Introductory Course for MCS

Dennis Frisch

22th of September - 02nd of October 2008

Introduction

This script is for MCS Students at the TU Darmstadt. The goal of this introductory course is to bring all students to a common knowledge. This implies that some of the stuff (maybe most of it) isn't new for you, but you still can learn about it.
Still when you think to know this from school, it is no fault to come to the course and the exercises. To study mathematics is most of the time very different from the school class called Mathematics.

Since all students begin there study in english and have to switch after one year to german, I give a translation for many mathematical expressions. You see it in corner brackets after the english word, for example *[Beispiel]*.

This cours has two parts. First part is the lecture, which explained new concepts and give examples. Most of the time it isn't enough to hear or read this. You have to think about it and work with it. This would be in part two, the exercises, where you get an exercise sheet and try to solve the given problems. You should work in small groups and try to find out what is right. Discuss your problems with the other students and with the tutor. This is a skill you should need your whole life, especially during your studies.

I have taken the parts about series and Integration from the Analysis Script of Prof. Grosse-Brauckmann from the Winterterm 2006. Special thanks to him.

If you find any mistakes in this text or have some advice, please feel free to send them at frisch@mathematik.tu-darmstadt.de.

Dennis Frisch

Contents

1 Numbers

1.1 The Natural Numbers

The natural numbers *[natürliche Zahlen]* count objects, e.g., 3 eggs, 160 students, about 10^{70} atoms in the universe.

The set of natural numbers is denoted by $\mathbb{N}$. Two natural numbers can be added and multiplied:

$$3 + 5 = 8 \qquad 12 \cdot 11 = 132 \qquad 7^2 = 49.$$

There are many interesting subsets of $\mathbb{N}$, three of them are

$2, 4, 6, 8, 10, \ldots$	the even numbers	*[gerade Zahlen]*
$1, 3, 5, 7, 9, 11, \ldots$	the odd numbers	*[ungerade Zahlen]*
$1, 4, 9, 16, 25, \ldots$	the perfect squares	*[Quadratzahlen]*

Any two natural numbers can be compared and they can either be equal or one can be smaller than the other. We say that the natural numbers are equipped with a *total order [totale Ordnung]*. For any two natural numbers m and n we have that

$m < n$	m is less than n	*m ist kleiner als n*
$m > n$	m is greater than n	*m ist grösser als n*
$m = n$	m is equal to n	*m ist gleich n*

Furthermore, if $a < b$ and $m < n$, then $a + m < b + n$ and $a \cdot m < b \cdot n$.

The elements of any subset of $\mathbb{N}$ can be put in increasing order starting with the smallest element. Each non-empty subset of $\mathbb{N}$ has a unique smallest element. However, subsets of $\mathbb{N}$ need not have a largest element.

Exercise 1. *Find an example of a subset of $\mathbb{N}$ that does not have a largest element. Describe the subsets of $\mathbb{N}$ that have a largest element!*

Definition 1.1.1. *We say that a natural number n is* divisible [teilbar] by *a natural number d if there exists a natural number m in $\mathbb{N}$ such that*

$$d \cdot m = n.$$

If this is the case, we also say that d divides [teilt] *n. We write $d|n$. A natural number d that divides n is also called a* divisor [Teiler] *of n. Vice versa, n is a* multiple [Vielfaches] *of d.*

Examples

- The number 12 is divisible by 4.

  Proof: We need to use the definition above. Here we have that $n = 12$ and $d = 4$. We have to find a natural number m such that $12 = 4 \cdot m$. This is easy since $m = 3$ is such a number (in fact the only one).

- The number 12 is not divisible by 7.

  Proof: We need to show that there is no number m such that $7 \cdot m = 12$. (If there was such an m, then 7 would divide 12). In other words, we need to show that no multiple of 7 is equal to 12. The first few multiples of 7 are 7, 14, 21 which shows that 12 is not a multiple of 7.

Exercise 2.

- *Prove: If d is a divisor of n, then $d = n$ or $d < n$.*
- *List all divisors of 12, 140 and 1001. Prove for 12 that there are no other divisors.*
- *Show that 7 is not a divisor of 100.*
- *Show that each natural number n is divisible by 1 and by n.*
- *Show: If d divides m and n, then d also divides $m + n$ and $m - n$ and d^2 divides mn.*

Definition 1.1.2. *A natural number different from 1 that is divisible by 1 and itself only, is called a* prime number [Primzahl].

Examples: Examples of primes are: 2, 3, 5, 7, 2003, $2^{13} - 1$.

The definition of primes raises the question how one can find primes. This is a difficult problem in general. There is an algorithm which, in principle, can find all the primes, although it is impractical for large prime numbers.

The Sieve of Eratosthenes

This procedure finds all primes up to a given bound. It works as follows: Choose a number N, e.g., $N = 20$. List the natural numbers up to N:

$$2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12 \quad 13 \quad 14 \quad 15 \quad 16 \quad 17 \quad 18 \quad 19 \quad 20.$$

We iterate the following procedure: The next number which is not crossed out is a prime. We record it and cross out all its multiples:

So, 2 is a prime. Cross out its multiples, the even numbers:

$$2 \quad 3 \quad \cancel{4} \quad 5 \quad \cancel{6} \quad 7 \quad \cancel{8} \quad 9 \quad \cancel{10} \quad 11 \quad \cancel{12} \quad 13 \quad \cancel{14} \quad 15 \quad \cancel{16} \quad 17 \quad \cancel{18} \quad 19 \quad \cancel{20}.$$

The next prime is 3. Cross out all multiples of 3:

$$2 \quad 3 \quad \cancel{4} \quad 5 \quad \cancel{6} \quad 7 \quad \cancel{8} \quad \cancel{9} \quad \cancel{10} \quad 11 \quad \cancel{12} \quad 13 \quad \cancel{14} \quad \cancel{15} \quad \cancel{16} \quad 17 \quad \cancel{18} \quad 19 \quad \cancel{20}.$$

The next prime is 5. At this point we notice that all multiples of 5 have already been crossed out. The same is true for 7 and all the remaining numbers. Therefore, all the remaining numbers are primes.

Trial division

How does one check if a natural number n is a prime? One way is to try if it is divisible by any smaller number. To do that one has to carry out $n-2$ divisions if n is a prime.

The following theorem helps in reducing the number of trial divisions because it shows that one only has to do trial divisions with smaller *prime* numbers. We have already used this fact in Eratosthenes' Sieve because we had declared a number a prime if it is not a multiple of any smaller prime.

Theorem 1.1.3. *Any natural number n is divisible by a prime.*

Proof. Consider all divisors of n different from 1. There is a smallest element q among these. Let m be a natural number such that $q \cdot m = n$.

We will show that q is a prime. Suppose that q is not a prime. Then q has a divisor $1 < d < q$ and $q = d \cdot m'$. We get

$$n = q \cdot m = (d \cdot m') \cdot m = d \cdot (m' \cdot m).$$

We see that d is a divisor of n. But d is smaller than q which contradicts the choice of q. Therefore, it is impossible that q has a proper divisor. Hence, q is a prime. □

Theorem 1.1.4 (without a proof)**.** *Each natural number is a product of primes. This product is unique up to permuting the factors.*

Theorem 1.1.5. *There are infinitely many primes.*

Proof. We assume that there are only finitely many primes and show that this assumption leads to a contradiction.

Let k be the number of primes and let p_1, p_2, p_3, $\ldots$, p_k be the finitely many primes. Consider $M = p_1p_2p_3 \ldots p_k + 1$. Clearly, p_j divides $p_1p_2p_3 \ldots p_k$. If p_j divides M, then

p_j also divides $M - p_1 p_2 p_3 \dots p_k = 1$. But no prime is a divisor of 1. Therefore, M is not divisible by any of the k primes above. This contradicts our theorem that every natural number is divisible by a prime. $\square$

We will now leave the prime numbers and turn our attention back to using natural for counting.

Counting

How many ways are there to put k objects out of n different objects into a row?

Example: Consider the five vowels A E I O U. Here are all three-letter arrangements (without repitition of letters):

AEI AEO AEU AIE AIO AIU AOE AOI AOU AUE AUI AUO EAI EAO EAU
EIA EIO EIU EOA EOI EOU EUA EUI EUO IAE IAO IAU IEA IEO IEU
IOA IOE IOU IUA IUE IUO OAE OAI OAU OEA OEI OEU OIA OIE OIU
OUA OUE OUI UAE UAI UAO UEA UEI UEO UIA UIE UIO UOA UOE UOI

If we want to write down all three-letter words, then we have 5 choices for the first letter. Once the first letter is fixed we have 4 choices for the second letter and after that 3 choices for the last letter. This gives $5 \cdot 4 \cdot 3 = 60$ different choices each of which produces a different word.

The general argument goes like this: For the first object we have n choices. For the second object we have $n-1$ choices. As each choice of the first object can be combined with each choice of the second object, this gives $n(n-1)$ possibilities. For the third choice we have $n-2$ possibilities. Therefore there are $n(n-1)(n-2)$ possibilities to place 3 objects out of n objects in a row. In general, there are $n(n-1)(n-2)\dots(n-(k-1))$ possibilities to place k out of n objects in a row.

If $k = n$, then this give $n(n-1)(n-2)\dots 3 \cdot 2 \cdot 1$ possibilities to arrange n different objects in a row. We denote the number $n(n-1)(n-2)\dots 3 \cdot 2 \cdot 1$ by $n!$, which is pronounced *n factorial [n Fakultät]*. We set $0! := 1$. This can be interpreted as saying that there is one way to arrange no objects.

Example: The four symbols $+ - \cdot /$ can be arranged in $24 = 4 \cdot 3 \cdot 2 \cdot 1$ ways:

$$
\begin{array}{cccccccc}
+-\cdot/ & +-/\cdot & +\cdot-/ & +\cdot/- & +/-\cdot & +/\cdot- & -+\cdot/ & -+/\cdot \\
-\cdot+/ & -\cdot/+ & -/+\cdot & -/\cdot+ & \cdot+-/ & \cdot+/- & \cdot-+/ & \cdot-/+ \\
\cdot/+- & \cdot/-+ & /+-\cdot & /+\cdot- & /-+\cdot & /-\cdot+ & /\cdot+- & /\cdot-+
\end{array}
$$

With the factorial notation we can write the number $n(n-1)(n-2)\dots(n-(k-1))$ as

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

This counts the number of arrangements of k objects out of n objects.

If we would like to know how many ways there are to choose k objects out of n objects, then the order in which objects are chosen is unimportant. The words AEI and EIA consist of the same letters and would not be considered different choices of three vowels.

Example There are 10 ways to choose 3 vowels from A E I O U:

AEI AEO AEU AIO AIU AOU EIO EIU EOU IOU

In general, we need to take the number $\frac{n!}{(n-k)!}$ and divide by the number of arrangements of k objects. This gives

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

This expression is abbreviated by

$$\binom{n}{k}$$

which is pronounced as *n choose k [n über k]*. Note that $\binom{n}{0} = 1$. This means that there is one way to choose no object out of n.

Exercise 3.

- *Show that* $\binom{n}{k} = \binom{n}{n-k}$
- *Show that* $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$

The last property can be used to compute these numbers in the form of a triangle, known as Pascal's Triangle *[Pascalsches Dreieck]*:

$$\begin{array}{ccccccccccccccccccc}
&&&&&&&&&1&&&&&&&&&\\
&&&&&&&&1&&1&&&&&&&&\\
&&&&&&&1&&2&&1&&&&&&&\\
&&&&&&1&&3&&3&&1&&&&&&\\
&&&&&1&&4&&6&&4&&1&&&&&\\
&&&&1&&5&&10&&10&&5&&1&&&&\\
&&&1&&6&&15&&20&&15&&6&&1&&&\\
&&1&&7&&21&&35&&35&&21&&7&&1&&\\
&1&&8&&28&&56&&70&&56&&28&&8&&1&\\
1&&9&&36&&84&&126&&126&&84&&36&&9&&1
\end{array}$$

Each number is the sum of the two numbers above. The number $\binom{n}{k}$ is the k-th element in row n (counted from top to bottom).

The numbers $\binom{n}{k}$ are called *binomial coefficients [Binomnialkoeffizient]*. The reason for this name becomes clear from the following. Consider the powers of the expression $x + y$. The first few are:

n	$(x+y)^n$
1	$x+y$
2	$x^2+2xy+y^2$
3	$x^3+3x^2y+3xy^2+y^3$
4	$x^4+4x^3y+6x^2y^2+4xy^3+y^4$

Compare the numbers in the expressions above with the numbers in Pascal's Triangle.

For future convenience we make a short stop at this point and fill in an usefull abbreviation which will ouccur during your whole studies.

The Sigma Sign

A huge part of mathematics is based on sequences and series. We will study them later on in Chapter 5. Most of the time sequences have a structure. For example take

$$a_1 := 1,\ a_2 := \frac{1}{2},\ a_3 := \frac{1}{3},\ a_4 := \frac{1}{4},\ \ldots$$

Then we can write $a_i := \frac{1}{i}$. If we want to calculate the sum

$$a_1 + a_2 + a_3 + a_4 + a_5 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5},$$

we can use the sigma sign as an abbreviation. Then we write

$$\sum_{i=1}^{5} a_i := a_1 + a_2 + a_3 + a_4 + a_5 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}.$$

With this usefull tool, we can rewrite the tabular above as

n	$(x+y)^n$	
1	$x+y$	$= \sum_{k=0}^{1} \binom{1}{k} x^{1-k} y^k$
2	$x^2+2xy+y^2$	$= \sum_{k=0}^{2} \binom{2}{k} x^{2-k} y^k$
3	$x^3+3x^2y+3xy^2+y^3$	$= \sum_{k=0}^{3} \binom{3}{k} x^{3-k} y^k$
4	$x^4+4x^3y+6x^2y^2+4xy^3+y^4$	$= \sum_{k=0}^{4} \binom{4}{k} x^{4-k} y^k$

From the right side we can derive a formula, which allows us to compute all expressions $(x + y)^n$ for a natural number n. This fact is known as

Theorem 1.1.6 (Binomial Theorem *[Binomischer Lehrsatz]*). *For a natural number n we have the following expression:*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$$

The expression on the right hand side is the abbreviation of the sum

$$\binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \ldots + \binom{n}{n-1} x y^{(n-1)} + \binom{n}{n} y^n.$$

This theorem is not only true for natural numbers x and y, it is also true for real numbers. We give a proof of the theorem, but we will see another proof later on to demonstrate another proof technique called *Induction [Induktion]*.

Proof. We consider the coefficient of the expression $x^{n-k}y^k$ and how it arises from the product

$$\underbrace{(x+y)(x+y)\ldots(x+y)(x+y)}_{n-\text{times}}.$$

Expanding the brackets, we have to multiply each occurrence of x or y with each other occurrence of x or y. To obtain $x^{n-k}y^k$ we have to choose y exactly k-times. Since we are choosing k times y out of n occurrences of y, we can do this in $\binom{n}{k}$ ways. Therefore, the term $x^{n-k}y^k$ occurs $\binom{n}{k}$ times. $\square$

1.2 The Integers

The set $\mathbb{N}$ is closed *[abgeschlossen]* under taking sums and products of natural numbers. However the difference of two natural numbers need not be a natural number: $7 - 13 = ?$ In other words, there is no solution to the equation $7 = x + 13$ in the set of natural numbers.

This leads to the set $\mathbb{Z}$ of integers:

$$\ldots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \ldots$$

The integers are closed under taking sums, products and differences.

The notions of order, divisibility and primes defined above can be extended to the set of integers in a natural way with little changes:

Theorem 1.2.1 (without proof). *Each integer is a product of primes and ± 1. This product is unique up to permuting the factors.*

1.3 The Rational Numbers

The quotient of two integers need not be an integer. In fact, the quotient of an integer m and an integer d is an integer if and only if m is divisible by d. In other words, for integers m and n the equation $m \cdot x = n$ need not have a solution for x in the set of integers.

This leads to the set $\mathbb{Q}$ of rational numbers *[rationale Zahlen]*. It consists of all fractions *[Brüche]* $\frac{a}{b}$ where a is an integer and b a non-zero integer. The integer a is called the *numerator [Zähler]* and b is called the *denominator [Nenner]*. The rationals are closed under addition, subtraction, multiplication and division. A set of numbers in which those four arithmetic operations can be performed is called a *field [Körper]*, $\mathbb{Q}$ is called the field of rationals numbers.

Arithmetic of Rational Numbers

Definition 1.3.1. Let a, b, c and d be integers with b and d not 0.

Addition Two fractions are added by finding a common denominator (you may want to look for their smallest common denominator):

$$\frac{a}{b} + \frac{c}{d} = \frac{ad}{bd} + \frac{bc}{bd} = \frac{ad+bc}{bd}.$$

Multiplication Two fractions are multiplied by multiplying numerator and denominators:

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

Division A fraction is divided by another fraction by multiplying with the *reciprocal [Kehrwert]* of the second fraction ($c \neq 0$):

$$\frac{a}{b} : \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}.$$

Between any two different rational numbers lie infinitely many rational numbers. For this it is enough to show that there is always a rational number lying strictly between any pair of different rational numbers. For example, a rational number lying between the rational numbers x and y is the number $\frac{x+y}{2}$.

Equality Two fractions $\frac{a}{b}$ and $\frac{c}{d}$ are equal if and only if $ad = bc$.

This definition implies that canceling common factors in the numerator and denominator of a fraction does not change the value of the fraction: Let a, b and c be integers with b

and c different from 0.
$$\frac{ac}{bc} = \frac{a}{b} \quad \text{because} \quad ac \cdot b = bc \cdot a.$$

However, certain equations do not have a solutions in the set of rational numbers. For example, the equation $x^2 = 2$.

Theorem 1.3.2. *A solution of the equation $x^2 = 2$ is not a rational number.*

Proof. Let $\frac{a}{b}$ be a rational number with $(\frac{a}{b})^2 = 2$. We may assume that a and b have no common factor.
Then $a^2 = 2b^2$. Therefore a^2 is an even number. The square of an integer is even if and only if the integer is even. Therefore, a is even and can be written as $a = 2d$. This gives $2b^2 = 4d^2$ and dividing by two gives $b^2 = 2d^2$. By the same reasoning as above, b is even. Hence a and b contain the common factor 2 contrary to our assumption. $\square$

Exercise 4. *1. Let n be a natural number. Show that n^2 is even if and only if n is even.*

2. Show that $x^2 = 6$ does not have a rational solution.

3. Show that $1 + \sqrt{2}$ is not a rational number.

4. Show that $x^3 = 2$ does not have a rational solution.

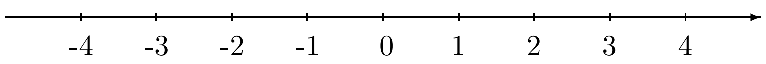
1.4 The Real Numbers

The set of real numbers *[reelle Zahlen]*, denoted by $\mathbb{R}$ is an extension of the rational numbers containing all limits *[Grenzwerte]* of rational sequences *[Folgen]* such as

$$\begin{aligned}
\sqrt{2} &= 1,4142135623730950488016887242096980785696718753769480731766797379907324784621\ldots \\
e &= 2.7182818284590452353602874713526624977572470936999595749669676277240766303535\ldots \\
\pi &= 3.1415926535897932384626433832795028841971693993751058209749445923078164062862\ldots
\end{aligned}$$

and the solutions to equations of the form $x^5 + x + 1 = 0$ and many more. The real numbers are much more complicated than the rational numbers. Most real numbers cannot be written down explicitly.

The set of real numbers is often visualized by a line, called *the real line.*



Definition 1.4.1 (Roots *[Wurzel]*). Let a be a non-negative real number and n a natural number. The n-th root of a is a non-negative real number r such that $r^n = a$.

Note that the n-th root is in general only defined for non-negative real numbers. Also the n-th root of a non-negative real number is always a non-negative real number. Taking the root of a positive number is the inverse operation to raising a real number to the n-th power.

If x is a *negative* number, then taking the square root is not the inverse operation of squaring x because the square root is positive: $x \neq \sqrt{x^2} = -x$. The same is true for even powers n. If n is an uneven number, then the n-th root is declared for all real numbers x.

Example

$$\sqrt[3]{-8} = -2.$$

Definition 1.4.2. Let a be a real number. We define the following function:

$$|a| = \begin{cases} a & \textit{if} \quad a > 0 \\ 0 & \textit{if} \quad a = 0 \\ -a & \textit{if} \quad a < 0 \end{cases}$$

The non-negative real number $|a|$ is called the absolute value of a *[Betrag]*.

Exercise 5. *Let a, b and c be real numbers and ε a positive real number.*

- *Show that $|a| \leq c$ is the same as saying $-c \leq a \leq c$.*
- *Show that $a \leq |a|$ and $-|a| \leq a$.*
- *Prove the triangle inequality: $|a + b| \leq |a| + |b|$. Hint: Use the previous two inequalities.*
- *Prove the inequality $|a| - |b| \leq |a - b|$.*
- *Show that $|x - a| \leq \varepsilon$ is the same as saying $a - \varepsilon \leq x \leq a + \varepsilon$. Interpret this geometrically! What is the set of all x satisfying this condition?*
- *Determine the solutions of the inequalities $|4 - 3x| > 2x + 10$ and $|2x - 10| \leq x$.*

1.5 The Complex Numbers

The real numbers allow us to solve many more equations than the rational numbers, which in turn allow solving more equations than the integers. Still, there are some simple equations we can not solve. In particular, the equation $x^2 + 1 = 0$ has no solution over the reals. A solution to this would be $\sqrt{-1}$ if it was defined.

When faced with the problem of not being able to divide by arbitrary non-zero numbers, we simply introduced new symbols (namely fractions). We do the same with the square root of -1 by defining the symbol i (the imaginary unit) such that $i^2 = -1$.

This leads to the set $\mathbb{C}$ of complex numbers *[komplexe Zahlen]*. It consists of all terms of the form $a + bi$, where a and b are real numbers. We call a the *real part [Realteil]*, and b the *imaginary part [Imaginär Teil]*. The complex numbers form a field with the real numbers naturally embedded in them. Unlike the number sets we saw so far, the complex numbers do not permit a natural total order.

Arithmetic of Complex Numbers

Equality Two complex numbers $a + bi$ and $c + di$ are equal if and only if their real and imaginary parts are equal, i.e. if $a = c$ and $b = d$.

Addition Two complex numbers are added as one might expect:

$$(a + bi) + (c + di) = (a + c) + (b + d)i.$$

Multiplication Two complex numbers are multiplied by following the normal rules of multiplication and treating i like a variable (and using that $i^2 = -1$):

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

Division A complex number is divided by another (non-zero) complex number by multiplying with the *inverse* of the second number. The inverse is computed as follows:

$$(a + bi)^{-1} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}i.$$

Exercise 6. *Verify that the inversion formula is correct.*

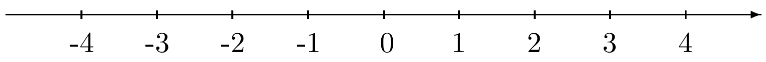
We define the *complex conjugate [komplex konjugierte]* of the complex number $c = a + bi$ as $\bar{c} := a - bi$. We now define the absolute value for a complex number c as

$$|c| := \sqrt{c\bar{c}} = \sqrt{a^2 + b^2}$$

. Note that over the real numbers this coincides with the previous definition of absolute value. Using these notations, we can write c^{-1} as $\frac{\bar{c}}{|c|^2}$.

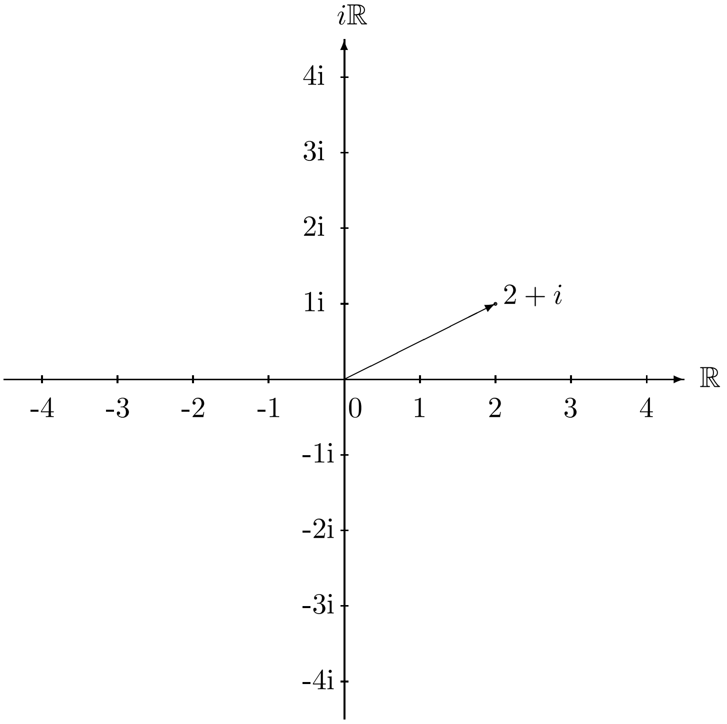
Complex Numbers from a geometrical Point of View

When introducing the real numbers, we introduced the real line too.



The real line is a geometric way to visualize the real numbers. We can try to find out, if the arithmetic operations have a meaning in this geomtric environment. We can see easily, that additon is a translation and multiplication is a dilation. If the number is negativ, then the delation changes the direction.

The complex numbers are the union of two real lines. One describing the real part and one describing the imaginary part. We can visualize complex numbers in a coordinate system:



Now we can view complex numbers as part of a two dimensional plane, the so called *complex plane [komplexe Zahlenebene].*

There is another possibility to describe complex numbers:

Definition 1.5.1. *Each complex number $z = a + bi$ can be expressed by*

$$z = |z|(\cos\phi + i \sin\phi),$$

where ϕ is a real number called the argument of z [Argument] and the absolute value

$$|z| := \sqrt{a^2 + b^2}.$$

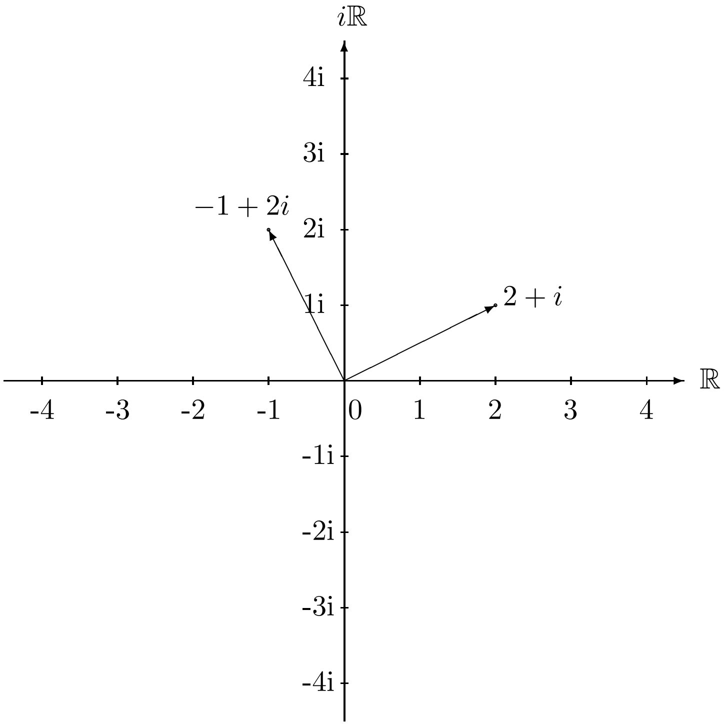
If we take $-\pi < \phi \leq \pi$, then ϕ is unique determined.

We know from above, that adding a real number to a real number is a translation. This is still true for complex numbers.
Looking at the multiplication, we saw that multiplicating a real number with a real number is a dilatation.
This is still correct, if we multiplicate a real number to a complex number. But what is about a complex number with nonzero imaginary part?

We take the example above $z := 2 + i$. If we multiplicate this number with i we get

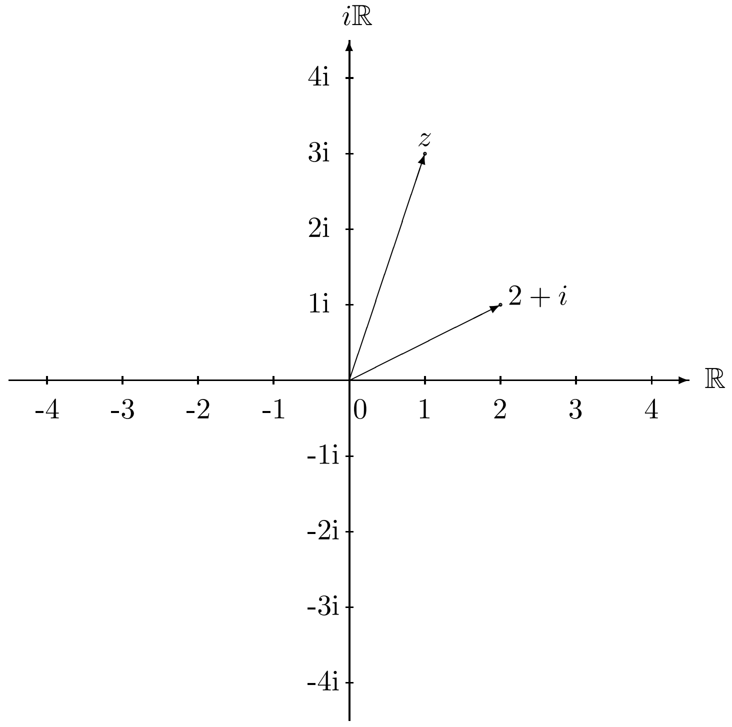
$$(2 + i)\dot{i} = 2i + i^2 = -1 + 2i.$$

We see, that $|2 + i| = | - 1 + 2i|$. If we draw this number in the complex plane, we get



We see, that multiplicating with a purely imaginary complex number of absolute value 1 is a rotation (in this case a rotation of 90° counterclockwise, the angle of i with the positive real line).
So if the complex number has absolute value different from 1, we get a dilation too. If we take the product $z := (2 + i)\dot{(}1 + i)$, we see, that $1 + i$ includes an angle of 45° with the positive real line, meaning that multiplication includes a rotation by 45° counterclockwise. The absolute value $\sqrt{2}$, meaning a dilatation of $\sqrt{2}$.

Putting this in the coordinate system, we get


$i\mathbb{R}$
4i
3i
2i
1i
z
$2+i$
$\mathbb{R}$
-4
-3
-2
-1
0
1
2
3
4
-1i
-2i
-3i
-4i

2 Propositional Logic

Examples of propositions: 5 is not a number. Darmstadt is in Germany. Mathematics is a science. 7 divides 12.

A *proposition [Aussage]* is a grammatically correct statement which it can be decided of whether it is true or false.

More interesting than deciding wether one proposition is true or false is to decide whether a proposition ist true under certain circumstances. This process is fundamental in mathematics.

We now have a look how to combine given propositions to new propositions and under which circumstances the new proposition is true.

2.1 Logical Operators

Negation The negation *[Verneinung]* of a proposition A is false when A is true and vice versa (written $\neg A$):

A	$\neg A$
t	f
f	t

And Two propositions A and B can be combined by *and* to give a new proposition $A \wedge B$ which is true precisely when both A and B are true:

A	B	$A \wedge B$
t	t	t
t	f	f
f	t	f
f	f	f

Or Two propositions A and B can be combined by *or* to give a new proposition $A \vee B$ which is true precisely at least one of A and B is true:

A	B	$A \vee B$
t	t	t
t	f	t
f	t	t
f	f	f

Implication If we want to determine, if a proposition B is true under the condition of another proposition A, then we have an implication *[Implikation]*:

A	B	$A \implies B$
t	t	t
t	f	f
f	t	t
f	f	t

Equivalence A proposition A is equivalent *[äquivalent]* to a proposition B (written $A \Leftrightarrow B$) if A is true precisely when B is true and A is false precisely when B is false (also written A *iff* B, which means A is true if and only if B is true).

A	B	$A \Leftrightarrow B$
t	t	t
t	f	f
f	t	f
f	f	t

We give another characterisation for an equivalence. And we take this as an example for a typical proof of such logical propostional statements:

Theorem 2.1.1. *Let A and B two propositions. Then is equivalent:*

a) $((A \Rightarrow B) \wedge (B \Rightarrow A))$

b) $(A \Leftrightarrow B)$

Proof.

A	B	$A \Rightarrow B$	$B \Rightarrow A$	$(A \Rightarrow B) \wedge (B \Rightarrow A)$
t	t	t	t	t
t	f	f	t	f
f	t	t	f	f
f	f	t	t	t

$\square$

Implications that are not equivalences

Here are some examples for implications, which are only true in one direction:

a) For all $x \in \mathbb{R}$: $x > 0 \Rightarrow x^2 > 0$.

b) If x and y are negative real numbers, then $x \cdot y > 0$.

To see, that these propositions are wrong in the other direction, we need to reverse the proposition. What does this mean in the cases above?

a) If $x^2 > 0$ then $x > 0$.

b) If $x \cdot y > 0$ then x and y are negative real numbers.

In both cases we find easily a counterexample to prove that these propositions are wrong.

2.2 Quantors

Sometimes it is usefull to abreviate logical statements. Here are some usefull abreviations, called *quantors [Quantoren]*:

For all If for each element e of a set S, a proposition $A(e)$ is given, then

$$\forall e \in S : A(e)$$

is a proposition which is true if $A(e)$ is true for each $e \in S$. *Read: For all e in S is true: $A(e)$*

There exists If for each element e of a set S, a proposition $A(e)$ is given, then

$$\exists e \in S : A(e)$$

is a proposition which is true if $A(e)$ is true for at least one $e \in S$. *Read: It exists an element e, such that $A(e)$ is true.*

2.3 Negation of Propositions

If you have a proposition and you don't believe it's trueness, the easiest way to see that it is wrong is to find a counterexample. To find a counterexample, you need to know the negation of the proposition.
Take the following examples:

Examples

a) All sheeps are black.

b) It exists a male student at the TU Darmstadt.

c) An animal is a lion or a duck.

d) A real number is positiv and negative.

What are the negations of that?

a) If all sheeps are black, then there is no sheep with another colour. So the negation is: *There exists a sheep, which is not black.*

b) This proposition is true, if only one student at the TU Darmstadt is male. So the negation is *All students at the TU Darmstadt are not male.*

c) To be true, each animal has to be a duck or a lion. So the negation is *There is an animal, which is wether a duck nor a lion.*

d) The proposition is true, when all real numbers are both, negative and positiv. So the negation is *There is a real number, which is not negative or not positive.*

We summarize these facts with our known symbols, where A and B are propositions:

a) $\neg(\forall e \in S : A(e)) = \exists e \in S : \neg A(e)$

b) $\neg(\exists e \in S : A(e)) = \forall e \in S : \neg A(e)$

c) $\neg(A \vee B) = \neg A \wedge \neg B$

d) $\neg(A \wedge B) = \neg A \vee \neg B$

3 Proof Techniques

We have an assumption A and we would like to conclude from A that a proposition B is true.

To conclude means, that we make little logical steps. If all steps are true, so will be proposition B.

Examples

The sum of two even numbers is an even number.

Proof. Assumption: We have two numbers x, y which are even. This means that there exists numbers a and b, such that $x = 2a$ and $y = 2b$. Then

$$x + y = (2a) + (2b) = 2(a + b).$$

So $2|(x + y)$. $\square$

This is an example for a direct proof:

3.1 Techniques

Direct Proof

A direct proof *[direkter Beweis]* is straightforward. You take the assumptions A and try to conclude since the proposition B is proved.

$A \implies B$

Examples

- Sum of two even integers is even
- If a divides b and a divides c then a divides b + c.

Proof by contradiction

A proof by contradiction *[Beweis durch Widerspruch]* uses the fact, that

$$(A \implies B) \iff (\neg B \implies \neg A).$$

Examples

- For $a, b \in \mathbb{R}$: $\frac{a+b}{2} \geq \sqrt{ab}$.
- $\sqrt{3}$ is irrational.
- There is no smallest positive (i.e. > 0) rational number.

Proof by induction

A proof by induction *[Beweis durch vollständige Induktion]* is a usefull tool to prove a proposition $B(n)$ which is stated for all natural numbers n.

We need two parts for a proof by induction:

1. Induction start *[Induktionsanfang]*: Prove $B(1)$
2. Induction step *[Indusktionsschritt]*: Assume that $B(n)$ is true and show that $B(n+1)$ is true.

Imagine a domino chain. The induction step assures that a domino (here $n+1$) falls, if domino n falls. But this is not enough. Look at this example:

Proposition: $1+n > 2+n$ for all $n \in \mathbb{N}$.

This is obviously wrong, but we had no problem to show the induction step:

Assume, that the proposition is true for n, so $1+n > 2+n$ is true. Then

$$\begin{aligned} & 1+(n+1) > 2+(n+1) \\ \Leftrightarrow \quad & 1+n+1 > 2+n+1 \\ \Leftrightarrow \quad & 1+n > 2+n. \end{aligned}$$

The last row is true by our assumption and so is the first proving the induction step. Bet we found no beginning, since we found no natural number n, for which the proposition is true.

Otherwise it could be, that the indction start is not by $n = 1$. Look at

Example: $2^n > n+1$.

A typical example for a proposition which is proved by induction is:

Propositoin: $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

Proof. We prove it by induction.

I-Start: For $n = 1$, the proposition is

$$\sum_{k=1}^{1} k = 1 = \frac{1(2)}{2}.$$

So it is true for $n = 1$.

I-Step: Now we assume, that the proposition is true for n. That means

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

is true. With the help of this assumption we try to prove the proposition

$$\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2},$$

which is the propostition formulated for $n+1$. We get

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

□

3.2 Be carefull

Some proofs look good at the first view, but sometimes a subtle error is inside. Look at this proof:

Proposition: 2=1

Proof: Let a and b be nonzero real numbers with $a = b$. Then

$$\begin{aligned}
a &= b & &| \cdot a \\
a^2 &= ab & &| - b^2 \\
a^2 - b^2 &= ab - b^2 & & \\
(a+b)(a-b) &= ab - b^2 & &| : (a-b) \\
a+b &= \frac{ab-b^2}{a-b} & & \\
a+b &= \frac{(a-b)b}{a-b} & & \\
a+b &= b & &| \; a = b \\
b+b &= b & & \\
2b &= b & &| : b \\
2 &= 1. & &
\end{aligned}$$

What is wrong???

A fake induction proof

Theorem: All sheep have the same colour.

Proof. We proof inductively that any set of sheep consists of only sheeps of a single color, i.e. is equicolored.
Induction start: A set containing one sheep is obviously equicolored.
Induction step: Assume that any set of n sheep is equicolored. Now consider a set of $n+1$ sheep. The set formed by the first n sheep are equicolored. But so is the set formed by the last n sheep. Hence the whole set must be equicolored. □

3.3 Existence, Construction and Uniqueness

In mathematics, there are often propositions, which had a solution. For example:

Theorem 3.3.1. *The numbers 12 and 18 have a great common divisor* [grösster gemeinsamer Teiler].

Most of the times, we are interested in the existence of a solution to the given problem and most of the time this is all we can do. But sometimes we need the solution or we need to know if it is unique.

Existence proofs

Like the name promised, an existence proof *[Existenzbeweis]* showed the existence of something. Let's look at Theorem 3.3.1:

Proof. A divisor of a natural number has to be less or equal the number. Since there are only finite many natural numbers, which are less than 12 or 18 only finite many divisors can exist.
Take 1, it is a natural number and divides both 12 and 18. So 1 is a common divisor. Since there are only finite other divisors, a greatest common divisor exists. □

At the end of the proof, we know the theorem holds, but we can't say what the greatest common divisor is.

Constructive proof

A constructive proof *[konstruktiver Beweis]* is a proof which delivers a solution too. We look again at our Theorem 3.3.1:

Proof. The divisors of 12 are $1, 2, 3, 4, 6, 12$.
The divisors of 18 are $1, 2, 3, 6, 9, 18$.
So the greatest common divisor is 6. □

Uniqueness proof

Sometimes it is important to know if a solution is unique. Then we have to maintain a uniqeness proof *[Eindeutigkeitsbeweis]*. We look again at Theorem 3.3.1:

Proof. Assume, there are two different greatest common divisors a and b of 12 and 18. Since a and b are common divisors and a is the greatest common divisor, this leads to $a \geq b$. The same is true for b leading to $b \geq a$ and so $a = b$. □

4 Functions

Definition 4.0.2. Let X and Y be sets. A *function* f from the set X to the set Y is a rule that assigns to each element of X exactly one element of Y.

The element of Y assigned to a particular element $x \in X$ is denoted by $f(x)$ and $f(x)$ is called *the image of x under f [Bild von x unter f]* . Vice versa, x is called *a preimage of $y = f(x)$ [Urbild].* Note that an element $y \in Y$ can have more than one preimage under f or may not a have a preimage at all.

The set X is called *the domain of f [Definitionsbereich]* and Y is called *the range of f [Wertebereich].* The set $\{f(x)|x \in X\}$ of all images is called the *image of f [Bild].*

It is important to understand that the domain and the range are an essential part of the definition of a function. For example, consider the functions

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$$
$$g : \mathbb{R} \to \mathbb{R}_{>0}, x \mapsto x^2$$

Strictly speaking, these are two different functions. One obvious difference is that all elements in the range of g do have a preimage, while there are elements in the range of f which do not have a preimage (-1 for example). So the statement "All elements in the range of g have a preimage" is true for g and false for f.

Examples

1. Let $c \in Y$ a constant element in Y. Then the function

$$f: \begin{array}{ccc} X & \to & Y \\ x & \mapsto & c \end{array}$$

   is called a constant function *konstante Funktion.* It maps each element of X to the same value c.

2. The function

$$\mathrm{id}_X : \begin{array}{ccc} X & \to & X \\ x & \mapsto & x \end{array}$$

   is called the identity function of X *[Identität].* It maps each element of X to itself.

4.1 Properties of functions

Definition 4.1.1. For a given function $f : X \to Y$ we define:

injective The function f is called *injective [injektiv]*, iff for all $x_1, x_2 \in X$

$$x_1 = x_2 \Leftrightarrow f(x_1) = f(x_2).$$

surjective The function f is called *surjectiv [surjektiv]*, iff for all $y \in Y$ there exists a $x \in X$, such that $f(x) = y$.

bijective If the function is injective and surjective, then it is called *bijective [bijektiv]*.

Examples

1. The function id_X is bijective.
2. The constant function $f : x \mapsto c$ for a fixed c is injective if and only if X has exactly one element. It is surjective if and only if Y has exactly one element.
3. The function
$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & x \cdot (x-1)(x+1) \end{array}$$
is not injective because $f(-1) = f(0) = f(1) = 0$. The function is surjective because the equation $f(x) = c$ is equivalent to the equation $x^3 - x - c = 0$ of degree three which has a zero in $\mathbb{R}$.

4.2 Algebra with functions

Definition 4.2.1. We consider functions $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$. Then we can construct other functions

1. *$f \pm g : x \mapsto f(x) \pm g(x)$ for $x \in X \cap Y$*
2. *$f \cdot g : x \mapsto f(x) \cdot g(x)$ for $x \in X \cap Y$*
3. *$f/g : x \mapsto f(x)/g(x)$ for $x \in X \cap Y$ and $g(x) \neq 0$*
4. *$g \circ f : x \mapsto g(f(x))$ if $f(X)$ is contained in Y.*

   This is called the *composition [Hintereinanderausführung]* of functions. The function f is the *inner* function and the function g is the *outer* function.

Example We take the function $f : \mathbb{R} \to \mathbb{R}, f(x) = \sqrt{x^2 + 1}$ and decompose it as follows: Let $\mathbf{1}_\mathbb{R} : x \mapsto 1$ and $\sqrt{\cdot} : x \mapsto \sqrt{x}$. Then

$$f = \sqrt{\cdot} \circ (\mathrm{id}_\mathbb{R} \cdot \mathrm{id}_\mathbb{R} + \mathbf{1}_\mathbb{R})$$

Theorem Let $f : X \to Y$ be a bijective function. Then there is a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.

The function g is called the *inverse function [Umkehrfunktion]* of f. We write f^{-1} for g. If $f(x) = y$, then $f^{-1}(y) = x$.

4.3 Types of functions on $\mathbb{R}$

constant functions Let $c \in \mathbb{R}$. Then a function $f(x) = c$ is a constant function.

power functions The function $f(x) = x^n$ for a natural number n is called a *power function [Potenzfunktion]*.

polynomials A function of the form $f(x) = c_n x^n + c_{n-1} x^{n-1} + \ldots + c_1 x + c_0$ is called a *polynomial function [Polynom]*. Polynomial functions are built from the identity function $\mathrm{id}_\mathbb{R}$ and the constant functions using $+, -, \cdot$.

rational functions A function of the form $f(x) = p(x)/q(x)$ with polynomials p and q is called a *rational function [rationale Funktion]*. Note that its domain is $\mathbb{R} \setminus \{x \in \mathbb{R} | q(x) = 0\}$.

algebraic functions Algebraic functions are constructed from polynomials (or, equivalently from the identity function and the constant functions) by using $+$, $-$, $\cdot$, $/$ and taking root.

4.4 Zeros

It is often important to find zeroes of a given function. For example many surfaces can be described as the zeroset of a function. Take the function

$$\mathbb{R}^2 \to \mathbb{R}^2, \qquad x \mapsto \|x\| - 1.$$

Then the zeroset are all elements $x \in \mathbb{R}^2$ wich have absolute value 1. This zeroset desribes the circle of radius 1.

In general, a surface which is given by a zeroset of a algebraic function is called an *algebraic surface*. If you google "algebraic surface" then you find many intresting examples for algebraic surfaces[1].

[1] e.g. http://www1-c703.uibk.ac.at/mathematik/project/bildergalerie/gallery.html

So we have a look how to find zeroes for a given function. Because this can be very difficult (e.g. it is proved, that the Riemanian ζ-function has zeros, but it is still an open problem to find an explicit one) we look at some easy functions.

Zeros of polynomnials (Degree 1)

Finding zeros in this case is very easy. A polynomial of degree 1 is

$$p(x) = mx + b.$$

A solution of the equation

$$p(x) = 0$$

is easily given by

$$x = \frac{-b}{m}.$$

Zeros of polynomials (Degree 2)

A polynomial of degree 2 looks like

$$p(x) = ax^2 + bx + c.$$

If we want to find zeros, this is equivalent to solve

$$p(x) = ax^2 + bx + c = 0 \quad \Leftrightarrow \quad x^2 + \frac{b}{a}x + \frac{c}{a} = 0.$$

Substituting $p := \frac{b}{a}$ and $q := \frac{c}{a}$ yields to

$$x^2 + px + q = 0. \tag{4.1}$$

Theorem 4.4.1. *The solutions (if they exist) of equation* (4.1) *are given by*

$$x_{1/2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}.$$

Proof. Substitution of x in (4.1) leads to

$$\begin{aligned}
x^2+px+q &= \left(-\frac{p}{2}\pm\sqrt{\left(\frac{p}{2}\right)^2-q}\right)^2+p\left(-\frac{p}{2}\pm\sqrt{\left(\frac{p}{2}\right)^2-q}\right)+q \\
&= \left(-\frac{p}{2}\right)^2\mp p\sqrt{\left(\frac{p}{2}\right)^2-q}+\left(\frac{p}{2}\right)^2-q+p\left(-\frac{p}{2}\pm\sqrt{\left(\frac{p}{2}\right)^2-q}\right)+q \\
&= \frac{p^2}{4}\mp p\sqrt{\left(\frac{p}{2}\right)^2-q}+\frac{p^2}{4}-\frac{p^2}{2}\pm p\sqrt{\left(\frac{p}{2}\right)^2-q} \\
&= \frac{p^2}{4}+\frac{p^2}{4}-\frac{p^2}{2} \\
&= 0.
\end{aligned}$$

□

Zeros of polynomials (Degree n)

Now we give some hints to find zeros of polynomials of degree n. The first hint is to guess a zero. If we find a zero then we can make a polynomial division to decomposite the give polynomial $p(x)$ of degree to

$$p(x)=(x-a)q(x),$$

where $q(x)$ is a polynomial of degree $n-1$. Now we can guess again till the polynomial q has degree 2.

If there is an integer zero, then this integer is a divisor of the absolute term. So if we had a polynomial of degree n with n zeros $a_1,\dots,a_n$, then the polynomial cn be expressed by

$$(x-a_1)(x-a_2)\cdots(x-a_n)=x^n+bx^{n-1}+\ldots+a_1a_2\cdots a_n.$$

So any zero is a divisor of the absolute term $a_1a_2\cdots a_n$. But don't forget that in this context a divisor can be negative. For example

$$x^2-2x+1=0.$$

The absolute term is 1, so possible zeros are $\{1,-1\}$. We try -1 and find that this is a zero, but 1 isn't.

5 Infinite Sequences of Real Numbers

The mathematical concept of a sequence [Folge] is easy to understand. First we look at a few examples.

Examples

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \ldots \quad \text{the sequence of natural numbers}$$

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots, \quad \text{a sequence of rational numbers}$$

$$-1, 1, -1, 1, -1, 1, -1, \ldots \quad \text{a sequence of 1s and } -1\text{s}$$

$$\pi, \frac{2}{3}, 15, \log 2, \sqrt{15}, \ldots \quad \text{a sequence of random reals number}$$

The characteristic feature of a sequence of numbers is the fact that there is a first term of the sequence, a second term, and so on. In other words, the numbers in a sequence come in a particular order. This gives rise to the following formal definition:

Definition 5.0.2. An sequence of real numbers is a map from the natural numbers $\mathbb{N}$ to the set of real numbers $\mathbb{R}$. This means that for each natural number n there is an element of the sequence, which we denote by a_n. In this notation, the elements of the sequence can be listed as

$$a_1, a_2, a_3, \ldots$$

More concisely, we write $(a_n)_{n\in\mathbb{N}}$ for the sequence.

Examples Here are examples of infinite sequences:

1. Let c be a fixed constant real number. Then the sequence $a_n = c$ for $n \in \mathbb{N}$ is called constant sequence.

$$c, \quad c, \quad c, \quad c, \quad c, \quad c, \quad c, \ldots$$

2. $a_n = \frac{1}{n}$ for $n \in \mathbb{N}$. The term a_{17} is $\frac{1}{17}$. A sequence like this is explicitly defined. It is given by a formula which can be used directly to compute an arbitrary term of the sequence.

3. Here is another example of an explicitly given sequence: $(\frac{n}{n+1})_{n\in\mathbb{N}}$.

4. Define $a_1 = 1$ and $a_{n+1} = a_n + (2n + 1)$. This is a recursively defined sequence. To compute a_{n+1} we need to know a_n, for which we need to know a_{n-1} and so on. Sometimes it is not difficult to find an explicit description for a recursively defined sequence. Here, we have $a_n = n^2$.

5. A famous (and more difficult) example for a recursively defined sequence is the Fibonacci sequence: $f_1 = 1$, $f_2 = 1$ and $f_{n+1} = f_n + f_{n-1}$ for $n > 2$, $n \in \mathbb{N}$. The first few terms of the sequence are

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

There is the following closed form:

$$f_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right).$$

We will now look at the second and the third sequence.

$$1, \frac{1}{2} = 0.5, \frac{1}{3} = 0.3333\ldots, \frac{1}{4} = 0.25, \frac{1}{5} = 0.2, \ldots, \frac{1}{200} = 0.005, \ldots$$

$$\frac{1}{2} = 0.5, \frac{2}{3} = 0.6666\ldots, \frac{3}{4} = 0.75, \frac{4}{5} = 0.8, \ldots, \frac{199}{200} = 0.995, \ldots$$

While the terms of the first sequence get closer and closer to 0, the term of the second sequence get closer and closer to 1. Although no term of either sequence ever reaches 0 or 1, respectively, we would like to be able to express the fact that both sequences approach a certain number and get arbitrarily close.

Definition 5.0.3. A sequence $(a_n)_n \in \mathbb{N}$ has a *limit [Grenzwert]* $a \in \mathbb{R}$ if for every $\varepsilon > 0$ there is a $N \in \mathbb{N}$ such that

$$|a_n - a| < \varepsilon \text{ for } n \geq N.$$

If the sequence $(a_n)_{n\in\mathbb{N}}$ has a limit b, then $(a_n)_{n\in\mathbb{N}}$ is called *convergent [konvergent]*. We write

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

Read: The limit of a_n as n goes to ∞ is b.

If a sequence is not convergent it is called *divergent [divergent]*.

It is worthwhile to think about this definition for a while and understand what the different parts of the definition mean. One way to interpret it is to say that b is a limit of a sequence $(a_n)_{n\in\mathbb{N}}$ if the distance of b to all except a finite number of terms of the sequence is smaller than ε. The finite number of terms which may be further away from b than ε are

$$a_1, a_2, a_3, \ldots, a_{N-1}.$$

Note that N depends on ε, although we do not say this explicitly in the definition. This is because we have to choose N appropriately, depending on the given ε.

Example Let us consider the sequence $a_n = \frac{1}{n}$ for $n \in \mathbb{N}$. We would like to show that the sequence has limit 0. We will follow the definition of a limit and need to show that for each $\varepsilon > 0$ there is a N such that

$$\left|\frac{1}{n}\right| < \varepsilon \text{ for all } n \geq N$$

We take ε as given. The condition $\frac{1}{n} < \varepsilon$ is equivalent to the condition $n > \frac{1}{\varepsilon}$. So let us try to choose N to be the next natural number larger than $\frac{1}{\varepsilon}$. The we have that $\frac{1}{N} < \varepsilon$. With this we get the following chain of inequalities for $n \geq N$:

$$\frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

In particular, we see that $\frac{1}{n} < \varepsilon$ for all $n \geq N$. Hence we have shown that 0 is the limit of the sequence $(\frac{1}{n})_{n\in\mathbb{N}}$.

Example The sequence $1, -1, 1, -1, \ldots$ is divergent. It is interesting to prove this using the definition of limit. It requires working (implicitly or explicitly) with the negation of the defining property including the various quantors.

Theorem 5.0.4 (Algebra with sequences)**.** *Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be convergent sequences. Then:*

1. *$(a_n \pm b_n)_{n\in\mathbb{N}}$ is convergent and*

$$\lim_{n\to\infty}(a_n \pm b_n) = \lim_{n\to\infty} a_n \pm \lim_{n\to\infty} b_n$$

2. *$(a_n \cdot b_n)_{n\in\mathbb{N}}$ is convergent and*

$$\lim_{n\to\infty}(a_n \cdot b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n$$

3. *If $b_n \neq 0$ and $\lim_{n\to\infty} b_n \neq 0$: $(\frac{a_n}{b_n})_{n\in\mathbb{N}}$ is convergent and*

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

What is if one of the sequences (for example sequence $(a_n)_{n\in\mathbb{N}}$) is divergent? What can we say about

$$\begin{gathered}(a_n + b_n)_{n\in\mathbb{N}} \\ (a_n \cdot b_n)_{n\in\mathbb{N}} \\ \left(\frac{a_n}{b_n}\right)_{n\in\mathbb{N}}\end{gathered} \qquad ?$$

6 Series

A series is another name for an infinite sum. Later we shall introduce many functions as infinite sums: the exponential function, trigonometric functions, etc. Thus we want to investigate series in general.

6.1 Partial sums and convergence

The most prominent example of a series is perhaps the exponential function $\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$. For each $x \in \mathbb{C}$, we regard it the limit of the sequence (s_n) of numbers $s_1 = 1$, $s_2 = 1 + x$, $s_3 = 1 + x + \frac{x^2}{2!}$, ... Similarly in general:

Definition 6.1.1. *(i) Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. Then a* series *[Reihe] is the sequence $(s_n)_{n\in\mathbb{N}}$ of* partial sums

$$s_n := a_1 + \ldots + a_n.$$

Usually we write $\sum_{n=1}^{\infty} a_n$ for the sequence $(s_n)_{n\in\mathbb{N}}$, and call a_n its terms [Summanden].
(ii) In case the series (s_n) converges to $s \in \mathbb{R}$ we write

$$\sum_{n=1}^{\infty} a_n := \lim_{n\to\infty} \sum_{k=1}^{n} a_k = \lim_{n\to\infty} s_n = s.$$

Remark: In the convergent case, the notation $\sum_{n=1}^{\infty} a_n$ has two different meanings:
• The sequence of partial sums $(a_1 + \ldots + a_n)_{n\in\mathbb{N}}$,
• a number $s \in \mathbb{R}$, namely the limit of the partial sums; it is also called the *value [Wert]* of the series.

Examples: 1. *Decimal expansion*: $3.14\ldots = 3 + \frac{1}{10} + \frac{4}{100} + \ldots$. We will study these series in more detail below.

2. We claim $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$, that is, we claim for the partial sums

$$s_n := \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{n(n+1)} \quad \to \quad 1 \quad \text{as} \quad n \to \infty.$$

Proof: Writing

$$\frac{1}{n(n+1)} = \frac{-(n^2-1)+n^2}{n(n+1)} = -\frac{n-1}{n} + \frac{n}{n+1}, \qquad \text{for } n \in \mathbb{N},$$

we see we can apply a telescope sum trick:

$$\begin{aligned}
s_n &= \Big(-0+\frac{1}{2}\Big)+\Big(-\frac{1}{2}+\frac{2}{3}\Big)+\Big(-\frac{2}{3}+\frac{3}{4}\Big)+\ldots+\Big(-\frac{n-1}{n}+\frac{n}{n+1}\Big)\\
&= -0+\frac{n}{n+1}=\frac{1}{1+\frac{1}{n}} \quad\to\quad 1 \qquad \text{as } n\to\infty.
\end{aligned}$$

If we are careless, we can easily run into contradictions:

$$0=(1-1)+(1-1)+\ldots=1+(-1+1)+(-1+1)+\ldots=1$$

In naive language, infinite sums are not associative. Thus only manipulations stipulated by the limit theorems for sequences are admissable.

Theorem 6.1.2. *If $\sum_{n=1}^{\infty} a_n$ converges then $a_n \to 0$ as $n\to\infty$.*

Proof. We have $a_n = s_n - s_{n-1}$ for $n \geq 2$ and thus, using $s_n=\sum_{k=1}^{n} a_k \to s$,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}(s_n-s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = s-s=0. \qquad \square$$

Does the converse of the theorem hold? This is not the case:
Examples: The *harmonic series*

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots.$$

has terms $\frac{1}{n}$ forming a null sequence. Nevertheless, the partial sums are not bounded. Indeed, for a subsequence,

$$\begin{aligned}
s_{2^n} &= 1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{2^n}\\
&= 1+\frac{1}{2}+\Big(\underbrace{\frac{1}{3}+\frac{1}{4}}_{\geq 1/2}\Big)+\Big(\underbrace{\frac{1}{5}+\ldots+\frac{1}{8}}_{\geq 1/2}\Big)+\ldots+\Big(\underbrace{\frac{1}{2^{n-1}+1}+\ldots+\frac{1}{2^n}}_{\geq 1/2}\Big)\\
&\geq 1+\frac{n}{2}\to\infty.
\end{aligned}$$

This unboundedness means the harmonic series cannot converge. Moreover (s_n) is increasing, and hence our argument shows that $\sum \frac{1}{n}$ diverges to infinity; as for sequences we denote this symbolically by $\sum \frac{1}{n}=\infty$.

The most important series will turn out to be the following:

Theorem 6.1.3. *Let $x\in\mathbb{R}$. The* geometric series $1+x+x^2+x^3+\ldots$ *converges for all $|x|<1$ to*

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x},$$

while for $|x|\geq 1$ the series diverges.

Proof. The geometric sum gives

$$s_n = \sum_{j=0}^{n} x^j = 1 + x + x^2 + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x} \qquad \text{for } x \neq 1. \tag{6.1}$$

When $|x| < 1$ we see that $x^n \to 0$ as $n \to \infty$; hence $\lim s_n = \frac{1}{1-x}$.

For $|x| \geq 1$ also $|x^n| = |x|^n \geq 1$, and so (x^n) is not a null sequence. Theorem 6.1.2 gives that $\sum x^n$ cannot converge. □

Example: $|\frac{1}{2}| < 1$ and hence

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = \frac{1}{1 - \frac{1}{2}} = 2$$

Remark: A periodic decimal expansion is, up to an additive constant, a geometric series; it always defines a rational number. For example,

$$\begin{aligned} 2.\overline{34} := 2.343434\ldots &= 2 + \frac{34}{10^2} + \frac{34}{10^4} + \frac{34}{10^6} + \cdots = 2 + \frac{34}{100}\left(1 + \frac{1}{100} + \frac{1}{100^2} + \ldots\right) \\ &= 2 + \frac{34}{100} \cdot \frac{1}{1 - \frac{1}{100}} = 2 + \frac{34}{100} \cdot \frac{100}{99} = 2 + \frac{34}{99} = \frac{232}{99}. \end{aligned}$$

6.2 Series of real numbers

There are two useful tests for convergence of real series. The first one can deal with series whose sign alternates:

Theorem 6.2.1 (Leibniz)**.** *Let $(a_n)_{n\geq 0}$ be a monotone decreasing null sequence, $a_0 \geq a_1 \geq a_2 \geq \ldots \geq 0$. Then the* alternating sum *$\sum_{n=0}^{\infty} (-1)^n a_n$ converges.*

Example: The *alternating harmonic series*

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \ldots.$$

converges.

Proof. The idea is to see the alternating series defines an interval nesting whose common point is the limit.

To see this, consider odd and even partial sums,

$$A_n := s_{2n+1} = a_0 - a_1 + \ldots + a_{2n} - a_{2n+1} \quad \text{and} \quad B_n := s_{2n} = a_0 - a_1 + \ldots + a_{2n},$$

where $n \in \mathbb{N}_0$. The equations

$$A_n = A_{n-1} + \underbrace{a_{2n} - a_{2n+1}}_{\geq 0}, \quad B_n = B_{n-1} \underbrace{-a_{2n-1} + a_{2n}}_{\leq 0} = A_n + \underbrace{a_{2n+1}}_{\searrow 0} \qquad \text{for } n \in \mathbb{N}$$

prove the following facts: (A_n) increases monotonously, (B_n) decreases monotonously, $A_n \leq B_n$, and $(B_n - A_n)$ is a null sequence.

Hence $[A_n, B_n]$ is a sequence of nested intervals, containing a common point s, and so

$$s = \lim_{n\to\infty} A_n = \lim_{n\to\infty} B_n = \lim_{n\to\infty} s_n = \sum_{n=0}^{\infty} (-1)^n a_n. \qquad \square$$

A second test applies to real series whose terms all have the same sign:

Theorem 6.2.2. *A series $\sum_{n=1}^{\infty} a_n$ with $a_n \geq 0$ converges if and only if its partial sums are bounded.*

Proof. The assumption $a_n \geq 0$ means that the sequence of partial sums (s_n) is increasing. Therefor $s_{n+1} \geq s_n$ and so $s_n \leq s$. $\square$

Example: Consider a decimal expansion $0.d_1d_2d_3\ldots = \sum_{n=1}^{\infty} \frac{d_n}{10^n}$ with $d_n \in \{0, 1, \ldots, 9\}$. The partial sums

$$s_n = \frac{d_1}{10} + \frac{d_2}{100} + \ldots + \frac{d_n}{10^n}$$

are increasing in n and are bounded by

$$s_n \leq \frac{9}{10} + \frac{9}{100} + \ldots + \frac{9}{10^n} \overset{\text{geom.series}}{=} \frac{9}{10} \cdot \frac{1 - (\frac{1}{10})^n}{1 - \frac{1}{10}} < \frac{9}{10} \cdot \frac{1}{1 - \frac{1}{10}} = \frac{9}{10} \cdot \frac{10}{9} = 1$$

(our estimate says that $0.99\ldots 9$, with n digits, is indeed less than 1). Theorem 6.2.2 gives that each decimal expansion converges.

The boundedness criterion can be used for a comparison test for convergence:

Theorem 6.2.3 (Majorization of real series)**.** *Suppose $(x_n)_{n\in\mathbb{N}}$ is a real sequence for which there exists a convergent series $\sum_{n=1}^{\infty} a_n$ of real numbers $a_n \geq 0$ with*

$$0 \leq x_n \leq a_n \qquad \textit{for all } n \in \mathbb{N}.$$

Then $\sum_{n=1}^{\infty} x_n$ also converges and $\sum_{n=1}^{\infty} x_n \leq \sum_{n=1}^{\infty} a_n$.

We call a_n a *majorant* of x_n.

Proof. We consider partial sums. By the theorem, $\sum_{k=1}^{n} a_k \leq C$ for some $C \in \mathbb{R}$ and so

$$0 \leq \sum_{k=1}^{n} x_k \leq \sum_{k=1}^{n} a_k \leq C. \tag{6.2}$$

But applying the theorem once again, we see that $\sum x_k$ must converge.

□

Exercise 7. (Minorization) *Suppose for a real series $\sum a_n$ there exists a sequence (x_n) with $a_n \geq x_n \geq 0$ such that $\sum x_n$ is divergent. Prove that $\sum a_n$ diverges as well.*

6.3 Decimal expansions

In antiquity, the only numbers that could be represented arithmetically were rational numbers $\mathbb{Q}$ or *proportions*. Geometry was considered superior to algebra as it could deal with "all" numbers. Since its invention in medieval time, decimal representations have changed the view of mathematics. Nowadays many people believe that real numbers and decimal representations are identical, so that the nonuniqueness of the type $1 = 0.999\ldots$ poses a problem. This problem is easy to resolve once decimal expansions are regarded as series.

Let us first deal with real numbers which are not negative.

Definition 6.3.1. *A* decimal expansion *is a series*

$$\sum_{n=0}^{\infty} \frac{d_n}{10^n} = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \ldots,$$

with $d_0 \in \mathbb{N}_0$ and digits $d_n \in \{0, \ldots, 9\}$ for $n \geq 1$.

Then each decimal expansion defines some real number, and for each real number there is at least one decimal representation:

Theorem 6.3.2. *1. Each decimal expansion $\sum_{n=0}^{\infty} \frac{d_n}{10^n}$ converges to a number*

$$x = \sum_{n=0}^{\infty} \frac{d_n}{10^n} \in [0, \infty). \tag{6.3}$$

2. For each real number $x \in [0, \infty)$ there exists $d_0 \in \mathbb{N}_0$ and a sequence $d_n \in \{0, \ldots, 9\}$ for $n \geq 1$ such that (6.3) holds.

Proof. 1. We can majorize:

$$0 \leq \sum_{n=0}^{\infty} \frac{d_n}{10^n} \leq d_0 + \sum_{n=1}^{\infty} \frac{9}{10^n} = d_0 + \frac{9}{10} \frac{1}{1 - \frac{1}{10}} = d_0 + 1$$

(Our estimate says that $0.99\ldots 9$, with n digits, is indeed less than 1). Theorem 6.2.3 gives the claim.

2. For a given x, let us define the digits d_n by an interval nesting: There exists $d_0 \in \mathbb{N}$ with $d_0 \leq x < d_0 + 1$. Then we define recursively: Suppose $d_1, \ldots, d_n$ are constructed, such that

$$a_n := d_0 + \frac{d_1}{10} + \ldots + \frac{d_n}{10^n} \leq x < d_0 + \frac{d_1}{10} + \ldots + \frac{d_n + 1}{10^n} =: b_n \tag{6.4}$$

Then subdivide the interval $I_n = [a_n, b_n)$ into the 10 halfopen disjoint intervals

$$[a_n, a_n + \frac{1}{10^{n+1}}),\ [a_n + \frac{1}{10^{n+1}}, a_n + \frac{2}{10^{n+1}}),\ \ldots [a_n + \frac{9}{10^{n+1}}, b_n),$$

whose union gives I_n. One of these ten intervals contains x; call it $I_{n+1} = [a_{n+1}, b_{n+1})$; this constructs d_{n+1} such that (6.4) holds. But $a_1 \leq \ldots \leq a_n \leq x < b_n \leq \ldots \leq b_1$ and so By the interval nesting property, $\bigcap [a_n, b_n]$ contains x, and so $\lim_{n\to\infty} \sum_{k=0}^{n} \frac{d_n}{10^n} = x$, meaning that (6.3) holds.

$\square$

Remark: A *rearrangement of a series [Umordnung]* is a change in the order of summation. For absolutely convergent series, the limit remains unchanged upon rearrangement. For convergent series which are not absolutely convergent (*conditionally convergent series*) the limit surprisingly can change, however. Let us rephrase this fact by saying that the commutative law is not automatic for convergent series. (See exercises).

7 Continuous Functions

We talked about functions in Chapter 4. Functions assigend an element x of a given set X to an element y of a given set Y. Let's have a look if there is more than this assignement only.

Examples:

1. If we drive a car wich accelarates with a constant acceleration a. Then we know from physics, that the velocity v after a time t is equals
$$v = v_0 + a \cdot t,$$
were v_0 is the velocity at time $t = 0$. We can also say, that v depends on t, which means that v is a function of the time.
$$\begin{aligned} v : [0, \infty) &\longrightarrow \mathbb{R} \\ t &\longmapsto v_0 + a \cdot t \end{aligned}$$
If we now look at small changes of the variable t, then we see that the velocity $v(t)$ changes small too.

2. When we use the train from Darmstadt to Frankfurt and we want to catch another train in Frankfurt, then we are interested in the delay of the first train. We look at the function
$$w : [0, \infty) \longrightarrow [0, \infty)$$
which assignes to each delay t the time we have to wait in Frankfurt $w(t)$. If we assume the first train arrives at $xx : 48$ and the next train departes at $xx : 56$, then a delay of 3 minutes means that we have to wait 5 minutes. And a delay of 8 minutes means that we have to wait 0 minutes. But if we had a delay of $8 + \varepsilon$, then we will not catch the train and we have to wait for the next one (assume it will depart in 30 minutes).

   So a delay of 8 minutes has no waiting time, but if it was a little bit more, than we have to wait for almoust 30 minutes. This means a small change in the variable t can result in a big change of variable w.

Definition 7.0.3 (ε-δ-condition)**.** *A function $f : X \to Y$ is called* continous [stetig] *in a point $x_0 \in X$, if for all $\varepsilon > 0$ there is a $\delta > 0$, such that*
$$|f(x) - f(x_0)| < \varepsilon \qquad \textit{for all } x \in X \textit{ with } |x - x_0| < \delta. \tag{7.1}$$

Now lets have a look again at our example:

1. If we had a look at the values of $v(t)$, then we get

   $$|v(t_0) - v(t)| = |v_0 + a \cdot t_0 - (v_0 + a \cdot t)| = |a \cdot t_0 - a \cdot t| = |a(t_0 - t)| = |a|\ |t_0 - t|.$$

   Now we take $\varepsilon > 0$, and assumed that $|v(t_0) - v(t)| < \varepsilon$. Now we have to find $\delta > 0$, such that $|v(t_0) - v(t)| < \varepsilon$ holds for all t with $|t_0 - t| < \delta$. From the inequality

   $$|v(t_0) - v(t)| = |a|\ |t_0 - t| < \varepsilon$$

   we can pick $\delta := \frac{\varepsilon}{|a|}$ and this will work. So this function is continous and we see, that the δ could depend on ε.

2. We look at the value $t_0 = 8$ minutes. We assume, that $|w(8) - w(t)| < \varepsilon$. We choose for example $\varepsilon = 1$. Then we had to find a $\delta > 0$, such that for all t in $|8 - t| < \delta$ the inequality

   $$|w(8) - w(t)| < \varepsilon$$

   holds. But since we can pick for every $\delta > 0$ a value $t \in |t_0 - t| < \delta$, wich is strictly bigger than 8, that means $t = 8 + s$ with $s > 0$, we see that

   $$|w(8) - w(t)| \geq |w(8) - w(t + s)| = |0 - (30 - s)| = 30 - s.$$

   So for $\varepsilon = 1$ we found no δ, which implies that this function is not continous.

If we had a function $\mathbb{R} \to \mathbb{R}$ (in general from a metric space in another metric space), then we had another characterization of continous:

Theorem 7.0.4 (The limit test)**.** *Let f be a function defined on a neighbourhood U of x_0 but possibly not defined in x_0. Then is equivalent:*

(a) f is continous in x_0

(b) for every convergent sequence $(x_n)_{n \in \mathbb{N}} \subset X$ with limit x_0 is true:

$$\lim_{n \to \infty} f(x_n) = f\left(\lim_{n \to \infty} x_n\right) = f(x_0).$$

Proof. "$\Rightarrow$" We assume the ε-δ-condition in x_0. We need to show, that for any given sequence $x_n \to x_0$ in the domain, the image sequence satisfies $f(x_n) \to f(x_0)$. Let $\varepsilon > 0$ be arbitrary, and pick a $\delta > 0$ from (7.1). Since $x_n \to x_0$, we can choose $N \in \mathbb{N}$, such that

$$|x_n - x_0| < \delta \qquad \text{for all } n \geq N.$$

But then (7.1) implies

$$|f(x_n) - f(x_0)| < \varepsilon \text{ for all } |x_n - x| < \delta$$

which is equivalent to

$$|f(x_n) - f(x_0)| < \varepsilon \text{ for all } n \geq N,$$

which implies that $f(x_n)$ converges to $f(x_0)$.

"$\Leftarrow$" Assume, the limit condition holds. We prove the continuity of f by contradiction.

Assume, there exists an $\varepsilon > 0$, were we can't find $\delta > 0$, such that (7.1) holds. In particular, (7.1) could not be satisfied for any $\delta = \frac{1}{n}$ where $n \in \mathbb{N}$. Thus, there exists x_n in $|x_n - x_0| < \frac{1}{n}$, such that

$$|f(x_n) - f(x_0)| > \varepsilon.$$

Therefor we had $x_n \to x$, but not $f(x_n) \to f(x_0)$, contradicting the limit test.

$\square$

Example

1. Consider the function $f(x) = x^2$ defined on $\mathbb{R}$. Choose $x_0 = 0$. Then the function has the limit 0 in x_0. For a proof, we need to take any sequence $(x_n)_{n \in \mathbb{N}}$ with $\lim_{n\to\infty} x_n = 0$. Now we have to show that the sequence $f_n = x_n^2$ for $n \in \mathbb{N}$ converges to 0. This, however, is not difficult using our theorem about algebra with sequences:
$$\lim_{n\to\infty} x^2 = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} x_n = 0 \cdot 0 = 0$$

2. Now let us look at a complicated example, which does not have any limit in 0. We take the function $f(x) = \sin(\frac{1}{x})$ for $x \in \mathbb{R}$, $x \neq 0$.

   We take the sequence $x_n = \frac{1}{\pi \cdot n}$ for $n \in \mathbb{N}$. Then
$$f(x_n) = \sin(\frac{1}{x_n}) = \sin(\pi \cdot n) = 0$$

   Now we take the sequence $y_n = \frac{2}{\pi \cdot (2n+1)}$. This again gives a sequence of function values:
$$f(y_n) = \sin(\frac{1}{y_n}) = \sin(\pi \cdot (2n+1)/2) = \sin(n\pi + \frac{\pi}{2}) = 1$$

   So this time we get 1 as the limit of our sequence.

   We have taken two different sequences and have obtained two different limits. This contradicts the limit test. Therefor, this function does not have a limit in 0 in so it is not continous in 0.

Theorem 7.0.5 (Theorem (Algebra with continuous functions)). *Let $f, g : U \to \mathbb{R}$ be two continuous functions. Then*

- $f \pm g$
- $f \cdot g$
- $\frac{f}{g}$
- $f \circ g$

is continous (if defined).

Proof See the exercises.

Examples of continuous functions

1. All polynomials are continuous. This follows easily from the theorem about algebra with continuous functions and from the fact that the constant functions and the identity function on $\mathbb{R}$ are continuous.
2. We define the following functions:

$$\begin{aligned} \exp : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ \sin : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ \cos : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \end{aligned}$$

   These functions are continous.
3. Rational functions are continuous on the subset of $\mathbb{R}$ where the denominator is different from 0.

Now we prove an important theorem about continous functions:

Theorem 7.0.6 (Intermediate Value Theorem). *Let $f : [a, b] \to \mathbb{R}$ be continuous and let c be strictly between $f(a)$ and $f(b)$. Then there is an x strictly between a and b such that $f(x) = c$.*

We will not give a proof, because it is very technical.

8 Differentiable Functions

At a fixed point x_0 of a given function we would like to construct the tangent t to the graph of f. As t goes through the point $(x_0, f(x_0))$, it suffices to determine the slope of t. For this we draw a line l_x through the point $(x_0, f(x_0))$ and the point $(x, f(x))$. This lines intersects the graph of f and is not the required tangent yet. But note what happens when the point x is moved towards x_0 on the x-axis. The line l_x becomes more and more like the tangent t. In the limit $x \to x_0$ (if it exists), the line l_x and the tangent t coincide.

Now lets have a look at the slope of l_x. It is

$$\frac{f(x) - f(x_0)}{x - x_0}.$$

As l_x becomes t as x moves to x_0, the slope of t is the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Definition 8.0.7. *Let $f : (a, b) \to \mathbb{R}$ be a function and $x_0 \in (a, b)$. Then the f is called differentiable in x_0, if the limit*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists for all sequences $x_n \to x_0$ and coincides. The derivative of f in x_0 is denoted by $f'(x_0)$ (speak: f prime of x_0).

If f is differentiable in each point of (a, b) then it is called differentiable on (a, b). In this case, f' is a function on (a, b).

(The last step is more abstract than it seems. It takes us in one stride from a single value $f'(x_0)$ to a function f')

Theorem 8.0.8. *If $f : (a, b) \to \mathbb{R}$ is differentiable in x_0, then f is continuous in x_0.*

Proof. Let $x_0 \in (a, b)$. Then the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists since f is differentiable in x_0. As x goes to x_0, the numerator goes to zero. The limit can only exist, if the denominator goes to zero at the same time. If the denominator goes to zero, then $f(x)$ goes to $f(x_0)$. In other words,

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

This is our definition of continuity. $\square$

On the other hand, the continuity of f doesn't implies the differentiability. Take for example:

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto |x| \end{aligned}$$

We look at $x_0 = 0$ and show that f is continous in x_0. We take a sequence $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ wich goes to 0 from above (we denote it by $x_n \searrow 0$). Then this yields to

$$\lim_{x_n \searrow 0} f(x_n) = \lim_{x_n \searrow 0} |x_n| \overset{x_n \geq 0}{=} \lim_{x_n \searrow 0} x_n = 0.$$

If we take a sequence x_n which goes to 0 from bottom (we denote it by $x_n \nearrow 0$), then this yields to

$$\lim_{x_n \nearrow 0} f(x_n) = \lim_{x_n \nearrow 0} |x_n| \overset{x_n \leq 0}{=} \lim_{x_n \nearrow 0} -x_n = 0.$$

So we have shown, that f is continous in x_0. Now we look, if f is differentiable. Again we take a sequence $x_n \searrow 0$. This yields to

$$\lim_{x_n \searrow 0} \frac{f(x_n) - f(0)}{x_n - 0} = \lim_{x_n \searrow 0} \frac{f(x_n)}{x_n} = \lim_{x_n \searrow 0} \frac{|x_n|}{x_n} \overset{x_n \geq 0}{=} \lim_{x_n \searrow 0} \frac{x_n}{x_n} = 1.$$

If we take again a sequence $x_n \nearrow 0$ this yields to

$$\lim_{x_n \nearrow 0} \frac{f(x_n) - f(0)}{x_n - 0} = \lim_{x_n \nearrow 0} \frac{f(x_n)}{x_n} = \lim_{x_n \nearrow 0} \frac{|x_n|}{x_n} \overset{x_n \leq 0}{=} \lim_{x_n \nearrow 0} \frac{-x_n}{x_n} = -1.$$

We see that both limits did not coincide, which means that f is not differentiable in $x_0 = 0$.

While a differentiable function is continous, the derivative of a continous function need not to be continous.

A similar way to the definition

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

is to write the sequence x_n with limit x_0 as $x_0 + h$ and look at the limit $h \to 0$. In this case we can rewrite the definition as

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

Now we look at a few examples and determine some derivatives:

Examples:

1. $f : \mathbb{R} \to \mathbb{R}, x \mapsto c \cdot x$ with $c \in \mathbb{R}$.

$$f'(x_0) = \lim_{x_n \to x_0} \frac{cx_n - cx_0}{x_n - x_0} = \lim_{x_n \to x_0} \frac{c(x_n - x_0)}{x_n - x_0} = c.$$

2. $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^k$.

$$f'(x_0) = \lim_{x_n \to x_0} \frac{x_n^k - x_0^k}{x_n - x_0} = \lim_{x_n \to x_0} (x_n^{k-1} + x_n^{k-2} x_0 + \ldots + x_0^{n-1}) = kx_0^{k-1}.$$

Theorem 8.0.9 (Algebra with differentiable functions)**.** *Let $f, g : (a, b) \to \mathbb{R}$ be two functions differentiable in x_0. Then*

- $f \pm g$
- $f \cdot g$
- $\frac{f}{g}$
- $f \circ g$

is differentiable and the derivative is

- $(f \pm g)' = f' \pm g'$
- $(fg)' = f'g + fg'$*(* product rule*)*
- $\frac{f}{g} = \frac{f'g - fg'}{g^2}$ *(*quotient rule*)*
- $(f \circ g)' = f' \circ g \cdot g'$ *(*chain rule*)*

Proof.

- $f \pm g$ see exercise.

- $f \cdot g$:

$$\begin{aligned}
& \lim_{x_n \to x_0} \frac{(fg)(x_n) - (fg)(x_0)}{x_n - x_0} \\
= & \lim_{x_n \to x_0} \frac{f(x_n)g(x_n) - f(x_0)g(x_0)}{x_n - x_0} \\
= & \lim_{x_n \to x_0} \frac{f(x_n)g(x_n) - f(x_0)g(x_0) + \overbrace{f(x_n)g(x_0) - f(x_n)g(x_0)}^{=0}}{x_n - x_0} \\
= & \lim_{x_n \to x_0} \frac{f(x_n)g(x_n) - f(x_n)g(x_0) + f(x_n)g(x_0) - f(x_0)g(x_0)}{x_n - x_0} \\
= & \lim_{x_n \to x_0} \frac{f(x_n)(g(x_n) - g(x_0)) + (f(x_n) - f(x_0))g(x_0)}{x_n - x_0} \\
= & \lim_{x_n \to x_0} \frac{f(x_n)(g(x_n) - g(x_0))}{x_n - x_0} + \frac{(f(x_n) - f(x_0))g(x_0)}{x_n - x_0} \\
= & \lim_{x_n \to x_0} f(x_n)\frac{(g(x_n) - g(x_0))}{x_n - x_0} + g(x_0)\frac{(f(x_n) - f(x_0))}{x_n - x_0} \\
= & \lim_{x_n \to x_0} f(x_n) \lim_{x_n \to x_0} \frac{(g(x_n) - g(x_0))}{x_n - x_0} + \lim_{x_n \to x_0} g(x_0) \lim_{x_n \to x_0} \frac{(f(x_n) - f(x_0))}{x_n - x_0} \\
= & f(x_0)g'(x_0) + g(x_0)f'(x_0).
\end{aligned}$$

- $\frac{f}{g}$ see exercises

- $f \circ g$ *(we do the proof in the case, that g is injective).* Write

$$\frac{(f \circ g)(x_n) - (f \circ g)(x_0)}{x_n x_0} = \frac{(f \circ g)(x_n) - (f \circ g)(x_0)}{g(x_n) - g(x_0)} \frac{g(x_n) - g(x_0)}{x_n - x_0}$$

  Here we used the injectivity of g, which assures that $g(x_n) - g(x_0) \neq 0$. Now we can determine the limit:

$$\begin{aligned}
\lim_{x_n \to x_0} \frac{(f \circ g)(x_n) - (f \circ g)(x_0)}{x_n - x_0} &= \lim_{x_n \to x_0} \frac{(f \circ g)(x_n) - (f \circ g)(x_0)}{g(x_n) - g(x_0)} \frac{g(x_n) - g(x_0)}{x_n - x_0} \\
&= \lim_{x_n \to x_0} \frac{f(g(x_n)) - f(g(x_0))}{g(x_n) - g(x_0)} \frac{g(x_n) - g(x_0)}{x_n - x_0} \\
&= \lim_{x_n \to x_0} \frac{f(g(x_n)) - f(g(x_0))}{g(x_n) - g(x_0)} \lim_{x_n \to x_0} \frac{g(x_n) - g(x_0)}{x_n - x_0} \\
&= f'(g(x_0))g'(x_0)
\end{aligned}$$

□

Now we have a look at a usefull application from everdays life. At first we have the following definition:

Definition 8.0.10. *Let $f : (a, b) \to \mathbb{R}$ be a function and $x_0 \in (a, b)$. Then x_0 is called*

- local minimum *[lokales Minimum] if there exists an $\varepsilon > 0$, such that for all $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$ $f(x_0) \leq f(x)$*
- local maximum *[lokales Maximum] if there exists an $\varepsilon > 0$, such that for all $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$ $f(x_0) \geq f(x)$.*

Now we can formulate

Theorem 8.0.11 (Local extrema)**.** *Let $f : (a, b) \to \mathbb{R}$ be a differentiable function and $x_0 \in (a, b)$. If f has a local extremum in x_0, then $f'(x_0) = 0$.*

Proof. Without loss of generality we assume that we had a local maximum in x_0, that means $f(x_0) \geq f(x)$ for all $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$. Then

$$f'(x_0) = \lim_{h \to 0, h<0} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0$$

since $f(x_0 + h) - f(x_0) \leq 0$ and $h < 0$. But on the other hand

$$f'(x_0) = \lim_{h \to 0, h>0} \frac{f(x_0 + h) - f(x_0)}{h} \leq 0$$

since $f(x_0 + h) - f(x_0) \leq 0$ and $h > 0$. Since f is differentiable these two limits have to coincide which yields to $f'(x_0) = 0$. $\square$

Here is another way to find limits:

Theorem 8.0.12 (L'Hospitals Rule)**.** *Let $f, g : (a, b) \to \mathbb{R}$ be differentiable functions and $x_0 \in (a, b)$. Furthermore, let $\lim_{x \to x_0} f(x) = 0$ and $\lim_{x \to x_0} g(x) = 0$. We consider the function $\frac{f(x)}{g(x)}$. Note that it is not defined in x_0 because $\lim_{x \to x_0} g(x) = 0$. But if $\frac{f'(x)}{g'(x)}$ exists, then $\frac{f(x)}{g(x)}$ exists and*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}.$$

Remark: This rule is only defined for limits which approach to a real number. So if we had for example

$$\lim_{n \to \infty} \sin(\frac{1}{n})n,$$

then we can't apply L'Hopitals rule. First we have to substitute the sequence by (for example) $k := \frac{1}{n}$. As n goes to infinity, k goes to 0. This yields to

$$\lim_{n \to \infty} n \sin(\frac{1}{n}) = \lim_{k \to 0} \frac{\sin(k)}{k}.$$

Now we can use the rule and we get

$$\lim_{n \to \infty} n \sin(\frac{1}{n}) = \lim_{k \to 0} \frac{\sin(k)}{k} = \lim_{k \to 0} \frac{\cos(k)}{1} = 1.$$

9 Integral

In antiquity, Archimede determined the volume of special bodies such as the cone, sphere, and cylinder. To calculate areas or volumes in general is the main task of integration. The first attempt for a systematic treatment of integration goes back to Cavalieri in the 17th century.

The integral of a function of one variable is the oriented area content bounded by the graph. Two questions arise:
- For which functions can we declare the integral?
- How do we compute integrals?

The answer to the second question will be deferred until Section 9.3: The Fundamental Theorem of Calculus will turn out to be crucial.

The first question has less practical impact; in fact, all functions of daily life are integrable. It is, however, an interesting mathematical problem. We will approach it as follows. We set off with step functions, for which integration is obvious, and then use a limit process to extend the integral to a large class of functions. Suprisingly, this class is not explicit, and so, in a second step, we will show that, for instance, continuous functions belong to this class.

9.1 Step functions

A function $\phi\colon [a,b] \to \mathbb{R}$ is a *step function [Treppenfunktion]*, if there is a *partition [Zerlegung]* $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ of $[a,b]$, such that ϕ is constant on each interval (x_{k-1}, x_k) for $k = 1, \ldots, n$. Note that the number n of steps is finite, and we do not constrain the values $\phi(x_k)$.

Let us denote the set of step functions on $[a,b]$ by $S[a,b]$. The sole point of introducing these functions is that their integrals are obvious, as we know the area of a rectangle:

Definition 9.1.1 (Integral of step functions)**.** *Let $\phi \in S[a,b]$ with $\phi(x) = c_k$ on (x_{k-1}, x_k) for $k = 1, \ldots, n$. Then we set*

$$\int_a^b \phi(x)\,dx := \sum_{k=1}^{n} c_k (x_k - x_{k-1}). \tag{9.1}$$

We also admit $a = b$, in which case the sum is empty and $\int_a^a \phi(x)\,dx := 0$.

The same step function can be described with respect to many different partitions, for instance, we can always include additional support points into a given partition. Then $\int_a^b \phi(x)\,dx$ remains invariant:
• For just one additional support point, this is seen as follows: If $\phi(x) := c$ on $[a,b]$ and $\xi \in (a,b)$, then

$$c(b-a) = c(\xi - a) + c(b - \xi), \tag{9.2}$$

just as rectangle areas add.
• For the general case, if X is a partition $a = x_0 < x_1 < \ldots < x_i = b$ and Y is $a = y_0 < y_1 < \ldots < y_j = b$ then their union forms a partition Z of form $a = z_0 < z_1 < \ldots < z_k = b$, having $k \le i + j$ points. Appealing to (9.2), we see that the sums with respect to X and Z are equal, and so are the sums with respect to Y and Z. Consequently, the sums for X and Y are also equal, which means the integral is well-defined.

Theorem 9.1.2. *Let $\phi, \psi \in S[a,b]$ and $\lambda \in \mathbb{R}$, then:*

(i) $\int_a^b \lambda\phi + \psi\,dx = \lambda \int_a^b \phi\,dx + \int_a^b \psi\,dx$ *(linearity)* [Linearität]

(ii) For $a \le \xi \le b$ *we have* $\int_a^\xi f + \int_\xi^b f = \int_a^b f$. *(monotonicity)* [Monotonie]

(iii) $\phi \le \psi \implies \int_a^b \phi\,dx \le \int_a^b \psi\,dx$.

In (iii), the notation $\phi \le \psi$ is shorthand for $\phi(x) \le \psi(x)$ for all $x \in [a,b]$.

9.2 The Riemann integral

Definition 9.2.1 (Lower and upper integral *[Unter- und Oberintegral]*). *Suppose $f\colon [a,b] \to \mathbb{R}$ is an arbitrary bounded function. Then we set*

$$L := \left\{ \int_a^b \phi(x)\,dx : \phi \in S[a,b],\ \phi \le f \right\}, \quad U := \inf \left\{ \int_a^b \phi(x)\,dx : \phi \in S[a,b],\ \phi \ge f \right\}.$$

$$\underline{\int_a^b} f(x)\,dx := \sup L, \qquad \overline{\int_a^b} f(x)\,dx := \inf U$$

Since we assume $|f| \le C$ the set U contains the constant step function C and is nonempty. Moreover, U is bounded from below by $-C(b-a)$ and so $\inf U$ exists. Likewise for L.

Due to monotonicity it is immediate that $\underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx$.
Examples:
1. For $\phi \in S[a,b]$ we have $\underline{\int_a^b}\phi(x)\,dx = \overline{\int_a^b}\phi(x)\,dx = \int_a^b \phi(x)\,dx$ (why?).
2. Let $\chi_{\mathbb{Q}}\colon [0,1] \to \mathbb{R}$ be the *characteristic function of* $\mathbb{Q}$ with $f(x) = 1$ for $x \in \mathbb{Q}$, and 0

otherwise. Since the rational numbers are dense in the irrational ones, each step function $\phi \geq f$ satisfies $\phi \geq 1$ (except, perhaps, at the partition points), and so $\overline{\int_a^b} \chi_{\mathbb{Q}}(x)\, dx = 1$. Similarly, $\underline{\int_a^b} \chi_{\mathbb{Q}}(x)\, dx = 0$.

For the second example, the "area" of the graph is a doubtful quantity: Is it 0, 1, or any intermediate value? However, when upper and lower integral coincide, these should represent "the" area:

Definition 9.2.2 (Riemann 1854)**.** *A bounded function $f\colon [a,b] \to \mathbb{R}$ is* (Riemann) integrable *[(Riemann)-integrierbar], if $\underline{\int_a^b} f(x)\, dx = \overline{\int_a^b} f(x)\, dx$. In that case we write*

$$\int_a^b f(x)\, dx := \underline{\int_a^b} f(x)\, dx = \overline{\int_a^b} f(x)\, dx.$$

The definition leaves open how to assert that a function is integrable. We will show below that functions which are continuous or monotone are integrable. The following reformulation of the integrability definition will be our test for integrability:

9.3 The Fundamental Theorem of Calculus

A differentiable function $F\colon [a,b] \to \mathbb{R}$ is called a *primitive* or *antiderivative [Stammfunktion]* of $f\colon [a,b] \to \mathbb{R}$, if $F' = f$.

Examples:
(i) For $f(x) = x^2$ the function $F(x) = \frac{1}{3}x^3$ is a primitive.
(ii) For $f(x) = e^{ix}$ the function $F(x) = -ie^{ix}$ is a primitive.

Often, the explicit form of a primitive can only be guessed. Nevertheless it always exists for f continuous:

Theorem 9.3.1. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous. Then the* indefinite integral *[unbestimmtes Integral]*

$$I(x) := \int_a^x f(t)\, dt$$

gives a differentiable function $I\colon [a,b] \to \mathbb{R}$. Moreover, I is a primitive of f, that is, $I'(x) = f(x)$.

Problem: If f is merely integrable then I is only Lipschitz.

Proof. We first suppose f is real-valued and compute the difference quotient of $I(x)$. Suppose $a \leq x < b$. Then for sufficiently small $h > 0$ we have $x + h < b$. For such h

follows

$$\frac{I(x+h)-I(x)}{h} = \frac{1}{h}\Big(\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt\Big) = \frac{1}{h}\int_x^{x+h} f(t)\,dt. \tag{9.3}$$

Moreover, by the Mean Value Theorem of Integration[1] there exists $\xi_h \in [x, x+h]$ with

$$\frac{1}{h}\int_x^{x+h} f(t)\,dt = f(\xi_h).$$

Now as $h \to 0$ we have $\lim_{h\to 0}\xi_h = x$, and so the limit of (9.3) exists:

$$I'(x) = \lim_{h\to 0} f(\xi_h) \overset{f \text{ continuous}}{=} f(x)$$

In case $a < x \leq b$ we can similarly consider $h < 0$ and proceed as before: Then $\frac{I(x+h)-I(x)}{h} = -\frac{1}{h}\int_{x+h}^x f(t)\,dt = -\frac{1}{h}|h|f(\xi_h) = f(\xi_h)$ for some $\xi_h \in [x+h, x]$ which again implies $I'(x) = f(x)$. □

Let us rephrase the statement, which presents, perhaps, the most important fact of calculus. The equation $\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$ means that indefinite integration and differentiation are inverse operations, cancelling one another. This is not at all clear from the definition of integral and derivative!

For f constant, $f(x) \equiv c$, this is immediate to see: $I(x) = (x-a)c$ and so $I'(x) = c$.

Obviously, when F is a primitive of f, then so is $F + c$ for c constant. Conversely, any two primitives $F, G\colon [a,b] \to \mathbb{R}$ of the same function f satisfy

$$(F-G)' = F' - G' = f - f = 0;$$

This implies that $F - G$ is constant. That is, a primitive of f is well-defined up to a constant. Making use of this property we see that the integral of f can be computed using any of its primitives F:

Theorem 9.3.2 (Fundamental theorem)**.** *Suppose a continuous function $f\colon [a,b] \to \mathbb{R}$ has a primitive $F\colon [a,b] \to \mathbb{R}$. Then*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Proof. By Thm. 9.3.1, the function $I(x) := \int_a^x f(t)\,dt$ is a primitive of f. Hence $F(x) - I(x)$ is constant, say equal to $c \in \mathbb{R}$, and

$$\int_a^b f(x)\,dx = I(b) - \underbrace{I(a)}_{=0} = \big(F(b) - c\big) - \big(F(a) - c\big) = F(b) - F(a).$$ □

[1] see for example Analysis I by K.H. Hofmann (2000), Thm 4.29

The Fundamental Theorem allows us to integrate most functions introduced so far. It will be convenient to write $F(x)\big|_a^b := F(b) - F(a)$.
Examples: From the examples for differentiation, the following is immediate:

$$\int_a^b x^n = \frac{1}{n+1} x^{n+1}\Big|_a^b$$

Thanks to the linearity of the integral this formula suffices to integrate polynomials.

$$\int_a^b e^x\,dx = e^x\Big|_a^b, \qquad \int_a^b e^{ix}\,dx = -ie^{ix}\Big|_a^b.$$

Invoking the Euler formula and taking real and imaginary parts of the second integral or (**??**) we find

$$\int_a^b \cos x\,dx = \sin x\Big|_a^b, \qquad \int_a^b \sin x\,dx = -\cos x\Big|_a^b.$$

Moreover,

$$\int_a^b \frac{1}{1+x^2}\,dx = \arctan x\Big|_a^b,$$

and, provided $[a,b]$ does not contain a zero of cosine,

$$\int_a^b \frac{1}{\cos^2 x}\,dx = \tan x\Big|_a^b.$$

9.4 Rules for integration

Each law of differentiation yields a law for integration, via the Fundamental Theorem.

Let us call a function *continuously differentiable [stetig differenzierbar]* if its derivative is continuous.

We consider the product law first.

Theorem 9.4.1 (Integration by parts)**.** *If $f, g\colon [a,b] \to \mathbb{R}$ are continuously differentiable, then*

$$\int_a^b f'(x)g(x)\,dx = f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x)\,dx.$$

Note the two integrals on the right hand side exist in view of our assumptions on f, g.

Proof. The function $h := fg$ can be differentiated using product law: $h' = f'g + fg'$. In particular, h' is continuous, and so

$$\int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx = \int_a^b h'(x)\,dx \overset{\text{Fund'l Thm.}}{=} h(x)\Big|_a^b = f(x)g(x)\Big|_a^b. \qquad \square$$

Examples:

$$\int_a^b \cos(x)x \, dx = \sin(x)x\Big|_a^b - \int_a^b \sin(x) \cdot 1 \, dx = \sin(x)x\Big|_a^b + \cos(x)\Big|_a^b$$

Exercise 8.

$$\int_{-\pi/2}^{\pi/2} \sin^2 x \, dx = \int_{-\pi/2}^{\pi/2} \cos^2 x \, dx = \frac{\pi}{2}. \tag{9.4}$$

We now discuss the Chain Rule. Let us first introduce some more notation. Suppose $F, f = F' \colon [a,b] \to \mathbb{R}$ and $x, y \in [a,b]$. Then the Fundamental Theorem gives $F(y) - F(x) = \int_x^y f(t)\, dt$. The same formula will hold for $x > y$ as well provided we set

$$\int_x^y f(t)\, dt := -\int_y^x f(t)\, dt. \tag{9.5}$$

Theorem 9.4.2 (Substitution)**.** *Let $f \colon [\alpha,\beta] \to \mathbb{R}$ be continuous and $\phi \colon [a,b] \to [\alpha,\beta]$ be continuously differentiable. Then*

$$\int_a^b f\big(\phi(t)\big)\phi'(t)\, dt = \int_{\phi(a)}^{\phi(b)} f(x)\, dx. \tag{9.6}$$

Proof. Let $F \colon [\alpha,\beta] \to \mathbb{R}$ be a primitive of f. According to the Chain Rule,

$$\big(F \circ \phi\big)'(t) = F'\big(\phi(t)\big)\phi'(t) = f\big(\phi(t)\big)\phi'(t),$$

and so (9.6) follows from

$$\int_a^b f\big(\phi(t)\big)\phi'(t)\, dt \overset{\text{Fund'l Thm.}}{=} \big(F\circ\phi\big)(t)\Big|_a^b = F\big(\phi(b)\big) - F\big(\phi(a)\big) \overset{\text{Fund'l Thm.}}{=} \int_{\phi(a)}^{\phi(b)} f(x)\, dx.$$

□

Examples:

1. Integration is invariant under translation in the domain: For $c \in \mathbb{R}$,

$$\int_a^b f(\underbrace{t+c}_{\phi(t)})\, dt \overset{(9.6)}{=} \int_{\phi(a)=a+c}^{\phi(b)=b+c} f(x)\, dx \qquad (\phi'(t) = 1).$$

2. For $c \in \mathbb{R}$ and $\phi(t) := ct$ we have

$$\int_a^b f(ct)c\, dt \overset{(9.6)}{=} \int_{ca}^{cb} f(x)\, dx \quad \overset{c\neq 0}{\Longrightarrow} \quad \int_a^b f(ct)\, dt = \frac{1}{c}\int_{ca}^{cb} f(x)\, dx.$$

3. Let us now discuss a classical problem: the area of the unit disk. The area of the upper half disk is the integral $\int_{-1}^{1}\sqrt{1-x^2}\,dx$. We want to substitute x by $\phi(t) := \sin t$ in order to take advantage of the identity $\sin^2 t + \cos^2 t = 1$. Note that $\phi\colon [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$ is continuously differentiable and invertible. Substitution gives

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx \;=\; \int_{\phi^{-1}(-1)}^{\phi^{-1}(1)} \underbrace{\sqrt{1-\sin^2 t}}_{\sqrt{\cos^2 t}}\, \underbrace{(\sin t)'}_{\cos t}\; dt \;=\; \int_{-\pi/2}^{\pi/2} \cos^2 t\, dt \overset{Ex.}{=} \frac{\pi}{2}.$$

Here, we used the fact $\cos t \geq 0$ for $t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Thus the unit disk has area π.