Introduction to Algebra 1

Erdélyi Márton
www.math.bme.hu/~merdelyi/bevalg1/
merdelyi@math.bme.hu
Building H, 6th floor, room 667

Based on the slides of Ferenc Wettl and Erzsébet Lukács

1. Introduction

2. Elementary arithmetic of integers

3. Modular arithmetic – Computing with residues

4. Complex numbers

5. Polynomials

6. Systems of linear equations

7. Vectorspaces

8. Matrices and linear maps

Introduction

Number systems

Reminder

- $\mathbb{N} = \{1, 2, 3, \dots\}$ is the set of **N**atural numbers
- $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$ is the set of integers (**Z**ahlen)

Remark

In general $\mathbb{N}$ might denote $\{0, 1, 2, 3, \dots\}$ as well, but we will stick to the terminology above. If we want to emphasize which one we use we can write $\mathbb{N}_+$ or $\mathbb{N}_0$.

Reminder

- $\mathbb{R}$ is the set of **R**eal numbers.
- $\mathbb{Q} = \{p/q \in \mathbb{R} | p, q \in \mathbb{Z}, q \neq 0\}$ is the set of rational numbers (**Q**uotient numbers).
- The set $\mathbb{R} - \mathbb{Q}$ containts the irrational numbers.

Remark

- It is easy to show that a number x is rational: it is enough to show that there exists p, q such that $x = p/q$.
- It is also not hard to show that "most of" the real numbers are irrational.
- But it is relatively hard to prove that a given number x is irrational (as we should verify infinitely many pairs (p, q)).

Theorem

$\sqrt{2}$ is irrational.

Remark

$\sqrt{n}$ is either irrational or an integer for $n \in \mathbb{N}$.

Theorem

- e (the Euler's number) is irrational (Euler, $\sim$ 1737).
- π (the length of the half perimeter of the unit circle) is irrational (Lambert, 1761).

For the proof of the irrationality of $\sqrt{2}$ we introduce the notion of ordering of a set:

Definition

Let X be a set.

- A relation $<$ on X is an **(total) ordering** if
  - it is trichotomic (for any $x, y \in X$ exactly one of the following holds: 1. $x = y$, 2. $x < y$ and 3. $y < x$) and
  - it is transitive ($x < y, y < z$ implies $x < z$)
- A subset $S \subset X$ is **well ordered** if for any $T \subset S$, $T \neq \emptyset$ has a least element.

Example

$(\mathbb{R}, <)$ is an ordered set.
$\mathbb{N}$ is well ordered – well ordering principle, but $\mathbb{Z}$ is not.

Definition

- $x \in \mathbb{R}$ is **algebraic** if there exists $n \in \mathbb{N}$ and $a_0, a_1, \ldots, a_n \in \mathbb{Z}$ (a polynomial $f \in \mathbb{Z}[x]$) such that

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0 = 0$$

- $x \in \mathbb{R}$ is **transcendental** if it is not algebraic.

Example

- All $x = p/q \in \mathbb{Q}$ is algebraic: $qx - p = 0$ $(n = 1)$.
- $\sqrt{2}$ is algebraic: $x^2 - 2 = 0$ $(n = 2)$.

Again "most" of the real numbers are transcendental, but it is much harder to prove that for a given x:

Theorem

- e is transcendental (Hermite, 1873).
- π is transcendental (von Lindemann, 1882).

Induction and recursion

Theorem (The technique of mathematical induction)

Let $P(n)$ be statements for all $n \in \mathbb{N}$. If

1. $P(1)$ is true and
2. $P(n) \implies P(n+1)$ for all n,

then $P(n)$ is true for all n.

$\iff$

Let $S \subseteq \mathbb{N}$ such that

1. $1 \in S$ and
2. $n \in S \implies n+1 \in S$ for all n,

then $S = \mathbb{N}$.

Example

The following identity holds for all $n \in \mathbb{N}$:

$$1 \cdot 2 + 2 \cdot 3 + \cdots + n \cdot (n+1) = \sum_{k=1}^{n} k(k+1) = \frac{n(n+1)(n+2)}{3}.$$

Definition

A **recursive sequence** is a sequence of numbers defined by

- the exact value of the first (few) entries $a_1, a_2, \ldots, a_{n_0}$ and
- a formula f for the othes containing the earlier entries: $a_n = f(a_1, \ldots, a_{n-1})$.

Example

- The Fibonacci numbers are $f_1 = f_2 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 3$. The first few entries are: $1, 1, 2, 3, 5, 8, 13, 21, 34$.
- The Catalan numbers $C_0 = 1$ and $C_n = \sum_{k=0}^{n-1} C_k C_{n-k-1}$. The first few entries are: $1, 1, 2, 5, 14, 42, 132, 429, 1430$.
- If the following number contains n fractions then

$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\ddots + \cfrac{1}{1+1}}}} = \frac{f_{n+3}}{f_{n+2}}$$

Approximation with rationals

Definition

For $x \in \mathbb{R}$ let

- $[x] = \lfloor x \rfloor = \max(n \in \mathbb{Z} | n \leq x)$ be the **integral part** of x,
- $\lceil x \rceil = \min(n \in \mathbb{Z} | x \leq n)$ and
- $\{x\} = x - [x]$ be the **fractional part** of x.

Example

$[0] = 0$, $[\pi] = 3$, $[-\pi] = -4$, $\{4/3\} = \{1/3\} = 1/3$.

Remark

Note that the best approximation of x with integers is either $\lfloor x \rfloor$ or $\lceil x \rceil$.

Theorem (Dirichlet's approximation theorem)

For all $x \in \mathbb{R}$ and $n \in \mathbb{N}$ there exists $a, b \in \mathbb{Z}$, $1 \leq a \leq n$ such that $|ax - b| < 1/n$.

Corollary

There is no "best approximation", in other words $\mathbb{Q}$ is dense in $\mathbb{R}$.

Theorem (Diophantine approximation)

For all irrational number x there exist infinitely many pairs $a, b \in \mathbb{Z}$, $a \neq 0$ such that

$$\left| x - \frac{b}{a} \right| < \frac{1}{a^2}.$$

Remark

- We can get a fraction only from finitely many equivalent way.
- A stronger statement is true: we can write $\frac{1}{2a^2}$ instead of $\frac{1}{a^2}$.

Definition

- Let $a_0, a_1, \ldots, a_n \in \mathbb{R} - \{0\}$. The **continued fraction**

$$[a_0, a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}$$

- For $x \in \mathbb{R}$ let $a_0 = [x]$ and $x_1 = \{x\}$. If $x_1 \neq 0$ then let $a_1 = \left[\frac{1}{x_1}\right]$ and $x_2 = \left\{\frac{1}{x_1}\right\}$, so $x = [a_0, a_1 + x_2]$. Similarly if $x_n \neq 0$, then let $a_n = \left[\frac{1}{x_n}\right]$ and $x_{n+1} = \left\{\frac{1}{x_n}\right\}$, so $x = [a_0, a_1, \ldots, a_n + x_{n+1}]$.
- The **continued fraction form** of x is $[a_0, a_1, \ldots, a_n]$ if $x_{n+1} = 0$ for some n, and the infinite continued fraction $[a_0, a_1, a_2, \ldots]$ if $x_n \neq 0$ for all n.

Example

- $\sqrt{2} = [1, 2, 2, 2, \ldots]$
- $\sqrt{15} = [3, 1, 6, 1, 6, 1, 6, \ldots]$
- $e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, \ldots]$
- $\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, \ldots]$
- $\dfrac{1+\sqrt{5}}{2} = [1, 1, 1, \ldots]$

Theorem

The continued fraction form of x is finite $\iff$ $x \in \mathbb{Q}$.

Theorem

Let $x = [a_0, a_1, a_2, \ldots]$ be an irrational number and for any $n \in \mathbb{N}$ let $\dfrac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$, with $\gcd(p_n, q_n) = 1$. Then for all $n \in \mathbb{N}$

a) $\left| x - \dfrac{p_n}{q_n} \right| < \dfrac{1}{q_n^2}$, and

b) either $\left| x - \dfrac{p_n}{q_n} \right| < \dfrac{1}{2q_n^2}$ or $\left| x - \dfrac{p_{n+1}}{q_{n+1}} \right| < \dfrac{1}{2q_{n+1}^2}$.

Elementary arithmetic of integers

Rings

Reminder

Arithmetic properties of integers:

For any $a, b, c \in \mathbb{Z}$ we have:

	addition	multiplication
associativity	$(a+b)+c = a+(b+c)$	$(ab)c = a(bc)$
commutativity	$a+b = b+a$	
identity	$\exists 0 \in \mathbb{Z} : 0+a = a$	$\exists 1 \in \mathbb{Z} : 1 \cdot a = a$
inverse	$\exists(-a) : a+(-a) = 0$	
distributivity:	$a(b+c) = ab+ac.$	

Definition

Let R be a set and $+, \cdot$ be binary operations on it (i. e. $+, \cdot$ are $R \times R \to R$ functions). We call $(R, +, \cdot)$ a **ring** if the above properties hold for all $a, b, c \in R$.

Example

- $\mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$ are rings with the usual operations.
- $\mathbb{N}$ is not a ring with the usual operations.
- For a ring R the set $R[x]$ of polynomials with variable x and coefficients in R and the usual operations is also a ring. For example $\mathbb{Z}[x]$ is the ring of polynomials with integral coefficients.
- The set of even numbers is a so called non-unital ring with the usual operations.
- The set Z_n of modulo n residue classes with the modular operations form a ring for any $n \in \mathbb{N}$.
- The set of $n \times n$ matrices with the usual operations (if you know it) is a noncommutative ring for $n > 1$.

Remark

Parts of what we will do for integers can be done generally for commutative rings.

Divisibility

Definition

The integer a is a **divisor** of the integer b – notation: $a|b$ – if there exists an integer d such that $b = ad$.
a is a **proper divisor** of b if $a|b$ and $a \neq \pm 1, \pm b$.

Theorem (Basic properties of divisibility)

For all integers a, b, c, m, n we have

1. $a|a$,
2. $a|b, b|c \implies a|c$,
3. $a|b, a|c \implies a|(mb + nc)$ and
4. $a|b, m|n \implies am|bn$.

Example

- Either $4|n^2$ or $8|n^2 - 1$ for any $n \in \mathbb{N}$.
- $7|(3^{2n+1} + 2^{n+2})$ for all $n \in \mathbb{N}$.

Units

Definition

Let R be a ring. $a \in R$ is a **unit** if for all $r \in R$ $a|r$.

Example

What are the units in

- $\mathbb{Z}$,
- $\mathbb{Q}$,
- the nonunital ring of even numbers and
- $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} | a, b \in \mathbb{Z}\} \subset \mathbb{R}$?

Theorem

$a \in R$ is a unit $\iff$ $a|1 \in R$.

Division with remainder

Theorem

For all $a, b \in \mathbb{Z}$, $b \neq 0$ there exists a unique pair of integers q, r, such that

$$a = qb + r \text{ and } 0 \leq r < b.$$

Remark

- q is the quotient, r is the residue
- in the theorem we might write $-\left[\frac{|b|}{2}\right] < r \leq \left[\frac{|b|}{2}\right]$ instead of $0 \leq r < b$. Then we get the least residue in absolute value instead of the least nonnegative residue.

Example

Divide ± 20 by ± 7 with remainders!

Numeral systems

Theorem

Let $b > 1$ be an integer. Then for any $m \in \mathbb{N}$ there exist unique $n \in \mathbb{N}$ and $a_k \in \{0, 1, \ldots, b-1\}$ for $k = 0, 1, \ldots, cn$ such that

$$m = a_n b^n + a_{n-1} b^{n-1} + \cdots + a_1 b + a_0 \text{ and } a_n \neq 0.$$

Definition

The representation of an integer m in **base** b is $(a_n a_{n-1} \ldots a_1 a_0)_b$. We can omit the brackets or b if it does not lead to confusion.

Remark

If $10 < b \leq 36$ then most commonly the capital English letters $(A, B, C, \ldots)$ are used as numerals $10, 11, 12, \ldots$.

Example

$$251_{10} = 2001_5 = 373_8 = FB_{16} = 3323_4 = 11111011_2.$$

Horner's method

Theorem

A polynomial of degree n can be evaluated at $x = c$ by n multiplications and n additions:

$$a_n c^n + a_{n-1} c^{n-1} + \cdots + a_1 c + a_0 =$$
$$(\ldots((a_n \cdot c + a_{n-1}) \cdot c + a_{n-2}) \cdot c + \cdots + a_1) \cdot c + a_0.$$

Remark

The evaluation can be arranged in a table:

a_n	a_{n-1}	a_{n-2}	$\ldots$	a_1	a_0
b_{n-1}	b_{n-2}	b_{n-3}	$\ldots$	b_0	b_{-1}

,

where $b_{n-1} = a_n$ and $b_i = c \cdot b_{i+1} + a_{i+1}$ for $i < n - 1$. The value of the polynomial at $x = c$ is b_{-1}.

Example

For $f(x) = x^3 + 2x^2 - 3$, $c = 2$:

1	2	0	−3
↓ 1	↓ +2	↓ +0	↓ -3
1	4	8	13
	2·	2·	2·

Greatest common divisor

Definition

For $a, b \in \mathbb{Z}$ an integer d is

- a **common divisor** (of a and b) if $d|a$ and $d|b$.
- the **greatest common divisor** (of a and b) if
  - $d = 0$ in case $a = b = 0$ and otherwise
  - d is a common divisor and
  - it is the greatest among the common divisors, i. e. for any $c \in \mathbb{Z}$, $c|a$, $c|b$ we have $c \leq d$.

  The notation is $\gcd(a, b) = (a, b) = d$.

Remark

- $\gcd(a, b)$ exists for any a and b.
- For the definition we used the ordering of $\mathbb{Z}$.

Example

$\gcd(12, 18) = 6$ as the common divisors of 12 and 18 are $\pm 1, \pm 2, \pm 3$ and ± 6.

Definition

For $a, b \in \mathbb{Z}$ the **greatest common divisor** of a and b is the integer $\gcd(a, b) = d \geq 0$ for which

- d is a common divisor, i. e. $d|a$ and $d|b$ and
- for any $c \in \mathbb{Z}$, $c|a$, $c|b$ we have $c|d$.

Theorem

$\gcd(a, b)$ exists and is unique for any $a, b \in \mathbb{Z}$ and $\gcd(a, b) = \gcd(a, b)$.

Remark

- Thus we can always use the black gcd.
- After dropping the condition $d \geq 0$ this latter definition generalizes for any (nonunital) commutative ring. However, in general neither the existence nor the uniqueness are true any more.

Euclidean algorithm

Proof of the theorem and a method for computing $\gcd(a, b)$.

If $\gcd(a, b)$ exists, then it is unique and equals $\gcd(a, b)$.
Method (Euclidean algorithm):

1. If $b = 0$, then set $d = |a|$. If $b|a$, then set $d = |b|$.
   Otherwise set $r_0 = a$ and $r_1 = b$.

k. If $r_{k-1} = 0$, then $\gcd(a, b) = r_{k-2}$.
   Otherwise divide r_{k-2} by r_{k-1} with remainders:

$$r_{k-2} = q_{k-1} r_{k-1} + r_k$$

   Repeat this until $r_k = 0$.

Claims:

- This process is finite, since $r_1 > r_2 > \ldots$ are nonnegative numbers. Let the last step be the n-th, and set $d = r_{n-1}$.
- d is a common divisor of a, b.
- For any common divisor c we have $c|d$.
- Thus $d = \gcd(a, b)$.

Example

A graphical illustration of the Euclidean algorithm computing $\gcd(24, 17)$:

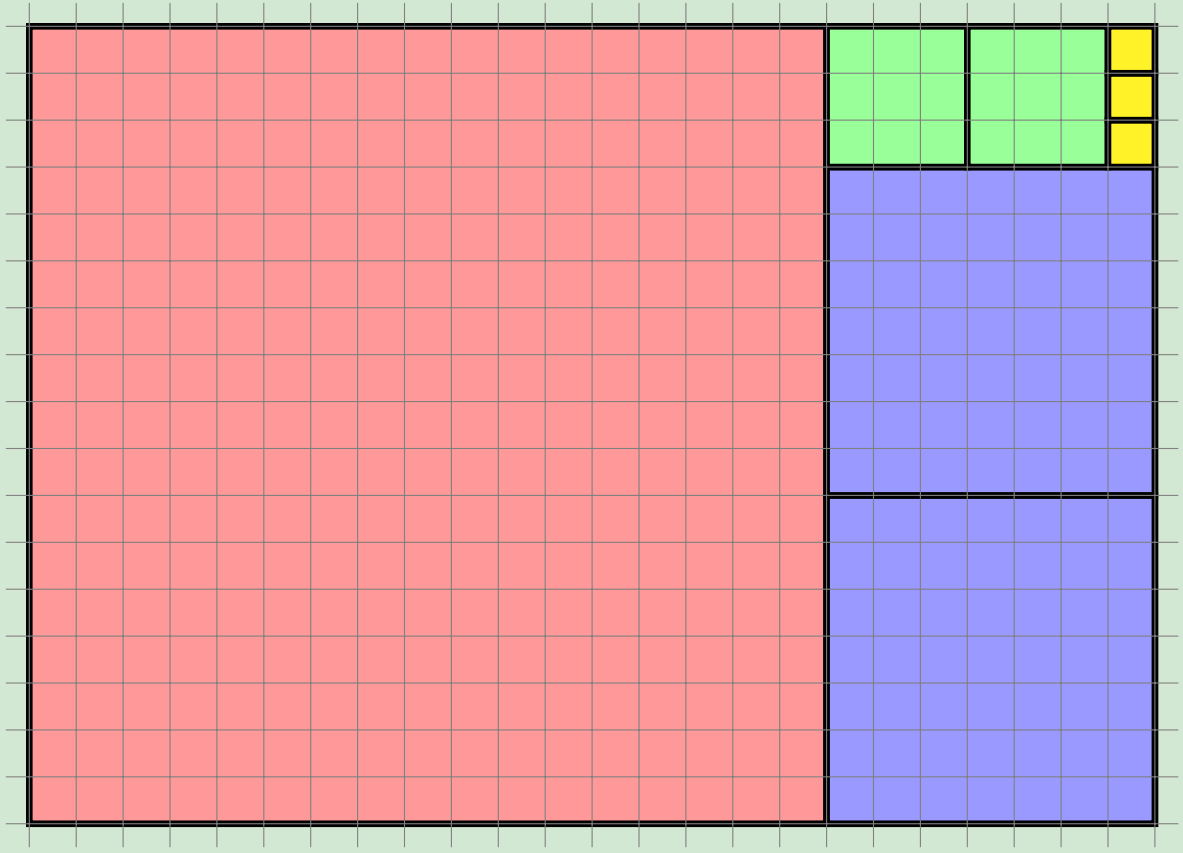
$$24 = 1 \cdot 17 + 7$$

$$17 = 2 \cdot 7 + 3$$

$$7 = 2 \cdot 3 + 1$$

$$3 = 3 \cdot 1 + 0$$

$$\gcd(24, 17) = 1$$



Remark

The Euclidean algorithm works also with the least residue in absolute value as well.

Definition

$a, b \in \mathbb{Z}$ are **relatively prime** if $\gcd(a, b) = 1$.

Theorem (Basic properties of the greatest common divisor)

For all $a, b, n \in \mathbb{Z}$ and $c \in \mathbb{N}$

1. $\gcd(ca, cb) = c \gcd(a, b)$,
2. if $d = \gcd(a, b) \neq 0$, then $\gcd\left(\frac{a}{d}, \frac{b}{d}\right) = 1$,
3. $\gcd(a + nb, b) = \gcd(a, b)$ and
4. if $c|ab$ and $\gcd(a, c) = 1$, then $c|b$.

Corollary

- $\{ma + nb | m, n \in \mathbb{Z}\} = \{c \cdot \gcd(a, b) | c \in \mathbb{Z}\}$.
- In particular there exist $m, n \in \mathbb{Z}$ such that
$$\gcd(a, b) = ma + nb.$$
In our method we can keep track of this in each step – that is called Extended Euclidean algorithm.

Diophantine equations

Definition

A **Diophantine equation** is a polynomial equation in 2 or more variables such that only the integer solutions are sought.

Example

- Pythagorean triples: $x^2 + y^2 = z^2$.
- Fermat's last theorem: If $x^n + y^n = z^n$ for fixed $n \geq 3$, then $xyz = 0$. (Proven by Wiles in 1995)
- Two-square problem: $x^2 + y^2 = n$ for a fixed $n \in \mathbb{N}$.
  For which n-s do we have a solution?
- Four-square problem: $x^2 + y^2 + z^2 + w^2 = n$ for a fixed $n \in \mathbb{N}$.
  This is solvable for all n.
- Pell's equation: $x^2 - ny^2 = 1$ for a fixed $n \in \mathbb{N}$.
  It has infinitely many solutions if n is not a perfect square.

Theorem (Solutions of linear Diophantine equations)

Let a, b and c be integers, $a, b \neq 0$ and $d = \gcd(a, b)$. The linear Diophantine equation

$$ax + by = c$$

is solvable if and only if $d|c$. If an integer solution is x_0, y_0 then all of the solution is in the form

$$x = x_0 + \frac{b}{d}t, y = y_0 - \frac{a}{d}t \qquad \text{for some } t \in \mathbb{Z}.$$

Remark (More variables)

The linear Diophantine equation

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k = c$$

is solvable if and only if $d|c$ for

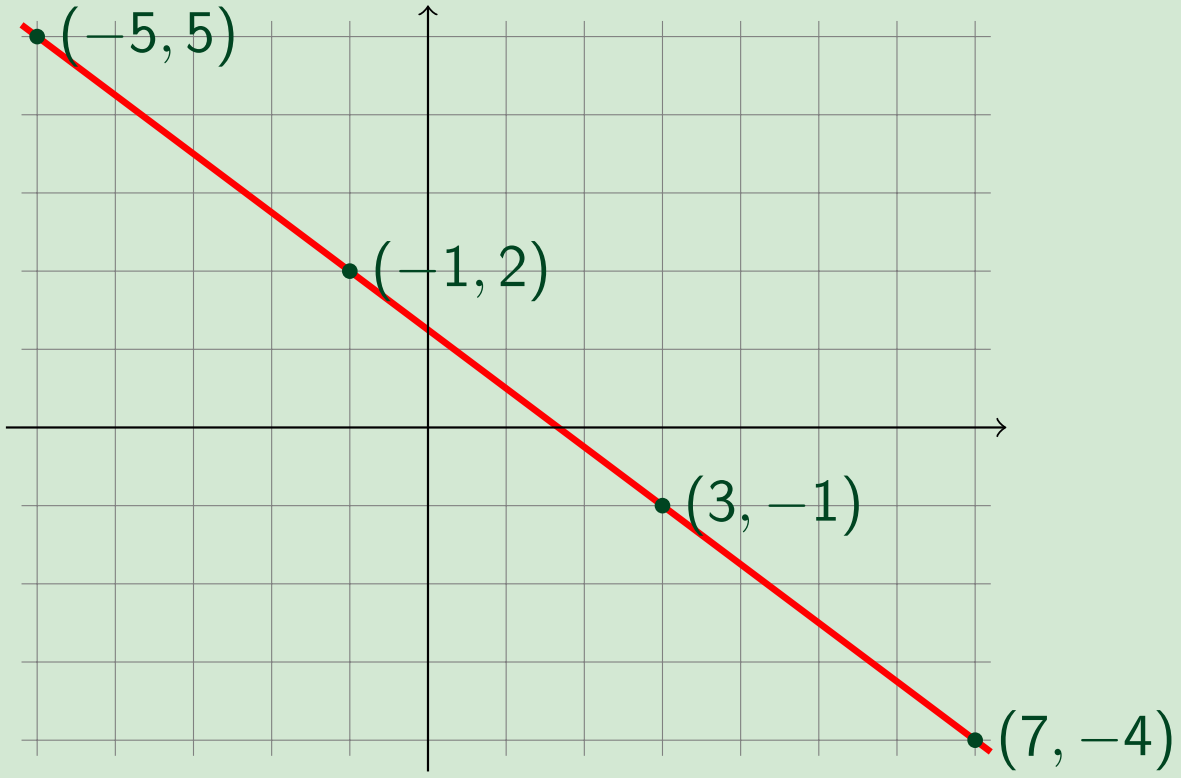
$$d = \gcd(a_1, a_2, \ldots, a_k) = \gcd(\ldots(\gcd(\gcd(a_1, a_2), a_3), \ldots), a_k).$$

and the solutions can be parametrized by some $t_1, t_2, \ldots t_{k-1} \in \mathbb{Z}$.

Example (Solutions of $3x + 4y = 5$)

It is solvable since $d = \gcd(3, 4) = 1|5$. Write $d = 4 - 3$ (by the first glimpse or by the extended Euclidean algorithm), so $5 = 3 \cdot (-5) + 4 \cdot 5$ and hence $x_0 = -5$, $y_0 = 5$. So the general solution is $x = -5 + 4t, y = 5 - 3t \quad (t \in \mathbb{Z})$.

It can be interpreted as finding lattice points on the line $y = (5 - 3x)/4$:



Irreducible numbers

Definition

An integer p is irreducible if p is not a unit and $p = ab \in \mathbb{Z}$ implies that either a or b is a unit.

Remark

- This makes sense for any (nonunital) commutative ring.
- Usually it is said that a natural number p is irreducible if $p \neq 1$ and $p = ab \in \mathbb{N} \implies a = 1$ or $b = 1$. $n \in \mathbb{N}$ is a composite number if $n = ab$ for some $a, b \in \mathbb{N}$, $a, b < n$.

Example

- $2, 3, 5, 7, 11$ are irreducibles $4, 6, 8, 9$ are composite numbers.
- $2, 7 \in \mathbb{Z}[\sqrt{2}]$ are not irreducibles but $2 + \sqrt{2} = \sqrt{2}(1 + \sqrt{2})$ is.

Lemma

For any $n > 1$ be an integer there exists $d \in \mathbb{N}$ such that $d|n$ and d is irreducible.

Prime numbers

Definition

An integer p is a prime number if

- $p \neq 0$ and p is not a unit,
- $p|ab \in \mathbb{Z}$ implies either $p|a$ or $p|b$.

Remark

- This makes sense for any (nonunital) commutative ring.
- Usually it is said that a natural number p is irreducible if $p > 1$ and $p|ab \in \mathbb{N} \implies p|a$ or $p|b$.

Theorem

$p \in \mathbb{Z}$ is irreducible $\iff$ $p \in \mathbb{Z}$ is prime.

Remark

- What happens in E – the nonunital ring of even numbers?
- $\impliedby$ is true in a more general setting, $\implies$ is not.

Theorem (Euclid, ∼300BC.)

In $\mathbb{N}$ (or in $\mathbb{Z}$) there exist infinitely many prime numbers.

Lemma

If n is a composite number, then it has a prime divisor $p \leq \sqrt{n}$.

Remark

Sieve of Eratosthenes - see gif file.

Remark (Irreducibility tests and factorization)

- The AKS (Agrawal-Kayal-Saxena, 2002) test is a "fast" (polynomial time) primality test.
- No "fast" algorithm is known for factorization.
- If there were some, then it would make lot of trouble: most of open-key encryption systems (such as RSA) are based on that factorization is "slow".

Remark (Theorems about the distribution of prime numbers)

- (Dirichlet, 1837) Let $a, b \in \mathbb{N}$ be relatively primes. There exist infinitely many primes in the arithmetic progression $an + b$.
- (Green, Tao, 2006) For any $n \in \mathbb{N}$ there exists an arithmetic progression of length n containing prime numbers.
- Prime number theorem (Hadamard, de la Vallée-Poussin, 1896) Let $\pi(x)$ denote the number of primes $2 \leq p \leq x$, then
$$\pi(x) \sim \frac{x}{\ln x}, \quad \text{i. e. } \lim_{x \to \infty} \frac{\pi(x)}{x/\ln x} = 1.$$
- Twin prime conjecture: There are infinitely many primes p such that $q = p + 2$ is also prime.
- (Zhang, Maynard, Tao, 2013-2014) There are infinitely pairs of primes $p < q$ such that $q - p \leq 246$.
- Goldbach's conjecture: Every even integer greater than 2 is the sum of two primes.
- Goldbach's weak conjecture: Every integer larger than 5 is the sum of three primes. Proven for sufficiently large integers (Vinogradov, 1937) and in general (Helfgott, 2013).

The fundamental theorem of Number Theory

Theorem

Any natural number $n > 1$ decomposes as the product of finitely many primes. This decomposition is unique up to the order of the factors.

Definition

The **canonical representation** of an integer n is the above decomposition with collecting the same primes to powers.

Example

$30 = 2 \cdot 3 \cdot 5$, $720 = 2^4 \cdot 3^2 \cdot 5$ and $2020 = 2^2 \cdot 5 \cdot 101$.

Theorem

Any integer n, which is nonzero and nonunit, decomposes as a product of primes. This decomposition is unique up to the order of the factors and multiplying with units.

Definition

Let a and b be nonzero integers. The **least common multiple** of a and b is the least positive integer m such thet $a|m$ and $b|m$. The notation is $\operatorname{lcm}(a,b) = [a,b] = m$.

Theorem

Let $a, b > 1$ integers and let $\{p_1, p_2, \ldots, p_r\}$ be the set of prime divisors of ab. Write $a = p_1^{a_1} p_2^{a_2} \ldots p_r^{a_r}$ and $b = p_1^{b_1} p_2^{b_2} \ldots p_r^{b_r}$ for some $a_i, b_i \in \mathbb{N} \cup \{0\}$, $i = 1, 2, \ldots, r$. Then

1. $a|b \iff a_i \leq b_i$ for all $i = 1, 2, \ldots, r$,
2. $\gcd(a,b) = p_1^{\min(a_1,b_1)} p_2^{\min(a_2,b_2)} \ldots p_r^{\min(a_r,b_r)}$,
3. $\operatorname{lcm}(a,b) = p_1^{\max(a_1,b_1)} p_2^{\max(a_2,b_2)} \ldots p_r^{\max(a_r,b_r)}$ and
4. $\gcd(a,b) \cdot \operatorname{lcm}(a,b) = ab$.

Theorem (Legendre's formula)

In the canonical representation of $n!$ the exponent of a prime p is

$$\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \ldots$$

Modular arithmetic – Computing with residues

Congruences and residue classes

Definition

Let $a, b \in \mathbb{Z}$ and $m \in \mathbb{N}$.

- The remainder of the division of a by m is called "a **modulo** m" and denoted by $a \mod m$.
- a is **congruent** to b modulo m if $m|a-b$. The notation is $a \equiv b \pmod{m}$.

Theorem

The modulo m congruence relation ($\equiv \pmod{m}$) is an equivalence relation, i. e. the following hold for all $a, b, c \in \mathbb{Z}$:

1. Reflexivity: $a \equiv a \pmod{m}$,
2. Symmetry: $a \equiv b \pmod{m} \iff b \equiv a \pmod{m}$ and
3. Transitivity: $a \equiv b, b \equiv c \pmod{m} \implies a \equiv c \pmod{m}$.

Remark

In other words $\equiv \pmod{m}$ partitions $\mathbb{Z}$ to disjoint subsets (classes).

Definition

- The equivalnece classes of the above relation are called the modulo m **residue classes**.
- A set of integers R is called a **complete residue system** or **CRS** modulo m if it contains exactly one of each residue classes modulo m.

Example

- $\{0, 1, 2, \ldots, m-1\}$ is a CRS modulo m.
- $\{100, 123, 116\}$ is a CRS modulo 3.

Lemma

1. $\{r_1, r_2, \ldots, r_m\}$ is a CRS mod m $\iff$ $r_i \not\equiv r_j \pmod{m}$ for all $1 \leq i < j \leq m$.
2. $\{r_1, r_2, \ldots, r_m\}$ is a CRS mod m $\implies$ for any $a, b \in \mathbb{Z}$, $\gcd(a, m) = 1$ $\{ar_1 + b, ar_2 + b, \ldots, ar_m + b\}$ is also a CRS mod m.

Operations with congruences

Theorem

Let $a, b, c, d \in \mathbb{Z}$, $n \in \mathbb{N}$ such that $a \equiv c, b \equiv d \pmod{m}$. Then

1. $a + b \equiv c + d \pmod{m}$, especially $a + b \equiv c + b \pmod{m}$,
2. $a - b \equiv c - d \pmod{m}$, especially $a - b \equiv c - b \pmod{m}$ and
3. $ab \equiv cd \pmod{m}$, especially $ab \equiv cb \pmod{m}$.

Remark

So we can add, subtract or multiply congruences or residue classes without changing the modulus.

Example

(1) : $5 \equiv 11 \pmod{3}$ and (2) : $2 \equiv 14 \pmod{3}$. Hence

- (1) + (2) : $7 \equiv 25 \pmod{3}$,
- (1) − (2) : $3 \equiv -3 \pmod{3}$ and
- (1) · (2) : $10 \equiv 156 \pmod{3}$.

Dividing congruences

Example (Be careful with division!)

- $4 \equiv 10 \pmod 3$ divided by 2 is $2 \equiv 5 \pmod 3$,
- $6 \equiv 15 \pmod 9$ divided by 3 is $2 \equiv 5 \pmod 3$ and **not** $2 \equiv 5 \pmod 9$ – which is false.

Theorem

Let $a, b, c \in \mathbb{Z}$, $m \in \mathbb{N}$ and $d = \gcd(c, m)$ such that $ac \equiv bc \pmod m$. Then

1. $a \equiv b \pmod{m/d}$ and
2. if $d = \gcd(c, m) = 1$, then $a \equiv b \pmod m$.

Theorem

Let $m_1, m_2, \ldots m_r \in \mathbb{N}$ and $m = \operatorname{lcm}(m_1, m_2, \ldots, m_r)$. Then $a \equiv b \pmod{m_k}$ for $k = 1, 2, \ldots, r \implies a \equiv b \pmod m$.

Computing powers of congruences

Theorem

$m, n \in \mathbb{N}, a \equiv b \pmod{m} \implies a^n \equiv b^n \pmod{m}$.

Remark

For computing $a^n \pmod{m}$ efficiently see the following scheme:

- If n is even, say $n = 2k$, then $a^n \equiv (a^k)^2 \pmod{m}$ and
- If n is odd, say $n = 2k + 1$, then $a^n \equiv (a^k)^2 \cdot a \pmod{m}$.

Theorem

Let $a, n \in \mathbb{N}$ such that $a^n - 1$ is prime. Then $a = 2$ and n is prime.

Definition

A prime $p \in \mathbb{N}$ is a **Mersenne prime**, if $p = 2^n - 1$ for some $n \in \mathbb{N}$.

Linear congruences

Theorem

Let $a, b \in \mathbb{Z}$, $m \in \mathbb{N}$ and $d = \gcd(a, m)$. Then

$$ax \equiv b \pmod{m} \text{ is solvable } \iff d|b.$$

If the congruence is solvable, then the number of incongruent solutions modulo m is d.

Example

The solutions of $12x \equiv 15 \pmod{21}$ are $x \equiv 3, 10, 17 \pmod{21}$.

Definition

If the congruence $ax \equiv 1 \pmod{m}$ is solvable, then a solution is called the **modular inverse** of $a \pmod{m}$. The notation is $a^{-1} \pmod{m}$.

Example

A modular inverse of $3 \pmod{10}$ is 7, but $2 \pmod{10}$ has no modular inverse.

Example

Are the following systems of linear congruences solvable?

a) $x \equiv 1 \pmod 3$
$x \equiv 1 \pmod 5$
$x \equiv 1 \pmod 7$

b) $y \equiv 2 \pmod 3$
$y \equiv 4 \pmod 5$
$y \equiv 6 \pmod 7$

c) $z \equiv 2 \pmod 3$
$z \equiv 3 \pmod 5$
$z \equiv 4 \pmod 7$

Theorem (Chinese remainder theorem)

Let $m_1, m_2, \ldots, m_r \in \mathbb{N}$ pairwise relatively prime, $b_1, b_2, \ldots, b_r \in \mathbb{Z}$ and let $M = m_1 \cdot m_2 \cdot \ldots \cdot m_r$. Then the system of congruences

$$\begin{aligned} x &\equiv b_1 \pmod{m_1} \\ x &\equiv b_2 \pmod{m_2} \\ &\vdots \\ x &\equiv b_r \pmod{m_r} \end{aligned}$$

has a unique solution modulo M.

Some applications

- Division rule for 9. $x \equiv$ the sum of digits of $x \pmod 9$:
  $x = \sum_{k=0}^{n} x_k 10^k \equiv \sum_{k=0}^{n} x_k \pmod 9$.
- Same for 11. $x \equiv$ the alternating sum of digits of $x \pmod{11}$:
  $x = \sum_{k=0}^{n} x_k 10^k \equiv \sum_{k=0}^{n} (-1)^k x_k \pmod{11}$.
- ISBN-13 and EAN (European Article Number):
  $x_1 + 3x_2 + x_3 + 3x_4 + x_5 + 3x_6 + \ldots + 3x_{12} + x_{13} \equiv 0 \pmod{10}$
- Perpetual calendar: What day is a date $(d - m - y)$?

$$w = d + \left\lfloor \frac{13m - 32}{5} \right\rfloor + y + \left\lfloor \frac{y}{4} \right\rfloor - \left\lfloor \frac{y}{100} \right\rfloor + \left\lfloor \frac{y}{400} \right\rfloor \pmod 7 + 1$$

$w = 1$ is Monday, $w = 2$ is Tuesday, $\ldots$ and $w = 7$ is Sunday.

The ring of modulo m residue classes

Definition

Let $m > 1$ be an integer, $\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}$ and define the (binary) operations $\oplus : a \oplus b = a + b \pmod{m}$ and $\otimes : a \otimes b = a \cdot b \pmod{m}$.

Theorem

$(\mathbb{Z}_m, \oplus, \otimes)$ is a ring for all integer $m > 1$.

Example (Operations on $\mathbb{Z}_2$)

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

Definition

We use $+$ and $\cdot$ instead of $\oplus$ and $\otimes$ – if it does not lead to confusion – and denote $\mathbb{Z}_m = (\mathbb{Z}_m, +, \cdot)$ the ring of modulo m residue classes.

Domains and fields

Lemma

Let R be a ring and $a \in R$. Then $a \cdot 0 = 0 = 0 \cdot a$.

Definition

Let R be a commutative ring and $a \in R$.

- $a \in R - \{0\}$ is a **zero divisor** if there exists $b \neq 0$ such that $ab = 0$,
- R is a **domain** if it has no zero divisors,
- R is a **field** if $0 \neq 1 \in R$ and there exists multiplicative inverse: for all $a \in R - \{0\}$ there exists $a^{-1} \in R$ such that $a \cdot a^{-1} = 1$.

Example

$\mathbb{Q}, \mathbb{R}$ are fields, but $\mathbb{Z}$ is not. All three are domains.

Theorem

R is a field $\implies$ R is a domain.

Example (The operations on $\mathbb{Z}_5$ and $\mathbb{Z}_6$)

$\mathbb{Z}_5$:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\times$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

$\mathbb{Z}_6$:

$+$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$\cdot$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Theorem

Let $m > 1$ be an integer. The following are equivalent:

1. m is a prime number,
2. $\mathbb{Z}_m$ is a field and
3. $\mathbb{Z}_m$ is a domain.

Reduced residue systems and Euler's totient function

Lemma

If $a \equiv b \pmod{m}$, then $\gcd(a, m) = \gcd(b, m)$.

Definition

Let $m \in \mathbb{N}$, $m > 1$. Then

- a modulo m residue class is **reduced** if an (or every) element of it is relatively prime to m,
- a set of integers R is a **reduced residue system** or **RRS** modulo m if it contains exactly one of each reduced residue classes modulo m,
- the number of reduced residue classes modulo m is denoted $\varphi(m)$ and φ is called **Euler's totient function**.

Example (Some RRS-s)

$\{1, 2\} \pmod{3} \implies \varphi(3) = 2$, $\{1, 11\} \pmod{6} \implies \varphi(6) = 2$ and $\{n \in \mathbb{N} | 1 \leq m < n, \gcd(m, n) = 1\} \pmod{m}$.

Lemma

1. $R = \{r_1, r_2, \ldots, r_k\}$ is a RRS mod m $\iff$ $|R| = \varphi(m)$, $\gcd(r_i, m) = 1$ and $r_i \not\equiv r_j \pmod{m}$ for all $1 \leq i < j \leq k$.

2. $R = \{r_1, r_2, \ldots, r_k\}$ is a RRS mod m $\implies$ for any $a \in \mathbb{Z}$, $\gcd(a, m) = 1$ $\{ar_1, ar_2, \ldots, ar_k\}$ is also a RRS mod m.

Example

- $\mathbb{Z}_m^* = \{a \in \mathbb{Z}_m | (a, m) = 1\} \subset \mathbb{Z}_m$ is the subset of reduced residue classes. This is not closed under the operation $+$, so $(\mathbb{Z}_m^*, +, \cdot)$ is **not** a ring.
- However the operation $\cdot$ is well defined is associative, has unit and inverse. $\mathbb{Z}_m^* = (\mathbb{Z}_m^*, \cdot)$ and such structures will be called **groups**.

Theorem

φ is a multiplicative arithmetic function, i. e. if $\gcd(m, n) = 1$ then $\varphi(mn) = \varphi(m)\varphi(n)$.

Remark

It is necessary, that $\gcd(m, n) = 1$. For example
$6 = \varphi(9) \neq \varphi(3)\varphi(3) = 2 \cdot 2 = 4$.

Example (Chinese remainder theorem revisited)

Let $m_1 = 4$ and $m_2 = 3$. The Chinese remainder theorem produces a mapping from pairs of modulo 3 and modulo 4 residue classes to modulo 12 residue classes:

		0	**1**	2	**3**	$\leftarrow$	$\mathbb{Z}_4$
	0	0	9	6	3		
$\mathbb{Z}_3$ $\rightarrow$	**1**	4	**1**	10	**7**	$\leftarrow$	$\mathbb{Z}_{12}$
	2	8	**5**	2	**11**		

The operations and the reduced residue classes are also preserved.

Homomorphisms and isomorphisms

Definition

Let $(R, +, \cdot), (R', \oplus, \otimes)$ be rings and $f : R \to R'$ be a function.

- f is a **homomorphism** if $f(1_R) = 1_{R'}$ and it preserves operations, i. e. $f(a + b) = f(a) \oplus f(b)$ and $f(a \cdot b) = f(a) \otimes f(b)$ for all $a, b \in R$.
- f is an **isomorphism** if it is a bijective homomorphism.

Example

Let $m, n > 1$ be integers.

- The map $\mathbb{Z} \to \mathbb{Z}_m$, $n \mapsto (n \mod m)$ is a homomorphism.
- If $\gcd(m, n) = 1$, then the Chinese remainder theorem gives an isomorphism $\mathbb{Z}_m \times \mathbb{Z}_n \to \mathbb{Z}_{mn}$ (here $\times$ is the Cartesian product on the sets and the operations work elementwise).
- If $\gcd(m, n) > 1$, then there is no isomorphism $\mathbb{Z}_n \times \mathbb{Z}_m \to \mathbb{Z}_{mn}$ – check that the image of a unit should be a unit and count units in both rings.

Lemma

Let $q = p^{\alpha}$ be a prime power. Then $\varphi(q) = (p-1)p^{\alpha-1}$.

Theorem (The canonical form of $\varphi(m)$)

Let $m = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ be the canonical representation of $1 < m \in \mathbb{N}$. Then

$$\varphi(m) = (p_1-1)p^{\alpha_1-1}(p_2-1)p^{\alpha_2-1}\dots(p_r-1)p^{\alpha_r-1} = n\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\dots\left(1-\frac{1}{p_r}\right)$$

or using $\prod$ sign : $\varphi(m) = \prod_{k=1}^{r}(p_k-1)p_k^{\alpha_k-1} = n\prod_{k=1}^{r}\left(1-\frac{1}{p_k}\right)$.

Example

- $\varphi(1200) = \varphi(2^4 \cdot 3 \cdot 5^2) = (2-1)2^3(3-1)(5-1)5 = 320$,
- $\varphi(2020) = \varphi(2^2 \cdot 5 \cdot 101) = (2-1)2(5-1)(101-1) = 800$.

Euler-Fermat theorem

Theorem (Euler-Fermat theorem)

Let $m > 1$ be an integer, and $a \in \mathbb{Z}$ such that $\gcd(a, m) = 1$. Then $a^{\varphi(m)} \equiv 1 \pmod{m}$.

Example

$(\pm 1)^4 \equiv (\pm 5)^4 \equiv 1 \pmod{12}$. But $2^4 \equiv 4, 3^4 \equiv 9, 4^4 \equiv 4$ and $6^4 \equiv 0 \pmod{12}$.

Corollary (Fermat's little theorem)

Let $p > 0$ be a prime and $a \in \mathbb{Z}$. Then

- First form: If $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$ and
- Second form: $a^p \equiv a \pmod{p}$.

Example

$1^4 = 1 \equiv 2^4 = 16 \equiv 3^4 = 81 \equiv 4^4 = 256 \pmod{5}$.

Corollary

If $\gcd(a, m) = 1$, then $a^n \equiv (a \mod m)^{n \mod \varphi(m)} \pmod{m}$.

Example

- $$18^{53} \equiv (18 \mod 13)^{53 \mod \varphi(13)} \equiv 5^5 = (5^2)^2 \cdot 5 \equiv (-1)^2 \cdot 5 \equiv 5 \pmod{13}.$$
- $$91^{89} \equiv (91 \mod 11)^{89 \mod \varphi(11)} = 3^9 \equiv 3^{-1} = \frac{1}{3} \equiv \frac{12}{3} = 4 \pmod{11}.$$

Remark

For the method for computing powers modulo m we do not need the prime factorization of m, but for computing $\varphi(m)$ we do. This makes the latter "slower", since no efficient algorithm is known for prime factorization.

Theorem (Wilson's theorem)

$(p-1)! \equiv -1 \pmod{p}$ for all prime numbers $p > 0$.

Complex numbers

Definition and algebraic properties

Definition (The set)

- $i = \sqrt{-1}$,
- $\mathbb{C} = \{a + bi | a, b \in \mathbb{R}\}$ is the set of **complex numbers**,
- the **real part** of $z = a + bi$ is $\operatorname{Re} z = a$ and
- the **imaginary part** of $z = a + bi$ is $\operatorname{Im} z = b$.

Definition (The opertaions)

- Addition: $(a + bi) + (c + di) = (a + c) + (b + d)i$,
- Multiplication: $(a + bi)(c + di) = (ac - bd) + (bc + ad)i$

Example

$(1 + 2i) + (3 + 4i) = (4 + 6i)$, $(1 + 2i)(3 + 4i) = -5 + 10i$.

Theorem

$z_1 = z_2 \iff \operatorname{Re} z_1 = \operatorname{Re} z_2$ and $\operatorname{Im} z_1 = \operatorname{Im} z_2$.

Theorem

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

Remark

- Complex numbers can be represented as a plane of numbers. $z = a + bi$ corresponds to the point with coordinates (a, b). Addition corresponds to addition of position vectors.
- The division goes as follows:
$$\frac{a+bi}{c+di} = \frac{a+bi}{c+di} \cdot \frac{c-di}{c-di} = \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2} i \in \mathbb{C}$$

Example

$$\frac{1}{3+4i} = \frac{3-4i}{3^2+4^2} = \frac{3}{25} - \frac{4}{25} i$$
$$\frac{3+4i}{1+i} = \frac{(3+4i)(1-i)}{1^2+1^2} = \frac{7+i}{2} = \frac{7}{2} + \frac{1}{2} i$$

Remark

The set $\mathbb{C}$ is **not** ordered: the relation $z_1 < z_2$ makes no sense in $\mathbb{C}$.

The conjugate and the absolute value

Definition

- the **conjugate** of $z = a + bi$ is $\overline{z} = a - bi$ and
- the **absolute value** of $z = a + bi$ is $|z| = \sqrt{a^2 + b^2}$.

Theorem

If $z = a + bi, z_1, z_2 \in \mathbb{C}$ then

- $z = \overline{z} \iff z \in \mathbb{R}$
- $z + \overline{z} = 2a \in \mathbb{R}$,
- $z\overline{z} = a^2 + b^2 = |z|^2 \in \mathbb{R}$,
- $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2$ és
- $\overline{z_1 z_2} = \overline{z}_1 \overline{z}_2$
- $|z| \in \mathbb{R}, |z| \geq 0$
- $|z| = 0 \iff z = 0$,
- $|z_1 + z_2| \leq |z_1| + |z_2|$ és
- $|z_1 z_2| = |z_1||z_2|$.

Lemma

If $p(x)$ is polynomial with real coefficients, then $\overline{p(z)} = p(\overline{z})$, thus $p(z) = 0 \implies p(\overline{z}) = 0$.

The fundamental theorem of Algebra

Definition

A field $\mathbb{F}$ is **algebraically closed** if any nonconstant polynomial with coefficients in $\mathbb{F}$ has a solution in $\mathbb{F}$.

Theorem (The fundamental theorem of Algebra)

$\mathbb{C}$ is algebraically closed.

Theorem

Every field $\mathbb{F}$ can be embedded into an algebraically closed field.

Definition

The smallest algebraically closed extension of a field $\mathbb{F}$ (with respect to inclusion) is called the **algebraic closure** of $\mathbb{F}$.

Remark

- The algebraic closure of $\mathbb{R}$ is $\mathbb{C}$,
- but the algebraic closure of $\mathbb{Q}$ is **not** $\mathbb{C}$.

Theorem

If $w \in \mathbb{C}$ is a nonzero complex number, then the equation $z^2 = w$ has exactly two solutions.

Definition

Let $z \in \mathbb{C}$ and $n \in \mathbb{N}$. The solutions of $u^n = z$ in $\mathbb{C}$ are called the **n-th roots** of z and the set of them is denoted $\sqrt[n]{z}$.

Remark (Reasons for defining $\sqrt{z}$ to be multivalued)

What is wrong:
$1 = 1 \cdot 1 = \sqrt{1}\sqrt{1} = \sqrt{1 \cdot 1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = i \cdot i = -1$?

- There is no natural way to choose one of the roots and
- then $\sqrt{z_1 z_2} = \sqrt{z_1}\sqrt{z_2}$ holds.

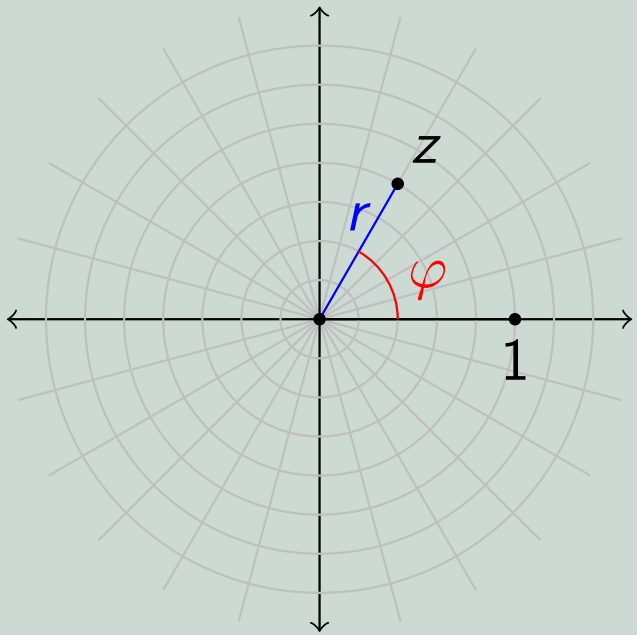
Theorem (Solution of quadratic equations in $\mathbb{C}$)

The complex solutions of $az^2 + bz + c = 0$ $(a, b, c \in \mathbb{C})$ are $\dfrac{-b + \sqrt{b^2 - 4ac}}{2a}$ ($+$ is written instead of $\pm$ intentionally).

Trigonometric form and geometric properties

Definition

- The **argument** of $z \in \mathbb{C}$ is the angle of the line segments $\overline{0z}$ and $\overline{01}$ in the plane (modulo 360° or 2π). $\arg 0$ is not defined. The notation is $\arg z$.
- The **trigonometric form** of $z \neq 0$ is $z = r(\cos\varphi + i\sin\varphi)$, where $r = |z|$ and $\varphi = \arg z$.
- The (original) **algebraic form** of z is $z = \mathrm{Re}(z) + \mathrm{Im}(z)i$.



Remark

The notation $z = re^{i\varphi} = r(\cos\varphi + i\sin\varphi)$ is commonly used, you can understand the meaning of this with Taylor series (Calculus). Here φ must be in radian.

Theorem (The trigonometric form's non-uniqueness)

The following are equivalent:

1. $r(\cos\varphi + i\sin\varphi) = r'(\cos\varphi' + i\sin\varphi')$,
2. $r = r'$ and $\varphi = \varphi' + 2k\pi$ for some $k \in \mathbb{Z}$.

Theorem

- If $z \neq 0$ has trigonometric form $z = r \cdot (\cos\varphi + i\sin\varphi)$, then its algebraic form is $z = a + bi$, where $a = r \cdot \cos\varphi$ and $b = r \cdot \sin\varphi$.
- If $z \neq 0$ has algebraic form $z = a + bi$, then its trigonometric form is $z = r(\cos\varphi + i\sin\varphi)$, where $r = |z| = \sqrt{a^2 + b^2}$ and

$$\varphi = \begin{cases} \pi/2, & \text{if } a = 0 \text{ and } b > 0 \\ -\pi/2, & \text{if } a = 0 \text{ and } b < 0 \\ \arctan(b/a), & \text{if } a > 0 \\ \arctan(b/a) + \pi, & \text{if } a < 0 \text{ and } b \geq 0 \\ \arctan(b/a) - \pi, & \text{if } a < 0 \text{ and } b < 0 \end{cases}$$

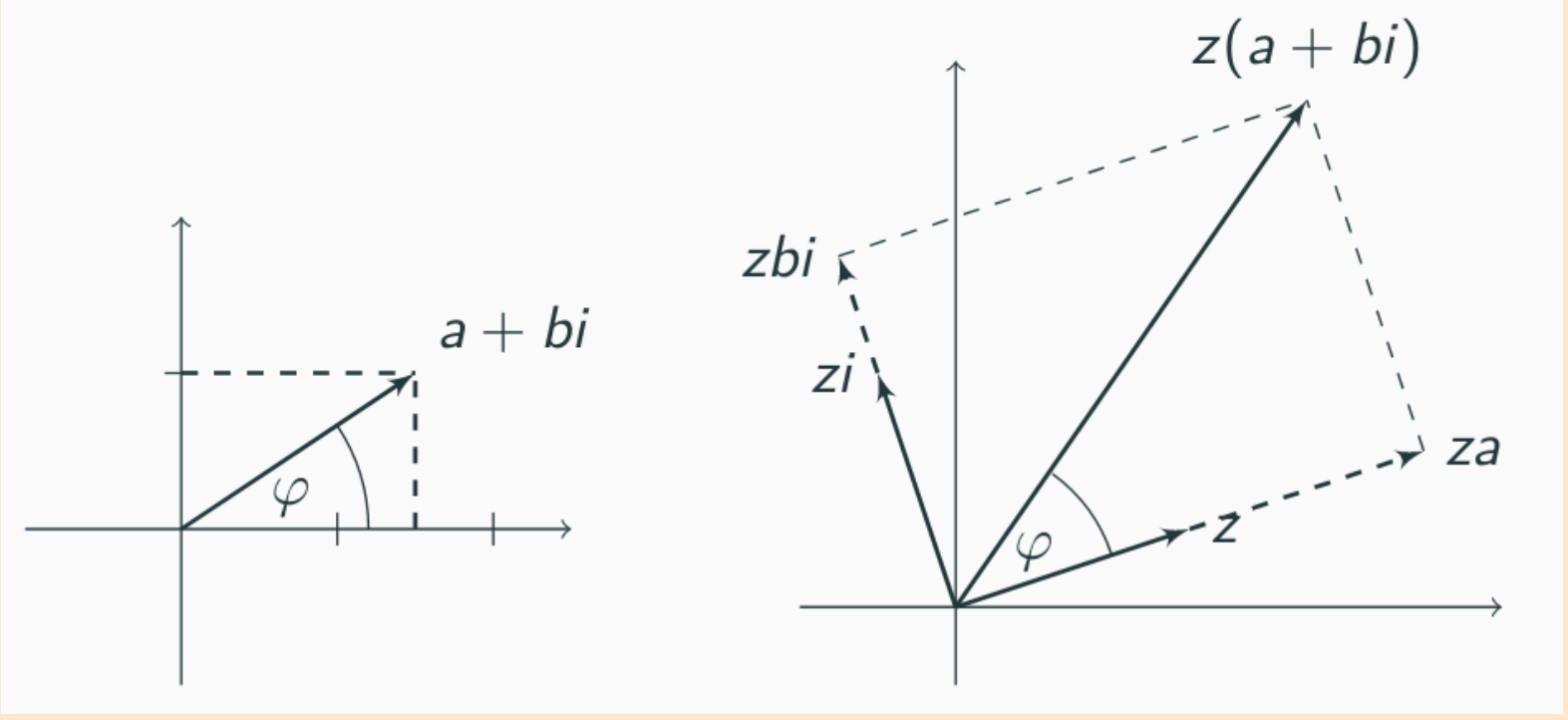
The geometric meaning of multiplication

Theorem

$\arg(z \cdot z') = \arg(z) + \arg(z')$ and $|z \cdot z'| = |z| \cdot |z'|$.

Remark

Thus "multiplying complex number we have to multiply the length and add the argument".



Powers and roots

Corollary

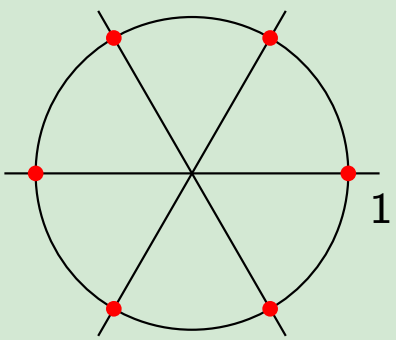
- $\arg(1/z) = -\arg(z)$ and $\arg(z/z') = \arg(z) - \arg(z')$,
- $(r(\cos\varphi + i\sin\varphi))^n = r^n(\cos(n\varphi) + i\sin(n\varphi))$,
- The n-th roots of $z = r(\cos\varphi + i\sin\varphi) \neq 0$ are

$$\sqrt[n]{z} = \sqrt[n]{r}\left(\cos\left(\frac{\varphi + 2k\pi}{n}\right) + i\sin\left(\frac{\varphi + 2k\pi}{n}\right)\right)$$

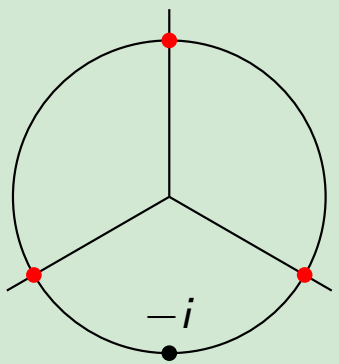
for $k = 0, 1, \ldots, n-1$.

Example

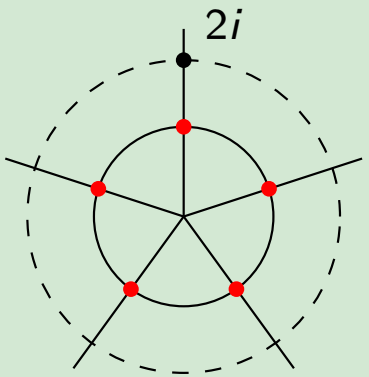
$\sqrt[6]{1}$:



$\sqrt[3]{-i}$:



$\sqrt[5]{2i}$:



Remark (Which form is "comfortable" for different operations)

	$z_1 + z_2$ $z_1 - z_2$	$z_1 z_2$ z_1 / z_2	z^n $\sqrt[n]{z}$
algebraic form	OK	OK	X
trigonometric form	X	OK	OK

Lemma

If $w = r(\cos\varphi + i\sin\varphi)$, then the map $z \mapsto z \cdot w$ corresponds to the following transformation of the plane: scaling with factor r and rotation around 0 with angle φ counterclockwise.

Example

- Rotating around 0 with angle φ is $z \mapsto z(\cos\varphi + i\sin\varphi)$.
- Rotating around u with angle φ is $z \mapsto (z-u)(\cos\varphi + i\sin\varphi) + u$.
- Reflecting to the axis $\arg z = \varphi$ is $z \mapsto \overline{z}(\cos(2\varphi) + i\sin(2\varphi))$.

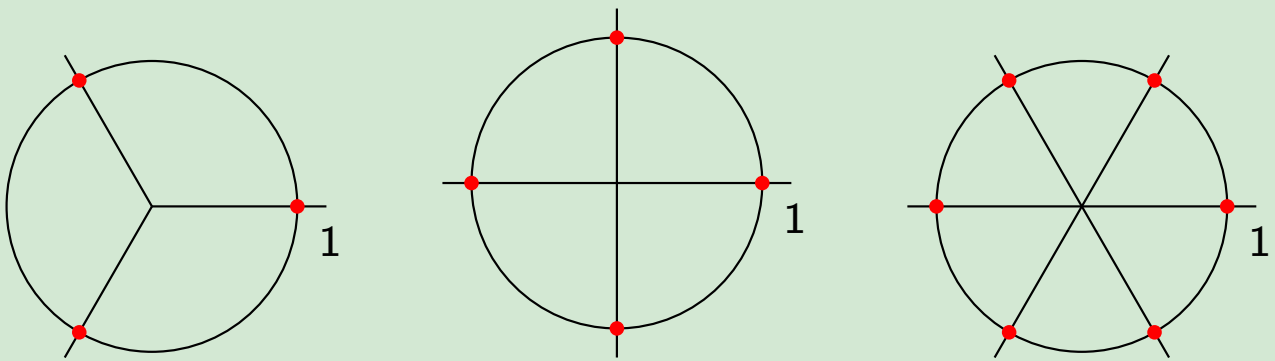
Roots of unity

Definition

$\varepsilon \in \mathbb{C}$ is a n-**th root of unity** (for $n \in \mathbb{N}$) if $\varepsilon^n = 1$.

Example

- The 2nd roots of unity are 1 and -1.
- The 3rd roots of unity are 1, $-\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $-\frac{1}{2} - \frac{\sqrt{3}}{2}i$.
- The 4th roots of unity are 1, i, -1 and $-i$.
- The 6th roots of unity are ± 1, $\pm\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$.



Remark

A d-th root of unity is also an nth root of unity if $d|n$.

Definition

- The **(multiplicative) order** $\varepsilon \in \mathbb{C}$ is
$$o(\varepsilon) = \min(n \in \mathbb{N} | \varepsilon^n = 1)$$
if there is such n and ∞ if not.
- $\varepsilon \in \mathbb{C}$ is a **primitive n-th root of unity** if $o(\varepsilon) = n$.

Lemma

If ε is an n-th root of unity, then $o(\varepsilon) | n$.

Example

$2i$ and $\cos(\sqrt{2}\pi) + i \sin(\sqrt{2}\pi)$ are **not** roots of unity.

Theorem

1. ε is a root of unity $\iff$ $|\varepsilon| = 1$ and $\arg(\varepsilon) = 2\pi r$ for some $r \in \mathbb{Q}$,
2. the order of ε is q, where $r = p/q \in \mathbb{Q}$ with $\gcd(p, q) = 1$.

Theorem

Let ε be a primitive n-th root of unity. Then

1. the n-th roots of unity are $1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}$ and
2. the primitive n-th roots of unity are ε^k for $k \in \mathbb{N}$ such that $1 \leq k < n$ and $\gcd(n, k) = 1$. That is the number of primitive n-th roots of unity is $\varphi(n)$.

Corollary

- ε is a primitive n-th root of unity $\iff$ all n-th roots of unity is a power of ε.
- $\sum_{d|n} \varphi(d) = n$ for all $n \in \mathbb{N}$.

Theorem

Let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ be the n-th roots of unity. Then

$$\sum_{k=1}^{n} \varepsilon_k = \begin{cases} 1, & \text{if } n = 1 \\ 0, & \text{if } n > 1 \end{cases} \quad \text{and} \quad \prod_{k=1}^{n} \varepsilon_k = \begin{cases} 1, & \text{if } n \text{ is odd} \\ -1, & \text{if } n \text{ is even} \end{cases}$$

Binomial coefficients

Definition

For $n, k \in \mathbb{Z}$, $0 \leq k \leq n$ the **binomial coefficient** (n choose k) is

$$\binom{n}{k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{1 \cdot 2 \cdot 3 \dots k} = \frac{n!}{k!(n-k)!},$$

that is the number of k element subsets of an n element set.

Remark

$0! = 1$ and it is worth to define $\binom{n}{k} = 0$ for $k > n$.
In fact one can define $\binom{x}{k}$ for any $x \in \mathbb{R}$.

Theorem

1. $\binom{n}{0} = \binom{n}{n} = 1$,
2. $\binom{n}{k} = \binom{n}{n-k}$,
3. $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$ and $\binom{n}{k} = \frac{n-k+1}{k}\binom{n}{k-1}$,
4. $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.

Example (Pascal's triangle)

$$
\begin{array}{ccccccccccccccccccc}
 & & & & & & \binom{0}{0} & & & & & & & & 1 & & & & \\
 & & & & \binom{1}{0} & & & & \binom{1}{1} & & & & & & 1 & & 1 & & \\
 & & & \binom{2}{0} & & & \binom{2}{1} & & & & \binom{2}{2} & & & & 1 & & 2 & & 1 \\
 & & \binom{3}{0} & & & \binom{3}{1} & & & \binom{3}{2} & & & \binom{3}{3} & & & 1 & & 3 & & 3 & & 1 \\
\binom{4}{0} & & & \binom{4}{1} & & & \binom{4}{2} & & & \binom{4}{3} & & & \binom{4}{4} & 1 & & 4 & & 6 & & 4 & & 1
\end{array}
$$

Theorem (Binomial theorem)

$$(a+b)^n = a^n + na^{n-1}b + \cdots + nab^{n-1} + b^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Remark

This is true in any commutative ring.

Corollary

- $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots = \sum_{k=0}^{n} \binom{n}{k} = 2^n,$
- $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots = \sum_{k=0}^{n} (-1)^k \binom{n}{k} = \begin{cases} 0, & \text{if } n > 0 \\ 1, & \text{if } n = 0 \end{cases}$

and

- $\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = \sum_{k=0}^{[n/2]} \binom{n}{2k} = \begin{cases} 2^{n-1}, & \text{if } n > 0 \\ 1, & \text{if } n = 0. \end{cases}$

Lemma

1. $\binom{n}{0} - \binom{n}{2} + \binom{n}{4} - \binom{n}{6} + \cdots = (\sqrt{2})^n \cos \frac{n\pi}{4},$
2. $\binom{n}{1} - \binom{n}{3} + \binom{n}{5} - \binom{n}{7} + \cdots = (\sqrt{2})^n \sin \frac{n\pi}{4}.$

Remark

Or in an other form:

$$\binom{n}{0} - \binom{n}{2} + \binom{n}{4} - \cdots = \begin{cases} 2^{(n-1)/2}, & \text{if } n \equiv \pm 1 \pmod 8 \\ -2^{(n-1)/2}, & \text{if } n \equiv \pm 3 \pmod 8 \\ 0, & \text{if } n \equiv \pm 2 \pmod 8 \\ -2^{n/2}, & \text{if } n \equiv 4 \pmod 8 \\ 2^{n/2}, & \text{if } n \equiv 0 \pmod 8 \end{cases}$$

Example (Extra problem)

What is the explicit value of

$$\binom{n}{0} + \binom{n}{4} + \binom{n}{8} + \binom{n}{12} + \cdots = \sum_{k=0}^{[n/4]} \binom{n}{4k}?$$

Polynomials

Basic notions

Definition

Let R be a commutative ring, $n \in \mathbb{N}$ and $a_0, a_1, \ldots, a_n \in R$. The formal expression $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ is called **polynomial** (in one variable x).
Let $p(x) = \sum_{l=0}^{m} a_l x^l$ and $q(x) = \sum_{l=0}^{m} b_l x^l$ be polynomials.

- The **degree** of p is $\max(m \leq n | a_m \neq 0)$. The notation is $\deg(p)$. $\deg(0) = -\infty$.
- The **leading coefficient** of p is $a_{\deg(p)}$. The leading coefficient of 0 is 0. The **constant term** of p is a_0,
- p and q are **equal** (or $p = q$) if $a_k = b_k$ when both sides are defined and the others are 0.
- The set of polynomials with coefficient in R is denoted $R[x]$.

Remark

The notion of equality is an equivalence relation and $R[x]$ is the set containing the cosets of the relation $=$.

Polynomial functions

Definition

Let $p \in R[x]$ and $r \in R$

- the **value** of p at r is $p(r) = a_n r^n + a_{n-1} r^{n-1} + \cdots + a_1 r + a_0 \in R$.
- r is a **root** of p if $p(r) = 0$.
- $f : R \to R$ function is a **polynomial function** over R if there exists $p \in R[x]$ such that $p(r) = f(r)$ for all $r \in R$.

Remark

The set of polynomial functions might not be equivalent to the set of polynomials, for example if $R = \mathbb{Z}_p$ for some prime $p \in \mathbb{N}$, then by Fermat's little theorem we have $x^p = x$ as functions.

Example

What is the number of roots of the polynomials $x^4 - 2$ and $x^2 - 1$ over $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ and $\mathbb{Z}_8$.

The ring of polynomials

Definition (The operations)

$$+ \left(\sum_{k=0}^{m} a_k x^k\right) + \left(\sum_{l=0}^{m} b_l x^l\right) = \sum_{j=0}^{\max(m,n)} (a_j + b_j)x^j \text{ (set } a_k = 0$$
if $k > \deg(a)$ and $b_l = 0$ if $l > \deg b$).

$$\cdot \left(\sum_{k=0}^{m} a_k x^k\right) \cdot \left(\sum_{l=0}^{m} b_l x^l\right) = \sum_{j=0}^{m+n} (c_j)x^j, \text{ where } c_j = \sum_{k+l=j} a_k b_l.$$

Theorem

1. $\deg(p+q) \leq \max(\deg(p), \deg(q))$ and
2. if R is a domain, then $\deg(pq) = \deg(p) + \deg(q)$.

Example

$(x^2 + 2x) + (-x^2 + x) = 3x - 1$, so there is no equality in the first. $3x \cdot (2x + 1) = 3x$ over $\mathbb{Z}_6$, so the second is not true in rings with zero divisors.

Theorem

1. $R[x] = (R[x], +, \cdot)$ is a commutative ring and
2. if R is a domain, then $R[x]$ is also a domain.

Remark

- For a fixed $r \in R$ the map $R[x] \to R$, $p \mapsto p(r)$ is a homomorphism.
- So is $R \to R[x]$, $r \mapsto r$, where the second r is the constant polynomial.
- For a fixed $p \in R[x]$ the map $R \to R$, $r \mapsto p(r)$ is not a homomorphism in general: if $R = \mathbb{R}$ and $p(x) = x + 1$, then $3 = p(1 + 1) \neq p(1) + p(1) = 4$.

Horner's method revisited

Lemma

Let r be a root of a polynomial $p \in R[x]$. Then there exists $q \in R[x]$ such that $p(x) = (x - r)q(x)$.

Example

A root of $p(x) = x^3 - 1$ is $r = 1$ and $x^3 - 1 = (x - 1)(x^2 + x + 1)$.

Theorem (Horner's method – second form)

Let $a(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in R[x]$ polynomial and $r \in R$. Consider the following table:

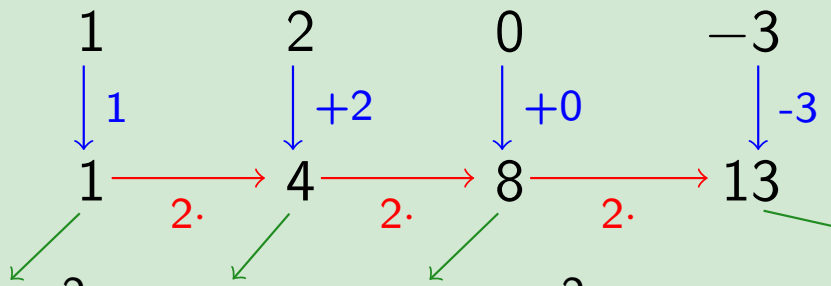
a_n	a_{n-1}	a_{n-2}	$\ldots$	a_1	a_0
b_{n-1}	b_{n-2}	b_{n-3}	$\ldots$	b_0	b_{-1}

,

where $b_{n-1} = a_n$ and $b_i = r \cdot b_{i+1} + a_{i+1}$ for $i < n - 1$.
Let $q(x) = b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + b_1 x + b_0$.
Then $a(x) = (x - r)q(x) + b_{-1}$, so $f(r) = b_{-1}$.
In particular r is a root $\Longleftrightarrow b_{-1} = 0$ and then $a(x) = (x - r)q(x)$.

Example

For $a(x) = x^3 + 2x^2 - 3$ and $r = 2$:



és ekkor $q(x) = 1x^2 + \quad 4x + \quad 8 = x^2 + 4x + 8$, $f(2) = 13$ és $a(x) = (x - 2)q(x) + a(2)$.

Theorem

If R is a domain and $p \in R[x]$ is nonzero, then p has at most $\deg(p)$ roots.

Example (It is necessary, that R is a domain)

In $\mathbb{Z}_8$ the polynomial $x^2 - 1$ has 4 roots: $x^2 - 1 = (x - 1)(x + 1)$, so ± 1 are roots. But ± 3 are also roots, since $x^2 - 1 = (x - 3)(x + 3)$ and we don't have a contradiction since $(3 - 1)(3 + 1) = 0 \in \mathbb{Z}_8$.

Number theory of polynomials over fields

Reminder

The fundamental theorem of number theory states that any integer decomposes as product of primes and this decomposition is unique up to order and units.
The structure of the proof was as follows:

1. Introduction the notions of divisibility, units, primes and irreducibles,
2. Development a method for division with remainder,
3. Proof of the existence of the greatest common divisor with the help of the Euclidean algorithm and
4. Proof of "irreducibles = primes" and the theorem.

Remark

Our next goal is to reproduce this for polynomials.
Those rings, for which this can be done are called Euclidean rings.

Step 1 – The notions

Lemma

Let R be a domain. Then $p \in R[x]$ is a unit $\iff$ p is a constant, and $p(0) \in R$ is a unit.

Example

If R has zero divisors this is not true: $(2x+1)^2 = 1 \in \mathbb{Z}_4[x]$.

Lemma

Let $\mathbb{F}$ be a field and $p \in \mathbb{F}[x]$.

1. If $\deg(p) = 1$, then p is irreducible and
2. If $\deg(p) = 2$ or 3, then p is irreducible $\iff$ p has no roots in $\mathbb{F}$.

Example

If $\deg(p) \geq 4$ this 2 is not true: $(x^2+1)^2 \in \mathbb{R}[x]$ is not irreducible and it has no roots in $\mathbb{R}$.

Step 2 – Division with remainder

Reminder

For integers: dividing a by b we get a quotient q, and a remainder r, such that $0 \leq r < b$ and $a = qb + r$.

Theorem

Let $\mathbb{F}$ be a field and $a, b \in \mathbb{F}[x]$ such that $b \neq 0$. Then there exist unique $q, r \in \mathbb{F}[x]$ for which $\deg(r) < \deg(b)$ and $a = qb + r$.

Remark

- It is necessary that we have a field: in $\mathbb{Z}[x]$, for example if $a(x) = x^2$, $b(x) = 2x$, we have no such q and $r \in \mathbb{Z}[x]$.
- Horner's method is actually a division with remainder: a and q are as above, $b(x) = x - r$ and the remainder is the constant $r(x) = b_{-1}$.

Step 3 – The gcd and the Euclidean algorithm

Reminder

$\gcd(a, b)$ is the positive integer d such that it is a common divisor and it is divisible by all of the common divisors.

Theorem

Let $a, b \in \mathbb{F}[x]$. Then there exists a $d \in \mathbb{F}[x]$ unique up to multiplying with units (=constants), such that

- $d|a$ and $d|b$ and
- if $c|a$ and $c|b$ for some $c \in \mathbb{F}[x]$ then $c|d$.

Definition

Let $a, b \in \mathbb{F}[x]$. The **greatest common divisor** of a and b is those d which has leading coefficient 1.

Proof (Euclidean algorithm for polynomials).

Method:

1. If $b = 0$, then set $d = a$. If $b|a$, then set $d = b$.
   Otherwise set $r_0 = a$ and $r_1 = c \cdot b$.
2. k. If $r_{k-1} = 0$, then $\gcd(a, b) = r_{k-2}$.
   Otherwise divide r_{k-2} by r_{k-1} with remainders:
   $$r_{k-2} = q_{k-1} r_{k-1} + r_k$$
   Repeat this until $r_k = 0$.

Claims:

- This process is finite, since $\deg(r_1) > \deg(r_2) > \dots$ are nonnegative integers or $-\infty$. Let the last step be the n-th, and set $d = r_{n-1}$.
- d is a common divisor of a and b.
- For any common divisor c we have $c|d$.
- d is unique up to units, so $\gcd(a, b) = c \cdot d$ for some $c \in \mathbb{F}$.
- The extended algorithm also works: there exist $x, y \in \mathbb{F}[x]$ such that $\gcd(a, b) = ax + by$.

Step 4 – Proof of "irreducibles = primes" and the theorem

Theorem

Let $p \in \mathbb{F}[x]$. Then p is irreducible $\iff$ p is prime.

Theorem

Let $a \in \mathbb{F}[x]$ such that $\deg(a) > 0$. Then a decomposes as a product of irreducibles and this decomposition is unique up to the order of the factors and multiplying with units.

Example (The decomposition of $x^4 + 1$)

$x^4 + 1 = (x^2 + 1)^2 - (\sqrt{2}x)^2 = (x^2 - \sqrt{2}x + 1)(x^2 + \sqrt{2}x + 1)$.
In $\mathbb{R}[x]$ these are irreducibles, but in $\mathbb{C}[x]$:

$$x^4+1 = \left(x - \frac{1+i}{\sqrt{2}}\right)\left(x - \frac{1-i}{\sqrt{2}}\right)\left(x - \frac{-1+i}{\sqrt{2}}\right)\left(x - \frac{-1-i}{\sqrt{2}}\right).$$

Example

- It is necessary that R is a field. In $(x-1)(x+1) = (x-3)(x+3) \in \mathbb{Z}_8$.
- Moreover $x = (2x+3)(3x+2) \in \mathbb{Z}_6$.

Remark

The notion of least common multiple also makes sense in $\mathbb{F}[x]$. The same statements about gcd and lcm is true as for integers.

Example

- What are the irreducibles of degree 2 and 3 over $\mathbb{F}_2$?
- Decompose $x^5 + 3x - 1$ to product of irreducibles over $\mathbb{Z}_2$ and $\mathbb{Z}$!

Remark

Let $p \in \mathbb{N}$ be a prime. $\mathbb{Z}$ and $\mathbb{Z}_p[x]$ are very similar in some sense, but "life is much nicer" in $\mathbb{Z}_p[x]$. The analogues of main theorems and conjectures about primes (such as Prime number theorem and Riemann hypothesis) can be proven in a simple way.

The case of $\mathbb{C}[x]$ and $\mathbb{R}[x]$

Theorem

1. $p \in \mathbb{C}[x]$ is irreducible $\iff$ $\deg(p) = 1$,
2. $p \in \mathbb{R}[x]$ is irreducible if and only if
   a) $\deg(p) = 1$ or
   b) $\deg(p) = 2$ and p has no real roots.

Corollary

1. (Fundamental theorem of Algebra, second form)
   Every $p \in \mathbb{C}[x]$ decomposes as $p(x) = c\prod_{j=1}^{n}(x - \alpha_j)$
   for some $c, \alpha_j \in \mathbb{C}$.
2. Every $p \in \mathbb{R}[x]$ decomposes as
   $p(x) = c\prod_{j=1}^{n}(x - \alpha_j)\prod_{k=1}^{m}(x^2 + \beta_k x + \gamma_k)$ for some
   $c, \alpha_j, \beta_k, \gamma_k \in \mathbb{R}$ such that $\beta_k^2 - 4\gamma_k < 0$ for all k.

The case of $\mathbb{Q}[x]$ and $\mathbb{Z}[x]$

Remark

- In $\mathbb{Q}[x]$ every polynomial p can be written as $p = cq$, where c is a constant (unit) and $q \in \mathbb{Z}[x]$.
- In $\mathbb{Z}[x]$ the whole machinery fails, since there is no division with remainder.

Definition

$p(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$ is **primitive** if $\gcd(a_0, a_1, \ldots, a_n) = 1$.

In other words there is no prime $p_0 \in \mathbb{Z}$ such that $p_0 | a_k$ for all k.

Lemma

For any $p \in \mathbb{Q}[x] - \{0\}$ there are $c \in \mathbb{Q}$ and $q \in \mathbb{Z}[x]$ primitive such that $p = cq$. Moreover the pair c and q is unique up to the sign.

Example

$\frac{6}{5}x^3 + \frac{14}{5}x - \frac{2}{3} = \frac{1}{15}\left(18x^3 + 14x - 10\right) = \frac{2}{15}\left(9x^3 + 7x - 5\right)$.

Gauss lemma

Lemma (Gauss lemma)

If p and q are primitive, then pq is also primitive.

Corollary

If $p \in \mathbb{Z}[x]$ is decomposes as a nontrivial product in $\mathbb{Q}[x]$ (i. e. the terms has lower degree), then so does in $\mathbb{Z}[x]$.

Example

$1/2$ is a root of $2x^3 + 5x^2 - x - 1$, so we can divide by $(x - 1/2)$:

$$2x^3+5x^2-x-1 = \left(x - \frac{1}{2}\right)(2x^2+6x+2) = (2x-1)(x^2+3x+1).$$

Corollary

For any $p \in \mathbb{Z}[x]$ nonconstant primitive we have
p is reducible in $\mathbb{Q}[x] \iff p$ is reducible in $\mathbb{Z}[x]$.

Theorem

1. $p \in \mathbb{Q}[x]$ is irreducible if p is a unit (= nonzero constant) multiple of a primitive irreducible polynomial.
2. $p \in \mathbb{Z}[x]$ is irreducible if
   a) p is a constant and $p \in \mathbb{Z}$ is prime, or
   b) p is primitive and irreducible in $\mathbb{Q}[x]$.

Example

5 is irreducible in $\mathbb{Z}[x]$, but not in $\mathbb{Q}[x]$ (as it is a unit).
$2x - 4 = 2(x - 2)$ is irreducible in $\mathbb{Q}[x]$, but not in $\mathbb{Z}[x]$.

Theorem (Fundamental theorem of Number theory in $\mathbb{Z}[x]$)

In $\mathbb{Z}[x]$ every nonconstant polynomial decomposes as product of irreducibles and this decomposition is unique up to the order of the terms and multiplying with units (sign).

Remark

Thus in $\mathbb{Z}[x]$ there is gcd without Euclidean algorithm. It worked the other way round: we get it from the fundamental theorem.

Schönemann-Eisenstein criterion

Reminder

In $\mathbb{C}[x]$ the irreducible polynomials have degree 1.
In $\mathbb{R}[x]$ the irreducible polynomials have degree at most 2.
Question: is there $D \in \mathbb{N}$ such that if $a \in \mathbb{Q}[x]$ (or $\mathbb{Z}[x]$) is irreducible, then $\deg(a) \leq D$?

Theorem (Schönemann-Eisenstein criterion)

Let $a(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$ and $p \in \mathbb{Z}$ prime such that

$$p \nmid a_n, p | a_{n-1}, a_{n-2}, \dots a_0 \text{ and } p^2 \nmid a_0.$$

Then a is irreducible in $\mathbb{Q}[x]$.

Corollary

- If a is also primitive, then it is irreducible in $\mathbb{Z}[x]$.
- For any $d \in \mathbb{N}$ there are irreducible polynomials of degree d in $\mathbb{Q}[x]$ and $\mathbb{Z}[x]$: for example $a(x) = x^d - 2$.

Corollary (Reversed Schönemann-Eisenstein criterion)

Let $a(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$ and $p \in \mathbb{Z}$ prime such that

$$p \nmid a_0, p | a_1, a_2, \ldots a_n \text{ and } p^2 \nmid a_n.$$

Then a is irreducible in $\mathbb{Q}[x]$.
If moreover a is primitive, then it is irreducible in $\mathbb{Z}[x]$.

Lemma

Let $\mathbb{F}$ be a field, $a, b \in \mathbb{F}$, $a \neq 0$ and $f \in \mathbb{F}[x]$. Then $f(x)$ is irreducible $\iff$ $f(ax + b)$ is irreducible.

Example

Show that $x^5 - 4 \in \mathbb{Z}[x]$ is irreducible!

Remark (Factorization in $\mathbb{Q}[x]$ and $\mathbb{Z}[x]$)

The Schöneman-Eisenstein criterion gives a sufficient (but not necessary!) condition for irreducibility.
The LLL (Lenstra-Lenstra-Lovász) algorithm is a "fast" (polynomial time) tool to factorize a polynomial.

Cyclotomic polynomials

Definition

For $n \in \mathbb{N}$ the n-th **cyclotomic polyinomial** is

$$\Phi_n(x) = \prod_{k=1}^{\varphi(n)} (x - \varepsilon_k),$$

where $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{\varphi(n)}$ are the primitive n-th roots of unity.

Lemma

1. $x^n - 1 = \prod_{d|n} \Phi_d(x)$ and
2. $\Phi_n(x) \in \mathbb{Z}[x]$.

Example

$\Phi_p(x)$ is irreducible in $\mathbb{Z}[x]$ for any prime $p \in \mathbb{Z}$.

Theorem

$\Phi_n(x)$ is irreducible for any $n \in \mathbb{N}$.

Roots of polynomials

Remark (Finding roots of polynomials)

- Linear: $ax + b = 0 \implies x = -b/a$.
- Quadratic: $ax^2 + bx + c = 0 \implies x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.
- Cubic: $ax^3 + bx^2 + cx + d = 0 \implies$ one of its roots is

$$x = \sqrt[3]{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right) + \sqrt{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right)^2 + \left(\frac{c}{3a} - \frac{b^2}{9a^2}\right)^3}}$$

$$+ \sqrt[3]{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right) - \sqrt{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right)^2 + \left(\frac{c}{3a} - \frac{b^2}{9a^2}\right)^3}} - \frac{b}{3a}.$$

- Quartic:
  I did not put it here, since there is not enough space.
- Quintic:
  There is no general formula (with basic algebraic operations)!

Rational root test

Theorem

Let $a(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$. If $p/q \in \mathbb{Q}$ with $\gcd(p, q) = 1$ is a root of a, then $p|a_0$ and $q|a_n$.

Corollary

Let $a \in \mathbb{Z}[x]$ with leading coefficient 1 (then a is called **monic**). The rational roots of a are integer and are divisors of the constant term.

Example

Find the roots of $2x^4 - 5x^3 - 8x^2 + 17x - 6$!
Hint: this theorem and Horner's method makes it very quick!

Remark

The theorem does not guarantee the existence of rational roots.

Multiple roots

Definition

Let R be a domain, $r \in R$ and $p \in R[x]$.

- r is an m-**fold root** of p for some $m \in \mathbb{N}$, if $(x-r)^m | p$,
- the root r has **multiplicity** m if there is $q \in \mathbb{R}[x]$ such that $p(x) = (x-r)^m q(x)$ and $q(r) \neq 0$ and
- r is a multiple root if its multiplicity is at least 2.

Example

If $p(x) = c\prod_{j=1}^{k}(x-r_j)^{m_j}$, then the multiplicity of r_j is m_j.

Definition

Let $\mathbb{F}$ be a field and $p(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{F}[x]$.

The **(formal) derivative** of p is $p'(x) = \sum_{k=1}^{n} k a_k x^{k-1} \in \mathbb{F}[x]$.

Theorem (Properties of the formal derivative)

Let $p, q \in \mathbb{F}[x]$, $c, r \in \mathbb{F}$ and $n \in \mathbb{N}$. Then

1. $c' = 0$ (the constants have 0 derivative),
2. $(p+q)' = p' + q'$,
3. $(cp)' = cp'$,
4. $(pq)' = p'q + pq'$ and
5. $((x-r)^n)' = n(x-r)^{n-1}$.

Theorem

r is a multiple root of $p \iff p(r) = p'(r) = 0$.

Corollary

The multiple roots of p are the common roots of p and p'.

Example

What are the multiple roots of $x^5 + 2x + 1 \in \mathbb{C}[x]$ and $x^3 + 1 \in \mathbb{Z}_3[x]$?

Relations between roots and coefficients - Vieta's formulas

Theorem

Let $p \in \mathbb{F}[x]$ be a polynomial. If p decomposes as product of linear factors, say

$$p(x) = \sum_{k=0}^{n} a_k x^k = a_n \prod_{k=1}^{n} (x - \alpha_k),$$

for some $a_k, \alpha_k \in \mathbb{F}$, then

$$\begin{aligned}
-\frac{a_{n-1}}{a_n} &= \alpha_1 + \alpha_2 + \cdots + \alpha_n, \\
\frac{a_{n-2}}{a_n} &= \alpha_1\alpha_2 + \cdots + \alpha_1\alpha_n + \alpha_2\alpha_3 + \cdots + \alpha_{n-1}\alpha_n, \\
&\vdots \\
(-1)^n \frac{a_0}{a_n} &= \alpha_1\alpha_2 \ldots \alpha_n.
\end{aligned}$$

Or for $1 \leq k \leq n$

$$(-1)^k \frac{a_{n-k}}{a_n} = \sum_{0 \leq i_1 < i_2 < \cdots < i_k \leq n} \left(\prod_{l=1}^{k} \alpha_{i_l} \right)$$

Remark

If $\mathbb{F}$ is algebraically closed (especially if $\mathbb{F} = \mathbb{C}$), then the condition always holds.

Definition

$e_k(x_1, x_2, \ldots, x_n) = \sum_{0 \le i_1 < i_2 < \cdots < i_k \le n} \left(\prod_{l=1}^{k} x_{i_l} \right)$ is the kth **elementary symmetric polynomial** in n variable.

Corollary

When p is monic, then $(-1)^k a_{n-k} = e_k(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Example

- What is the sum, product, sum of squares of the roots of $2x^4 - x^3 + 3x^2 - 5$? Show that it has not only real roots! Determine the number of the real roots!
- What is the sum and product of the nth roots of unity?
- What is the sum and product of 15th primitive roots of unity?

Solution of cubic equations

Remark

Any cubic equation can be reduced to one in the form $x^3 + px + q = 0$: we can divide by the leading coefficient and then make a substitution $y = x + a_2/3$, where a_2 is the coefficient of x^2.

Example

Transform $2x^3 - 18x^2 + 72x = 56$ to such form!

Theorem (Cardano's formula)

The solutions of $x^3 + px + q = 0$ are in the form

$$x = u + v, \text{ where } u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \text{ and } v = -\frac{p}{3u}.$$

Example

Solve the equation $y^3 + 9y + 26 = 0$!

Remark

Let $f(x) = x^3 + px + q \in \mathbb{R}[x]$ and $D = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$. It can be shown that

- if $D > 0$, then $f(x) = 0$ has 1 real and 2 non-real roots,
- if $D = 0$, then $f(x) = 0$ has a multiple root (that implies 3 real roots),
- if $D < 0$, then $f(x) = 0$ has 3 real roots, but for computing them complex numbers must be used in Cardano's formula.

Example

Solve the equation $x^3 - 7x - 6 = 0$!

Definition

D is the **discriminant** of the cubic polynomial $f(x)$.

Corollary

f has multiple roots $\iff$ $D = 0$.

Symmetric polynomials

Definition

Let $R[x_1, x_2, \ldots, x_n] = (\ldots((R[x_1])[x_2])\ldots)[x_n]$ be the polynomial ring in n variables $x_1, x_2, \ldots x_n$ for a commutative ring R.

Example

$x^3y^2 + 2x^2y^2 - x^3y + 2xy - y^2 + 6 =$
$(x^3 + 2x^2 - 1)y^2 + (-x^3 + 2x)y + 6 \in \mathbb{Z}[x, y]$.

Theorem

The units of $R[x_1, x_2, \ldots, x_n]$ are the constant units (in R).
If R is a domain, then so is $R[x_1, x_2, \ldots, x_n]$.

Definition

- The **degree** of a monomial $ax_1^{i_1}x_2^{i_2}\ldots x_n^{i_n}$ is $i_1 + i_2 + \cdots + i_n$ for $a \in R - \{0\}$ and $i_k \in \mathbb{N} \cup \{0\}$.
- The **degree** of a polynomial $p \in R[x_1, x_2, \ldots, x_n]$ is the maximum of the degree of the monomials.

Definition

A $p \in R[x_1, x_2, \ldots, x_n]$ is **symmetric** if for all permutations $p(x_1, x_2, \ldots, x_n) = p(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$.

Example

- $x + y \in R[x, y]$ is symmetric, but $x - y$ is not,
- $e_k(x_1, x_2, \ldots, x_n)$ (defined at Vieta's formulas) is symmetric.

Theorem

Any symmetric polynomial $p \in R[x_1, x_2, \ldots, x_n]$ can be written as a polynomial of elementary symmetric polynomials. In other words there exists $q \in R[x_1, x_2, \ldots, x_n]$ such that

$$p(x_1, x_2, \ldots, x_n) = q(e_1(x_1, x_2, \ldots x_n), \ldots, e_n(x_1, x_2, \ldots x_n)).$$

Example

Let $p(x) = x^3 - 3x + 1$

- What is the sum of the cubes of the roots of p?
- What is the monic polynomial which has roots $\alpha\beta, \beta\gamma$ and $\gamma\alpha$, where α, β and γ are the roots of p?

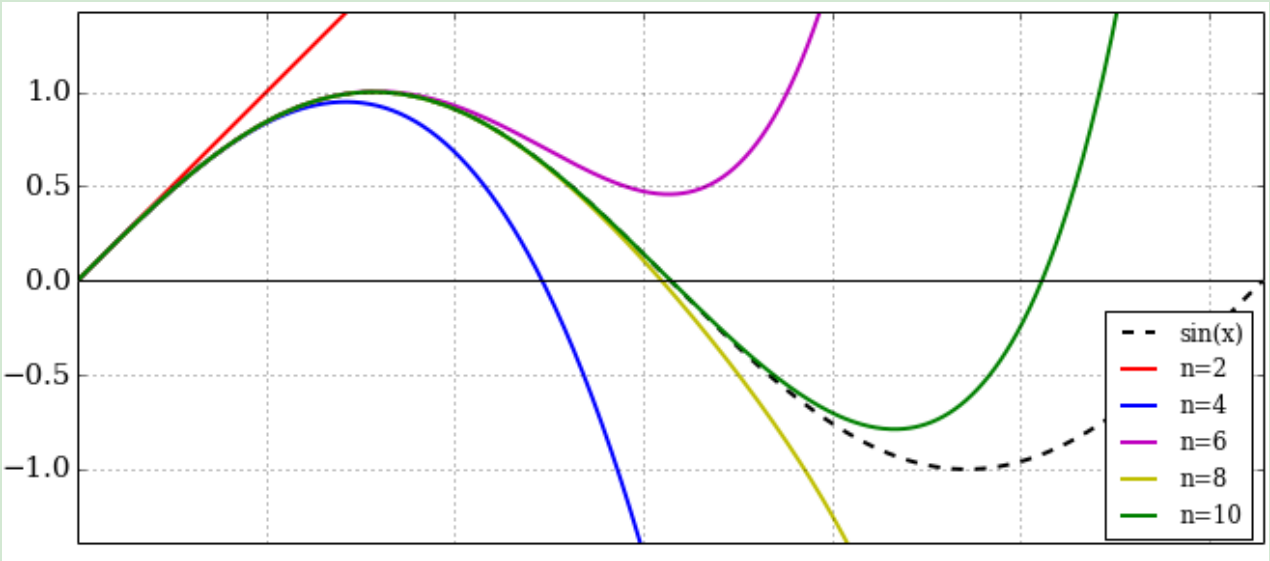
Polynomial interpolation

Theorem

Let $\mathbb{F}$ be a field, and $x_k, y_k \in \mathbb{F}$ for $k = 1, 2, \ldots, n$ such that $x_k \neq x_l$ for any $k \neq l$. Then there exists a unique polyinomial $p \in \mathbb{F}[x]$ of degree at most $n - 1$ such that $p(x_k) = y_k$ for all k.

Example

- Interpolate $\sin x$ with low degree polynomials! ($x_k = k/100$)



- What is the lowest degree polynomial $p \in \mathbb{R}[x]$ such that $p(-1) = -5$, $p(0) = 5$, $p(1) = 5$ and $p(2) = 7$?

1st method for constructing p: Lagrange interpolation.

Let $L_k(x) = \prod_{\substack{j=1 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j} \in \mathbb{F}[x]$, then $L_k(x_j) = \begin{cases} 1, & \text{if } j = k \\ 0, & \text{if } j \neq k \end{cases}$

Thus $p(x) = \sum_{k=1}^{n} y_k L_k(x)$ is a good choice. □

2nd method for constructing p: Newton interpolation.

Let $N_1(x) = y_1$ and for $k \underset{k-1}{\geq} 1$

$$N_k(x) = N_{k-1}(x) + c \prod_{j=1}^{} \frac{x - x_j}{x_k - x_j}, \quad \text{such that } N_k(x_k) = y_k.$$

Then $N_k(x_j) = y_j$ for $j \leq k$, thus $p = N_n$ is a good choice. □

Remark (Which is better to use?)

- Lagrange interpolation is more efficient when you have to interpolate several data sets on the same data points.
- Newton interpolation is more efficient when you have to interpolate data incrementally.

A cryptographic application

Example (SSS (Shamir's Secret Sharing) algorithm)

Task: Share a secret integer $S \in \mathbb{Z}$ between n people such that

1. any k of them can recover S together, but
2. any $k-1$ of them does not know anything about S!

Method: Choose a polynomial $p \in \mathbb{Z}[x]$ such that $p(0) = S$ and $\deg(p) < k$. Share the value $p(k)$ with the kth person.

Verification: This works by the interpolation theorem, since

1. any k of the people can find p (uniqueness) and
2. based on the knowledge of any $k-1$ of the people the value of $p(0)$ can be any integer (existence).

Systems of linear equations

Linear equations

Definition

Let $\mathbb{F}$ be a field, $m, n \in \mathbb{N}$.

- A **linear equation** in n variables $x_1, x_2, \ldots, x_n$ is an equation
$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$
where $a_1, a_2, \ldots, a_n, b \in \mathbb{F}$.
- A **system of linear equations** in n variables $x_1, x_2, \ldots, x_n$ is a system of equations
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$
were $a_{jk}, b_k \in \mathbb{F}$ for $1 \leq j \leq m$ and $1 \leq k \leq n$.
- The a_{jk}-s are the **coefficients**, the x_k-s are the **variables** and the b_j-s are the **constants**.

Example (Some systems of linear equations)

$$\begin{aligned} x+y &= 3 \\ x-y &= 1 \end{aligned} \qquad \begin{aligned} x+y &= 3 \\ 3x-y &= 5 \\ 2y &= 2 \end{aligned} \qquad \begin{aligned} 2x+z &= 4 \\ y+z &= 1 \end{aligned}$$

Definition

- A **solution** of the above system of equations is a tuple $(x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$ (i. e. $x_k \in \mathbb{F}$ for $1 \leq k \leq n$) such that it satisfies each linear equations.
- The system of the linear equations is **homogeneous** if $b_j = 0$ for all j and **inhomogeneous** otherwise.

Lemma

The following are equivalent:

1. $x_1 = x_2 = \cdots = x_n = 0$ is a solution and
2. the system of linear equations is homogeneous.

Definition

Two systems of linear equations are **equivalent** if the sets of solutions are the same.

Example

The first and second systems of the previous example are equivalent, but not the third.

Lemma (Row operations)

The following operations results equivalent systems of linear equations:

1. Swap the positions of two rows,
2. Multiply a row by a non-zero $\lambda \in \mathbb{F}$ and
3. Add to one row a multiple of another.

Example

$$\begin{cases} 3x + 4y = 5 \\ x + 2y = 8 \end{cases} \Rightarrow \begin{cases} x = -11 \\ x + 2y = 8 \end{cases} \Rightarrow \begin{cases} x = -11 \\ 2y = 19 \end{cases} \Rightarrow \begin{cases} x = -11 \\ y = 8.5 \end{cases}$$

Definition

The **matrix** resp. **augmented matrix** of the system of linear equations $\sum_{k=1}^{n} a_{jk}x_k = b_j$ for $j = 1, 2, \ldots, m$ is

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \quad \text{resp.} \quad \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right)$$

Remark

- We will define the notion of matrices later.
- This enables us to give the solution in a more compact form.

Example (The previous equations in augmented matrix form)

$$\left(\begin{array}{cc|c} 3 & 4 & 5 \\ 1 & 2 & 8 \end{array}\right) \overset{(1)-2(2)}{\Longrightarrow} \left(\begin{array}{cc|c} 1 & 0 & -11 \\ 1 & 2 & 8 \end{array}\right) \overset{(2)-(1)}{\Longrightarrow}$$

$$\left(\begin{array}{cc|c} 1 & 0 & -11 \\ 0 & 2 & 19 \end{array}\right) \overset{\frac{1}{2}(2)}{\Longrightarrow} \left(\begin{array}{cc|c} 1 & 0 & -11 \\ 0 & 1 & 8.5 \end{array}\right)$$

Row echelon form

Definition

- A matrix is in **row echelon form** if
  1. its zero rows (if there are) are the lasts and
  2. for any two adjacent nonzero rows the latter begins with at least one more 0 than the previous.
- The first nonzero element in a row is called **pivot** of the row.

Example (Which of the following are in row echelon form?)

$$\begin{pmatrix} 2 & 3 \\ 0 & 4 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & -3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

Theorem

Any matrix can be transformed to row echelon form with elementary row operations.

Gaussian elimination

Example (Gaussian elimination = solving equations this way)

$$\begin{aligned} x+\ \ y+2z&=0\\ 2x+2y+3z&=2\\ x+3y+3z&=4\\ x+2y+\ \ z&=5 \end{aligned} \longrightarrow \left(\begin{array}{ccc|c} 1&1&2&0\\ 2&2&3&2\\ 1&3&3&4\\ 1&2&1&5 \end{array}\right) \xrightarrow[(3)-(1)]{(2)-2(1)\\ (4)-(1)} \left(\begin{array}{ccc|c} 1&1&2&0\\ 0&0&-1&2\\ 0&2&1&4\\ 0&1&-1&5 \end{array}\right)$$

$$\xrightarrow{(2)\leftrightarrow(3)} \left(\begin{array}{ccc|c} 1&1&2&0\\ 0&2&1&4\\ 0&0&-1&2\\ 0&1&-1&5 \end{array}\right) \xrightarrow{(4)-\frac{1}{2}(2)} \left(\begin{array}{ccc|c} 1&1&2&0\\ 0&2&1&4\\ 0&0&-1&2\\ 0&0&-\frac{3}{2}&3 \end{array}\right) \xrightarrow{(4)-\frac{3}{2}(3)}$$

$$\left(\begin{array}{ccc|c} 1&1&2&0\\ 0&2&1&4\\ 0&0&-1&2\\ 0&0&0&0 \end{array}\right) \longrightarrow \begin{aligned} x+\ \ y+2z&=0\\ 2y+\ \ z&=4\\ -\ \ z&=2\\ 0&=0 \end{aligned} \longrightarrow \begin{aligned} x&=1\\ y&=3\\ z&=-2 \end{aligned}$$

Thus the only solution is $(x, y, z) = (1, 3, -2)$.

Definition

A matrix is in **reduced row echelon form** if

1. it is in row echelon form,
2. each pivot element is 1 and
3. the other elements in the columns of pivot elements are 0.

Theorem

1. Any matrix can be transformed to reduced row echelon form with elementary row operations.
2. The process always ends up at the same form not depending on which elementary row operations were used.

Definition

The reduced row echelon form of a matrix A is the unique result of the above process, and is denoted by $\mathrm{rref}(A)$.

Gauss-Jordan elimination

Example (Gauss-Jordan elimination=solving equations this way)

$$\begin{aligned} x+\ \ y+2z&=0\\ 2x+2y+3z&=2\\ x+3y+3z&=4\\ x+2y+\ \ z&=5 \end{aligned} \longrightarrow \left(\begin{array}{ccc|c} 1&1&2&0\\ 2&2&3&2\\ 1&3&3&4\\ 1&2&1&5 \end{array}\right) \xrightarrow{\cdots} \left(\begin{array}{ccc|c} 1&1&2&0\\ 0&2&1&4\\ 0&0&-1&2\\ 0&0&0&0 \end{array}\right) \xrightarrow{\frac{1}{2}(2)}$$

$$\left(\begin{array}{ccc|c} 1&1&2&0\\ 0&1&1/2&2\\ 0&0&-1&2\\ 0&0&0&0 \end{array}\right) \xrightarrow{(1)-(2)} \left(\begin{array}{ccc|c} 1&0&3/2&-2\\ 0&1&\frac{1}{2}&2\\ 0&0&-1&2\\ 0&0&0&0 \end{array}\right) \xrightarrow{-(3)} \left(\begin{array}{ccc|c} 1&0&\frac{3}{2}&-2\\ 0&1&\frac{1}{2}&2\\ 0&0&1&-2\\ 0&0&0&0 \end{array}\right)$$

$$\xrightarrow[(2)-\frac{1}{2}(1)]{(1)-\frac{3}{2}(3)} \left(\begin{array}{ccc|c} 1&0&0&1\\ 0&1&0&3\\ 0&0&1&-2\\ 0&0&0&0 \end{array}\right) \longrightarrow \begin{aligned} x&=1\\ y&=3\\ z&=-2\\ 0&=0 \end{aligned}$$

Thus the only solution is $(x, y, z) = (1, 3, -2)$.

Definition

The variable x_k in a system of linear equations is **free** if the corresponding column in the reduced row echelon form are no pivots and **bounded** otherwise.

Theorem

Consider a system of linear equations and the reduced row ecelon form of its augmented matrix.

1. If there is a row, which is zero in the matrix, but the constant in the augmented matrix is nonzero, then there is no solution.
2. Otherwise there is exactly one solution for each possible values of the free variables.

Corollary

- If $\mathbb{F}$ is infinite, then the number of the solutions is either 0, or 1, or ∞.
- If $|\mathbb{F}| = q < \infty$, then the number of solutions is either 0 or q^f, where $f = \#(\text{free variables})$.

Example (No solution)

$$\begin{array}{rcl} x+y & = & 3 \\ 3x+4y & = & 10 \\ 2y & = & 4 \end{array} \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 3 \\ 3 & 4 & 10 \\ 0 & 2 & 4 \end{array}\right) \xrightarrow{(2)-3(1)} \left(\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 2 & 4 \end{array}\right)$$

$$\xrightarrow[(3)-2(2)]{(1)-(2)} \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{array}\right) \longrightarrow \begin{array}{rcll} x & = & 2 & \\ y & = & 1 & \\ 0 & = & 2 & \lightning \end{array}$$

Example (Many solutions)

$$\begin{array}{rcl} 2x+z & = & 4 \\ y+z & = & 1 \end{array} \longrightarrow \left(\begin{array}{ccc|c} 2 & 0 & 1 & 4 \\ 0 & 1 & 1 & 1 \end{array}\right) \xrightarrow{\frac{1}{2}(1)} \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & 2 \\ 0 & 1 & 1 & 1 \end{array}\right)$$

$$\longrightarrow \begin{array}{rcl} x+\frac{1}{2}z & = & 2 \\ y+z & = & 1 \end{array} \longrightarrow \begin{array}{rcl} x & = & 2-\frac{1}{2}z \\ y & = & 1-z \\ z & = & z \end{array} \text{, for any } z \in \mathbb{F}.$$

The **vectorial form** of the solution is $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -\frac{1}{2} \\ -1 \\ 1 \end{pmatrix}$

Vectorspaces

The vectorspace $\mathbb{F}^n$

Definition

Let $\mathbb{F}$ be a field and $n \in \mathbb{N}$. Consider the set $\mathbb{F}^n = \{(v_1, v_2, \ldots, v_n) | v_k \in \mathbb{F}\}$ of ordered n-tuples of $\mathbb{F}$.

- The elements of $\mathbb{F}$ are the **scalars**. The elements of $\mathbb{F}^n$ are the **vectors**. The notation is $\underline{v} = (v_1, v_2, \ldots, v_n)$. We often write $\underline{w}$ as a column vector $\underline{v} = (v_1, v_2, \ldots, v_n)^T = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$.
- $\underline{0} = (0, 0, \ldots 0) \in \mathbb{F}^n$ is the **zero vector**.
- the **addition** is $\underline{u} + \underline{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$ for any $\underline{u}, \underline{v} \in \mathbb{F}$ and
- the **scalar multiplication** is $\lambda \underline{v} = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n)$ for any $\lambda \in \mathbb{F}$ and $\underline{v} \in \mathbb{F}^n$.

Remark

This is consistent to the standard Euclidean vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Lemma (Basic properties of the operations)

Let $\underline{u}, \underline{v}, \underline{w} \in \mathbb{F}^n$ and $\lambda, \mu \in \mathbb{F}$. Then

A1 $\underline{u} + \underline{v} = \underline{v} + \underline{u}$

A2 $(\underline{u} + \underline{v}) + \underline{w} = \underline{u} + (\underline{u} + \underline{w})$

A3 $\underline{v} + \underline{0} = \underline{v}$

A4 $\exists(-\underline{v}) \in \mathbb{F}^n : \underline{v} + (-\underline{v}) = \underline{0}$

M1 $\lambda(\underline{u} + \underline{v}) = \lambda\underline{u} + \lambda\underline{v}$

M2 $(\lambda + \mu)\underline{v} = \lambda\underline{v} + \mu\underline{v}$

M3 $(\lambda\mu)\underline{v} = \lambda(\mu\underline{v})$

M4 $1\underline{v} = \underline{v}$

Remark

- Thus $(\mathbb{F}^n, +)$ is an abelian (= commutative) group, but $(\mathbb{F}^n, +, \cdot)$ is **not** a ring.
- In general $(V, +, \cdot)$ is a **vectorspace** over $\mathbb{F}$, if V is a set containing an element $\underline{0}$, $+$ is an operation and $\cdot : \mathbb{F} \times V \to V$ is a function satisfying the above properties.

Theorem (Some useful properties and notation)

1. $\lambda\underline{v} = \underline{0} \iff \lambda = 0$ or $\underline{v} = \underline{0}$.
2. $-\underline{v} = (-1)\underline{v}$ and $\underline{u} - \underline{v} := \underline{u} + (-\underline{v})$.

Definition

A **linear combination** of vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k \in \mathbb{F}^n$ is $\lambda_1\underline{v}_1 + \lambda_2\underline{v}_2 + \cdots + \lambda_k\underline{v}_k$ for some $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$.

Example

In $\mathbb{F}^n$ every vector is a linear combination of $\underline{e}_1 = (1, 0, \ldots, 0)$, $\underline{e}_2 = (0, 1, \ldots, 0)$, $\ldots$, $\underline{e}_n = (0, 0, \ldots, 1)$: $(v_1, v_2, \ldots, v_n) = v_1\underline{e}_1 + v_2\underline{e}_2 + \cdots + v_n\underline{e}_n$.

Definition

A subset $V \subseteq \mathbb{F}^n$ is a **subspace** if 1) $V \neq \emptyset$, 2) $\underline{u}, \underline{v} \in V \implies \underline{u} + \underline{v} \in V$ and 3) $\lambda \in \mathbb{F}, \underline{v} \in V \implies \lambda\underline{v} \in V$. The notation is $V \leq \mathbb{F}^n$.

Corollary

The following are equivalent for a nonempty subset $V \subseteq \mathbb{F}^n$:

1. $V \leq \mathbb{F}^n$ (is a subspace),
2. V is closed under the operations and
3. V is closed under linear combinations.

Corollary

If $U, V \leq \mathbb{F}^n$ are subspaces, then so is $U \cap V$.

Definition

The subspace **spanned (or generated)** by $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k \in \mathbb{F}^n$ is $\mathrm{Span}(\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k) = \{\lambda_1\underline{v}_1 + \lambda_2\underline{v}_2 + \cdots + \lambda_k\underline{v}_k \mid \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}\}$.

$\mathrm{Span}(\emptyset) = \{\underline{0}\}$.

Remark

The spanned subspace of a finite subset is indeed a subspace in $\mathbb{F}^n$.

Example

Question: What are the subspaces in $\mathbb{R}^3$?
Answer: After identifying points of $\mathbb{R}^3$ with their position vectors

- the origin,
- the lines passing through the origin,
- the planes passing through the origin and
- the whole space.

The nullspace of systems of linear equations

Theorem

Let $(A|0)$ be the augmented matrix of a homogeneous system of linear equations in n variables over $\mathbb{F}$. Then the set of solutions is a subspace in $\mathbb{F}^n$.

Definition

The set of solutions of a homogeneous system of linear equations with matrix A is called the **nullspace** of A and denoted by $\mathcal{N}(A)$.

Example

What is the nullspace corresponding to the following equations:

$$x + y = 0$$

$$\begin{cases} x + y = 0 \\ y + z = 0 \end{cases}$$

$$\begin{cases} x + y = 0 \\ y + z = 0 \\ z + x = 0 \end{cases}$$

Definition

A subset $U \subseteq \mathbb{F}^n$ is an **affine subspace** if there exists $\underline{u} \in \mathbb{F}^n$ and $V \leq \mathbb{F}^n$ such that $U = \underline{u} + V = \{\underline{u} + \underline{v} | \underline{v} \in V\}$.

Example

Question: What are the affine subspaces of $\mathbb{R}^3$?
Answer: Points, lines, planes and the whole space.

Lemma

Let $\underline{u}_1, \underline{u}_2 \in \mathbb{F}^n$ and $V_1, V_2 \leq \mathbb{F}^n$. Then
$\underline{u}_1 + V_1 = \underline{u}_2 + V_2 \implies V_1 = V_2$ and $\underline{u}_1 - \underline{u}_2 \in V_1 = V_2$.

Theorem

Let $(A|\underline{b})$ be the augmented matrix of a system of linear equations in n variables over $\mathbb{F}$. Then the set of solutions is an affine subspace in $\mathbb{F}^n$. More precisely if $\underline{x}_0$ is a solution then the set of solutions is $\underline{x}_0 + \mathcal{N}(A)$.

Hyperplanes

Definition

Let $\underline{a} \in \mathbb{F}^n - \{\underline{0}\}$ and $b \in \mathbb{F}$.

- The set of solutions of the linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ is called a **hyperplane**.
- The above equation is the **implicit equation**. The solutions parametrized by the free variables is the **explicit equation**.

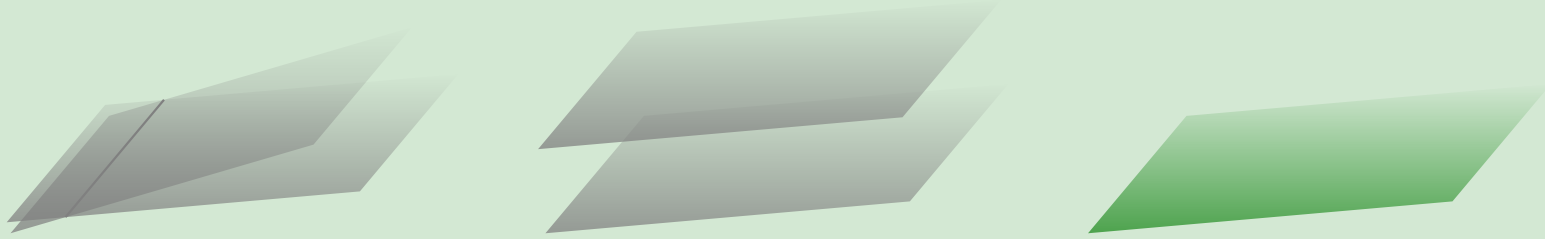
Example

- The hyperplanes of $\mathbb{R}^2$ (resp. $\mathbb{R}^3$) are the lines (resp. planes).
- The explicit equation of the hyperplane corresponding to $x + 2y - z = 3$ is $\begin{cases} x = 3 - 2s + t \\ y = s \\ z = t \end{cases}$ or in vectorial form

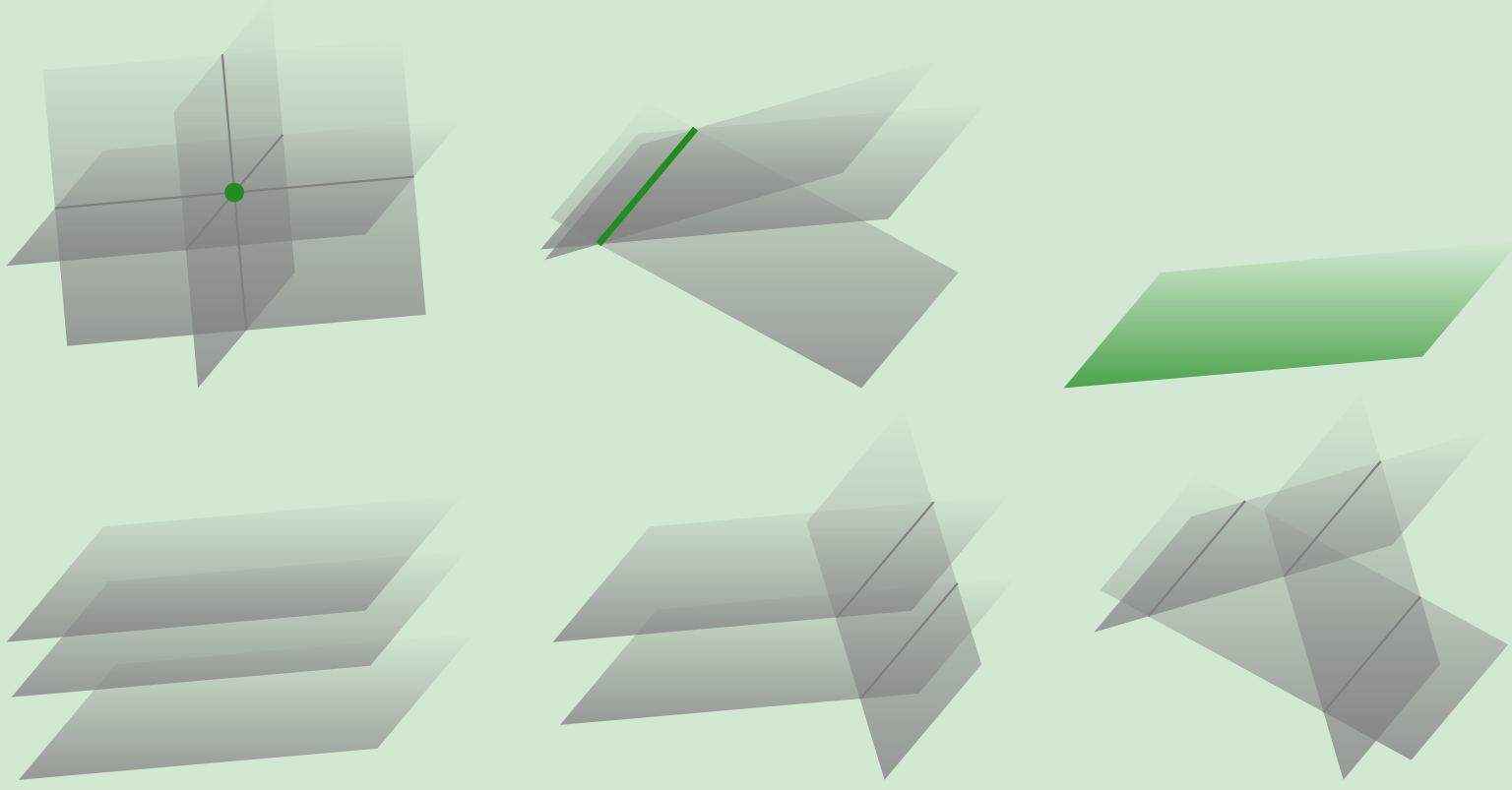
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Example

- Possible configurations of two hyperplanes in $\mathbb{R}^3$:



- Possible configurations of three hyperplanes in $\mathbb{R}^3$:



Remark

We can think about the solutions of a system of linear equations

- as the intersection of hyperplanes (point of view: rows),
- as possible coefficients of the linear combinations for the vector of constants (point of view: columns).

We can solve several systems of linear equations simultaneously if only the constants are different.

Example

Solve the following systems of linear equations:

$$\begin{cases} x+ & y+ & z= & 3 \\ x+ & 2y+ & 4z= & 7 \\ x+ & 4y+ & 10z= & 15 \end{cases} \qquad \begin{cases} x+ & y+ & z= & 5 \\ x+ & 2y+ & 4z= & -2 \\ x+ & 4y+ & 10z= & 1 \end{cases}$$

What does it mean in the above senses?

Linear independence

Definition

The vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k \in \mathbb{F}^n$ are **linearly independent** if

$$\lambda_1\underline{v}_1 + \lambda_2\underline{v}_2 + \cdots + \lambda_k\underline{v}_k = \underline{0} \implies \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0.$$

That is, that only the **trivial** linear combination gives $\underline{0}$.

Otherwise the vectors are **linearly dependent**.

Example

- $\underline{0}$ is linearly dependent,
- $\underline{u}, \underline{v} \in \mathbb{R}^3$ are linearly dependent $\iff$ they are collinear and
- $\underline{u}, \underline{v}, \underline{w} \in \mathbb{R}^3$ are linearly dependent $\iff$ they are complanar.

Theorem

1. $\underline{v}_1 \in \mathbb{F}^n$ is linearly independent $\iff$ $\underline{v}_1 \neq \underline{0}$.
2. $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k \in \mathbb{F}^n$ are linearly independent for $k > 1$ if and only if none of the vectors is the linear combination of the others.

Definition

Let X be a set and $H \subseteq \mathcal{P}(X)$, where $\mathcal{P}(X)$ is the power set (containing the subsets of X).

- $A \in H$ is **maximal** if $A \subseteq B \in H \implies A = B$,
- $A \in H$ is the **largest** in H if $B \subseteq A$ for all $B \in H$,
- $A \in H$ is **minimal** if $A \supseteq B \in H \implies A = B$ and
- $A \in H$ is the **smallest** in H if $B \supseteq A$ for all $B \in H$.

Example

- $\mathrm{Span}(\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k)$ is the smallest subspace containing $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k$.
- In the set $\{(1,0),(0,1),(1,1)\}$ the subsets $\{(1,0),(1,1)\}$ and $\{(0,1),(1,1)\}$ are maximal linearly independent subsets, but there is no largest linearly independent subset.

Definition

Let $V \leq \mathbb{F}^n$ and $U \subseteq V$ a possibly infinite subset.

- The subspace **spanned** by U is

$$\mathrm{Span}(U) = \bigcap_{U \subseteq W \leq V} W.$$

- U is a **generating set** in V if $\mathrm{Span}(U) = V$,
- U is an **independent set** if $\underline{u} \notin \mathrm{Span}(U - \{\underline{u}\})$ for all $\underline{u} \in U$.

Remark (This is consistent with the earlier definitions)

If $U = \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k\}$, then

- $\mathrm{Span}(U) = \mathrm{Span}(\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k)$,
- U is independent $\iff$ $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k$ are linearly independent.

Definition

U is a **basis** of V if it is independent and generating in V.

Lemma (Basic properties of independent and generating sets)

Let $U \subseteq W \subseteq V \leq \mathbb{F}^n$. Then

1. U is generating in $V \implies W$ is generating in V,
2. if U is generating in V and $\underline{u} \in \mathrm{Span}(U - \{\underline{u}\})$, then $U - \{\underline{u}\}$ is generating in V,
3. W is independent $\implies U$ is independent,
4. W is independent and $\underline{w} \notin \mathrm{Span}(W) \implies W \cup \{\underline{w}\}$ is independent.

Theorem (Properties of bases)

Let $U \subseteq V \leq \mathbb{F}^n$. Then the following are equivalent:

1. U is a basis of V,
2. U is a minimal generating set of V,
3. U is a maximal independent set in V,
4. all $\underline{v} \in V$ is a linear combination of the elements of U and the coefficients in the combination are unique.

Theorem (Existence of bases)

1. For any $V \leq \mathbb{F}^n$ there exists a basis.
2. The cardinality of the basis is same for all bases and it is at most n.

Example

- The basis of $\{\underline{0}\}$ is $\emptyset$,
- $\underline{e}_1 = (1, 0, \ldots, 0)$, $\underline{e}_2 = (0, 1, \ldots, 0)$, $\ldots$, $\underline{e}_n = (0, 0, \ldots, 1)$ is a basis of $\mathbb{F}^n$. This is called the **standard basis** of $\mathbb{F}^n$.

Remark

A similar statement is not true for rings (instead of vectorspaces), for example $\{1\}$ and $\{2, 3\}$ are both minimal generating sets in the ring $\mathbb{Z}_6$.

Definition

The **dimension** of a subspace $V \leq \mathbb{F}^n$ is the cardinality of a basis. The notation is $\dim(V)$.

Lemma (Basic properties of the dimension)

- $\dim(\mathbb{F}^n) = n$, thus all of the bases has n elements.

If $V \leq \mathbb{F}^n$ such that $\dim(V) = d$, then

- every generating set of V has at least d elements,
- every independent set in V has at most d elements and
- if $U \subset V$ has m elements, then U is a basis $\iff$ U is generating in V $\iff$ U is independent.

If moreover $U \leq V$, then

- $\dim(U) \leq \dim(V)$,
- $U = V \iff \dim(U) = \dim(V)$.

Definition

Let $B = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k\}$ be a basis of $V \leq \mathbb{F}^n$. For any $\underline{v} \in V$ the **coordinate vector** of $\underline{v}$ with respect to B is

$$[\underline{v}]_B = (\lambda_1, \lambda_2, \ldots, \lambda_k)^T, \text{ where } \underline{v} = \sum_{j=1}^{k} \lambda_j \underline{b}_j.$$

Matrices and linear maps

Matrices

Definition

Let $\mathbb{F}$ be a field and $n, m \in \mathbb{N}$.

- A rectangular array containing m rows and n columns of elements of $\mathbb{F}$ is called an $m \times n$ **matrix** (over $\mathbb{F}$).
- The set of $m \times n$ matrices over $\mathbb{F}$ is denoted by $\mathbb{F}^{m \times n}$.

Example

A matrix of a system of m linear equations with n variables is a matrix in $\mathbb{F}^{m \times n}$.

Remark

- We can view a matrix as a list of column (or row) vectors.
- We usually denote the matrices with latin capital letters and refer the entries of the matrix with the lowercase letters: if A is a matrix then a_{11} is its upper left entry. In most cases the underlined lowercase letters refer to the column vectors: $\underline{a}_1$ is its first column.

Definition

Let $A \in \mathbb{F}^{m\times n}$ be a matrix and $\underline{x} \in \mathbb{F}^n = \mathbb{F}^{n\times 1}$ a column vector. The **product** is the following column vector in $\mathbb{F}^m = \mathbb{F}^{m\times 1}$:

$$A\underline{x} = x_1\underline{a}_1 + x_2\underline{a}_2 + \cdots + x_n\underline{a}_n.$$

Example

- Let $B = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n\}$ be a basis of $\mathbb{F}^n$ and denote B also the matrix in $\mathbb{F}^{n\times n}$ with the corresponding column vectors. If $[\underline{v}]_B = \underline{x}$, then $\underline{v} = B\underline{x}$.
- If $(A|\underline{b})$ is the augmented matrix of a system of linear equations, then it can be written in matrix form: $A\underline{x} = \underline{b}$. Here $\underline{x} \in \mathbb{F}^m$ the vector of the variables. Then equations corresponds to the equations for the entries of the vectors.

Lemma

For any $A \in \mathbb{F}^{m\times n}$, $\underline{x}, \underline{y} \in \mathbb{F}^n$ and $\lambda \in \mathbb{F}$

1. $A(\underline{x} + \underline{y}) = A\underline{x} + A\underline{y}$
2. $\lambda(A\underline{x}) = A(\lambda\underline{x})$.

The column and row space of a matrix

Definition

Let $A \in \mathbb{F}^{m \times n}$.

- The **column space** of A is $\mathcal{C}(A) = \mathrm{Span}(\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n) \leq \mathbb{F}^m$ – the span of its colum vectors.
- Similarly the **row space** of A is $\mathcal{R}(A) \leq \mathbb{F}^n$ is the span of its row vectors.

Theorem

1. A basis of $\mathcal{C}(A)$ is the set of column vectors of A corresponding to the columns of the pivots of $\mathrm{rref}(A)$.
2. A basis of $\mathcal{R}(A)$ is the set of nonzero row vectors of $\mathrm{rref}(A)$.

Remark

The elementary row operations preserve the row space (but not the column space!), so for the row space any reduced echelon form suffices instead of the reduced one.

The rank of a matrix

Definition

The **rank** of a matrix $A \in \mathbb{F}^{m \times n}$ is $\mathrm{rk}(A) = \dim(\mathcal{C}(A))$.

Example

- Compute $\mathrm{rk}(A)$ and a basis of $\mathcal{C}(A)$ and $\mathcal{R}(A)$. What are the coordinates of the column vectors with respect to the previous basis?

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 2 \\ 2 & -1 & 0 & 1 & 3 \\ 0 & 1 & 2 & 1 & 3 \\ 1 & -1 & -1 & 1 & 1 \end{pmatrix}$$

- What is the rank of the following matrices?

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \iddots & \vdots & \vdots \\ 1 & \dots & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 & \dots & n \\ n+1 & n+2 & \dots & 2n \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & n^2 \end{pmatrix}$$

Theorem

$\mathrm{rk}(A) = \dim(\mathcal{C}(A)) = \dim(\mathcal{R}(A)) = \#(\text{pivots in } \mathrm{rref}(A)) = \#(\text{nonzero rows in } \mathrm{rref}(A))$.

Corollary

If $A \in \mathbb{F}^{m \times n}$, then $\mathrm{rk}(A) \leq \min(m, n)$.

Theorem (Solvability of SLE and rank)

Let $A \in \mathbb{F}^{m \times n}$ and $\underline{b} \in \mathbb{F}^m$ and consider the system of linear equations with augmented matrix $(A|\underline{b})$.

1. It is solvable $\iff \mathrm{rk}(A) = \mathrm{rk}(A|\underline{b})$ and
2. It has a unique solution $\iff \mathrm{rk}(A) = \mathrm{rk}(A|\underline{b}) = n$.

Linear maps of vectorspaces

Definition

Let V and W be vectorspaces over the field $\mathbb{F}$. A function $\varphi : V \to W$ is **linear**, if

1. $\varphi(\underline{u} + \underline{v}) = \varphi(\underline{u}) + \varphi(\underline{v})$ for all $\underline{u}, \underline{v} \in V$ and
2. $\varphi(\lambda \underline{v}) = \lambda \varphi(\underline{v})$ for all $\lambda \in \mathbb{F}$ and $\underline{v} \in V$.

Remark

$\varphi(\underline{0}) = \varphi(0 \cdot \underline{0}) = 0 \cdot \varphi(\underline{0}) = \underline{0}$.

Example

- The isometries of $\mathbb{R}^3$ which leave the origin fixed are linear. For example reflections to a plane containing the origin, rotations around an axis through the origin are linear.
- The map projection $\mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (x, y)$ is linear.
- For a fixed basis B of the vector space $\mathbb{F}^n$ the map $V \to V$, $\underline{v} \mapsto [\underline{v}]_B$ is also linear.

Lemma

For any $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n \in \mathbb{F}^m$ there exists a unique linear map $\varphi : \mathbb{F}^n \to \mathbb{F}^m$ such that $\varphi(\underline{e}_k) = \underline{a}_k$ for all $k = 1, 2, \ldots, n$ (here $\underline{e}_k \in \mathbb{F}^n$ is the standard basis vector).

Definition

The **(standard) matrix** of a linear map $\varphi : \mathbb{F}^n \to \mathbb{F}^m$ is $A = (\varphi(\underline{e}_k)_j)_{\substack{1 \le j \le m \\ 1 \le k \le n}} \in \mathbb{F}^{m \times n}$ - i. e. $\underline{a}_k = \varphi(\underline{e}_k)$. Then $\varphi(\underline{x}) = A \cdot \underline{x}$.

Remark

Thus $m \times n$ matrices are naturally in one-to-one correspondence with linear maps $\mathbb{F}^n \to \mathbb{F}^m$.

Example

What is the matrix of following linear maps?

- the rotation by the z-axis with angle α,
- the projection $(x, y, z) \mapsto (x, y)$ and
- the reflection to the plane $x - y + 2z = 0$ – hard to see now, we will return to it later.

Definition

Let $\varphi : V \to W$ be a linear map.

- The **image** of φ is $\mathrm{Im}(\varphi) = \{\varphi(\underline{v}) | \underline{v} \in V\}$.
- The **kernel** of φ is $\mathrm{Ker}(\varphi) = \{\underline{v} \in V | \varphi(\underline{v}) = \underline{0}\}$.
- The **rank** of φ is the rank of its matrix.

Remark (In the language of matrices:)

- $\mathrm{Ker}(\underline{x} \mapsto A\underline{x}) = \mathcal{N}(A)$ – the nullspace and
- $\mathrm{Im}(\underline{x} \mapsto A\underline{x}) = \mathcal{C}(A)$ – the column space.

Lemma

1. $\mathrm{Im}(\varphi) \leq W$ and $\mathrm{Ker}(\varphi) \leq V$ (are subspaces).
2. φ is surjective $\iff \mathrm{Im}(\varphi) = W$.
3. φ is injective $\iff \mathrm{Ker}(\varphi) = \{\underline{0}\}$.

Example

What is the image and the kernel of the projection to the plane $x + 2y + 3z = 0$?

Theorem (Dimension theorem)

If $\varphi : V \to W$ is linear then $\dim(V) = \dim(\mathrm{Ker}(\varphi)) + \dim(\mathrm{Im}(\varphi))$.

Remark

Thus if $A \in \mathbb{F}^{m \times n}$ then $\dim(\mathcal{N}(A)) + \dim(\mathcal{C}(A)) = n$.

Corollary

Let $A \in \mathbb{F}^{m \times n}$. Then a basis of $\mathcal{N}(A)$ is the set of coefficient vectors of the parameters in the solution of the homogeneous system of linear equations.

Example

Find a basis of $\mathcal{N}(A)$, where $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$!

Operations

Definition (Vector space operations)

Let $\mathbb{F}$ be a field, $\lambda \in \mathbb{F}$ and $A, B \in \mathbb{F}^{m \times n}$.

- $A + B \in \mathbb{F}^n$ the matrix such that $(A + B)_{jk} = a_{jk} + b_{jk}$ and,
- $\lambda A \in \mathbb{F}^n$ the matrix such that $(\lambda A)_{jk} = \lambda \cdot a_{jk}$ for all $1 \leq j \leq m$ and $1 \leq k \leq n$.

Remark

In other words $(\mathbb{F}^{m \times n}, +, \cdot)$ is an mn dimensional vectorspace over $\mathbb{F}$. A basis of it contains the matrices E_{jk} - which have a unique nonzero entry: in the jth row and kth column is a 1.

Definition

The **transpose** of $A \in \mathbb{F}^{m \times n}$ is the matrix A^T such $(A^T)_{jk} = a_{kj}$ for all $1 \leq j \leq m$ and $1 \leq k \leq n$.

Remark (This is compatible with the notation of vectors:)

$\underline{v}^T$ is the matrix transpose of the row vector.

Product of matrices

Definition

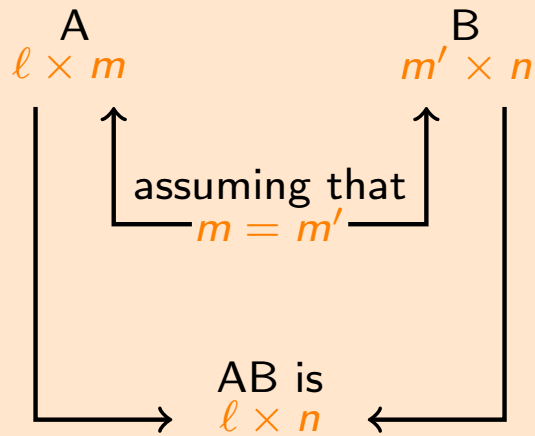
If $A \in \mathbb{F}^{\ell \times m}$ and $B \in \mathbb{F}^{m \times n}$, then the matrix $AB \in \mathbb{F}^{\ell \times n}$ is such that $(AB)_{jk} = \sum_{t=1}^{m} a_{jt} b_{tk}$.

Remark

- This is compatible with the earlier: $A\underline{b}$ equals to the matrix product for $A \in \mathbb{F}^{m \times n}$ and column vectors in $\underline{b} \in \mathbb{F}^{n}$.
- It is important that the product is defined only for matrices with "matching size":

$$\begin{pmatrix} & b_{1k} & \\ & b_{2k} & \\ & \vdots & \\ & b_{tk} & \end{pmatrix}$$

$$\begin{pmatrix} a_{j1} & a_{j2} & \dots & a_{jt} \end{pmatrix} \begin{pmatrix} & c_{jk} & \end{pmatrix}$$



- If $\varphi : \mathbb{F}^{m} \to \mathbb{F}^{\ell}, \underline{x} \mapsto A\underline{x}$ and $\psi : \mathbb{F}^{n} \to \mathbb{F}^{m}, \underline{y} \mapsto B\underline{y}$, then the compostion $\psi \circ \varphi : \mathbb{F}^{n} \to \mathbb{F}^{\ell}$ is the map $\underline{x} \mapsto (AB)\underline{x}$.

Example

Let $A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$, $B = (1\ 2\ 4)$, $C = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $D = \begin{pmatrix} 3 & 1 \\ -1 & 2 \\ 1 & 0 \end{pmatrix}$.

Compute those which are defined: BC, CB, $5A$, $A + D$, AD, DA, $D^T D + A$.

Lemma

Let $A \in \mathbb{F}^{\ell \times m}$ and $B \in \mathbb{F}^{m \times n}$.

1. The columns of AB are the colums A multiplied the columns of B: $A(\underline{b}_1\ \underline{b}_2\ \ldots \underline{b}_n) = (A\underline{b}_1\ A\underline{b}_2\ \ldots A\underline{b}_n)$.
2. Similarly the rows of AB are the rows of A multiplied by B.
3. The columns of AB are the linear combinations of the columns of A with coefficients of the corresponding column of B.
4. Similarly the rows of AB are the linear combinations of the rows of B with coefficients of the corresponding row of A.

Corollary

The map $X \mapsto AX$ performs row operations and $X \mapsto XB$ performs column operations.

Definition

- The $n \times n$ **identity matrix** is $I = I_n$ such that
$$i_{jk} = \begin{cases} 1, & \text{if } j = k \\ 0, & \text{otherwise} \end{cases}$$
- A matrix E is **elementary** if $X \mapsto EX$ performs an elementary row operation.

Lemma

1. $AI = A$ and $IB = B$ for all A and B where it makes sense
2. We can get a matrix of an elementary row operation if we perform it on the identity matrix.

Example

What is the elementary matrix corresponding to adding the double of the first row to the third in a matrix with 5 rows?

Theorem (Operations and rank)

Let $A, B \in \mathbb{F}^{\ell \times m}$ and $C \in \mathbb{F}^{m \times n}$. Then

1. $\mathrm{rk}(A^T) = \mathrm{rk}(A)$,
2. $\mathrm{rk}(A + B) \leq \mathrm{rk}(A) + \mathrm{rk}(B)$ and
3. $\mathrm{rk}(AC) \leq \min(\mathrm{rk}(A), \mathrm{rk}(C))$.

Definition

$A \in \mathbb{F}^{m \times n}$ has **full rank** if $\mathrm{rk}(A) = \min(m, n)$.

Theorem (Rank factorization)

Let $A \in \mathbb{F}^{m \times n}$ with $\mathrm{rk}(A) = r \geq 1$. Then there exist $C \in \mathbb{F}^{n \times r}$ and $R \in \mathbb{F}^{r \times m}$ of full rank such that $A = CR$:
C can be chosen to contain the columns of A corresponding to the pivots in $\mathrm{rref}(A)$ and R to be the nonzero rows of $\mathrm{rref}(A)$.

Example

What is the rank factorization of the matrix $\begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$?

Reminder

The scalar product is a matrix product: for $\underline{u}, \underline{v} \in \mathbb{F}^n = \mathbb{F}^{n\times 1}$ we have $\langle \underline{u}, \underline{v} \rangle = \underline{u}^T \cdot \underline{v} \in \mathbb{F} = \mathbb{F}^{1\times 1}$.

Definition

Let $\underline{u} \in \mathbb{F}^m$ and $\underline{v} \in \mathbb{F}^n$ (column vectors) and $A \in \mathbb{F}^{m\times n}$.

- The **dyadic** or **tensor** product is $\underline{u} \otimes \underline{v} = \underline{u} \cdot \underline{v}^T \in \mathbb{F}^{m\times n}$.
- A is **dyadic** if it is a dyadic product of two vectors.

Corollary

A is dyadic $\iff \mathrm{rk}(A) \leq 1$.

Theorem (Dyadic decomposition)

Any $A \in \mathbb{F}^{m\times n}$ can be written as a sum of r dyadic matrices.
This is minimal in the sense, that A is not the sum of $r-1$ dyadics.

Example

What is the dyadic decomposition of the previous matrix?

Properties of operations

Remark

$(\mathbb{F}^{m\times n}, +)$ is a vectorspace over $\mathbb{F}$, so it has the usual properties.

Theorem

Let A, B and C be matrices over $\mathbb{F}$ and $\lambda \in \mathbb{F}$. If the following expressions make sense then they are equal:

1. $\lambda(AB) = (\lambda A)B = A(\lambda B)$,
2. $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$,
3. $AI = A$ and $IA = A$ and
4. $(AB)C = A(BC)$.

Remark

- The product is not commutative:
$$\begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} -1 & -4 \\ 4 & 10 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 2 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$$
- The product might be zero even if the terms are nonzero:
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Remark

- $(\mathbb{F}^{n\times n}, +, \cdot)$ is a ring: the unit is $I = I_n$, it is not commutative and has zero divisors (if $n > 1$).
- $(\mathbb{F}^{m\times n}, +, \cdot)$ is **not** a ring if $n \neq m$, since $\cdot$ is not an operation on $\mathbb{F}^{m\times n}$.

Theorem

Let A and B be matrices over $\mathbb{F}$ and $\lambda \in \mathbb{F}$. If the following expressions make sense then they hold:

1. $(\lambda \cdot A)^T = \lambda \cdot A^T$,
2. $(A + B)^T = A^T + B^T$ and
3. $(AB)^T = B^T A^T$.

Remark

Thus $\cdot^T : \mathbb{F}^{m\times n} \to \mathbb{F}^{n\times m}$ is a linear map.

Inverse of matrices and linear maps

Definition

Let $f : X \to Y$ and $g : Y \to X$ be functions (or homomorphisms, or linear maps). g is an **inverse** of f if $f \circ g = \mathrm{id}_X$ and $g \circ f = \mathrm{id}_Y$, where id is the identity function (or homomorphism).

Remark (Both are needed:)

For $f : n \mapsto n+1$ and $g : n \mapsto \begin{cases} n-1, & \text{if } n > 1 \\ 1, & \text{if } n = 1 \end{cases}$ as functions $\mathbb{N} \to \mathbb{N}$ we have $f \circ g = \mathrm{id}_{\mathbb{N}} \neq g \circ f$.

Definition

Let A be a matrix over $\mathbb{F}$. Then

- the matrix B is an **inverse** of A if $AB = I$ and $BA = I$ and
- A is **invertible** if it has an inverse.

Example

Consider $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^T$. Are they inverses?

Lemma

If a matrix A is invertible then

1. its inverse is unique (and denoted by A^{-1}),
2. A is a square matrix of full rank
   (i. e. $A \in \mathbb{F}^{n\times n}$ for some $n \in \mathbb{N}$ and $\mathrm{rk}(A) = n$)
3. the linear map $\mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}$, $X \mapsto AX$ is invertible, its unique inverse is $Y \mapsto A^{-1}Y$.

Definition

If a matrix A is invertible, then its unique inverse is denoted by A^{-1}.

Theorem

Let $A \in \mathbb{F}^{n\times n}$. The following are equivalent:

1. A is invertible,
2. $\mathrm{rk}(A) = n$ and
3. $\mathrm{rref}(A) = I_n$.

Then A^{-1} can be computed by solving the systems of linear equations $(A|I)$ simoultaneously: the result is $(I|A^{-1})$.

Corollary

If the matrix A is invertible, the equations $A\underline{x} = \underline{b}$ or $AX = B$ can be solved by multiplying with A^{-1} **from left**: $\underline{x} = A^{-1}\underline{b}$ and $X = A^{-1}B$.

Example

Compute the inverse of the following matrices (if it exists):

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & 1 & 7 \end{pmatrix}$$

Solve the equations $A\underline{x} = \underline{e}_1$ and $CX = B$.

Lemma

Let $A, B \in \mathbb{F}^{n\times n}$ invertible matrices and $\lambda \in \mathbb{F} - \{0\}$. Then

1. $\left(A^{-1}\right)^{-1} = A$,
2. $(\lambda A)^{-1} = \lambda^{-1}A^{-1} = (1/\lambda)A^{-1}$,
3. AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$, in particular $(A^k)^{-1} = (A^{-1})^k$ for $k \in \mathbb{N}$ and
4. $(A^T)^{-1} = (A^{-1})^T$.

Linear transformations

Definition

A **linear transformation** is a linear map $\varphi : V \to V$ of vectorspaces.

Reminder

Let $B \subset \mathbb{F}^n$ be a basis.

- For a vector $\underline{v} \in \mathbb{F}^n$ the coordinate vector $[\underline{v}]_B$ is the unique vector such that $\underline{v} = B \cdot [\underline{v}]_B$.
- For a linear map $\varphi : \mathbb{F}^n \to \mathbb{F}^m$ the matrix of φ is the matrix $A = (\varphi(\underline{e}_1)|\varphi(\underline{e}_2)|\ldots|\varphi(\underline{e}_n)) \in \mathbb{F}^{m \times n}$, where $\underline{e}_k \in \mathbb{F}^n$ are the vectors in the standard basis. Then $\varphi(\underline{v}) = A\underline{v}$.

Remark

The goal is to do the latter in any basis.

Definition

Let $B = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n\}$ be a basis of $\mathbb{F}^n$ and $\varphi : \mathbb{F}^n \to \mathbb{F}^n$ linear. The **matrix** of φ with respect to B is $[\varphi]_B = ([\varphi(\underline{b}_1)]_B | [\varphi(\underline{b}_2)]_B | \ldots | [\varphi(\underline{b}_n)]_B)$.

Theorem

Then $[\varphi]_B [\underline{v}]_B = [\varphi(\underline{v})]_B$.

Example

- Let φ be the projection to the plane $x - 2y + z = 0 \subset \mathbb{R}^3$. Find a basis B, for which $[\varphi]_B$ is "nice"!
- What is the matrix of the reflection to this plane with respect to B?
- A linear transformation of $\mathbb{R}^3$ has matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ with respect to some basis. Is it a reflection?

Definition

Let $B, C \subset \mathbb{F}^n$ bases. The **transformation matrix** $T_{C \leftarrow B} \in \mathbb{F}^{n \times n}$ is the unique matrix for which $[\underline{v}]_C = T_{C \leftarrow B}[\underline{v}]_B$.

Example

- $T_{E \leftarrow B} = (\underline{b}_1 | \underline{b}_2 | \ldots | \underline{b}_n)$, where E is the standard basis.
- $T_{C \leftarrow B} = T_{B \leftarrow C}^{-1}$ for any bases B and C.

Theorem (Change of bases)

Let $\varphi : \mathbb{F}^n \to \mathbb{F}^n$ be a linear transformation, $B, C \subset \mathbb{F}^n$ bases and $P = T_{B \leftarrow C}$. Then $[\varphi]_C = P^{-1}[\varphi]_B P$.

Example

Let φ be the projection to the plane $x - 2y + z = 0 \subset \mathbb{F}^3$. What is the matrix of φ (in the standard basis)?

Special matrices

Definition

Let $A \in \mathbb{F}^{n\times n}$ and $\underline{a} \in \mathbb{F}$.

- The **diagonal** of A is the vector of elements in the form a_{jj}: $\mathrm{diag}(A) = (a_{11}, a_{22}, \ldots, a_{nn}) \in \mathbb{F}^n$.
- A is **diagonal** if $j \neq k \implies a_{jk} = 0$.
- The $\mathrm{diag}(\underline{a})$ is the diagonal matrix with diagonal $\underline{a}$.
- A is **upper triangular** (resp. **lower triangular**) if $j > k \implies a_{jk} = 0$ (resp $j < k \implies a_{jk} = 0$).

Lemma

Let $A, B \in \mathbb{F}^{n\times n}$ have property (P) from above and $\lambda \in \mathbb{F}$. Then

1. λA, $A + B$ and AB have property (P).
2. The diagonal elements are λa_{jj}, $a_{jj} + b_{jj}$ and $a_{jj} b_{jj}$ respectively.
3. A is invertible $\iff a_{11} a_{22} \ldots a_{nn} \neq 0$.
4. If A is invertible A^{-1} has also property (P).

Definition

$A \in \mathbb{F}^{m \times n}$ has **LU decomposition** if there exist a lower triangular matrix L and an upper triangular matrix U such that $A = LU$.

Example (Solving SLE-s with LU decomposition)

Using
$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 8 & 14 \\ 2 & 5 & 13 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$
solve the equation $A\underline{x} = \underline{b} = (1, 3, 6)^T$: write $\underline{y} = U\underline{x}$ and then solve $L\underline{y} = b$ and $U\underline{x} = \underline{y}$.

Remark

- Not all matrices have LU decomposition, for example $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has not, but "most" of them have.
- A has LU decompostion $\iff$ in the Gaussian elimination there was no need to swap rows. Then solving simoultaneously the systems of equations $(A|I)$ we get $(U|L^{-1})$, where U is in row echelon form (hence upper triangular) and L is lower triangular (because the rows were not swapped).In general this is "better".

Definition

$A \in \mathbb{F}^{n \times n}$ is a **permutation matrix** if the set of columns of A is the set of standard basis vectors (i. e. each column / row contains a unique entry equal to 1 and the others are 0).

Remark

Then the rows are permuted by $X \mapsto AX$ and the columns by $X \mapsto XA$.

Lemma

If $A, B \in \mathbb{F}^{n \times n}$ are permutation matrices, then so do AB and $A^{-1} = A^T$.

Theorem (LUP decomposition)

For any matrix $A \in \mathbb{F}^{n \times n}$ there exist

- a permutation matrix P,
- a lower triangular matrix L and
- an upper triangular matrix U

such that $A = LUP$.

Remark

Here the method is to clear lower entries of the columns without reaching row echelon form. Then with some column operations we get P: simoultaneously solving the systems of equations $(A|I)$ we arrive to $(UP|L^{-1})$.

Example

Find the LUP decomposition of $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 1 \end{pmatrix}$!

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & -3 & -1 & -2 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & -3 & 0 & -3 & 1 & 1 \end{array}\right) \longrightarrow P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

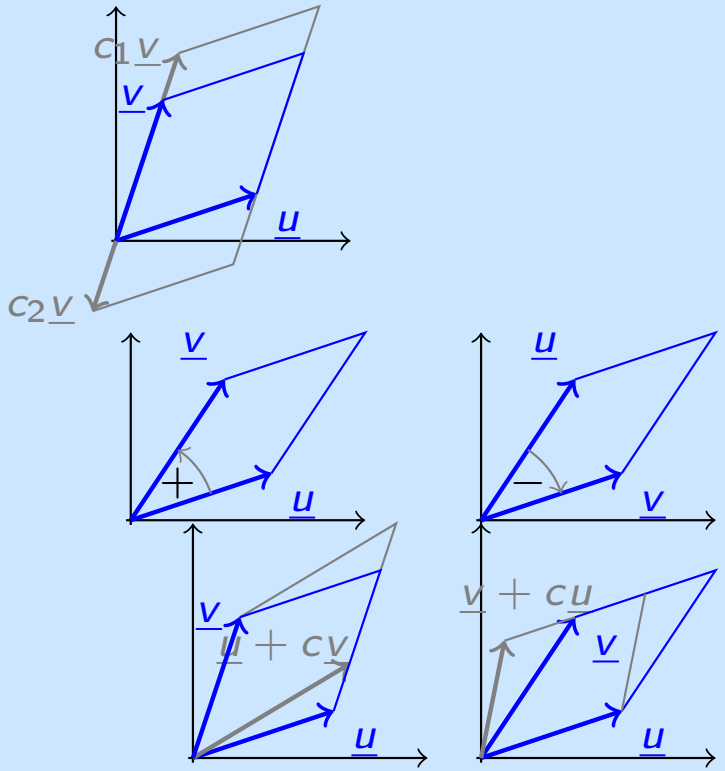
$$U = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} \text{ and } L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix}.$$

Motivation

Reminder (Signed area of parallelograms)

Let $f(\underline{u}, \underline{v})$ be the area of the parallelogram with sides $\underline{u}, \underline{v}$. Then

1. $f(c\underline{u}, \underline{v}) = cf(\underline{u}, \underline{v}) = f(\underline{u}, c\underline{v}) = cf(\underline{u}, \underline{v})$
2. $f(\underline{u}, \underline{v}) = -f(\underline{v}, \underline{u})$
3. $f(\underline{u}, \underline{v}) = f(\underline{u} + c\underline{v}, \underline{v}) = f(\underline{u}, \underline{v} + c\underline{u})$
4. $f(\underline{e}_1, \underline{e}_2) = 1$.



Example

What is the area of the paralelogram with sides $(2, 5), (-3, 1)$

The determinant as a function of the rows

Theorem

Let $\mathbb{F}$ be a field and $n \in \mathbb{N}$. There exists a unique map $\varphi : \mathbb{F}^{n \times n} \to \mathbb{F}$ such that if B is the matrix what is obtained from A

1. by multipliyng a row by $c \in \mathbb{F}$, then $\varphi(B) = c\varphi(A)$,
2. by swapping two rows of A, then $\varphi(B) = -\varphi(A)$,
3. by adding a constant multiple of a row of A to an other row, then $\varphi(B) = \varphi(A)$,
4. $\varphi(I) = 1$.

Remark

- the second is implied by the others,
- this could have been done with columns (instead of rows)

Definition

The above map is the **determinant** and its value is denoted by $\det(A)$ or $|A|$.

Lemma

If A is triangular (upper or lower), then $\det(A) = a_{11}a_{22}\dots a_{nn}$.

Corollary

- The determinant can be computed with Gaussian elimination (without multiplying the rows). If B is the resulting matrix in row echelon form, and the number of row swaps is s, then $\det(A) = (-1)^s \det(B)$.
- The rows can be multiplied as well, but then some bookkeeping must be done.
- The determinant map is unique (if it exists).

Example

Compute $\begin{vmatrix} 2 & 1 & 3 \\ 1 & -1 & 5 \\ 5 & 3 & 1 \end{vmatrix}$ and $\begin{vmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & 1 & 0 \\ 2 & -1 & 3 & 5 \\ 1 & 1 & 1 & 1 \end{vmatrix}$!

Corollary

The following are equivalent:

1. $\det(A) \neq 0$,
2. A is invertible,
3. $\mathrm{rk}(A) = n$,
4. $\mathrm{rref}(A) = I_n$,
5. the equation $A\underline{x} = \underline{0}$ has only the trivial solution and
6. the equation $A\underline{x} = \underline{b}$ is solvable for all $\underline{b} \in \mathbb{F}^n$.

Example

What is the determinant of

- the elementary matrices,
- diagonal matrices,
- permutation matrices and
- products of a diagonal and a permutation matrix?

Properties of the determinant

Lemma

Each invertible matrix $A \in \mathbb{F}^{n\times n}$ can be written as a product of elementary matrices.

Theorem

If $\lambda \in \mathbb{F}$ and $A, B \in \mathbb{F}^{n\times n}$, then

1. $\det(\lambda A) = \lambda^n \det(A)$,
2. $\det(AB) = \det(A)\det(B)$,
3. $\det(A^{-1}) = 1/\det(A)$ if A is invertible and
4. $\det(A^T) = \det(A)$.

Remark

$\det(A+B)$ can not be expressed from $\det(A)$ and $\det(B)$:
Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $A' = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and $B' = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.
Then $\det(A) = \det(B) = \det(A') = \det(B') = \det(A'+B') = 0$
but $\det(A+B) = 1$.

Theorem

$$\begin{vmatrix} \underline{a}_1 \\ \underline{a}_2 \\ \vdots \\ \underline{a}_k \\ \vdots \\ \underline{a}_n \end{vmatrix} + \begin{vmatrix} \underline{a}_1 \\ \underline{a}_2 \\ \vdots \\ \underline{b}_k \\ \vdots \\ \underline{a}_n \end{vmatrix} = \begin{vmatrix} \underline{a}_1 \\ \underline{a}_2 \\ \vdots \\ \underline{a}_k + \underline{b}_k \\ \vdots \\ \underline{a}_n \end{vmatrix}$$ (the rows are equal but the $k$ths)

Remark

- Thus the determinant is multilinear: it "preserves linear combinations" of rows.
- We can decompose the determinants as a sum of determinants of products of a diagonal and a permutation matrix.

Example

Compute $\begin{vmatrix} 4 & 2 \\ 3 & 1 \end{vmatrix}$ in the above way. How many nontrivial terms will be in the decomposition of the determinant of a 3×3 matrix?

The determinant as a function of the entries

Definition

Let π be a permutation of $\{1, 2, \ldots, n\}$ (i. e. a bijective function $\pi : S \to S$). We denote it by $\pi(1)\pi(2)\ldots\pi(n)$.

- The elements i and j are an **inversion** of π if $i < j$ and $\pi(i) > \pi(j)$.
- The **inversion number** of π is the number of inversions of π. The notation is $i(\pi)$.
- π is **even** (resp. **odd**) if $i(\pi)$ is even (resp. odd).

Example

- What is $i(3241)$?
- What is the maximum number of inversions of a permutation of $\{1, 2, \ldots, 6\}$?
- Find a permutation π of $\{1, 2, \ldots, 6\}$ such that $i(\pi) = 7$!

Remark

We can assign a permutation matrix P to a permutation π: the matrix where $p_{jk} = \begin{cases} 1, & \text{if } \pi(j) = k \\ 0, & \text{otherwise} \end{cases}$.

Theorem

Let P and π as above. Then $\det(P) = (-1)^{i(\pi)}$.

Definition

Let $det : \mathbb{F}^{n \times n} \to \mathbb{F}$ be the following

$$det(A) = \sum_{\pi} (-1)^{i(\pi)} a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)},$$

where the sum goes through all the permutations of $S = \{1, 2, \ldots, n\}$.

Remark

This can be defined over commutative rings (not only over fields).

Lemma

1. det is multilinear in the rows of the matrix.
2. If A has two equal rows, then $det(A) = 0$.

Theorem

det satisfies the defining properties of the determinant, so $det = \det$ and thus the determinant exists.

Corollary

If $A \in \mathbb{Z}^{n \times n} \subset \mathbb{Q}^{n \times n}$, then $\det(A) \in \mathbb{Z}$.

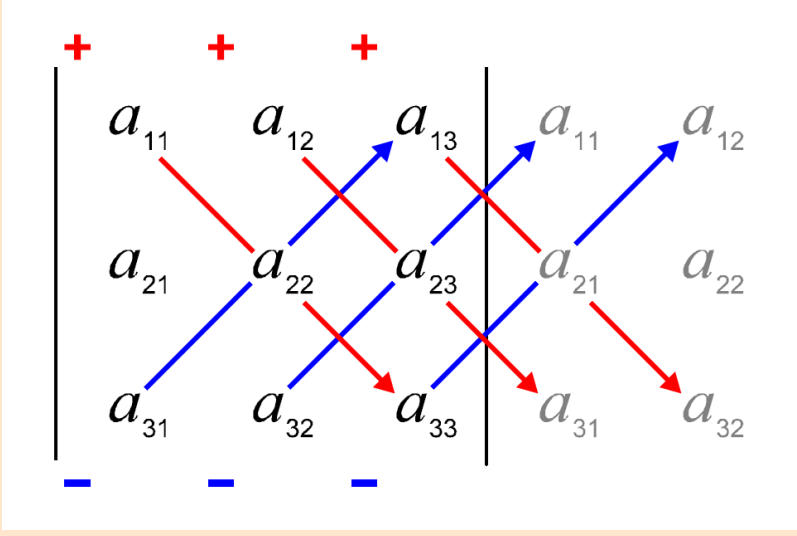
Example

What is $\begin{vmatrix} a & b & 0 & \dots & 0 \\ 0 & a & b & \dots & 0 \\ 0 & 0 & a & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & 0 & 0 & \dots & a \end{vmatrix}$?

Remark

- 2×2 determinants: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$.
- The rule of Sarrus:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$



$$(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31}).$$

Example

Compute $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ and $\begin{vmatrix} 8 & 1 & 3 \\ -1 & 2 & 1 \\ 4 & 5 & 1 \end{vmatrix}$!

Reducing the determinant to smaller ones

Definition

Let $A \in \mathbb{F}^{n\times n}$ and $1 \leq j, k \leq n$. The **cofactor** of A corresponding to a_{jk} is $(-1)^{j+k}$ times the $(n-1)\times(n-1)$ determinant which we get by omitting the jth row and kth column. The notation is A_{jk}.

Lemma

If in the jth row of A the single nonzero element is a_{jk}, then $\det(A) = a_{jk}A_{jk}$.

Remark

The above statement is true for columns instead of rows, as $\det(A^T) = \det(A)$.

Example

What is $\begin{vmatrix} 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 8 & 4 \\ 6 & 0 & 0 & 7 & 0 \\ 8 & 9 & 8 & 7 & 6 \\ 5 & 4 & 0 & 3 & 2 \end{vmatrix}$?

Theorem (Laplace expansion)

For any $A \in \mathbb{F}^{n\times n}$ and $1 \le j \le n$ we have

$$\det(A) = \sum_{k=1}^{n} a_{jk}A_{jk} = \sum_{k=1}^{n} a_{kj}A_{kj}.$$

Example

- Compute $\begin{vmatrix} a & b & 0 & \dots & 0 \\ 0 & a & b & \dots & 0 \\ 0 & 0 & a & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & 0 & 0 & \dots & a \end{vmatrix}$ again with Laplace expansion!
- It is worth combining the techniques we have learnt:

$$\begin{vmatrix} 1 & 1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 3 & 0 \\ 2 & 1 & 0 & 2 & 3 \\ 3 & 1 & 1 & 3 & 0 \\ -1 & 1 & 2 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & -1 & -3 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & -1 & 3 \\ 3 & 1 & 1 & 0 & 0 \\ -1 & 1 & 2 & -4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 & -3 & 1 \\ 2 & 0 & -1 & 3 \\ 3 & 1 & 0 & 0 \\ -1 & 2 & -4 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 4 & -1 & -3 & 1 \\ 2 & 0 & -1 & 3 \\ 0 & 1 & 0 & 0 \\ -7 & 2 & -4 & 1 \end{vmatrix} = -\begin{vmatrix} 4 & -3 & 1 \\ 2 & -1 & 3 \\ -7 & -4 & 1 \end{vmatrix} = -\begin{vmatrix} 0 & 0 & 1 \\ -10 & 8 & 3 \\ -11 & -1 & 1 \end{vmatrix} = -98.$$

Applications of the determinant

Definition

Let $x_1, x_2, \ldots, x_n \in \mathbb{F}$.

- The matrix $V(x_1, x_2, \ldots, x_n) = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix}$ is called the **Vandermonde matrix** and
- $\det(V(x_1, x_2, \ldots, x_n))$ is the **Vandermonde determinant**.

Theorem

$$\det(V(x_1, x_2, \ldots, x_n)) = \prod_{j<k} (x_k - x_j).$$

Corollary

- $\det(V(x_1, x_2, \ldots, x_n)) \neq 0 \iff$ the x_j-s are distinct
- This gives an other proof of the polynomial interpolation theorem.

Adjugate and inverse

Lemma ("Skew expansion")

$$\sum_{k=1}^{n} a_{jk} A_{lk} = \sum_{k=1}^{n} a_{kj} A_{kl} = 0 \text{ for any } A \in \mathbb{F}^{n \times n} \text{ and } j \neq l.$$

Definition

The **adjugate** of $A \in \mathbb{F}^{n \times n}$ is the matrix where the entry in the j-th row and k-th column is A_{kj}. The notation is $\operatorname{adj}(A)$.

Example

Compute $\operatorname{adj} \begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 1 \\ 1 & 1 & 5 \end{pmatrix}$. What is $A \cdot \operatorname{adj}(A)$?

Theorem

$A \cdot \operatorname{adj}(A) = \det(A) \cdot I$

Corollary

- if A is invertible, then $A^{-1} = \dfrac{1}{\det(A)} \operatorname{adj}(A)$
- in particular $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \dfrac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ if $\det = ad - bc \neq 0$.

Remark

This is not so efficient in general if $n \geq 4$.

Example

Show that if $A \in \mathbb{Z}^{n \times n}$ then

$$A^{-1} \in \mathbb{Z}^{n \times n} \iff \det(A) = \pm 1.$$

Cramer's rule

Definition

Let $A \in \mathbb{F}^{n\times n}$ with column vectors $\underline{a}_1, \underline{a}_2, \ldots \underline{a}_n$ and let $\underline{b} \in \mathbb{F}^n = \mathbb{F}^{n\times 1}$. Denote $A_{j,\underline{b}} = (\underline{a}_1, \underline{a}_2, \ldots \underline{a}_{j-1}, \underline{b}, \underline{a}_{j+1}, \ldots \underline{a}_n)$ - the matrix which we get by replacing the jth column with $\underline{b}$ in A.

Theorem (Cramer's rule)

Let $A \in \mathbb{F}^{n\times n}$ with $\det(A) \neq 0$ and $\underline{b} \in \mathbb{F}^n$. The equation $A\underline{x} = \underline{b}$ has a unique solution and $x_j = \dfrac{\det(A_{j,\underline{b}})}{\det(A)}$.

Example

Solve the following system of linear equations with Cramer's rule:

$$\begin{aligned} 2x + 5y &= 4 \\ 5x + 3y &= 6 \end{aligned}$$

Rank and determinant

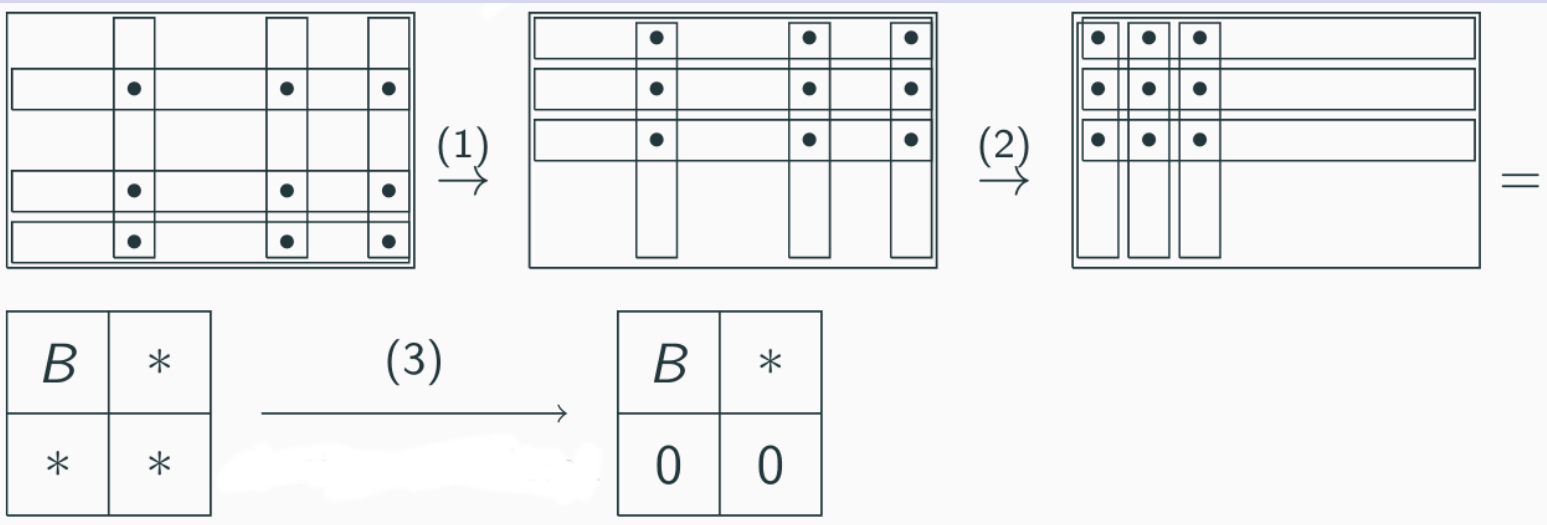
Theorem

Let $A \in \mathbb{F}^{m \times n}$. $\mathrm{rk}(A)$ is the maximal integer r such that A has a nonzero $r \times r$ subdeterminant.

Example

What is the rank of $\begin{pmatrix} 1 & 2 & 0 & -3 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 3 & 0 \\ 2 & 0 & 2 & 2 \end{pmatrix}$?

A figure for the proof.



Block matrices

Definition

A **block matrix** is a matrix partitioned to blocks with respect to a partition of the set of the rows and one of the columns.

Example

$\left(\begin{array}{cc|c} 1 & 2 & 3 \\ \hline 4 & 5 & 6 \\ 7 & 8 & 9 \end{array}\right) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where $A = (1\ 2)$, $B = (3)$, $C = \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}$ and $D = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$. The partition of the rows is $\{1\} \cup \{2, 3\}$ and the one of the columns is $\{1, 2\} \cup \{3\}$.

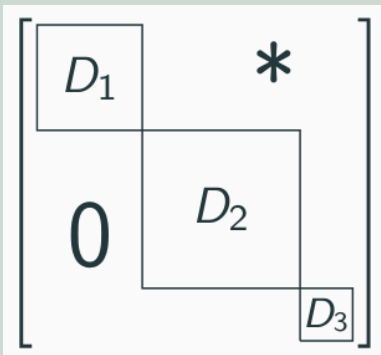
Lemma (Operations with block matrices)

1. If A and B are block matrices of the same size and partitions then in $A + B$ the corresponding blocks are summed.
2. If A and B are block matrices such that AB is defined and the column partition of A equals the row partition of B, then AB can be computed by multiplying the corresponding blocks.

Definition

Let $A \in \mathbb{F}^{n \times n}$ be a block matrix with the same row and column partition.

- A is **block diagonal** if all the blocks outside the diagonal are 0 matrices,
- A is **block upper triangular** if all the blocks below the diagonal are 0 matrices and
- A is **block lower triangular** if all the blocks above the diagonal are 0 matrices.



Lemma

Let A be a block triangular matrix with diagonal blocks $D_1, D_2, \ldots, D_k$. Then $\det(A) = \det(D_1)\det(D_2)\ldots\det(D_k)$.

Corollary

Let $A, B, C, D \in \mathbb{F}^{n \times n}$ such that $\det(A) \neq 0$ and $AC = CA$ and let $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Then $\det(M) = \det(AD - BC)$.

Table of contents

4. Complex numbers
    1. Definition and algebraic properties
    2. The fundamental theorem of Algebra
    3. Trigonometric form and geometric properties
    4. Roots of unity
    5. Binomial sums

5. Polynomials
    1. Basic notions and properties
    2. Number theory of polynomials over fields
    3. The case of $\mathbb{C}[x]$ and $\mathbb{R}[x]$
    4. The case of $\mathbb{Q}[x]$ and $\mathbb{Z}[x]$
    5. Roots of polynomials

6. Systems of linear equations

7. Vectorspaces
    1. The vectorspace $\mathbb{F}^n$
    2. Independent and generating subsets, bases and dimension

8. Matrices and linear maps
    1. Matrices and rank

2. Linear maps
3. Operations
4. Inverse of matrices and linear maps
5. Linear transformations
6. Special matrices
7. The determinant
8. Applications of the determinant