Algebra

(winter semester 08/09)

Emanuel Kieroński

1 Groups, rings, fields

1.1 Algebraic structures

Definition 1 An *algebraic structure* is a tuple $\mathcal{A} = (A, f_1, f_2, \ldots)$ consisting of a set A and operations $f_1, f_2, \ldots$ defined on A.

A is called the *universe* of $\mathcal{A}$. We often denote both algebra and its universe by the same symbol. Each of the operations $f_1, f_2, \ldots$ has a fixed *arity.* An operation of arity n (n-ary operation) is a function from A^n to A. Operations of arity 0 are called *constants*; a constant is identified with a distinguished element of the universe. We will usually work with algebraic structures with a finite set of (usually binary) operations.

A note about external operations (...).

Example 2 Examples of algebraic structures:

(a) $(\mathbb{Z}, +, -, \cdot, 0)$.

(b) $(\mathcal{P}(\{1, 2, \ldots, n\}, \cup, \cap, ')$.

In the following definition we distinguish some properties which may be enjoyed by binary operations.

Definition 3 (a) an operation $\cdot$ is *commutative*, if $\forall a, b \in A \;\; a \cdot b = b \cdot a$,

(b) an operations $\cdot$ is *associative*, if $\forall a, b, c \in A \;\; (a \cdot b) \cdot c = a \cdot (b \cdot c)$; for example, the power operation in the set of natural numbers is not associative,

(c) an element e is a *left identity* or a *left neutral element* of an operation $\cdot$, if $\forall a \in A \;\; ea = a$.

(d) an element e is a *right identity* or a *right neutral element* of an operation $\cdot$, if $\forall a \in A \;\; ae = a$.

(e) an element e is an *identity* or a *neutral element* of the operaton $\cdot$, if $\forall a \in A \;\; ea = ae = a$.

(f) for an operation with an identity e, an element b is a *left inverse* of a, if $b \cdot a = e$,

(g) for an operation with an identity e, an element b is a *right inverse* of a, if $a \cdot b = e$,

(h) an element a is an *inverse* of b, if $ab = ba = e$

(i) an operation $\cdot$ is *distributive* over an operation $+$, if $\forall a, b, c \in A \quad a \cdot (b + c) = a \cdot b + a \cdot c$ and $\forall a, b, c \in A \;\; (b + c) \cdot a = b \cdot a + c \cdot a$.

The following facts can be easily proved:

Fact 4 (i) *If an operation has a right identity e_l and a left identity e_r, then $e_l = e_r$. Thus, an operation has at most one identity.*

(ii) *If an operation $\cdot$ is associative and has an identity e, then for each element a, if b_l is a left inverse of a, and b_r is a right inverse of a, then $b_l = b_r$. Thus, each element has at most one inverse.*

Some classes of algebraic structures have a special significance. Now we introduce three of them.

Definition 5 (a) An algebraic structure $(A, \cdot)$, with a binary operation $\cdot$ is called a *group* if:

- $\cdot$ is associative,
- $\cdot$ has an identity,
- each element has an inverse.

additionally, if $\cdot$ is commutative, then the group is called *commutative* or *abelian.*

(b) An algebraic structure $(A, +, \cdot)$ with binary opeations $+$, $\cdot$ is called a *ring* if:

- $(A, +)$ is an abelian group,
- $\cdot$ is associative,

- $\cdot$ is distributive over $+$.

(c) $(A, +, \cdot)$ is a *field* if

- $(A, +, \cdot)$ is a ring,
- $(A \setminus \{0\}, \cdot)$ is a commutative group (by 0 we denote the identity of $+$);

A few examples of groups are given in the next section. Examples of fields are *number fields*, e.g. the set of rational numbers with naturally defined $+$ and $\cdot$. One can construct also *finite* fields. An example of a ring which is not a field: the set of integers with naturally defined $+$ and $\cdot$ (or the set of even integers – it even lacks an identity of $\cdot$). A very important role is played by the rings $\mathbb{Z}_n = (\{0, 1, \ldots n-1\}, +_n, \cdot_n)$, with addition and multiplication *modulo n*.

1.2 Groups - basic properties and examples

The notion of a group was introduced in the previous section. There are also some similar relaxed notions: a *semigroup* is a non-empty set with an associative operation, and a *monoid* is a semigroup with an identity.

Conventions. The group operation is often called *multiplication*; we write ab instead of $a \cdot b$; we use 1 to denote the identity; we use a^{-1} to denote the inverse of a; a^n denotes the result of multiplying a n-times by itself $a \cdot a \cdot \ldots \cdot a$ (the n-th power of a). This „style" of speaking about groups is called *multiplicative*. Alternatively, we can use the *additive* style: $+$ for the operation; 0 for the identity; $-a$ for the inverse of $-a$. In these notes we usually use the multiplicative style but denote the identity by e.

Remark. The notion of a group may be defined in a slightly different way: as a set with a binary operation $\cdot$, unary operation $^{-1}$ and a constant 1, with the appropriate properties of the operations.

1.3 Examples

Example 6 We give a few examples of groups (and check that they are groups indeed):

(a) $(\mathbb{Z}, +)$ – the set of integers with addition

(b) $(\mathbb{R} \setminus \{0\}, \cdot)$ – the set of non-zero real numbers with multiplication,

(c) the set of bijections from X to X with the operation of function composition, for an arbitrary non-empty set$X$,

(d) $(\mathbb{Z}_4, +_4)$ – $\{0, 1, 2, 3\}$ with addition modulo 4 ($a +_4 b$ is defined as the remainder from the division of $a + b$ by 4,

(e) $(\mathbb{Z}_5^*, \cdot_5)$ – $\{1, 2, 3, 4\}$ with multiplication modulo 5,

(f) $\{1, 3, 5, 7\}$ with multiplication modulo 8 (the Klein four-group),

(g) the group of rotations of a square (with composition),

(h) the group of symmetries of a square, with composition (a *symmetry* is a transformation preserving the distances between points). There are 8 symmetries of a square: 4 rotations (including identity) and 4 reflections (through the horizontal line, vertical line and two diagonals).

All the above groups, except (c) and (h) are abelian.

1.4 Group tables

To define a group we can use a *group table*. Below we present a table of a group from point (f) of Example 6:

$\cdot$	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

And here is a table of the group of symmetries of a square:

$\cdot$	id	r	r'	r''	h	v	d	d'
id	id	r	r'	r''	h	v	d	d'
r	r	r'	r''	id	d'	d	h	v
r'	r'	r''	id	r	v	h	d'	d
r''	r''	id	r	r'	d	d'	v	h
h	h	d	v	d'	id	r'	r	r''
v	v	d'	h	d	r'	id	r''	r
d	d	v	d'	h	r''	r	id	r'
d'	d'	h	d	v	r	r''	r'	id

Let us observe the properties of group tables:

Observation 7 (i) *there exists an element (identity) whose row and column are exactly the same as the row and column describing the elements of the universe*

(ii) *each row and each column contain the identity; the identities are placed symmetrically through the main diagonal of the table.*

(iii) *each row and each column is a permutation of the universe.*

The above properties are not sufficient conditions for being a group - we still have to ensure that the gropu operation is associative. The last observation from 7 is a consequence of the following lemma:

Lemma 8 *In a group, the equlities $ax = b$ and $ya = b$ have unique solutions.*

Corollary 9 *In a group, cancellations law hold. Left cancellation law: $ab = ac$ implies $b = c$. Right cancellation law: $ba = ca$ implies $b = c$.*

1.5 Group isomorphisms

We define the notion of a *group isomorphism.*[1]

Definition 10 We say that groups $(A, \cdot_1)$ and $(B, \cdot_2)$ are *isomorphic* if there exists a bijection $F : A \to B$, such that $\forall a, b \in A\ a \cdot_1 b = c$ if and only if $F(a) \cdot_2 F(b) = F(c)$. F is called an *isomorphism* between groups A and B.

It is not difficult to see that isomorphism preserves all the properties of the group operation. In particular:

Fact 11 *If F is an isomorphism from A to B, then F returns the identity of B for the identity of A, and the inverse of $F(a)$ in B for the inverse of a in A.*

This means that isomorphic groups have identical structures and essentially differs only in names of elements. The following observation is straightforward:

Fact 12 *The relation on the set of all groups, containing exactly those pairs of groups which are isomorphic, is an equivalence relation.*

Example 13 Groups (d), (e) and (g) from Example 6 are pairwise isomorphic. Groups (d) and (g) are not isomorphic.

Hence, the groups (d), (e) and (g) are in fact „incarnations" of the same abstract object. Soon we will see that, *up to isomorphism* there are only two four-element groups.

1.6 Orders of elements and orders of groups

We define the power of an element in a group in a natural way.

Definition 14 (a) $a^0 = e$, where e is the identity

(b) $a^m = a \cdot a^{m-1}$, for positive m

(c) $a^m = (a^{-1})^m$, for negative m

The following equalities hold:

Fact 15 (i) $a^r a^s = a^{r+s}$

[1] The notion of isomorphism can be naturally generalized to other algebraic structures.

(ii) $(a^r)^s = a^{rs}$.

Note however that $(ab)^n = a^n b^n$ is not necessarily true in a group (but it holds in *abelian* groups).

Definition 16 The *order of an element* a in a group is the smallest positive integer m such that $a^m = e$. If no such m exists, the group has an *infinite* order. The *order of a group* is its cardinality, i.e. the number of its elements.

A few examples were given during the lecture. Question: is it possible that an element of a finite group has infinite order? No.

1.7 Subgroups, generators, and cyclic groups

Definition 17 We say that B is a *subgroup* of a group A if $B \subseteq A$ and B is a group.

If $(A, \cdot)$ is a group then the set containing only the identity is its subgroup. According to the definition a whole group itself is also its own subgroup. These two special subgroups are called *trivial*.

Example 18 (a) The group of rotations of a square is a subgroup of the group of symmetries of this square.

(b) The set of even integers with addition is a subgroup of integers.

Note that B has to contain the identity; for every $a \in B$ the inverse of a a^{-1} has to belong to B; and B has to be closed under the operation $\cdot$ $(a, b \in B \Rightarrow ab \in B)$.

For finite groups we can even prove:

Lemma 19 *A non-empty subset H of a finite group G is its subgroup if and only if $\forall a, b \in H$ we have $ab \in H$ (in other words: H is closed under the operation $\cdot$).*

Proof: Let us tak an element $a \in H$. It has a finite order m in G, $a^m = e$. Hence, a^{m-1} is the inverse of a. □

Definition 20 (a) Let G be a group and X its non-empty subset. The smallest subgroup of G containing all elements of X is called a subgroup *generated* by X.

(b) If a group G is generated by a singleton set $\{a\}$ (for simplicity we also say that G is generated by the element a, or that a is a *generator* of G), then G is called *cyclic*.

Fact 21 (i) *The subgroup generated by X consists of all possible products of the form $x_1 x_2 \dots x_k$, where $x_i \in X$ or $x_i^{-1} \in X$, for $k \in \mathbb{N}$. Not all of this products are distinct of course.*

(ii) *If a is a generator of a finite (sub)group H, then $H = \{a, a^2, a^3, \dots a^m\}$, for the smallest $m > 0$, such that $a^m = e$. Moreover $a^i \neq a^j$ for $i \neq j, 0 < i, j \leq m$.*

Note that the number of elements of a cyclic group equals the order of its generator. A cyclic group may have more than one generator.

Example 22 (a) A generator of $(\mathbb{Z}, +)$ is 1.

(b) A generator of a group of rotations of a square is r – 90 degree rotation (or r'' – 270 degree rotation).

(c) A generator of the additive group $(\mathbb{Z}_4, +_4)$ is 1 (or 3).

(d) The Klein four-group is not cyclic. To generate this group we need at least two elements, e.g $\{3, 5\}$.

(e) The subgroup of the Klein four-group generated by 3 is $\{1, 3\}$.

(f) The subgroup of the group of symmetries of a square, generated by r (90 degree rotation) consists of all rotations (including identity).

(g) The group of symmetries of a square is not cyclic. It can be generated, e.g. by $\{r, d\}$.

Theorem 23 *If a group G is cyclic, and a its generator, then the order of a defines G up to isomorphism. More precisely, if the order of a is infinite then G is isomorphic to $(\mathbb{Z}, +)$, and if this order is k, then G is isomorphic to the additive group $(\mathbb{Z}_k, +_k)$.*

1.8 Groups of permutations

An important class of groups are *groups of permutations.* A permutation of a set X is a bijection of X to itself. In the example 6(c) we observed that the set of permutations of a set is a group with the operation of composition. This group if not abelian if $|X| > 2$.

We usually work with permutations of the sets $\{1, 2, 3, \dots, n\}$. The group of such permutations is denoted by S_n. A permutation f can be represented in a two-row form:

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \text{f(1)} & \text{f(2)} & \text{f(3)} & \dots & \text{f(n).} \end{pmatrix}$$

Remarks about the operation of composition (how to compute, S_n is closed under composition, inverses)

An important class or permutations are *cycles.*

Definition 24 A *k-cycle* (cycle of length k) is a permutation f of the set $X = \{1, 2, \dots, n\}$ $(k \leq n)$, such that there exist $1 \leq a_1, a_2, \dots a_k \leq n$ $(a_i \neq a_j$ dla $i \neq j)$, such that $f(a_1) = a_2$, $f(a_2) = a_3$, …, $f(a_k) = a_1$ and $f(a) = a$ for $a \notin \{a_1, \dots a_k\}$. A cycle is denoted by $(a_1, a_2, \dots, a_k)$. 2-cycles are also called *transpositions.* Cycles $(a_1, a_2, \dots a_{k_1})$ and $(b_1, b_2, \dots b_{k_2})$ are *disjoint* if $a_i \neq b_j$ for all i, j.

Note that disjoint f, g cycles commute, i.e. $fg = gf$.

The set of permutations S_n has the following properties:

Theorem 25 (i) *Any permutation can be uniquily expressed as a product (composition) of disjoint cycles.*

(ii) *Any permutation can be expressed as a product of transpositions (not necessarily disjoint; not uniquely)*

(iii) *Any transposition can be expressed as a product of transpositions of an odd number of transpositions of neighbouring elements. Hence, S_n is generated by the set of all transpositions of neighbouring elements.*

Proof:

(i) Easy.

(ii) Implied by $(a_1, a_2, \dots, a_k) = (a_1, a_k)(a_1, a_{k-1}) \dots (a_1, a_3)(a_1, a_2)$.

(iii) Implied by $(j, l) = ((j, j+1)(j+1, j+2) \dots (l-2, l-1)(l-1, l)(l-2, l-1) \dots (j+1, j+2)(j, j+1)$.

□

Fact 26 *In the group S_n:*

(i) *A k-cycle has order k,*

(ii) *the order of any permutation is the least common multiplier of the orders of the cycles appearing in the decomposition of the permutation into disjoint cycles.*

Definition 27 Let $f \in S_n$ and $P(f) = \prod_{1 \leq i < j \leq n}(f(j) - f(i))$. The sign of P (1 for positive P, -1 for negative P) is called a *sign of permutation* and denoted by $\mathrm{sgn}(f)$. Permutations with sign 1 are called *even*, with sign -1 – *odd.*

Note that $P(f)$ and $P(g)$ for $f, g \in S_n$ have the same absolute value, but this value does not play an important role in the definition of parity of a permutation. An alternative definition uses the notion of an *inversion*: elements $f(i)$ and $f(j)$, $i < j$ are an inversion in f if $f(i) > f(j)$; a permutation is even if and only if it has an even number of inversions. It is not difficult to see that this definition is equivalent to Definition 27.

Lemma 28 *Let f be a permutation, and t a transposition from S_n. Then $sgn(f) = -sgn(ft)$.*

Proof: We prove a lemma for t a transposition of neighbouring elements. Then we use Theorem 25, part (iii).

Lemma 29 *A permutation f is even if and only if in its any decomposition into transpositions the number of transpositions is even.*

Proof: Let us decompose f into a product of transpositions: $f = t_1 t_2 \ldots t_k$. We can thus write $f = i_n t_1 t_2 \ldots t_k$, where i_n is the identity in S_n. Clearly, $\mathrm{sgn}(i_n) = 1$. We use lemma 28. □

Using lemma 29 it is easy to see:

Fact 30 (i) *A product of two even permutations is an even permuation.*

(ii) *A product of two odd permutations is an even permuation.*

(iii) *A product of an even and an odd permutation is odd.*

(iv) *The inverse permutation of an even permutation is even.*

(v) *The inverse permutation of an odd permutation is odd.*

Points (i), (iv) (or point (i) and lemma 19) imply:

Fact 31 *The set of all even permutations of S_n is its subgroup.*

The above group is denoted by A_n and called an *alternating* group.

It is easy to show that $|S_n|$ (the order of S_n) equals $n!$. We prove that exactly one half of permutations in S_n is even.

Lemma 32 *For $n > 1$: $|A_n| = n!/2$.*

Proof: Let $f_1, f_2, \ldots, f_k$ be a list of all even permutations, and t – any transposition. We show that $tf_1, tf_2, \ldots, tf_k$ is a list of all odd permutations in S_n and that $tf_i \neq tf_j$ if $i \neq j$. □

Now we observe that every finite group is essentially a subgroup of a group of permutations.

Theorem 33 (Cayley) *Let G be a finite group of order n. Then G is isomorphic to a subgroup of S_n.*

Proof: W.l.o.g we can assume that $G = \{1, 2, 3, \ldots, n\}$. We construct a function $F : G \to S_n$. For $a \in G$, we define $F(a) = f_a$, where $f_a : G \to G$ is such that $f_a(b) = ab$ for all $b \in G$. It is easy to see that f_a is a bijection for any $a \in G$. Similarly, F is 1-1 (since $f_a \neq f_b$ for $a \neq b$). Now we show that the image $F(G)$ is a subgroup of S_n. By Lemma 19 it is sufficient to show that the product of two bijections from $F(G)$ is in $F(G)$. This is implied by the equality $f_a f_b = f_{ab}$. This equality proves also that F preserves the operation $\cdot$. □

1.9 Cosets and Lagrange's Theorem

Definition 34 We define an operation $\cdot$ on the set of subsets of a group. Let X and Y be subsets of a group G. $X \cdot Y := \{xy : \ x \in X, y \in Y\}$. We simplify the notation for singleton sets, and e.g. instead of $\{a\} \cdot Y$ we write aY.

Remark The operation defined above is associative and has an identity, but none of the sets with at least 2 elements has an inverse.

Every subgroup defines *cosets* of a group.

Definition 35 Let H be a subgroup of a group G. A *right coset* of H is each of the sets Ha, for $a \in G$. Similarly, a *left coset* is each of the sets aH, for $a \in A$. In particular H is its own right and left coset (since $H = He = eH$). The number of right coset is called an *index* of a subgroup H.

Lemma 36 (i) *If H is finite then each coset of H has $|H|$ elements.*

(ii) *For every pair W_1, W_2 of right (left) cosets of a subgroup H either $W_1 = W_2$ or $W_1 \cap W_2 = \emptyset$. Thus right (left) cosets of H for a division of G.*

Proof:

(i) Implied by canellation laws.

(ii) Let $x \in Ha$ and $x \in Hb$. Then $x = h_1a = h_2b$, for some $h_1, h_2 \in H$. Let $y \in Ha$. Then $y = h_3a = h_3h_1^{-1}x = h_3h_1^{-1}h_2b \in Hb$, since $h_3h_1^{-1}h_2 \in H$. Thus $Ha \subseteq Hb$. Similarly we show that $Hb \subseteq Ha$. Of course each element a is a member of the coset Ha.

□

The above Lemma implies in particular that each subgroup has the same number of left and right cosets.

Example 37 (a) Consider the group of symmetries of a square from Example 6. $H = \{i, h\}$ is its subgroup. It defines four left cosets: $Hi : \{i, h\}$, $Hr' : \{r', v\}$, $Hr : \{r, d\}$, and $Hr'' = \{r'', d'\}$.

(b) Let $G = S_6$ and H be the set of all f for which $f(1) = 1$. We have $6!/5! = 6$ left cosets (since $|H| = 5!$). Each of the coses is determined by the value of permutations on 1.

The consequence of Lemma 36 is the following theorem:

Theorem 38 (Lagrange) *The order of a finite group is a multiple of the order of its every subgroup.*

Corollary 39 *The order of an element of a finite group G divides $|G|$*

Proof: An element a generates a cyclic subgroup: $\{a, a^2, a^3, \ldots, a^m = e\}$ G.

Corollary 40 *In a finite group of order k, for every a we have $a^k = e$.*

Corollary 41 *Every group G, whose order is a prime number is cyclic.*

Thus we have, up to isomorphism, only one group of order 5 – $(\mathbb{Z}_5, +_{\bmod 5})$.

1.10 Groups acting on sets

Definition 42 We say that a group G acts on a set X if there exists a function from $G \times X$ into X, (notation: $(a, x) \to a.x$) such that:

(a) $e.x = x$ for all $x \in X$,

(b) $a.(b.x) = (ab)x$ for all $a, b \in G, x \in X$

Fact 43 *For every $a \in G$, the function $f_a(x) = a.x$ is a bijection of X into X.*

Definition 44 Assume that a group G acts on a set X.

(a) A *stabilizer* of an element $x \in X$ is the subgroup $stab(x) = \{a \in G : a.x = x\}$.

(b) An *orbit* of an element $x \in X$ is the set $orb(x) = G.x = \{a.x : a \in G\}$

Example 45 A subgroup H acts on the group G in a few ways:

(a) $h.a = ha$. In this case, orbits are right cosets of H. Stabilizer: $\{e\}$

(b) $h.a = ah^{-1}$. Orbits are left cosest of H.

Example 46 Every subgroup H of the permutation group S_n acts on the set $\{1, 2, \dots, n\}$ in a natural way: $f.x = f(x)$. $stab(x)$ is the set of permutations that do not affect x. $orb(x)$ is the set of those elements which x may be moved to by permutations from H. What are the orbits of the subgroup of S_8 generated by $(1, 7, 4, 2)(6, 8)$? (answer: {1,2,4,7}, {3}, {5}, {6,8})

Theorem 47 *The set of orbits forms a division of X. In other words: for all $x, y \in X$ we have either $orb(x) = orb(y)$ or $orb(x) \cap orb(y) = \emptyset$.*

Theorem 48 *Let G acts on a set X. Then for each $x \in X$ we have*

$$|G| = |stab(x)||orb(x)|.$$

Proof: We know that $stab(x)$ is a subgroup of G. By Lemma 36 it is enought to show the the number of the cosets of $stab(x)$ equals the number of the orbits of x. We define a bijection F from the set of left cosets of $stab(x)$ to $orb(x)$: $F(a \cdot stab(x)) = a.x$. ***Soundness of definition:*** If $astab(x) = bstab(x)$, then $b^{-1}a \in stab(x)$, so $b^{-1}a.x = x$, And thus $a.x = b.x$. F ***is „onto”:*** obvious. F **is „1-1”:** if $a.x = b.x$, then $b^{-1}a.x = x$, and thus $b^{-1}a \in stab(x)$, which implies that $a \in bstab(x)$.

□

Now we formulate an important combinatorial lemma. Let G acts on X. Let X/G denotes the set of orbits, and $fix(x) = \{x \in X : a.x = x\}$.

Lemma 49 (Burnside)

$$|X/G| = \frac{1}{|G|}\sum_{x \in X}|stab(x)| = \frac{1}{|G|}\sum_{a \in G}|fix(a)|.$$

Proof: We check first that both formulas give the same results. Let S denotes the set of pairs $(a, x) \in G \times X$ such that $a.x = x$. For a fixed a the number of such pairs is $|fix(a)|$. Thus $|S| = \sum_{a \in G}|fix(a)|$. On the other hand, for a fixed x the number of such pairs is $|stab(x)|$. Thus $|S| = \sum_{x \in X}|stab(x)|$. By theorem 48 the elements from the same orbit $orb(x)$ have the stabilizers of the same cardinality $\frac{G}{|orb(x)|}$. This gives: $\sum_{x \in X}|stab(x)| = \sum_{O \in G/X}(\sum_{x \in O}|stab(x)|) = \sum_{O \in X/G}|O| \cdot \frac{|G|}{|O|} = |G||X/G|$. □

Example 50 Let B be the set of boolean functions of three variables, i.e. function of the type $\{0, 1\}^{\{x_1, x_2, x_3\}} \to \{0, 1\}$. Such functions may be implemented as logical circuits with three inputs and one output. What is the minimal number of circuits which are necessary to implement an arbitrary function from B? Obviously, $|B| = 2^{2^3} = 256$, but we do not really need 256 circuits. - a given circuit may be used to compute several functions by reconnecting the wires with input signals. For example, using a circuit returning 1 if and only if its first input is set to 1 and the remaining inputs are 0s we can implement three boolean functions.

Let us define an action of S_3 on B. For $f \in S_3$ and $b \in B$, $(f.b)(y_1, y_2, y_3) = b(y_{f^{-1}(1)}, y_{f^{-1}(2)}, y_{f^{-1}(3)})$. Note, that the elements belonging to the same orbit can be implemented by the same circuit. So the question is: how many orbits is defined by the action of S_3 on B? We count the number of fixed points of all permutations:

(a) $|fix(id)| = 256$

(b) $|fix((1, 2))| = |fix((2, 3))| = |fix((1, 3))| = 2^6$, since, e.g. functions from $fix((1, 2))$ satisfy $f(x_1, x_2, x_3) = f(x_2, x_1, x_3)$, so to define them 6 values is required.

(c) $|fix((1, 2, 3))| = |fix(1, 3, 2)| = 2^4$, analogously, 4 values required.

Finally, by Burnside's Lemma we have $|B/S_3| = \frac{1}{6}(256 + 3 * 64 + 2 * 16) = 80$.

1.11 Euclidean Algorithm

For integers a, b we write $a|b$ if a divides b, i.e. if there exists an integer k such that $b = ka$. A number $n > 1$ is *prime* if its only positive divisors are 1 and n.

Fact 51 *For integers a and b, $b \neq 0$ there exists exactly one non-negative integer $r < b$ such that for some q we have $a = qb + r$; the number r is called the* reminder *of the division of a by b and is denoted by $a \mod b$.*

For any pair of integers a, b there exists their *greatest common divisor*, denoted $gcd(a, b)$. If $gcd(a, b) = 1$ then a and b are called *relatively prime*.

The greatest common divisor of positive numbers m and n can be computed by *Euclidean algorithm*:

```
(1) $m_0 := m, : n_0 = n$.
(2) $i = 0$
(3) If $m_i = 0$ return $n_i$; If $n_i = 0$ return $m_i$.
(4) If $m_i > n_i$, then $m_i := m_i \mod n_i$ else $n_i := n_i \mod m_i$
(5) $i := i + 1$. Go to 3.
```

Note first, that the algoritm always terminates: the numbers m_i and n_i are nonnegative and their sum decreases so the process cannot last infinitely. To see that the algorithm is correct we prove the following lemma:

Lemma 52 $gcd(a, b) = gcd(a \mod b, b)$.

Now it is easy to see that the $gcd(m_i, n_i)$ is preserved by the algorithm and so the value returned is correct. An important consequence of the algorithm is the following theorem:

Theorem 53 *Let m and n be integer numbers. Then there exist integers a and b such that $am + bn = gcd(m, n)$.*

Proof: We inductively prove that the values of m_i and n_i are in the set $\{am + bn : a, b \in \mathbb{Z}\}$. One of this values is returned as $gcd(m, n)$.

The values of a and b from Theorem 53 can be computed by the *extended* Euclidean algoritm.

Example 54 Computations for $m = 81$ and $n = 57$:

$$\begin{aligned} 81 &= 1 \cdot 57 + 24 \\ 57 &= 2 \cdot 24 + 9 \\ 24 &= 2 \cdot 9 + 6 \\ 9 &= 1 \cdot 6 + 3 \\ 6 &= 2 \cdot 3 + 0 \end{aligned}$$

So $gcd(81, 57) = 3$. To find a and b we „reverse” the computations:

$$\begin{gathered} 3 = 9 - 1 \cdot 6 \\ 3 = 9 - 1 \cdot (24 - 2 \cdot 9) = -1 \cdot 24 + 3 \cdot 9 \\ 3 = -1 \cdot 24 + 3 \cdot (57 - 2 \cdot 24) = 3 \cdot 57 - 7 \cdot 24 \\ 3 = 3 \cdot 57 - 7 \cdot (81 - 57) = -7 \cdot 81 + 10 \cdot 57 \end{gathered}$$

A particular corollary of 53 is:

Corollary 55 $gcd(m, n) = 1$ *if and only if* $am + bn = 1$ *for some integers a i b.*

1.12 Modular arithmetic

Recall that the result of $a +_m b$, for $a, b \in \mathbb{Z}$, is defined as the reminder of division of $a + b$ by m, i.e. $(a + b) \mod m$ (and similarly for $\cdot_m$). We observed that $\mathbb{Z}_m = \{0, 1, \ldots, m - 1\}$ forms a commutative group with $+_m$ for all $m > 0$.

We introduce a relation $\equiv_m$ in the set of integers: $a \equiv_m b$ if and only if $a - b = km$ for some $k \in \mathbb{Z}$ (i.e. when m divides $a - b$, which is equivalent to the fact that the reminders of a and b by m are equal). It is easy to check that $\equiv_m$ is an equivalence relation with m equivalence classes, determined by reminders of the division by m.

Fact 56 (i) *if $a \equiv_m b$ and $c \equiv_m d$, then $a + c \equiv_m b + d$*

(ii) *if $a \equiv_m b$ and $c \equiv_m d$, then $ac \equiv_m bd$*

(iii) *if $a \equiv_m b$, then $a^n \equiv_m b^n$ for $n \in \mathbb{N}$*

In particular $(a+b) \mod m = ((a \mod m)+(b \mod m)) \mod m$ oraz $(a \cdot b) \mod m = ((a \mod m) \cdot (b \mod m)) \mod m$ – because $a \equiv_m a \mod m$. The above fact imply that $(Z_m, +_m, \cdot_m)$ is a ring (commutative, with an identity): $+_m$ satisfies all conditions of a group operation and is commutative, all the remaining conditions of rings are also satisfied: We check e.g. that $a \cdot_m (b +_m c) = a \cdot_m b +_m a \cdot_m c$.

Lemma 57 *An element a in a ring $\mathbb{Z}_m$ has an inverse (with respect to $\cdot_m$) if and only if $gcd(a, m) = 1$.*

Proof: $\Rightarrow$ If $a \cdot_m b = 1$, i.e. $ab \mod m = 1$ then $ab = qm + 1$ for some q, so $ab - qm = 1$. By Corollary 55 $gcd(a, m) = 1$. $\Leftarrow$ Again by 55 we have that $ax + my = 1$ for some integers x i y. In other words: $ax = -ym + 1$, i.e. $ax \mod m = 1$, so $x \mod m$ is the inverse of a. □

The above lemma implies that:

Theorem 58 (i) *$(\{1, 2, \ldots m - 1\}, \cdot_m)$ is a group if and only if m is a prime number.*

(ii) *The ring $(\mathbb{Z}_m, +_m, \cdot_m\}$ is a field if and only if m is a prime number.*

For any field $(C, +, \cdot)$ we say that the group $(C, +)$ is its additive group, and that the group $(C \setminus \{0\}, \cdot)$ is its multiplicative group. For any ring $(P, +, \cdot)$ we say that $+$ is an additive operation, and $\cdot$ – a multiplicative operation Let $\mathbb{Z}_m^*$ denotes the subset of $\mathbb{Z}_m$ containing element relatively prime to m. $|\mathbb{Z}_m^*|$ is denoted as $\varphi(m)$ (this is called the *Euler's function*). It is easy to see that $\varphi(p) = p - 1$ for a prime p.

Theorem 59 *The set of invertible elements (with respect to the multiplicative operation) of a ring forms a group (with the multiplicative operation). In particular $(\mathbb{Z}_m^*, \cdot_m)$ is a group for any $m > 0$, as the set of invertible elements of the ring$(\mathbb{Z}_m, +, \cdot)$.*

Recall that, by Corollary 40, for any gropu G, $a^{|G|} = e$ is satisfied for any $a \in G$. This implies:

Theorem 60 (i) **(Fermat's little theorem)** *If p is a prime number, then for any $a \in \mathbb{Z}$, such that p does not dived a, we have $a^{p-1} \equiv_p 1$.*

(ii) **(Euler)** *If $gcd(m, n) = 1$, then $a^{\varphi(m)} \equiv_m 1$ for any $a \in \mathbb{Z}$.*

Example 61 (a) An example application of Fermat's theorem: what is the reminder of the division of $2^{1000000}$ by 101? We know that $2^{100} \equiv_{101} 1$ (since 101 is a prime). $2^{1000000} = (2^{100})^{10000} \equiv_{101} 1$.

(b) What is the last digit in the decimal representation of 3^{2009}? We are looking for the result of multiplying 3 by itself 2009 times in $\mathbb{Z}_{10}$. $3^2 = 9$, $3^3 = 7$, $3^4 = 7 \cdot 3 = 1$, $3^{2008} = 3^{4 \cdot 502} = (3^4)^{502} = 1$. So, the result is $3 \cdot 1 = 3$.

(c) The additive group $(\mathbb{Z}_p, +_p)$ for a prime p is cyclic (by Corollary 41). It can be also shown (but we skip the proof), that the multiplicative group $(\mathbb{Z}_p \setminus \{0\}, \cdot_p)$ is also cyclic. In this example we check that 2 is a generator of the multiplicative group $\mathbb{Z}_{101}$ (recall that 101 is a prime). We show that the order of 2 is 100. It is enaugh to see that 2^{20} and 2^{50} are not equal to 1. Since the order of 2 has to be a divisor of 100, and all divisor of 100 (less then 100) are divisors of 20 or 50 we have the required result.

In the above examples we observed that in a ring $\mathbb{Z}_m$ we can easily compute powers. In the case when m is prime we can also easily compute inverses (Euclidean algorithm). These to properties we be very important in applications to cryptography.

1.13 Chinese reminder theorem

Definition 62 Let $(G_1, \cdot_1), \ldots, (G_k, \cdot_k)$ be groups. Their *product* is the structure $((G_1 \times, \ldots, \times G_k), \cdot)$, whose universe is the cartesian product of G_i-s and the operation is defined as follows: $(g_1, \ldots, g_k) \cdot (g'_1, \ldots, g'_k) = (g_1 \cdot_1 g'_1, \ldots, g_k \cdot_k g'_k)$.

Fact 63 *The product of groups is a group.*

Example 64 (a) $\mathbb{Z}_6$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_3$ (with addition). Isomorphism: $F(x) = (x \mod 2, x \mod 3)$

(b) $\mathbb{Z}_8$ is not isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_4$. It can be checked that $\mathbb{Z}_2 \times \mathbb{Z}_4$ does not contain an element of order 8 and such an element exists in $\mathbb{Z}_8$ (1).

(c) $\mathbb{Z}_2 \times \mathbb{Z}_2$ is isomorphic to Klein four-group.

Example 64 (a) can be generalized to the following theorem:

Theorem 65 (Chinese reminder theorem). *Let $m_1, m_2, \ldots, m_k$ be pairwise coprime. Let $m = m_1 m_2 \ldots m_k$. Then $\mathbb{Z}_m$ is isomorphic to $\mathbb{Z}_{m_1} \times Z_{m_2} \times \ldots \times Z_{m_k}$. The isomorphism can be defined as follows: $F(x) = (x \mod m_1, x \mod m_2, \ldots x \mod m_k)$.*

Proof: (skipped during the lecture) F **is 1-1:** if $F(a) = F(b)$, then $a \mod m_i = b \mod m_i$ for all i, so $a \equiv_m b$. This means that $m_i|(a-b)$ for all i. Since m_i are pairwise coprime, so $m_1 \dots m_k|(a-b)$, [2] but $|a-b| < m$, and thus $a = b$. F **is „onto":** Obvious: both sets are finite and of equal cardinalities, and F is 1-1. F **preserves the operation:** $F(a +_m b) = (a +_m b \mod m_1, a +_m b \mod m_2, \dots a +_m b \mod m_k) =$ [3] $(a + b \mod m_1, \dots, a + b \mod m_k) =$[4]$(a \mod m_1, \dots, a \mod m_k) + (b \mod m_1, \dots, b \mod m_k)$. □

Corollary 66 *If m_1, m_2, ..., m_k are pairwise coprime, $m = m_1 m_2 \dots m_k$ and $a_i \in \{0, \dots m_i - 1\}$ then there exists exactly one x, $0 \le x < m$, such that:*

$$\begin{array}{c} x \mod m_1 = a_1 \\ x \mod m_2 = a_2 \\ \dots \\ x \mod m_k = a_k \end{array}$$

The system of congruences above can be effectively solved: $x = (a_1 z_1 y_1 + \dots a_k z_k y_k) \mod m$, where $z_i = m/m_i$, and y_i is a number such that $z_i y_i \equiv_{m_i} 1$ (which can be found using extended Euclidean algorithm). Proof of correctness: $x \mod m_1 = (a_1 z_1 y_1 + \dots a_k z_k y_k) \mod m_1 = ((a_1 z_1 y_1 \mod m_1) + \dots + (a_k z_k y_k \mod m_1)) \mod m_1 = a_1 + 0 + \dots + 0 = a_1$ (since m_1 is a divisor of z_i for $i > 1$). Similarly for all other m_i.

Example 67 Find the smallest positive integer which gives the reminder 1 when divided by 2, 2 when divided by 3, and 2 when divided by 7. Result: 23.

[2] This fact is implied by the following lemma: if $m_1, \dots, m_k$ are pairwise coprime and each m_i divides n, then the product of all m_i also divides n.

[3] In this point we use the following lemma: if $m_i|m$, then $a +_m b \equiv_{m_i} a + b$).

[4] And here the fact 56 and the definition of the product of groups

1.14 Rings and fields

Recall:

(a) $(A, +, \cdot)$ is called a *ring* if:

- $(A, +)$ is a commutative group,
- operation $\cdot$ is associative,
- operation $\cdot$ is distributive over $+$.

(b) $(A, +, \cdot)$ is a *field* if

- $(A, +, \cdot)$ is a ring
- $(A \setminus \{0\}, \cdot)$ is a commutative group (by 0 we denote the identity of $+$);

A ring is called *commutative* if its multiplicative operation is commutative. A ring is called a *ring with identity* if its multiplicative operation has an identity. Every field is a ring. We usually use 0 to denote the identity of $+$, 1 – the identity of $\cdot$ (if exists). $-a$ is the additive inverse (opposite) of a, a^{-1} – is the multiplicative inverse of a. Instead of $a + (-b)$ we often write $a - b$.

Fact 68 *In every ring:*

(i) $0 \cdot a = 0$.

(ii) $(-x)y = x(-y) = -(xy)$

Fact 69 *In every field:* $ab = 0 \rightarrow a = 0 \vee b = 0$.

The previous fact is not true in all rings. For example $2 \cdot 2 = 0$ in $\mathbb{Z}_4$. An element a of a ring R, for which there exists a non-zero b such that $ab = 0$ is called a *zero divisor*.

1.15 Some info about rings of polynomials

Definition 70 Let $(R, +, \cdot)$ be a ring. An expression $f = a_n x^n + a_{n-1}x^{n-1} + \ldots a_1 x^1 + a_0 x^0$, where $a_i \in R$, is called a *polynomial* over the ring R. Elements a_i are called *coefficients* of polynomial. The greatest i for which $a_i \neq 0$ is called the *degree* of the polynomial and is denoted by $deg(f)$. The coefficient of such i is called *leading*. A polynomial whose leading coefficient is a_0 is called a *constant polynomial.*

A special polynomial is the *zero polynomial*, $f = 0$. It does not has a leading coefficient, and we assume that its degree is $-\infty$, When presenting a polynomial we usually skip $a_i x^i$ for which $a_i = 0$ (with the exception of the zero polynomial). In the case of $f = a_n x^n + a_{n-1}x^{n-1} + \ldots a_1 x^1 + a_0 x^0$, we sometimes speak about coefficients a_i, for $i > n$. We assume that such $a_i = 0$. Two polynomials $f = a_n x^n + a_{n-1}x^{n-1} + \ldots a_1 x^1 + a_0 x^0$ and $f = b_m x^m + b_{m-1}x^{m-1} + \ldots b_1 x^1 + b_0 x^0$ are *equal* if for all i we have $a_i = b_i$.

With each polynomial f we associate in a natural way a function $f(x) : R \rightarrow R$ (the value for an argument $b \in R$ is computed by substituting the value of b for x in the expression describing f). Note however, that polynomialls and associated functions are formally different notions.

Fact 71 *It is not always the case that two distinct polynomials over the same ring describe distinct functions.*[5]

Proof: Let R be the field Z_{11}, $f = 0$ and $g = x^{11} - x$. Both are associate with the zero function (to see this for the second polynomial use Fermat's little theorem). Morover, observe that for any finite ring$R$, the set of polynomials over F is infinite, while the set R^R is finite. □

The set of polynomials over a ring R is denoted by $R[x]$. In the set $R[x]$ we define natural operations of $+$ and $\cdot$:

Definition 72 Let $f = a_n x^n + a_{n-1}x^{n-1} + \ldots a_1 x^1 + a_0 x^0$, $g = b_m x^n + b_{m-1}x^{m-1} + \ldots b_1 x^1 + b_0 x^0$. Then:

$$f + g = \sum_{0 \leq i \leq \max\{m,n\}} (a_i + b_i)x^i$$

$$f \cdot g = \sum_{0 \leq i \leq m+n} (\sum_{0 \leq j \leq i} a_j b_{i-j})x^i$$

Example 73 Examples in $\mathbb{Z}_6[x]$. Let $f = 3x^2 + 2x + 2$, $g = 5x + 4$. Then $f + g = 3x^2 + x$, and $fg = 3x^3 + 4x^2 + 2x + 2$.

[5] But, as we see later, distinct polynomial over *infinite* fields describe distinct functions.

Fact 74 *Let f and g be polynomials over a ring R of degrees m and n, respectively. Then*

(i) *The degree of $f+g$ is not greater than* $\max\{m,n\}$,

(ii) *The degree of $f \cdot g$ is not greater than $m+n$.*

(iii) *If R is a field, then the degree of $f \cdot g$ is $m+n$. In particular, if $f \neq 0$ and $g \neq 0$, then $fg \neq 0$.*

Proof: The first two points are obvious. The third is implied by the fact that in a field, $ab = 0$ implies $a = 0$ or $b = 0$. So the product of leading coefficients cannot be zero and thus becomes the leading cofficient of the product of the polynomials □

The proof of the following fact is routine:

Fact 75 *Let $(R, +, \cdot)$ be a ring. Then $R[x]$ with the operation of addition and multiplication of polynomials is a ring. If R is commutative, then $R[x]$ is also commutative. If R has a multiplicative identity then $R[x]$ also has one.*

1.16 Divisibility of polynomials

From this point we consider rings of polynomials over fields. We usually denote a field by F. We show that for polynomials we can develop a theory of divisibility, similar to the theory of divisibility of integers.

Fact 76 *For any pair of polynomials f, $g \in F[x]$, $g \neq 0$ there exists exactly one pair of polynomials q, r such that $deg(r) < deg(g)$ and $f = qg + r$. The polynomial r is called the* reminder *of division of f by g. In particular, the reminder of the division of f by $x - c$ is a constant.*

Proof: *Existence.* If $deg(f) < deg(g)$, then we take $q = 0$, $r = f$. In the other case we proceed inductively by the degree of f. Let $f = a_n x^n + a_{n-1}x^{n-1} + \ldots a_1 x^1 + a_0 x^0$, $g = b_m x^n + b_{m-1}x^{m-1} + \ldots b_1 x^1 + b_0 x^0$, $n \geq m$. Consider the polynomial $h = f - (a_n b_m^{-1} x^{n-m})g$. The coefficient of x^n of h is 0, so $deg(h) < deg(f)$ and we may apply the inductive assumption $h = q'g + r$. Now $f = h + (a_n b_m^{-1} x^{n-m})g = q'g + r + (a_n b_m^{-1} x^{n-m})g = (q' + (a_n b_m^{-1} x^{n-m}))g + r$. *Uniqueness.* If $f = qg + r = q'g + r'$, then $(q - q')g = r' - r$. The polynomial $r - r'$ has degree smaller than $deg(g)$, so $q = q'$. This implies that $r' - r$ zero, so $r' = r$. □

The proof suggests a method for dividing polynomials. Let us see an example:

Example 77 Let us divide $x^3 + 2x$ by $2x + 1$ in $\mathbb{Z}_7[x]$. (...)

Definition 78 We say that a polynomial f divides g if there exists a polynomial h such that $g = f \cdot h$. As in the case of integers we write then $f|g$.

Fact 79 (i) *If $fg = c$, where $c \neq 0$ is a constant, then both f and g are constants.*

(ii) *If $f|g$ and $g|f$, then $f = cg$ for some constant c.*

Definition 80 We say that a non-constant polynomial f is *irreducible* (or *prime*) in $F[x]$ if there exists no pair of polynomials g, h, of degree smaller than $deg(f)$ such that $f = gh$.

Note that all polynomials of degree 1 are irreducible.

Definition 81 A greatest common divisor of polynomials f and g is a polynomial h which is a common divisor of f and g of the smallest possible degree.

Soon we will show that every pair of non-zero polynomials has a greatest common divisor. It may be noted that *gcd* of polynomials is unique up to a constant factor. More precisely, if h and h' are greatest common divisors of f and g, then $h = ch'$ for some constant c.

Lemma 82 (i) *Every pair f, $g \in F[x]$ has a gcd.*

(ii) *If h is a gcd of f and g, then there exist such a and b, that $af + bf = h$.*

Example 83 An example: we compute gcd of $x^3 + x^2 + 5x + 5$ and $x^3 + 5x$ in $\mathbb{Z}_7$.

Lemma 84 *If $f|gh$ and f is irreducible, then $f|g$ or $f|h$.*

Fact 85 *Every polynomial f can be uniquely represented (up to the order of f_i) in the form $f = cf_1f_2\dots f_k$, where c is a constant, and every f_i is a irreducible polynomial with leading coefficient* 1.

Fact 86 *If f and g are irreducible and $f|h$ and $g|h$, then $fg|h$.*

Theorem 87 (Bezout) *The reminder of the division of f by $x - c$ equlas $f(c)$. In particular $x - c|f$ if and only if $f(c) = 0$.*

An element c such that $f(c) = 0$ is called a *root* of polynomial f.

Fact 88 *A polynomial $f \in F[x]$ of degree n has at most n roots.*

Proof: Assume to the contrary that there exist $n + 1$ roots $a_1, a_2, \dots a_{n+1}$. By the theorem of Bezouta f divides by $x - a_i$, for all i. Since $x - a_i$ are ireeducible, then f is divisible also by $(x - a_1)(x - a_2)\dots(x - a_{n+1})$. The degree of the resulting polynomial is $n + 1$, contradiction. □

Theorem 89 *Let $x_1, \dots, x_n$ and $y_1, \dots, y_n$ be elements of a field F. Then there exists exactly one polynomial f of degree smaller than n such that $f(x_i) = y_i$ for $1 \le i \le n$.*

Proof: We use Lagrange's interpolation formula to define f: $f = y_1g_1(x) + y_2g_2(x) + \dots + y_ng_n(x)$, where $g_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}$. Uniqueness is implied by the theorem about the number of roots of a polynomial.

2 Linear spaces

2.1 Definition

Definition 90 A set V is called a *linear (vector) space over a field K* if there are two operations defined: $+ : V \times V \to V$ i $\cdot : K \times V \to V$ and, for any $\alpha, \beta \in K$, $u, v \in V$ we have:

(a) $(V, +)$ is a commutative group

(b) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

(c) $\alpha \cdot (v + u) = \alpha \cdot v + \alpha \cdot u$.

(d) $(\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$

(e) $1 \cdot v = v$ (where 1 is the identity of the multiplicative operation in K).
Elements of V are called *vectors*, elements of K – *scalars*.

We usually use $+$ and $\cdot$ to denote operation in K and operation on vectors. The identity of $(V, +)$ is denoted by $\vec{\mathbf{0}}$.

The following properties can be derived from the definition of linear spaces:

Fact 91 *For any $\alpha \in K$ and $v \in V$:*

(i) $0 \cdot v = \vec{\mathbf{0}}$ *(0 is the identity of the additive operation in K)*

(ii) $\alpha \cdot \vec{\mathbf{0}} = \vec{\mathbf{0}}$

(iii) $\alpha v = \vec{\mathbf{0}}$ *if and only if* $\alpha = 0$ *or* $v = \vec{\mathbf{0}}$

(iv) $(-1)v$ *is the inverse of* v

2.2 Space K^n

Let K be an arbitrary field. By K^n we denote the set of n-element sequences of elements from K: $K^n = \{(\alpha_1, \dots, \alpha_n) : \alpha_i \in K\}$. We define operation $+ : K^n \times K^n \to K^n$:

$$(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n).$$

We also define multiplication of elements from K^n by elements of K:

$$\alpha \cdot (\alpha_1, \dots, \alpha_n) = (\alpha \cdot \alpha_1, \dots, \alpha \cdot \alpha_n).$$

Fact 92 *K^n with the operations defined above is a linear space over K.*

E.g. the space $\mathbb{R}^2$ may be identified with the space of vectors in the place.

2.3 Examples of linear spaces

Example 93 (a) A field K is a linear space over itself (this space is actually K^1).

(b) The set of infinite sequences over K (operations defined analogously to K^n).

(c) The set of polynomials over a field K, $K[x]$, over K (naturally defined operations).

(d) The set of functions from $\mathbb{R}$ to $\mathbb{R}$ with the following operations: $(f+g)(x) = f(x) + g(x)$, $(\alpha f)(x) = \alpha(f(x))$ (this space is over the field $\mathbb{R}$).

(e) The set of functions from X into a linear space V over K, with operations defined as above (this space is over K).

2.4 Subspaces

Definition 94 Let V be a linear space over a field K. We say that U is a *subspace* of V is $U \subseteq V$ and U is a linear space over K.

Example 95 (a) The set of polynomials of degree smaller than 10, over a field K is a subset of $K[x]$.

(b) $\{(a,b,c) : a+b+c=0\}$ is a subset of $\mathbb{R}^3$.

Fact 96 *A non-empty set S of vectors from V is a subspace of V if and only if it is closed under vector addition and multiplication by scalars.*

Fact 97 *The intersection of two subspaces is a subspace.*

2.5 Linear combinations of vectors

Definition 98 Let V be a linear space over K. Let $v_1, \dots v_k$ be vectors of V. A vector $v = \alpha_1 v_1 + \dots \alpha_k v_k$, $(\alpha_i \in K)$ is called a *linear combination* of $v_1, \dots, v_k$ with *coefficients* $\alpha_1, \dots, \alpha_k$. Let $A \subseteq V$. By $LIN(A)$ we denote the set of all linear combinations of A. $LIN(A)$ is called the *subspace generated by A*.

Fact 99 *$LIN(A)$ is a subspace of V. This is the smallest subspace containing A.*

2.6 Linear independance

Definition 100 A set of vectors $v_1, \dots, v_k$ of a space V is *linearly independent* if $\alpha_1 v_1 + \dots + \alpha_k v_k = \vec{\mathbf{0}}$ only if $\alpha_1 = \dots = \alpha_k = 0$. Vectors which are not linearly independent are called *linearly dependent.*

The definition implies that $\vec{\mathbf{0}}$ is not a member of any linearly independent set; every subset of a linarly indepents set is linearly independend; any superset of linearly dependent set is linearly dependent.

Example 101 In $\mathbb{R}^3$:

(a) $(1,2,3), (2,3,4), (6,10,14)$ are dependent,

(b) $(1,1,1), (1,1,0), (1,0,0)$ are independent,

Lemma 102 *A set of vectors $v_1, \dots, v_k$ is linearly independent if and only if one of vectors is a linear combination of the remaining.*

Lemma 103 (i) *The set of vectors $v_1, \dots, v_k$ is linearly dependent if and only if it contains a proper subset generating the same subspace*

(ii) *Any finite set of vectors contains a linearly independent subset generating the same subspace.*

2.7 Basis and dimension of a space

Definition 104 A set of vectors B is a *basis* of a space V, if $LIN(B) = V$ and B is linearly independent. A space is *finitely-dimensional,* if it has a finite basis.

Example 105 (a) Bases in $\mathbb{R}^3$: $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ and $(1,1,1), (1,1,0), (1,0,0)$.

(b) A basis in $\mathbb{R}^n$ is $(1,0,\dots,0), (0,1,0,\dots,0), \dots, (0,0,\dots,0,1)$. This basis is called *canonical* or *standard.*

(c) A basis in $\mathbb{R}[x]$: $1, x, x^2, x^3, \dots$.

Fact 106 *Let $v_1, \dots v_l$ be linearly independent and for all i: $v_i \in LIN(e_1, \dots e_k)$. Then $l \leq k$.*

Proof: Induction over l. □

Corollary 107 *Any set of more than n vectors in K^n is dependent*

Proof: $K^n = LIN((1,0,0,\ldots,0),(0,1,0,0,\ldots,0),\ldots,(0,0,\ldots,0,1))$ □

Corollary 108 *If $LIN(v_1,\ldots,v_k) = LIN(u_1,\ldots,u_l)$ and sets $v_1,\ldots,v_k$ and $u_1,\ldots,u_l$ are linearly independent then $k = l$.*

Corollary 109 *All bases of a finitely-dimensional space have the same number of elements.*

Definition 110 A *dimension* of a space V if the number of vectors in a basis of V.

Fact 111 (i) *Every maximal linearly independent set of vectors of V is a basis of V.*

(ii) *Every n-element linearly independent set of vectors of a n-dimensional space is a basis of V.*

(iii) *Every set of linearly independent vectors of a finitely-dimensional space may be extended to its basis.*

Theorem 112 *Every space of the form $LIN(v_1,\ldots,v_k)$ has a basis.*

It can be shown that arbitrary space has a basis (we skip the proof of this fact).

Theorem 113 *Every vector can be uniquely represented as a linear combination of vectors from a basis.*

If $e_1,\ldots,e_k$ is a basis of V, $v = \alpha_1 e_1 + \ldots + \alpha_k e_k$, then coesfficients α_i are called *coordinates* of v with respect to the basis $e_1,\ldots,e_k$.

2.8 Matrices

Formally, a matrix of dimensions $m \times n$ (a matrix with m *rows* and n *columns*), over a field K, is a fuction of type $\{1,\ldots,m\} \times \{1,\ldots,n\} \to K$. The set of matrices of dimensions $m \times n$ is denoted by $M_{mn}(K)$ (or M_{mn}, if the field is default). Matrices are usually represented as rectangular tables:

$$\begin{bmatrix} a_{1,1}, & a_{1,2}, & \ldots, & a_{1,n} \\ a_{2,1}, & a_{2,2}, & \ldots, & a_{2,n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ a_{m,1}, & a_{m,2}, & \ldots, & a_{m,n} \end{bmatrix},$$

where a_{ij} is the value reterned by the function for the pair (i,j). Notions of rows and columns are defined in a natural way.

Some special types of matrices are listed below:

(a) Zero matrix (of dimensions $m \times n$) – consisting of zeros only.

(b) Square matrices, i.e. matrices with n rows and n columns (in this case n is called the *degree* of a matrix):

- diagonal matrix: only elements on the main diagonal, i.e. $a_{1,1}, a_{2,2}, \ldots, a_{n,n}$ may not be equal to zero.
- identity matrix: a diagonal matrix withs 1s on the main diagonal.
- upper triangular matrix: zeros below the main diagonal.
- lower triangular matrix: zeros above the main diagonal.

2.9 Operations on matrices

2.9.1 Addition of matrices, multiplication by scalars

In the set of matrices with m rows and n columns we can define the operations of addition and multiplication by scalars from K:

$$\begin{bmatrix} a_{1,1}, & a_{1,2}, & \ldots, & a_{1,n} \\ a_{2,1}, & a_{2,2}, & \ldots, & a_{2,n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ a_{m,1}, & a_{m,2}, & \ldots, & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1}, & b_{1,2}, & \ldots, & b_{1,n} \\ b_{2,1}, & b_{2,2}, & \ldots, & b_{2,n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ b_{m,1}, & b_{m,2}, & \ldots, & b_{m,n} \end{bmatrix} = \begin{bmatrix} a_{1,1}+b_{1,1}, & a_{1,2}+b_{1,2}, & \ldots, & a_{1,n}+b_{1,n} \\ a_{2,1}+b_{2,1}, & a_{2,2}+b_{2,2}, & \ldots, & a_{2,n}+b_{2,n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ a_{m,1}+b_{m,1}, & a_{m,2}+b_{m,2}, & \ldots, & a_{m,n}+b_{m,n} \end{bmatrix}$$

$$\alpha \cdot \begin{bmatrix} a_{1,1}, & a_{1,2}, & \ldots, & a_{1,n} \\ a_{2,1}, & a_{2,2}, & \ldots, & a_{2,n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ a_{m,1}, & a_{m,2}, & \ldots, & a_{m,n} \end{bmatrix} = \begin{bmatrix} \alpha a_{1,1}, & \alpha a_{1,2}, & \ldots, & \alpha a_{1,n} \\ \alpha a_{2,1}, & \alpha a_{2,2}, & \ldots, & \alpha a_{2,n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ \alpha a_{m,1}, & \alpha a_{m,2}, & \ldots, & \alpha a_{m,n} \end{bmatrix}$$

Example 114 Some examples of addition and multiplication by scalars.

The proof of the fact below is routine:

Fact 115 *The set of matrices with m rows and n columns with the operations defined above is a linear space of dimension mn.*

Proof: Observe that M_{mn} with addition is a group. Check eg. that $\alpha(A+B) = \alpha A + \alpha B$.

2.9.2 Multiplication

A product of matrices A and B is defined only if the number of columns of A equals number of rows of B. Firstly, we define the product of a matrix of dimensions $1 \times n$ by a matrix of dimensions $n \times 1$:

$$[a_1, a_2, \ldots, a_n] \cdot \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{bmatrix} = [a_1 b_1 + a_2 b_2 + \ldots + a_n b_n].$$

The result may be treated either as an element of the field K or a matrix over K of dimensions 1×1. Now we define the product of A and B, of dimensions $m \times n$ and $n \times k$, respectively. The result is a matrix of dimensions $n \times k$. Let

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \ldots \\ A_m \end{bmatrix},$$

where A_i are rows of A, and

$$B = [B_1, B_2, \ldots B_k],$$

B_i are columns of B. AB is defined as follows:

$$AB = \begin{bmatrix} A_1B_1, & A_1B_2, & \ldots, & A_1B_k \\ A_2B_1, & A_2B_2, & \ldots, & A_2B_k \\ \ldots, & \ldots, & \ldots, & \ldots \\ A_mB_1, & A_mB_2, & \ldots, & A_mB_k \end{bmatrix}$$

Matrix multiplication is not commutative (even for square matrices), but

Fact 116 *Matrix multiplication is associative.*

Fact 117 *Let A, B, C be matrices over K, I_n – the identity matrix of degree n, and $\alpha \in K$. If the following operations are defined then:*

(i) $I_n A = A$, $AI_n = A$,

(ii) $A(B+C) = AB + AC$,

(iii) $(B+C)A = BA + CA$,

(iv) $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Note, that multiplication is an operation in the set of square matireces of a fixed degree n. This operation is associative and has the identity I_n. But M_{nn} with multiplication is not a group – there are matrices which do not have inverses (e.g. zero matrix).

2.10 Linear transformations

As in the case of groups and other algebraic structures, we can define a notion of homomorphism of linear spaces. Homomorphisms of linear spaces are called *linear transformations*.

Definition 118 Let V and U be linear spacec over field K. A function $J : V \to U$ is called a linear transformation if:

(a) $\forall v_1, v_2 \in V$ we have $L(v_1 + v_2) = L(v_1) + L(v_2)$

(b) $\forall \alpha \in K$, $v \in V$ we have $L(\alpha v) = \alpha L(v)$.

Example 119 Consider a map from $\mathbb{R}^3 \to \mathbb{R}^2$, defined as $L(x, y, z) = (x, y)$. It is linear.

The notions of the kernel and the image of a linear transformation are defined analogously to the case groups. $Ker(L) = \{v \in V : L(v) = \vec{\mathbf{0}}\}$, $Im(L) = \{L(v) : v \in V\}$.

Lemma 120 *If V and U are linear spaces over K, and $L : V \to U$ is a linear transformation, then $Ker(L)$ is a subspace of V, and $Im(L)$ is a subspace of U.*

Fact 121 *The image of a linear subspace is a subspace and the inverse image of a linear subspace is a linear subspace.*

Example 122 (a) A function returning always the zero vector is a linear transformation.

(b) Function: $L : \mathbb{R}^3 \to \mathbb{R}^2$, $L((x, y, z)) = (x + y, z)$ is a linear. The kernel consist of vectors of the form $(x, -x, 0)$; and the image is the whole $\mathbb{R}^2$.

(c) $L(x, y, z) = (x, x)$ is linear.

(d) Function $L : R^3 \to R^2$, $L((x, y, z)) = (xy, z)$ is not linear.

Theorem 123 $dim(V) = dim(im(L)) + dim(Ker(L))$.

$dim(Im(L))$ is sometimes also called the *rank* of a transformation.

Example 124 Check that the theorem holds for $L(a, b, c, d) = (a, 0, 0, b)$.

The composition of linear transformations is a linear transformation:

Fact 125 *Let L be a linear transformation of V into U, and M a linear transformation of U into W. Then the composition of L and M, $(ML)(v) = M(L(v))$ is a linear transformation of V into W.*

2.11 A connection between matrices and linear spaces

Lemma 126 *Let V and U be linear space over a field K. Let $e_1, \ldots, e_n$ be a basis of V. Let $w_1, \ldots, w_n$ be a sequence of vectors in U. Then there exists exactly one linear transformation which maps e_i into f_i for all i.*

Proof: $L(v) = L(\alpha_1 e_1 + \ldots + \alpha_n e_n) = \alpha_1 L(e_1) + \ldots + \alpha_n L(e_n)$. □

Let V and U be linear space over a field K. Let $E = \{e_1, \ldots, e_n\}$ be a basis of V, and $F = \{f_1, \ldots, f_m\}$ be a basis of U. Let $L : V \to U$ be a linear transformation. The image of every e_i is a linear transformation of vectors f_i:

$$L(e_i) = \alpha_{1i} f_1 + \ldots + \alpha_{mi} f_m.$$

We associate the following matrix with a linear transformation $L : V \to U$. The matrix will be denoted as $A_{EF}(L)$:

$$A_{EF}(L) = \begin{bmatrix} \alpha_{1,1}, & \alpha_{1,2}, & \ldots, & \alpha_{1,n} \\ \alpha_{2,1}, & \alpha_{2,2}, & \ldots, & \alpha_{2,n} \\ \ldots, & \ldots, & \ldots, & \ldots \\ \alpha_{m,1}, & \alpha_{m,2}, & \ldots, & \alpha_{m,n} \end{bmatrix}$$

$A_{EF}L$ is called the matrix of L in bases E, F. Note that the i-th column is the vector from K^n whose coordinates are coefficients of the image of e_i given in the basis F.

Example 127 We construct the matrix of $L : \mathbb{R}^3 \to \mathbb{R}^2$, $L((x, y, z)) = (x + y, z)$ in standard bases:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In the remaining part of this section we consider only linear transformations from R^n to R^m and matrices in standard bases. However all the results can be generalized to arbitrary spaces and bases.

Theorem 128 *There exists a natural bijective correspondence between the set of linear transformation form R^n to R^m and the set of matrices $M_{m\times n}$.*

Fact 129 *Let $A(L)$ be the matrix of L. Then*

$$A(L) \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \dots \\ \gamma_m \end{bmatrix},$$

where $\gamma_1, \dots, \gamma_m$ are coordinates of $L(v) \in U$.

Theorem 130 *The matrix of the composition of linear transformations L and M is the product of matrices $A(L)$ and $A(M)$.*

Example 131 Consider the transformatoin $L : \mathbb{R}^2 \to \mathbb{R}^2$ rotating the input vector about φ (we identify vectors with points of the plane. The matrix $A(L)$ is given below.

$$\begin{bmatrix} cos\varphi & -sin\varphi \\ sin\varphi & cos\varphi \end{bmatrix}$$

Consider also the rotation about ψ and the composition of these two rotations, ie. the rotation about $\varphi + \psi$. According to fact 130 the matrix of the composition can be computed as follows:

$$\begin{bmatrix} cos\psi & -sin\psi \\ sin\psi & cos\psi \end{bmatrix} \cdot \begin{bmatrix} cos\varphi & -sin\varphi \\ sin\varphi & cos\varphi \end{bmatrix}$$

Multiplying the matrices we obtain well known formulas:

$$\begin{bmatrix} cos(\psi + \varphi) & -sin(\psi + \varphi) \\ sin(\psi + \varphi) & cos(\psi + \varphi) \end{bmatrix} = \begin{bmatrix} cos\psi\, cos\varphi - sin\psi\, sin\varphi, & -cos\psi\, sin\varphi - sin\psi\, cos\varphi \\ sin\psi\, cos\varphi + cos\psi\, sin\varphi, & -sin\psi\, sin\varphi + cos\psi\, cos\varphi \end{bmatrix}$$

2.12 Determinants

2.12.1 Definition

Determinants of degree 2 should be known from school:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

This can be generalized.

Definition 132 Let A be a square matrix of degree n:

$$A = \begin{bmatrix} a_{1,1}, & a_{1,2}, & \dots, & a_{1,n} \\ a_{2,1}, & a_{2,2}, & \dots, & a_{2,n} \\ \dots, & \dots, & \dots, & \dots \\ a_{n,1}, & a_{n,2}, & \dots, & a_{n,n} \end{bmatrix}$$

The determinant of A ($|A|$ or $detA$) is defined as:

$$\sum_{f \in S_n} \mathrm{sgn}(f) a_{1,f(1)} a_{2,f(2)} \cdots a_{3,f(3)},$$

2.12.2 Properties

Definition 133 The *transpsition* of a matrix A is the matrix A^T whose rows are columns of A. Formally, the element on the position i, j in A^T is the elements from the position j, i in A.

Fact 134 $|A| = |A^T|$.

Fact 135 *If B is obtained from A by multiplying one the rows by α, then $|B| = \alpha|A|$.*

Fact 136 *If B is obtained from A by switching two rows then $|B| = -|A|$.*

Fact 137 *A determinant with two identical rows is equal to 0.*

Fact 138 *If B is obtained from A by adding to a row some other row multiplied by a scalar then $|B| = A|$.*

2.12.3 Laplace expansion

For a square matrix A we denote by M_{ij} the determinant of the matrix which is obtained from A by removing the i-th row and $j - th$ column. We define also $A_{ij} := (-1)^{i+j} M_{ij}$.

Fact 139 *For every i:*

$$detA = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots a_{in}A_{in}$$

and

$$detA = a_{1i}A_{1i} + a_{2i}A_{2i} + \ldots a_{ni}A_{ni}$$