quad ,$$

according to example (2), are bounded for each $r \ge 0$.

7.1.4 The Derivative of a Power Series

Let $S = \sum_{n=0}^{\infty} a_n x^n$ be a power series with real (or complex) coefficients a_n and with the radius of convergence ϱ.

$$DS \;=\; \sum_{n=1}^{\infty} n\, a_n x^{n-1}$$

is the term wise differentiated series of S and shall have the radius of convergence $\varrho\,'$.

a) $\varrho = \varrho\,'$ holds

b) For all z with $|z| < \varrho$ applies

$$\boxed{\lim_{\substack{h \to 0 \\ h \neq 0}} \left[\frac{1}{h}(S(z+h) \;-\; S(z)) \;-\; DS(z) \right] \;=\; 0}$$

Proof: of **a)**

1) Show: $\varrho' \leq \varrho$

Let $r < \varrho'$, then by definition of ϱ'

$$(n\,|a_n|\,r^{n-1})_{n\,\geq\,1} \text{ is bounded,}$$

$(|a_n|\,r^n)_{n\,\geq\,0}$ is therefore all the more bounded, that means $r < \varrho$ and therefore $\varrho' \leq \varrho$.

2) Let $q \in \,]0,1[$ then $q = \dfrac{1}{1+\delta}$ with $\delta > 0$ and using Bernoulli's inequality

$$n\,q^n \;=\; \frac{n}{(1+\delta)^n} \;\leq\; \frac{n}{1+n\,\delta} \;\leq\; \frac{1}{\delta}$$

Without restriction let $\varrho > 0$ (otherwise the claim $\varrho = \varrho'$ is clear), $0 < r < \varrho$. Auxiliary number r_1 with $r < r_1 < \varrho\,;\; q = \dfrac{r}{r_1}$

$$n\,|a_n|\,r^{n-1} \;=\; n\,q^n\,|a_n|\,r_1^n\,\frac{1}{r}$$

As a product of two bounded sequences this sequence is itself bounded, hence $r \leq \varrho'$, therefore $\varrho' \leq \varrho$ is false, so $\varrho \leq \; varrho'$.

Together it follows $\varrho = \varrho'$.

□

Proof: of **b)**

$$\sum_{n\,=\,1}^{\infty} n\,|a_n|\,r^{n-1}$$

According to part **a)** (and because of the remark on the theorem about the radius of convergence) this series converges if $r < \varrho = \varrho'$

1) For fixed z with $|z| < \varrho$ choose r with

$$|z| \;<\; r \;<\; \varrho \;=\; \varrho'$$

Restriction of h to $|h| \leq r - |z|\,$. Then $|z+h| \leq r\,$. For a given $\epsilon > 0$ there exists an index N_ϵ with

$$\sum_{n=N_\epsilon+1}^{\infty} n\,|a_n|\,r^{n-1} \;<\; \frac{\epsilon}{4}$$

2) $$a^n - b^n \;=\; (a-b)\left(a^{n-1} + a^{n-2}\,b + \ldots + a\,b^{n-2} + b^{n-1}\right)$$
$$\;=\; (a-b)\sum_{k=0}^{n-1} a^{n-1-k}\,b^k$$

yields for

$$T(h) \;=\; \left[\frac{1}{h}(S(z+h) - S(z)) - DS(z)\right]$$
$$\;=\; \sum a_n\left(\frac{(z+h)^n - z^n}{h} - n\,z^{n-1}\right)$$

for the error at a_n the expression

$$p_n(h) \;=\; \sum_{k=0}^{n-1}(z+h)^{n-1-k}\,z^k - n\,z^{n-1}$$

and the estimate

$$|p_n(h)| \;\le\; 2\,n\,r^{n-1}$$

3) $T(h)$ is (as an infinite series) the limit value of the sequence of partial sums t_n. Subtracting the variable t_{N_ϵ} which is independent of n yields a decomposition of T.

$$T \;=\; t_{N_\epsilon} + \widetilde{T} \qquad \text{with} \qquad \widetilde{T} \;=\; \sum_{n=N_\epsilon+1}^{\infty} a_n p_n(h)$$

Estimation of $\widetilde{T}$ via the partial sums:

$$\sum_{n=N_\epsilon+1}^{N_\epsilon+p} a_n p_n(h) \;\overset{\text{after } \mathbf{2)}}{\underset{\Downarrow}{\le}}\; 2\sum_{n=N_\epsilon+1}^{N_\epsilon+p} |a_n|\,n\,r^{n-1}$$
$$\le\; 2\sum_{n=N_\epsilon+1}^{\infty} |a_n|\,n\,r^{n-1} \;<\; \frac{\epsilon}{2}$$

Estimation of $t_{N_\epsilon}(h)$:

This is a finite sum of polynomials $a_n p_n(h)$ with value 0 at the origin ($h = 0$). In particular, $t_{N_\epsilon}(h)$ is continuous and dependent on h with value zero at zero. Therefore a $\delta_\epsilon > 0$ exists with

$$|t_{N_\epsilon}(h)| \;<\; \frac{\epsilon}{2} \qquad \text{with} \qquad |h| < \delta_\epsilon$$

Combined:

For every $\epsilon > 0$ exists a $\delta_\epsilon > 0$ with $|T(h)| < \epsilon$ if (in addition) $0 < |h| < \delta_\epsilon$.

$\square$

Remark:

If the power series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ has a positive radius of convergence $\varrho > 0$, then the series sum

$$f(x) \;:=\; \sum_{n=0}^{\infty} a_n(x-a)^n \qquad \text{for} \quad |x-a| \;<\; \varrho$$

defines an arbitrarily often differentiable function. Derivatives:

$$f^{(k)}(x) \;=\; \sum_{n=k}^{\infty} n\,(n-1)\cdots(n-k+1)a_n(x-a)^{n-k}$$

Evaluation at the reference point a

$$\frac{f^{(k)}(a)}{k!} \;=\; a_k \qquad (k \;=\; 0,1,2\ldots)$$

The power series is also the Taylor expansion of the function represented by the series.

Example:

$$\cos x \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \quad ; \quad \sin x \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

are the initially considered fundamental solutions of the oscillation equation:

$$y'' + y \;=\; 0$$

For $0 \le x \le 2$ those are the alternating series in the sense of the Leibniz criterion, resulting in the derivates:

$$1 - \frac{x^2}{2} \;\le\; \cos x \;\le\; 1 - \frac{x^2}{2} + \frac{x^4}{24}$$

$$1 - \frac{x^3}{6} \;\le\; \sin x \;\le\; x \;, \quad \text{if} \quad 0 \le x \le 2$$

In particular

$$\cos 2 \;\le\; 1 - \frac{4}{2} + \frac{16}{24} \;=\; -\frac{1}{3} \;<\; 0 \;<\; \cos 0.$$

Therefore $\cos x$ has a zero in the interval $]0,2[$ (exactly one zero , namely $\pi/2$).

7.2 Inverse Functions

The number π is declared as the smallest positive zero of $\sin x$! From the formula $\sin x = \sin(x/2)\cos(x/2)$ one can further recognize that $\pi/2$ is the smallest positive zero of $\cos x$.

For $\sin x$ in $0 < x < 2$: $\quad \sin x \;\geq\; x\left(1 - \dfrac{x^2}{6}\right) \;>\; 0$

$\cos' x = -\sin x$; hence $\cos x$ strictly decreases in $]0,2[$ and has exactly one zero: $\pi/2$

$$\sin(x+\pi/2) \;=\; \cos x \quad ; \quad \cos(x+\pi/2) \;=\; -\sin x \qquad (*)$$

$$\sin(x+\pi) \;=\; -\sin x \quad ; \quad \cos(x+\pi) \;=\; -\cos x$$

$$\sin(x+2\pi) \;=\; \sin x \quad ; \quad \cos(x+2\pi) \;=\; \cos x$$

In particular, all solutions of the oscillation equation $y'' + y = 0$ are periodic functions of period 2π!

It suffices to prove the first equation. Both sides are solutions of the oscillation equation. Because of $\sin'(x) = \cos x$ and $\cos x > 0$ for $0 \leq x < \pi/2$, $\sin x \geq 0$ for $0 \leq x < \pi/2$. Furthermore $\cos^2 x + \sin^2 x = 1$ for all x! Evaluate at $x = \pi/2$, $\sin^2 \pi/2 = 1$, $\sin \pi/2 = 1$.

On the left (as well as on the right) is that solution of the oscillation equation which has the value 1 at 0 and the derivative value 0 at 0! Solutions with the same initial conditions are generally the same.

The set of zeros for $\sin x$ is $\quad \pi\mathbb{Z} \;:=\; \{\pi k \;/\; k \text{ integer}\}$

The set of zeros for $\cos x$ is $\quad \dfrac{\pi}{2} + \pi\mathbb{Z} \;:=\; \left\{\dfrac{\pi}{2} + \pi k \;/\; k \text{ integer}\right\}$

In particular, the quotient can be defined:

$$\tan x \;:=\; \frac{\sin x}{\cos x} \qquad \text{for all} \qquad x \in \mathbb{R} \setminus \left\{\frac{2k-1}{2}\pi \;/\; k \text{ integer}\right\}$$

Because of the equation $(*)$, the tangent is a periodic function with period π.

As with all periodic functions, one has to restrict oneself to monotony intervals when forming the inverse function for sine, cosine, and tangent. With the label principal values (for sine and tangent) one denotes the inverse functions $\arcsin(y)$ or respectively $\arctan(y)$ which are obtained (with restrictions) from $\sin \mid [-\pi/2, \pi/2]$ or respectively from $\tan \mid\,]-\pi/2, \pi/2[$.

Construction of the derivative of these functions according to the theorem about the derivative of the inverse function

$$\sin' x \;=\; \cos x \;=\; \sqrt{1-\sin^2 x} \quad ; \quad -\frac{\pi}{2} < x < \frac{\pi}{2}$$

$$\tan' x \;=\; \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \;=\; 1 + \tan^2 x \quad ; \quad -\frac{\pi}{2} < x < \frac{\pi}{2}$$

Hence it follows

$$\arcsin'(y) \;=\; \frac{1}{\sqrt{1-y^2}} \qquad ; \qquad \arctan'(y) \;=\; \frac{1}{1+y^2}$$

Because of $\displaystyle\lim_{x \nearrow \pi/2} \frac{\sin x}{\cos x} \;=\; +\infty$ and $\tan(-x) \;=\; -\tan x$

$$\tan\big(\,]-\pi/2, \pi/2[\,\big) \;=\;]-\infty, \infty[;$$

this is the domain of arctan!

Correspondingly, the domain of arcsin is

$$\sin\big([-\pi/2, \pi/2]\big) \;=\; [-1, 1]$$

The derivative does not exist in the end points of the domain of arcsin because it cannot be differentiated there.

From the Taylor series of the derivative

$$\arctan' x \;=\; \frac{1}{1+x^2} \;=\; \sum_{n=0}^{\infty} (-1)^n x^{2n}$$

one obtains the Taylor series of arctan itself (around $x = 0$)

$$\boxed{\arctan x \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}}$$

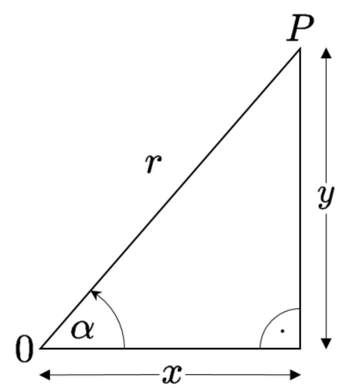
$$\boxed{\arcsin x \;=\; \sum_{n=0}^{\infty} \binom{-1/2}{n} (-1)^n \frac{x^{2n+1}}{2n+1}}$$

The proof arises from the Taylor expansion of the respective derivative function by specifying a "primitive function" (a function with the prescribed derivative) and adjusting at the zero point. Both have the radius of convergence $\varrho = 1$.

It is remarkable that $\arctan x$ is arbitrarily often differentiable for all $x \in \mathbb{R}$. The arctan series also converges for $x = 1$, its value becomes $\arctan 1 = {}^{\pi}/_{4}$

$$\frac{\pi}{4} \;=\; 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + - \ldots$$

7.2.1 Geometric Significance of the Angular Functions Sine and Cosine



The geometric significance of the angular functions sine and cosine becomes clear when one considers the aspect ratios in a right-angled triangle.

$$Sin\alpha \;=\; \frac{y}{r} \quad ; \quad Cos\alpha \;=\; \frac{x}{r}$$

The name "angle" stands for a number from the division of a circle. The full circumference is partitioned into 360° (in surveying also 400 gradians). In mathematical, physical and some technical applications, the division of a circle is based on the Euclidean length of the circumference $2\pi r$. $2r$ denotes the diameter of the considered circle! $1\,\text{rad} = 180\,\text{deg}\,/\pi \approx 57,296°$. The extension of the trigonometric functions to all $\alpha \in \mathbb{R}$ happens via the sign convention for the 4 quadrants of the plane: $x > 0, y > 0$; $x < 0, y > 0$; $x < 0, y < 0$; $x > 0, y < 0$;

$$\| \quad Cos^2\alpha + Sin^2\alpha \;=\; 1 \qquad \text{is an expression for the Pythagorean theorem.}$$

Projection of the corners of the oblique rectangle (see sketch for angle labels)

$$x \;=\; Cos\beta\, x' - Sin\beta\, y' \quad , \quad y \;=\; Sin\beta\, x' + Cos\beta\, y'$$

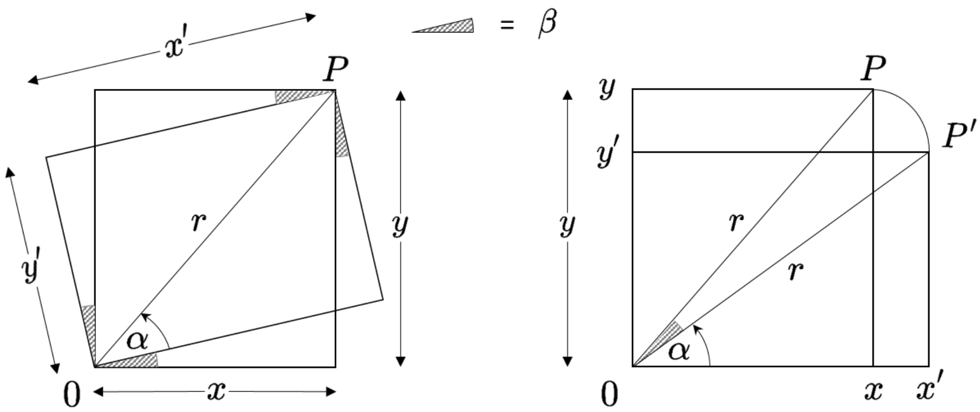
Back rotation of the oblique rectangle

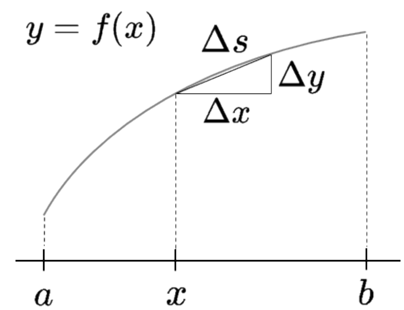
$$x \;=\; r\,Cos(\alpha+\beta) \quad , \quad y \;=\; r\,Sin(\alpha+\beta)$$

$$x' \;=\; r\,Cos\alpha \quad , \quad y' \;=\; r\,Sin\alpha$$

$$Cos(\alpha+\beta) \;=\; Cos\alpha\, Cos\beta - Sin\alpha\, Sin\beta$$

$$Sin(\alpha+\beta) \;=\; Sin\alpha\, Cos\beta + Cos\alpha\, Sin\beta$$





Euclidean length of an arc $s(x)$ of the form:

$$x \longmapsto \begin{pmatrix} x \\ f(x) \end{pmatrix} \quad , \quad x \in [a, b]$$

According to Pythagoras

$$(\Delta s)^2 \;=\; (\Delta x)^2 + (\Delta y)^2$$

In differential form:

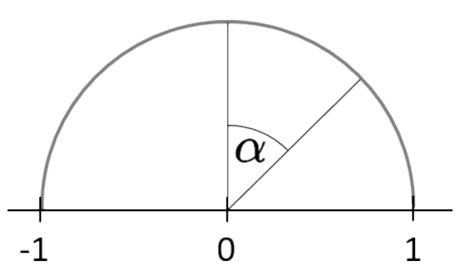
$$\frac{(\Delta s)^2}{(\Delta x)^2} \;=\; 1 + \frac{(\Delta y)^2}{(\Delta x)^2} \qquad \rightarrow \qquad s' \;=\; \sqrt{1 + f'(x)}$$

This heuristic formula for the derivative $f'(x)$ of the arc length $s(x)$ leads to the solution of the arc length question.

<u>Example:</u>

$$f(x) \;=\; \sqrt{1 - x^2}$$

describes the arc of the unit circle in the upper half-plane ($y > 0$).



$$f'(x) \;=\; \frac{-x}{\sqrt{1 - x^2}}$$

$$s'(x) \;=\; \sqrt{1 + \frac{x^2}{1 - x^2}} \;=\; \frac{1}{\sqrt{1 - x^2}}$$

According to the discussion of the oscillation equation, the primitive function with value 0 in $x = 0$ is $\arcsin x$.

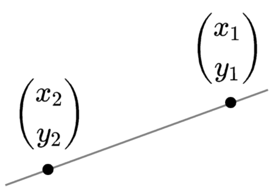
$$Sin\,\alpha \;=\; \sin\alpha \;=\; x$$

8. The Euclidean Plane and its Affine Mappings

When a system of coordinates is fixed, each point P in the plane is described by its pair of coordinates $\begin{pmatrix} x \\ y \end{pmatrix}$ $\{\vee\,(x/y)\}$.

Extended objects of the plane are also described by the coordinate systems of their points.

(1) Straight line through two different points $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$



$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} t + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} (1-t)\,;\; t \in \mathbb{R} \right\}$$

(2) Circle with center $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ and radius $r > 0$

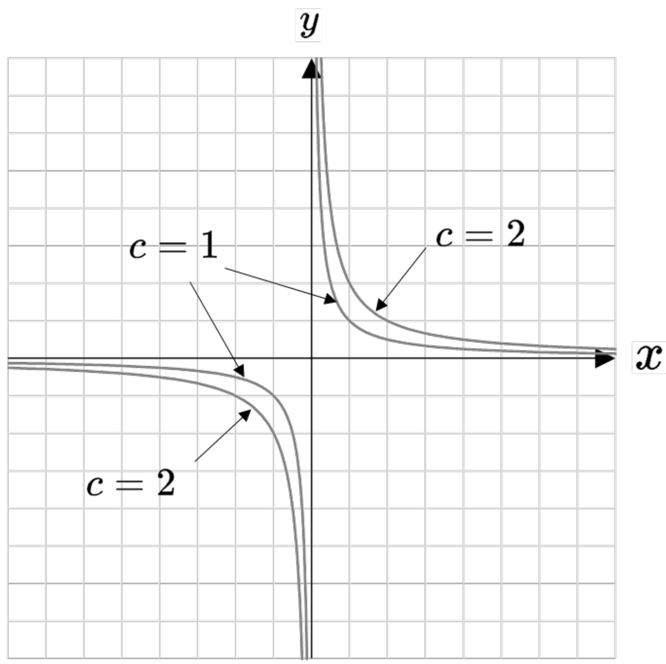
$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix};\; (x - x_0)^2 + (y - y_0)^2 = r^2 \right\}$$

(3) Open and closed disc!

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix};\; (x - x_0)^2 + (y - y_0)^2 \underset{(\leq)}{<} r^2 \right\}$$

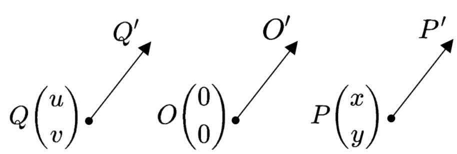
(4) Family of equilateral hyperbolas

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix};\; x\,y = c \right\}, \quad c \in \,]0, \infty[\quad \text{family parameter}$$



The pairs of numbers $\begin{pmatrix} x \\ y \end{pmatrix}$ with $x, y \in \mathbb{R}$ form the set $\mathbb{R}^2$. Objects in $\mathbb{R}^2$ can also have a different importance than that of the coordinate pair of a point. Let $\vec{a} = \binom{a_1}{a_2} \in \mathbb{R}^2$, with that comes the self-mapping of the plane called

Translation by $\vec{a}$:



$$t_{\vec{a}} \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} x + a_1 \\ y + a_2 \end{pmatrix}$$

Successive execution of two translations

$$t_{\vec{a}}\left(t_{\vec{b}} \begin{pmatrix} x \\ y \end{pmatrix} \right) = t_{\vec{a}} \begin{pmatrix} x + b_1 \\ y + b_2 \end{pmatrix} = \begin{pmatrix} x + a_1 + b_1 \\ y + a_2 + b_2 \end{pmatrix} = t_{\vec{a}+\vec{b}} \begin{pmatrix} x \\ y \end{pmatrix}$$

$t_{\vec{a}} \circ t_{\vec{b}} = t_{\vec{a}+\vec{b}}$ where the sum is declared as $\begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}$

Stretching (compression) by a factor $\lambda > 0$:

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \lambda\, x \\ \lambda\, y \end{pmatrix} =: \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$ The special case $\lambda = 1$ returns the identity!

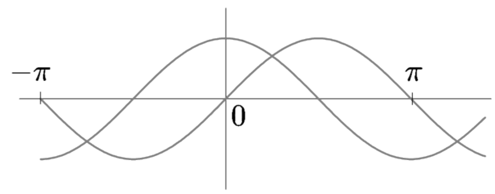
Point reflection at the origin:

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} -x \\ -y \end{pmatrix} \qquad \text{two-time application gives the identity.}$$

Summary of multiple formation:

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \lambda\, x \\ \lambda\, y \end{pmatrix} =: \lambda \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \lambda \in \mathbb{R}^* = \mathbb{R} \setminus \{0\}$$

8.1 Rotations Around the Origin



Let $c, s \in \mathbb{R}$ with $c^2 + s^2 = 1$; Then, exactly one $\varphi \in \,]-\pi, \pi]$ exists with $c = \cos\varphi$ and $s = \sin\varphi$. Namely, if $s = 0$, then $c = 1$, $\varphi = 0$ or $c = -1$, $\varphi = \pi$. Now let $s \neq 0$. Then $c \in \,]-1, 1]$. Because $\cos\varphi$ is an even function there exists a

$$\begin{array}{llll} \varphi_+ \in \,]0, \pi[& \text{with} & \cos\varphi_+ = c & \text{and a} \\ \varphi_- \in \,]-\pi, 0[& \text{with} & \cos\varphi_- = c & \text{namely } \varphi_- = -\varphi_+. \end{array}$$

Depending on the sign of s, $\varphi = \varphi_+$ or respectively $\varphi = \varphi_-$ is the only possibility. The mappings

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} c\, x - s\, y \\ s\, x + c\, y \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

are called rotations around the origin [rule for matrix multiplication: row times column]. Successive execution of two rotations! Let

$$\begin{pmatrix} \tilde{c} & -\tilde{s} \\ \tilde{s} & \tilde{c} \end{pmatrix} \quad \text{with} \quad \begin{array}{l} \tilde{c} = \cos\tilde{\varphi} \\ \tilde{s} = \sin\tilde{\varphi} \end{array}$$

be another rotation around the origin!

Calculation of $D(\tilde{\varphi})\, D(\varphi)$:

$$D(\varphi) D(\tilde{\varphi}) \begin{pmatrix} x \\ y \end{pmatrix} = D(\varphi) \left(D(\tilde{\varphi}) \begin{pmatrix} x \\ y \end{pmatrix} \right)$$

$$D(\varphi) \begin{pmatrix} \tilde{c}\, x - \tilde{s}\, y \\ \tilde{s}\, x + \tilde{c}\, y \end{pmatrix} = \begin{pmatrix} c(\tilde{c}\, x - \tilde{s}\, y) - s(\tilde{s}\, x + \tilde{c}\, y) \\ s(\tilde{c}\, x - \tilde{s}\, y) + c(\tilde{s}\, x + \tilde{c}\, y) \end{pmatrix}$$

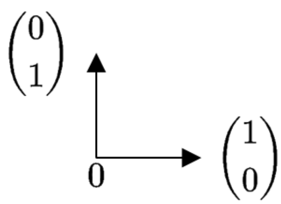
$$= \begin{pmatrix} (c\,\tilde{c} - s\,\tilde{s}) & -(c\,\tilde{s} + s\,\tilde{c}) \\ (c\,\tilde{s} + s\,\tilde{c}) & (c\,\tilde{c} - s\,\tilde{s}) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

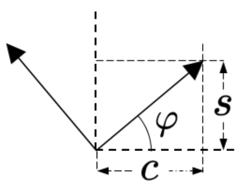
using the addition theorems:

$$= \begin{pmatrix} \cos(\varphi + \tilde{\varphi}) & -\sin(\varphi + \tilde{\varphi}) \\ \sin(\varphi + \tilde{\varphi}) & \cos(\varphi + \tilde{\varphi}) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Short: $D(\varphi) \circ D(\tilde{\varphi}) \;=\; D(\varphi + \tilde{\varphi})$

Application of the matrix notation to the special pairs $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ or respectively $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$





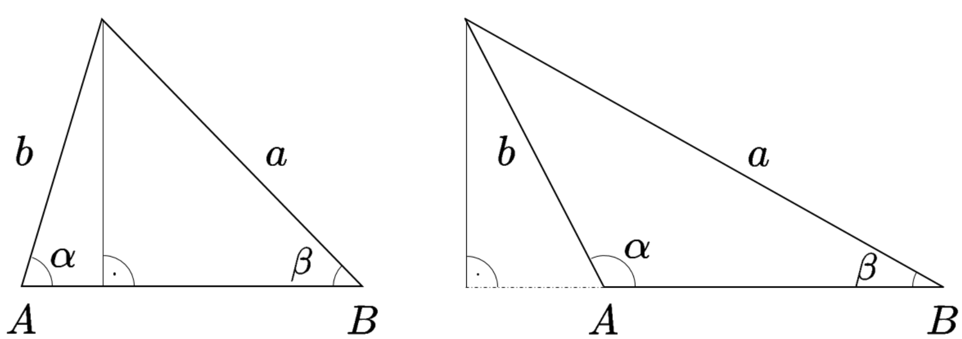
$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} cos\varphi \\ \sin\varphi \end{pmatrix}$$

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\varphi + \pi/2) \\ \sin(\varphi + \pi/2) \end{pmatrix}$$

Successive application of two rotations equals the rotation by the sum of the associated angles!

8.2 Basic Geometric Sine and Cosine Laws



$h \;=\; b\sin\alpha \;=\; a\sin\beta\;;$ Note: $\sin(\pi - \alpha) \;=\; \sin\alpha$

$$\frac{\sin\alpha}{a} \;=\; \frac{\sin\beta}{b} \;=\; \frac{\sin\gamma}{c}$$

Consider the partition of the base by the height

$$c \;=\; b\cos\alpha \;+\; a\cos\beta\;; \qquad \text{Note: } \cos\alpha < 0 \text{ for } \alpha > \pi/2$$

$$\begin{aligned} c^2 \;&=\; b^2(1-\sin^2\alpha) \;+\; a^2(1-\sin^2\beta) \;+\; 2ab\cos\alpha\cos\beta \\ &=\; a^2 \;+\; b^2 \;+\; 2ab(\cos\alpha\cos\beta - \sin\alpha\sin\beta) \end{aligned}$$

Because $\alpha + \beta + \gamma = \pi$

$$c^2 \;=\; a^2 \;+\; b^2 \;-\; 2ab\cos\gamma$$

8.3 The Scalar Product

is the inner product of two vectors $\vec{x}, \vec{y} \in \mathbb{R}^2$, $\vec{x} \;=\; \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$; $\vec{y} \;=\; \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$$\vec{x}\,\vec{y} \;:=\; x_1\,y_1 \;+\; x_2\,y_2$$

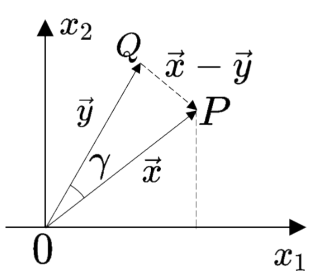
Through specialization one obtains the so-called Euclidean norm

$$\|\vec{x}\| \;:=\; \sqrt{\vec{x}\,\vec{x}} \qquad \text{norm of } \vec{x}$$

other common notations for the inner product:

$$\vec{x}\,\vec{y} \;=\; \langle\vec{x}\,\vec{y}\rangle \;=\; \langle\vec{x},\vec{y}\rangle$$

The scalar product (dot product) of two vectors is a number.



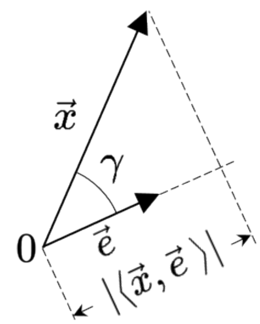
According to Pythagoras, the Euclidean norm is the distance of the point P with the position vector $\vec{x}$ from the origin 0.

To interpret the dot product, consider the norm square of the difference $\vec{x} - \vec{y}$

$$\begin{aligned} \|\vec{x}-\vec{y}\|^2 \;&=\; \langle\vec{x}-\vec{y},\vec{x}-\vec{y}\rangle \;=\; \langle\vec{x},\vec{x}-\vec{y}\rangle \;-\; \langle\vec{y},\vec{x}-\vec{y}\rangle \\ &=\; \langle\vec{x},\vec{x}\rangle \;+\; \langle\vec{y},\vec{y}\rangle \;-\; 2\langle\vec{x},\vec{y}\rangle \\ &=\; \|\vec{x}\|^2 \;+\; \|\vec{y}\|^2 \;-\; 2\langle\vec{x},\vec{y}\rangle \qquad \text{(law of cosines)} \\ &=\; \|\vec{x}\|^2 \;+\; \|\vec{y}\|^2 \;-\; 2\|\vec{x}\|\|\vec{y}\|\cos\gamma \end{aligned}$$

The dot (scalar) product of two vectors is equal to the product of their length times the cosine of the included angle!

$$\vec{x}\,\vec{y} \;=\; \|\vec{x}\|\|\vec{y}\|\cos\gamma$$



Sketch (right) for the case of a normalized

$$\vec{y} \;=\; \vec{e}$$

(that means: $\|\vec{y}\| \;=\; 1$)

To be highlighted:
The algebraic definition of the scalar product is simple in comparison to the formula $\langle\vec{x},\vec{y}\rangle \;=\; \|\vec{x}\|\|\vec{y}\|\cos\gamma$. In particular: $\vec{x}\perp\vec{y} \Leftrightarrow \langle\vec{x},\vec{y}\rangle \;=\; 0$.

Here it is specified that e.g. the zero vector $\begin{pmatrix}0\\0\end{pmatrix}$ is orthogonal to all vectors.

The principal properties of the norm!

(N1) $\|\vec{x}\| \;\geq\; 0 \;\forall\; \vec{x},$ "=" only for $\vec{x}=\vec{0}$

(N2) $\|\vec{x}+\vec{y}\| \;\leq\; \|\vec{x}\| + \|\vec{y}\|$ (triangle inequality)

(N3) $\|\lambda\cdot\vec{x}\| \;=\; |\lambda|\,\|\vec{x}\| \qquad \forall\; \vec{x},\vec{y}\in\mathbb{R}^2,\;\forall\;\lambda\in\mathbb{R}$

Proof of the triangle inequality!

LS and RS denote the left and the right side of the inequality (N2), respectively.

$$\begin{aligned}
&(RS)^2-(LS)^2\\
&\quad= x_1^2+x_2^2+y_1^2+y_2^2+2\|\vec{x}\|\|\vec{y}\|-(x_1+y_1)^2-(x_2+y_2)^2\\
&\quad= 2\|\vec{x}\|\|\vec{y}\|-2x_1y_1-2x_2y_2\\
&\quad= 2\|\vec{x}\|\|\vec{y}\|-2\langle\vec{x},\vec{y}\rangle\\
&\quad= 2\|\vec{x}\|\|\vec{y}\|(1-\cos\gamma)\;\geq\;0
\end{aligned}$$

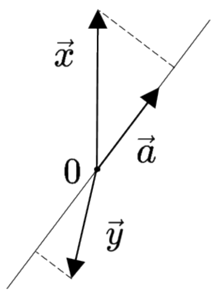
□

Addition: For vectors $\vec{x}\neq\vec{0}$, $\vec{y}\neq\vec{0}$,

$$\|\vec{x}+\vec{y}\| \;=\; \|\vec{x}\| + \|\vec{y}\|$$

applies if the enclosed angle $\gamma=0$.

Orthogonal projection of points onto a straight line through the origin



Let $\vec{a} \in \mathbb{R}^2$ be normalized, i.e. $\|\vec{a}\| = 1$

$$\{t\vec{a}\,;\ t \in \mathbb{R}\} \;=\; \mathbb{R}\,\vec{a}$$

describes the straight line through 0 in direction $\vec{a}$

$$p_{\vec{a}}(\vec{x}) \;:=\; \langle \vec{x}, \vec{a}\,\rangle\,\vec{a}$$

Justification:

The difference vector $\vec{x} - p_{\vec{a}}(\vec{x})$ is perpendicular to the vector $\vec{a}$:

$$\begin{aligned}\langle \vec{x} - p_{\vec{a}}(\vec{x})\,,\,\vec{a}\,\rangle &= \langle \vec{x} - \langle \vec{x}, \vec{a}\,\rangle\,\vec{a}\,,\,\vec{a}\,\rangle \\ &= \langle \vec{x}, \vec{a}\,\rangle - \langle \langle \vec{x}, \vec{a}\,\rangle\,\vec{a}\,,\,\vec{a}\,\rangle \\ &= \langle \vec{x}, \vec{a}\,\rangle - \langle \vec{x}, \vec{a}\,\rangle\,\underbrace{\langle \vec{a}, \vec{a}\,\rangle}_{=1} = 0\end{aligned}$$

Fundamental calculation rules for the scalar (dot) product

(SK1) $\langle \vec{x}, \vec{y}\,\rangle = \langle \vec{y}, \vec{x}\,\rangle$ (symmetry)

(SK2) $\langle \vec{x}, \vec{y} + \vec{z}\,\rangle = \langle \vec{x}, \vec{y}\,\rangle + \langle \vec{x}, \vec{z}\,\rangle$

(SK3) $\langle \vec{x}, \lambda\,\vec{y}\,\rangle = \lambda\langle \vec{x}, \vec{y}\,\rangle \qquad \forall\ \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^2,\ \forall\ \lambda \in \mathbb{R}$

From (SK2) and (SK3) follows linearity in the second argument. The linearity in the first argument follows from (SK1). (SK4) is a repetition of (N1):

(SK4) $\langle \vec{x}, \vec{x}\,\rangle \geq 0$, "=" only for $\vec{x} = \vec{0}$

Representation of the orthogonal projections onto a straight line $\vec{a}\,\mathbb{R}$ ($\|\vec{a} = 1\|$) with a matrix

$$\begin{aligned}p_{\vec{a}}(\vec{x}) = \langle \vec{x}, \vec{a}\,\rangle\vec{a} &= (x_1a_1 + x_2a_2)\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\ &= \begin{pmatrix} a_1^2 & a_1a_2 \\ a_1a_2 & a_2^2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\end{aligned}$$

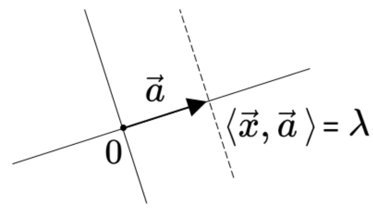
$p_{\vec{a}} \cdot p_{\vec{a}} = p_{\vec{a}}$ Projections are idempotent.

8.4 Hessian Normal Form of a Straight Line

Sei $\quad \vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2$

$\{\vec{x}\,;\ \langle \vec{x}, \vec{a}\,\rangle = 0\} \ = \ \vec{a}^{\perp}\ ;\ \|\vec{a}\,\| \ = \ 1$ consists of all vectors orthogonal to $\vec{a}$, i.e. a straight line through 0:

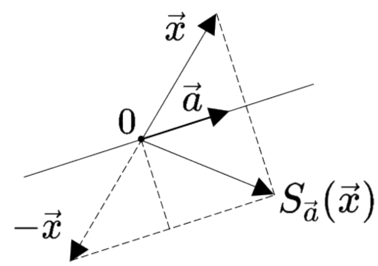
$$\{\vec{x}\,;\ \langle \vec{x}, \vec{a}\,\rangle = \lambda\}\,, \quad \lambda \in \mathbb{R}$$



contains the vector $\lambda\,\vec{a}$

$$x_1 a_1 \, + \, x_2 a_2 \ = \ \lambda$$

Point reflection across a straight line $\mathbb{R}\,\vec{a}$



Let $\vec{a} \in \mathbb{R}^2$ be normalized.

$$S_{\vec{a}}(\vec{x}) \ = \ -\vec{x} \, + \, 2p_{\vec{a}}(\vec{x})$$

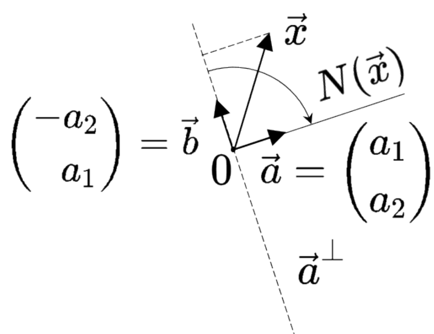
is called point reflection across a straight line $\mathbb{R}\,\vec{a}$

Matrix notation for $\ \vec{x} \longmapsto -\vec{x} \quad$ by $\quad \vec{x} \ \longmapsto \ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

With that

$$S_{\vec{a}}(\vec{x}) \ = \ \begin{pmatrix} -1 + 2a_1^2 & 2a_1 a_2 \\ 2a_1 a_2 & -1 + 2a_2^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad S_{\vec{a}} \circ S_{\vec{a}} \ = \ \text{Identity}$$

Example of a nilpotent linear mapping!



$\vec{a} \ = \ \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2$ be normalized.

Then $\ \vec{b} \ = \ \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix}$ also normalized and$\vec{b} \in \vec{a}^{\perp}$

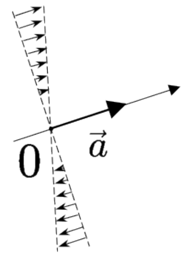
$$N(\vec{x}) \ = \ \langle \vec{x}, \vec{b}\,\rangle\,\vec{a}$$

is called a nilpotent linear mapping.

$$N(N(\vec{x})) \ = \ N(\langle \vec{x}, \vec{b}\,\rangle\,\vec{a}) \ = \ \langle \langle \vec{x}, \vec{b}\,\rangle\,\vec{a}\,,\,\vec{b}\,\rangle\,\vec{a} \ = \ \langle \vec{x}, \vec{b}\,\rangle\,\underbrace{\langle \vec{a}, \vec{b}\,\rangle}_{=\,0} \ = \ 0$$

$N \circ N$ is the zero operator!

Shear in $\mathbb{R}\,\vec{a}$



Given $\mu \in \mathbb{R}$

$$\vec{x} \;\longmapsto\; \underbrace{\vec{x} + \mu\, N(\vec{x})}_{S(\vec{x})}$$

Matrix notation for N

$$N(\vec{x}) \;=\; (-a_2 x_1 + a_1 x_2)\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \;=\; \begin{pmatrix} a_1 a_2 & a_1^2 \\ -a_2^2 & a_1 a_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Matrix notation for shear mapping

$$S(\vec{x}) \;=\; \begin{pmatrix} 1 - a_1 a_2\,\mu & a_1^2\,\mu \\ -a_2^2\,\mu & 1 + a_1 a_2\,\mu \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Compilation of uniform points of view!

$$\mathbb{R}^2 \ni \vec{x} \;=\; \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$\mathbb{R}^2$ is an example of a real “vector space”.

$$\vec{x} + \vec{y} \;=\; \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \qquad \text{addition}$$

$$\lambda\,\vec{x} \;=\; \begin{pmatrix} \lambda\, x_1 \\ \lambda\, x_2 \end{pmatrix} \qquad \text{multiplication with scalar } \lambda \in \mathbb{R}$$

The arithmetic laws in $\mathbb{R}$ result in rules in this vector space, for example:

$$(\vec{x} + \vec{y}) + \vec{z} \;=\; \vec{x} + (\vec{y} + \vec{z})$$

$$\vec{x} + \vec{y} \;=\; \vec{y} + \vec{x}$$

$$(\lambda\,\mu)\,\vec{x} \;=\; \lambda\,(\mu\,\vec{x})$$

$$1 \cdot \vec{x} \;=\; \vec{x} \qquad \forall\ \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^2,\ \ \lambda, \mu \in \mathbb{R}$$

Apart from the translations, the considered self-mappings $F : \mathbb{R}^2 \longmapsto \mathbb{R}^2$ have in common the properties of linearity. i.e.

$$F(\vec{x} + \vec{y}) \;=\; F(\vec{x}) + F(\vec{y})$$

$$F(\lambda\,\vec{x}) \;=\; \lambda\, F(\vec{x}) \qquad \forall\ \vec{x}, \vec{y} \in \mathbb{R}^2 \quad \forall\ \lambda \in \mathbb{R}$$

Such F map e.g. the straight lines

$$g \;=\; \vec{a} + \mathbb{R}\,\vec{b}\;; \qquad \vec{b} \neq \vec{0}$$

in the following manner:

$$F(g) \;=\; F(\vec{a} + \mathbb{R}\,\vec{b}) \;=\; F(\vec{a}) \,+\, \mathbb{R}\,F(\vec{b})$$

This is either a point (if $F(\vec{b}) = 0$) or a straight line. Furthermore, each such F is already completely determined by the images $F\left(\begin{pmatrix}1\\0\end{pmatrix}\right)$ and $F\left(\begin{pmatrix}0\\1\end{pmatrix}\right)$.

Let $\quad F\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}a\\c\end{pmatrix} \quad$ and $\quad F\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}b\\d\end{pmatrix}$

By writing these two image vectors next to each other, the matrix describing F is created

$$M \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$F(\vec{x}) \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;=\; \begin{pmatrix} a\,x_1 + b\,x_2 \\ c\,x_1 + d\,x_2 \end{pmatrix} \;=\; M\vec{x}$$

Now let $\widetilde{F}$ also be a linear self-mapping of $\mathbb{R}^2$, described by the matrix

$$\widetilde{M} \;=\; \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix}$$

<u>Questions:</u> a) Is the composite $\widetilde{F} \circ F$ also linear?
b) If so, which matrix does $\widetilde{F} \circ F$ describe?

$$\widetilde{F}(F(\vec{x} + \vec{y})) \underset{\substack{\uparrow \\ F \text{ additive}}}{=} \widetilde{F}(F(\vec{x}) + F(\vec{y})) \underset{\substack{\uparrow \\ \widetilde{F} \text{ additive}}}{=} \widetilde{F}(F(\vec{x})) \,+\, \widetilde{F}(F(\vec{y}))$$

that means

$$\widetilde{F} \circ F(\vec{x} + \vec{y}) \;=\; \widetilde{F} \circ F(\vec{x}) \;+\; \widetilde{F} \circ F(\vec{y})$$

$$\widetilde{F} \circ F(\lambda\,\vec{x}) \;=\; \widetilde{F}(F(\lambda\,\vec{x})) \underset{\substack{\uparrow \\ \text{Linearity} \\ \text{of } F}}{=} \widetilde{F}(\lambda\,F(\vec{x})) \underset{\substack{\uparrow \\ \text{Linearity} \\ \text{of } \widetilde{F}}}{=} \lambda\,\widetilde{F} \circ F(\vec{x})$$

The matrix for $\widetilde{F} \circ F$ has the first column

$$\widetilde{F} \circ F\left(\begin{pmatrix}1\\0\end{pmatrix}\right) \;=\; \widetilde{F}\begin{pmatrix}a\\c\end{pmatrix} \;=\; \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix}\begin{pmatrix}a\\c\end{pmatrix} \;=\; \begin{pmatrix} \tilde{a}\,a + \tilde{b}\,c \\ \tilde{c}\,a + \tilde{d}\,c \end{pmatrix}$$

and the second column

$$\widetilde{F} \circ F\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \;=\; \widetilde{F}\begin{pmatrix} b \\ d \end{pmatrix} \;=\; \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix}\begin{pmatrix} b \\ d \end{pmatrix} \;=\; \begin{pmatrix} \tilde{a}\, b + \tilde{b}\, d \\ \tilde{c}\, b + \tilde{d}\, d \end{pmatrix}$$

The resulting matrix for the compound $\widetilde{F} \circ F$ is

$$\widetilde{M}\, M \;=\; \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \;=\; \begin{pmatrix} \tilde{a}\, a + \tilde{b}\, c & \tilde{a}\, b + \tilde{b}\, d \\ \tilde{c}\, a + \tilde{d}\, c & \tilde{c}\, b + \tilde{d}\, d \end{pmatrix}$$

the so-called matrix product of $\widetilde{M}$ and M!

Final question: Which linear self-mappings are length-preserving?

$$(*) \qquad \|M\,\vec{x}\,\| \;=\; \|\vec{x}\,\| \qquad \forall\ \vec{x} \in \mathbb{R}^2 \qquad \text{(isometry condition)}$$

$$\text{Let} \quad M \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad ; \quad \vec{x} \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Three conditions:

$$\sqrt{a^2 + c^2} \;=\; 1 \quad , \quad \sqrt{b^2 + d^2} \;=\; 1 \quad , \quad \sqrt{(a+b)^2 + (c+d)^2} \;=\; \sqrt{2}$$

that means

$$a^2 + c^2 \;=\; 1 \quad , \quad b^2 + d^2 \;=\; 1 \quad , \quad 2\,(a\,b + c\,d) \;=\; 0$$

that means

$$\begin{pmatrix} a \\ c \end{pmatrix} \text{ and } \begin{pmatrix} b \\ d \end{pmatrix} \text{ are normalized to the length } 1\,; \quad \text{furthermore } \begin{pmatrix} b \\ d \end{pmatrix} \in \begin{pmatrix} a \\ c \end{pmatrix}^{\perp}$$

For a given normalized vector $\begin{pmatrix} a \\ c \end{pmatrix}$ there are exactly two possible second columns $\begin{pmatrix} b \\ d \end{pmatrix}$ namely $\begin{pmatrix} -c \\ a \end{pmatrix}$ or $\begin{pmatrix} c \\ -a \end{pmatrix}$:

$$M \;=\; \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \quad \text{(rotation)} \quad \text{or}$$

$$M \;=\; \begin{pmatrix} a & c \\ c & -a \end{pmatrix} \quad \text{(reflection across a straight line)}$$

$$M \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;,\quad a^2 + c^2 \;=\; b^2 + d^2 \;=\; 1 \;, \quad a\,b + c\,d \;=\; 0$$

$$\begin{aligned}
\langle M\,\vec{x}, M\,\vec{y}\rangle \;&=\; \left\langle \begin{pmatrix} a\,x_1 + b\,x_2 \\ c\,x_1 + d\,x_2 \end{pmatrix}, \begin{pmatrix} a\,y_1 + b\,y_2 \\ c\,y_1 + d\,y_2 \end{pmatrix} \right\rangle \\
&=\; (\underbrace{a^2 + c^2}_{=1})x_1y_1 \;+\; (\underbrace{b^2 + d^2}_{=1})x_2y_2 \;+\; (\underbrace{a\,b + c\,d}_{=0})(x_1y_2 + x_2y_1) \\
&=\; \langle \vec{x}, \vec{y}\,\rangle \;, \quad \text{for all} \quad \vec{x}, \vec{y} \in \mathbb{R}^2
\end{aligned}$$

Going further than $(*)$:

$$\langle M\,\vec{x}, M\,\vec{y}\,\rangle \;=\; \langle \vec{x}, \vec{y}\,\rangle \qquad \forall\; \vec{x}, \vec{y} \in \mathbb{R}^2$$

for the found matrices M.

Finally:

Each translation $t_{\vec{a}}\,,\;\vec{a} \in \mathbb{R}^2$ yields a length-preserving (but in the current sense non-linear) mapping

$$\begin{aligned}
\|t_{\vec{a}}(\vec{x}) - t_{\vec{a}}(\vec{y})\| \;&=\; \|(\vec{x} + \vec{a}) - (\vec{y} + \vec{a})\| \qquad \text{(by definition of } t_{\vec{a}}) \\
&=\; \|\vec{x} - \vec{y}\| \qquad \forall\; \vec{x}, \vec{y} \in \mathbb{R}^2
\end{aligned}$$

9. Complex Numbers

Choosing an origin and a pair of coordinate axes makes it possible to link points via (vectorial) addition. In addition, this also provides for multiplication with scalars $\lambda \in \mathbb{R}$. Significance of $\mathbb{C}$: In reality, this scalar multiplication is only the restriction to the real axis (x-axis) of a more general multiplication with any $\vec{a}$ of the plane! The geometric relevance of it being that of a similarity transformation called Drehstreckung (rotational stretch), whose angle of rotation is the angle of inclination of $\vec{a}$ against the x-axis, and whose stretching factor is $\|\vec{a}\|$! In this context, a $\begin{pmatrix} x \\ 0 \end{pmatrix}$ is identified with the real number $x \in \mathbb{R}$ which describes it. For the base vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ of the y axis, the character j (very often i) is chosen. Then the points of the plane are

$$z = x + jy \qquad x, y \in \mathbb{R}$$

If one continues to calculate distributively and 1 remains neutral for multiplications, then only $j \cdot j$ needs to be specified for fixing the multiplications. The geometric interpretation forces $j^2 = -1$. In this way the field of the complex numbers is created. The 9 field axioms are e.g. satisfied by $\mathbb{Q}$ or by $\mathbb{R}$! There is even a two-element model $\mathbb{F}_2 = \{O, L\}$ with the operations

$$O + O = O = L + L \quad , \quad O + L = L$$

$$O \cdot O = O = O \cdot L \quad , \quad L \cdot L = 1$$

Examine to which extent the 9 field axioms hold for the set $M2(\mathbb{R})$ of real 2×2 matrices.

Addition

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} a + a_1 & b + b_1 \\ c + c_1 & d + d_1 \end{pmatrix}$$

Then the axioms of addition apply with $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ as zero.

The interpretation of the matrices as linear transformations shows the validity of the distributive law in $M_2(\mathbb{R})$. As linear transformations the matrices deliver the associative law of multiplication.

Let F_1, F_2, F_3 be linear transformations of the plane.

Proposition: $\quad (F_1 \circ F_2) \circ F_3 \;=\; F_1 \circ (F_2 \circ F_3)$

This is an equation between transformations!

$$(F_1 \circ F_2) \circ F_3(\vec{x}) \;=\; (F_1 \circ F_2)(F_3(\vec{x})) \;=\; F_1(F_2(F_3(\vec{x})))$$

$$F_1 \circ (F_2 \circ F_3)(\vec{x}) \;=\; F_1((F_2 \circ F_3)(\vec{x})) \;=\; F_1(F_2(F_3(\vec{x})))$$

The identity element in $M_2(\mathbb{R})$ is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

> The product $M_1 M_2$ only sometimes agrees with $M_2 M_1$ (one then says M_1 and M_2 commute).

$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible only if the determinant $\det M = a\,d - b\,c \neq 0$. This is demonstrated by the multiplication theorem for determinants!

$$\det(M_2 M_1) \;=\; (\det M_1)\,(\det M_2) \qquad (\text{LS/RS} \;=\; \text{left/right side})$$

$$\text{LS: } (a_1a_2 + b_1c_2)(c_1b_2 + d_1d_2) \;-\; (a_1b_2 + b_1d_2)(c_1a_2 + d_1c_2)$$

$$\text{RS: } (a_1d_1 - b_1c_1)(a_2d_2 - b_2c_2)$$

□

If $\det M \;=\; a\,d - b\,c \neq 0$ then

$$M^{-1} \;=\; \frac{1}{a\,d - b\,c} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

<u>Summary:</u>
In $M_2(\mathbb{R})$ the first and the fourth axiom are only conditionally valid, the other axioms apply without restrictions!

Matrix representation of $\mathbb{C}$

$$z \;=\; x + jy \;\longmapsto\; D(z) \;=\; \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

For the addition operation the following applies

$$D(z_1) + D(z_2) \;=\; D(z_1 + z_2) \qquad \forall\; z_1, z_2 \in \mathbb{C}$$

The matrix representation of the multiplication of two complex numbers z_1, z_2

$$(x_1 + jy_1)(x_2 + jy_2) \;=\; x_1x_2 - y_1y_2 + j(x_1y_2 + y_1x_2)$$

is given by

$$\begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix}\begin{pmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{pmatrix} \;=\; \begin{pmatrix} x_1x_2 - y_1y_2 & -(x_1y_2 + y_1x_2) \\ x_1y_2 + y_1x_2 & x_1x_2 - y_1y_2 \end{pmatrix}$$

$$D(z_1) \cdot D(z_2) \;=\; D(z_1 \cdot z_2) \qquad \forall\; z_1, z_2 \in \mathbb{C}$$

Part of the significance of D is that it transports the addition and multiplication of $\mathbb{C}$ into $M_2(\mathbb{R})$ (and back).

In particular, the associative property of multiplication and the distributive law apply in $\mathbb{C}$, because they apply in $M_2(\mathbb{R})$.

$$\det D(z) \;=\; x^2 + y^2 \;=\; \|z\|^2 \;=\; |z|^2$$

The determinant of $D(z)$ is the square of the Euclidean norm of z, i.e. the absolute value of $|z|$.

The determinant multiplication theorem:

$$|z_1\, z_2| \;=\; |z_1|\,|z_2| \qquad \forall\; z_1, z_2 \in \mathbb{C}$$

The interpretation via the Euclidean norm provides the triangle inequality:

$$|z_1 + z_2| \;\leq\; |z_1| + |z_2|$$

For all $z \neq 0$, $D(z)$ is invertible. Given that, z is then invertible with the inverse

$$\frac{1}{z} \;=\; z^{-1} \;=\; \frac{1}{x^2 + x^2}\,(x - jy)$$

$$D(z) = |z|\begin{pmatrix} c & -s \\ s & c \end{pmatrix} \quad \text{with} \quad c^2+s^2=1$$

$$= |z|\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \quad \text{with} \quad \varphi \in \,]-\pi,\pi[$$

This form reveals the underlying geometric interpretation of multiplication as a Drehstreckung (rotational stretching). Moreover, from the interchangeability of each two rotation matrices

$$\begin{pmatrix} c_1 & -s_1 \\ s_1 & c_1 \end{pmatrix} \quad , \quad \begin{pmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{pmatrix}$$

follows the commutative property of the multiplication in $\mathbb{C}$.

The open disc

$$B_\epsilon(a) = \{z \in \mathbb{C} \,/\, |z-a| < \epsilon\}$$

serves as ϵ-neighborhood for $a \in \mathbb{C}$, $(\epsilon > 0)$!

With that, one can examine the diverse functions of $\mathbb{R}$ in $\mathbb{C}$, of $\mathbb{C}$ in $\mathbb{R}$ and of $\mathbb{C}$ in $\mathbb{C}$, for example for continuity!

$$f : X \to Y$$

is called continuous in $a \in X$ if for every (however small) neighborhood V of the pixel $b = f(a)$ a neighborhood U of a exists with $f(U) \subset V$.

Examples:

(1) The functions real part and imaginary part

$$\begin{array}{c|c} \mathrm{Re} : \mathbb{C} \to \mathbb{R} & \mathrm{Im} : \mathbb{C} \to \mathbb{R} \\ x + jy \longmapsto x & x + jy \longmapsto y \end{array}$$

are $\mathbb{R}$-linear and continuous

$$z_0 = x_0 + jy_0$$

$$|\mathrm{Re}(z) - \mathrm{Re}(z_0)| \le |z - z_0| \quad ; \quad |\mathrm{Im}(z) - \mathrm{Im}(z_0)| \le |z - z_0|$$

(2) The conjugation (conjugate formation) in $\mathbb{C}$

$$z = x + jy \longmapsto \bar{z} = x - jy$$

provides an involutive identity mapping of $\mathbb{C}$, interchangeable with addition and also with multiplication. In particular, it is continuous.

Proof:

$$\begin{aligned}
\overline{z_1 + z_2} &= \overline{(x_1 + x_2) + j(y_1 + y_2)} = x_1 + x_2 - j(y_1 + y_2) \\
&= (x_1 - jy_1) + (x_2 - jy_2) = \bar{z}_1 + \bar{z}_2 \\
\overline{z_1 \cdot z_2} &= \overline{x_1 x_2 - y_1 y_2 + j(x_1 y_2 + y_1 x_2)} \\
&= x_1 x_2 - y_1 y_2 + j(x_1(-y_2) - y_1 x_2) \\
&= (x_1 - jy_1)(x_2 - jy_2) = \bar{z}_1 \bar{z}_2
\end{aligned}$$

□

(3) Addition and multiplication on $\mathbb{C}$ are continuous.

In both cases it has to be shown: the result of addition and multiplication changes arbitrarily little if the arguments are changed only sufficiently little.

Proof:

$$\begin{aligned}
|(a+b) - (a_0 + b_0)| &= |(a - a_0) + (b - b_0)| \quad / \text{ triangle inequality} \\
&\leq |a - a_0| + |b - b_0|
\end{aligned}$$

Now let $\quad M \geq |a_0| + |b_0| + 1$

$$\begin{aligned}
|ab - a_0 b_0| &= |(ab - a_0 b) + (a_0 b - a_0 b_0))| \\
&\leq |a - a_0|\,|b| + |a_0|\,|b - b_0| \\
&\leq M\,(|a - a_0| + |b - b_0|)
\end{aligned}$$

if $\quad |b - b_0| \leq 1$

□

<u>Remark:</u>

According to example (3), each polynomial function becomes (through repeated application)

$$z \longmapsto \sum_{k=0}^{n} a_k z^{n-k}$$

continuous (coefficients a_k real or complex).

9.1 The Cauchy criterion applies in $\mathbb{C}$

Every Cauchy sequence $(a_n)_{n\geq 0}$ of complex numbers has a limit value in $\mathbb{C}$. (Cauchy sequence: For every $\epsilon > 0$ exists an index N_ϵ such that for all indices $k, m \geq N_\epsilon$ $|a_k - a_m| < \epsilon$ applies.)

Proof:

The two real sequences $(\operatorname{Re} a_n)_{n\geq 0}$, $(\operatorname{Im} a_n)_{n\ \ geq\,0}$ are also Cauchy sequences with $(a_n)_{n\geq 0}$, but real ones, and therefore do converge to $\alpha, \beta \in \mathbb{R}$.

$$a_n \;=\; \operatorname{Re} a_n + j \operatorname{Im} a_n \;=\; \alpha + j\beta$$

Example:

The exponential series in the complex

$$\exp(z) \;=\; \sum_{n=0}^{\infty} \frac{z^n}{n!} \;; \qquad z \in \mathbb{C}$$

According to the discussion of the power series, the partial sums of this series form a Cauchy sequence for each $z \in \mathbb{C}$.

According to the Cauchy criterion, the sequence of partial sums converges, the sign for the limit value is again $\exp(z)$.

And according to the theorem about the differentiation of power series in section 7.2 this series defines a function $y(z)$, for which the limit value

$$\lim_{\substack{h \to 0 \\ h \in \mathbb{C}^x}} \frac{y(z+h) - y(z)}{h} \quad \text{exists,}$$

the so-called complex derivative. According to the theorem, its value is the value of the term by term differentiated series!

$\exp(z)$ is therefore a $\mathbb{C}$ "holomorphic" solution of the differential equation

$$(*) \quad y' \;=\; y \quad \text{with initial value} \quad y(0) \;=\; 1$$

Note:

$$j^2 = -1 \quad ; \quad j^3 = -j \quad ; \quad j^4 = 1 \quad ; \quad \ldots\ldots$$

results in the decomposition of the restriction of exp onto the imaginary axis.

$$\exp(jt) \;=\; \sum_{k=0}^{\infty}(-1)^k\frac{t^{2k}}{(2k)!} + j\sum_{k=0}^{\infty}(-1)^k\frac{t^{2k+1}}{(2k+1)!}$$

$$\boxed{\exp(jt) \;=\; \cos t + j\sin t} \qquad \text{Euler's formula}$$

For $\mathbb{C}$-differentiability, the product rule and the chain rule apply again!

Now let $f(z)$ be any solution of $(*)$

Consider:

$$g(z) \;=\; \exp(-z)\,f(z)$$

$$g'(z) \;=\; \exp(-z)\,(-f(z) + f'(z)) \;=\; 0 \quad \forall\ z \in \mathbb{C}$$

From this follows: g is constant.

Application example: $\exp(z+w) \;=\; \exp(z) \;+\; \exp(w)$

9.1.1 Moivre's Theorem

For all natural n and all real φ:

$$(\cos\varphi + j\sin\varphi)^n \;=\; \cos(n\varphi) + j\sin(n\varphi)$$

Proof either from the addition theorem of the exponential function in the complex or from the addition theorems for cos and sin!

The equation $z^n = 1$ therefore has the solutions

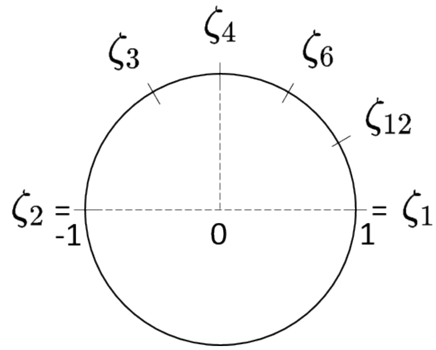
$$\zeta_n^k \;=\; \exp\left(\frac{2\pi j}{n}\,k\right) \qquad 1 \le k \le n$$

These are the so-called n-th roots of unity. Notable: Out of

$$\zeta_n \;=\; \exp\left(\frac{2\pi j}{n}\right)$$

the remaining n-th roots of unity arise from exponentiation.

$$\begin{aligned}
\zeta_2 &= -1 \\
\zeta_3 &= -\frac{1}{2} + j\,\frac{\sqrt{3}}{2} \\
\zeta_4 &= j \\
\zeta_6 &= \frac{1}{2} + j\,\frac{\sqrt{3}}{2}
\end{aligned}$$



The geometric meaning of multiplication shows directly that every complex w is a square:

$$w = |w|\,\mathrm{e}^{j\psi}$$

Therefore w is the square of $|w|^{1/2}\,\mathrm{e}^{j\psi/2}$.

Hence, every quadratic equation in $\mathbb{C}$ can be solved via the approach of completing the square.

$$z^2 + pz + q = \left(z + \frac{p}{2}\right) - \frac{1}{4}(p^2 - 4q)$$

9.2 The Fundamental Theorem of Algebra in $\mathbb{C}$

Any non-constant polynomial

$$P(z) = a_0 z^n + a_1 z^{n-1} + \ldots = \sum_{k=0}^{n} a_k z^{n-k}$$

has at least one root in $\mathbb{C}$!

<u>Conclusions from the fundamental theorem</u>

The factorization of the polynomials:

Every non-constant polynomial $P(z)$ with real or complex coefficients of degree $n > 0$ has exactly one decomposition

$$P(z) = a(z - z_1)^{r_1}(z - z_2)^{r_2} \cdot \ \ldots \ \cdot (z - z_p)^{r_p}$$

where a is the leading coefficient of P, $z_1, z_2, \ldots, z_p$ the different zeros in $\mathbb{C}$ of P and the exponents $r_1, r_2, \ldots, r_p \in \mathbb{N}$ with the sum $r_1 + r_2 + \ldots + r_p = n$.

Supplement: If P is a polynomial with real coefficients, then its factorization includes with every non-real zero z_k also its conjugate $\bar{z}_k$ as a zero; furthermore $(z - z_k)$ and $(z - \bar{z}_k)$ have the same exponent.

To the **proof** of the conclusion:

Repeated application of the fundamental theorem yields one of the given factorizations for P.

The uniqueness of the factorization is based on the observation that $\{z_1, z_2, \ldots, z_p\}$ is nothing other than the zero set of P.

For the supplement one considers $z = x$ real

$$P(x) \;=\; a(x - z_1)^{r_1}(x - z_2)^{r_2} \cdot \;\ldots\; \cdot (x - z_p)^{r_p}$$

Application of conjugation

$$\overline{P(x)} \;=\; P(x) \;=\; \bar{a}(x - \bar{z}_1)^{r_1}(x - \bar{z}_2)^{r_2} \cdot \;\ldots\; \cdot (x - \bar{z}_p)^{r_p}$$

The uniqueness of the factorization yields the assertion!

□

To the **proof** of the fundamental theorem:

Lemma 1:

Let

$$P(z) \;=\; \sum_{k=0}^{n} a_k z^{n-k}$$

be a non-constant polynomial. Furthermore, let $z_1 \in \mathbb{C}$, $P(z_1) \neq 0$. Then in every neighborhood of z_1 there exist numbers z_0 and z_2 with $|P(z_0)| < |P(z_1)| < |P(z_2)|$.

Lemma 2:

For every polynomial P over the complex numbers there is a position z_1 with $|P(z_1)| \leq |P(z)|$ for all $z \in \mathbb{C}$.

Together:

The z_1 of the second lemma cannot be the z_1 of the first lemma, hence for this z_1 the premise $P(z_1) \neq 0$ is not correct!

□

Proof of lemma 1:

With no restriction: $z_1 = 0$. Otherwise consider $Q(z) = P(z - z_1)$. Furthermore without restriction $P(0) = 1$. Then consider a polynomial

$$P(z) \;=\; 1 + a\, z^m (1 + R(z))$$

where $a \neq 0$, $m > 0$ and the remainder polynomial $R(z)$ vanishes at 0. In particular, there exists a $\varrho > 0$ with

$$|R(z)| \;\le\; \frac{1}{2} \qquad \text{if} \qquad |z| \;\le\; \varrho$$

choose $\varphi \in \mathbb{R}$ with $a = |a|\, e^{j\varphi}$

$z_2 = r\, e^{-j\varphi/m} \quad r \in \,]0, \varrho]\;,$ then

$$\big|P(z_2)\big| \;=\; \big|1 + |a|\, r^m + |a|\, r^m R(z_2)\big|$$

and with the inverse triangle inequality follows

$$\big|P(z_2)\big| \;\ge\; 1 + |a|\, r^m - |a|\, r^m \frac{1}{2} \;>\; 1$$

$z_0 = r\, e^{-j(\varphi+\pi)/m}\;,$ then

$$\big|P(z_0)\big| \;=\; \big|\,1 + |a|\,(-1)\, r^m - |a|\, r^m R(z_0)\,\big|$$

and with the ordinary triangle inequality follows

$$\Big|P(z_0)\Big| \;\ge\; 1 - |a| \cdot r^m + |a| \cdot r^m \cdot \frac{1}{2} \;>\; 1 \quad ,$$

if additionally r is so small that $|a|\, r^m < 1$. This is possible.

□

Proof of lemma 2:

The (real) function $z \longmapsto |P(z)|$ has an absolute minimum on $\mathbb{C}$. As for proving this for non-constant P:

1) $|P(z)|$ increases for $|z| \to \infty$ approximately as the absolute value of the highest term in P.

2) Every continuous function on a closed disc has a minimum in absolute value. Carrying out the proof will become routine later!

9.3 Partial Fraction Decomposition

For polynomials Q_0, Q_1 with coefficients in a field (e.g. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, etc.) repeated division with remainder can be performed in the case $Q_1 \neq 0$. Because of the constantly decreasing degrees of the remainder polynomials, eventually a remainder $Q_{n+1} = 0$ occurs.

$$\begin{array}{rcl@{\qquad\qquad}l}
Q_0 & = & Q_1 \cdot q_1 + Q_2 & \deg Q_2 < \deg Q_1 \\
Q_1 & = & Q_2 \cdot q_2 + Q_3 & \deg Q_3 < \deg Q_2 \\
\vdots & & \vdots & \qquad \vdots \\
Q_{n-2} & = & Q_{n-1} \cdot q_{n-1} + Q_n & \deg Q_n < \deg Q_{n-1} \\
Q_{n-1} & = & Q_n \cdot q_n + 0 &
\end{array}$$

$$Q_n \;=\; \text{gcd} \;\text{ of }\; Q_0 \;\text{ and }\; Q_0$$

The immediately preceding remainder Q_n is (with respect to degree) the greatest common divisor (gcd) of Q_0 and Q_1. Backward substitution results in any case in

$$Q_n \;=\; Q_0\widetilde{Q}_0 + Q_1\widetilde{Q}_1$$

In particular, if Q_0, Q_1 are relatively prime, then Q_n is constant $\neq 0$, i.e

$$1 \;=\; Q_0\widetilde{Q}_0 + Q_1\widetilde{Q}_1 \qquad (*)$$

For fractions P/Q with $Q = Q_0Q_1$ this results in a partial fraction decomposition:

$$\frac{P}{Q} \;=\; \frac{P_0}{Q_0} + \frac{P_1}{Q_1} \qquad \text{(multiply } (*) \text{ with } P)$$

The numerators P_0, P_1 are not yet unique. The process can be repeated for the summands if the denominator can be factored as a product of relatively prime polynomials! Result:

$$\frac{P}{Q} \;=\; p + \frac{p_1}{q_1} + \frac{p_2}{q_2} + \cdots + \frac{p_r}{q_r}$$

where each denominator q_i is a power of a prime polynomial and each $\deg p_i < \deg q_i$.

The fundamental theorem of algebra says:

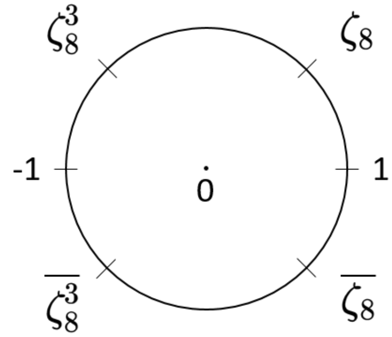
(1) Prime polynomials in $\mathbb{C}$ are: $X - c$, $c \in \mathbb{C}$

(2) Prime polynomials over $\mathbb{R}$ are: $X - a\,,\; a \in \mathbb{R}\,;\; X^2 + aX + b\,,\; (a^2 - 4b < 0)$

The usual form of partial fraction decomposition is created by expanding the numerator p_i by powers of the prime polynomial in q_i.

Example:

$$\frac{x^4}{x^4+1} \qquad \zeta_8 \;=\; -\frac{1+j}{\sqrt{2}}$$



$$\begin{aligned}
(x^4+1) &= (x-\zeta_8)\,(x-\overline{\zeta_8})\,(x-\overline{\zeta_8^3})\,(x-\zeta_8^3) \\
&= (x^2-\sqrt{2}\,x+1)\,(x^2+\sqrt{2}\,x+1) \\
(x^4+1) &= \Big(x-\frac{1+j}{\sqrt{2}}\Big)\Big(x-\frac{1-j}{\sqrt{2}}\Big)\Big(x-\frac{-1+j}{\sqrt{2}}\Big)\Big(x-\frac{-1-j}{\sqrt{2}}\Big) \\
&= (x^2-\sqrt{2}\,x+1)\,(x^2+\sqrt{2}\,x+1)
\end{aligned}$$

$$\frac{x^4}{x^4+1} = 1 - \frac{1}{x^4+1}$$

$$\frac{1}{x^4+1} = \frac{a_0x+a_1}{x^2-\sqrt{2}\,x+1} + \frac{b_0x+b_1}{x^2+\sqrt{2}\,x+1}$$

Numerator comparison:

$$\begin{aligned}
1 \;=\; &\underbrace{(a_0+b_0)}_{=0}\,x^3 + \underbrace{(a_1+b_1+\sqrt{2}\,(a_0-b_0))}_{=0}\cdot x^2 \\
&+ \underbrace{(a_0+b_0+\sqrt{2}\,(a_1-b_1))}_{=0}\,x + \underbrace{a_1 + b_1}_{=1}
\end{aligned}$$

$$\text{That means}\quad \left\{\begin{array}{lllll}
x^3: & a_0 & = & -b_0 & \\
x^1: & a_1 & = & b_1 & \\
x^0: & a_1 & = & b_1 & = \; 1/2 \\
x^2: & 1 & + & \sqrt{2}\,2\,a_0 & = \; 0
\end{array}\right\} \;\Rightarrow\; a_0 \;=\; -\frac{1}{2\sqrt{2}}$$

$$\frac{1}{x^4+1} = \frac{1}{2\sqrt{2}}\left(\frac{-x+\sqrt{2}}{x^2-\sqrt{2}\,x+1} + \frac{x+\sqrt{2}}{x^2+\sqrt{2}\,x+1}\right)$$

A·T·I·C·E
ATICE LLC, Albany NY

ISBN 978-1-951894-13-9