Course 1

0.0 Coordinates

- **Structure:**

  Chapter 1: Preliminaries

  Chapter 2: Vector Spaces

  Chapter 3: Matrices and Linear Systems

  Chapter 4: Introduction to Coding Theory

- **Bibliography:**

1. S. Crivei, *Basic Linear Algebra*, Presa Universitară Clujeană, Cluj-Napoca, 2022.

2. W. J. Gilbert, W. K. Nicholson, *Modern Algebra with Applications*, John Wiley, 2004.

3. J. S. Golan, *The Linear Algebra a Beginning Graduate Student Ought to Know*, Springer, Dordrecht, 2007.

4. P. N. Klein, *Coding the Matrix. Linear Algebra through Applications to Computer Science*, Newtonian Press, 2013.

5. R. Lidl, G. Pilz, *Applied Abstract Algebra*, Springer-Verlag, 1998.

6. I. Purdea, C. Pelea, *Probleme de algebră*, Eikon, Cluj-Napoca, 2008.

7. L. Robbiano, *Linear Algebra for Everyone*, Springer, Milan, 2011.

8. G. Strang, *Linear Algebra and its Applications*, Brooks/Cole, 1988.

- **Course:**

  Course materials will be available in the *Algebra* course on the Moodle platform (https://moodle.cs.ubbcluj.ro/). Enrolment key: Algebra-IE-2022

  Students may get up to 1 bonus point from course projects to the final grade: up to 5 projects, each for 0.2 points [you will receive details in due time...].

- **Seminar:**

  Minimum attendance: 75% for seminar classes in order to be allowed to participate in the second partial exam.

  Problems for the next week will be available in the *Algebra* course on the Moodle platform.

  Students may get up to 0.5 bonus points from seminar to the final grade: 5 problems solved during the seminar, each for 0.1 points [you will receive details during seminars...].

- **Exam:**

  Written partial exams in Week 8 (Chapters 1-2) and Week 14 (Chapters 3-4).

  The final grade is computed as follows:

$$G = 1 + P_1 + P_2 + B,$$

  where:

  G = the final grade

  P_1 = the grade from the first partial exam (max. 4)

  P_2 = the grade from the second partial exam (max. 5)

  B = bonus points from seminar or course (max. 1.5)

  Students may not pass the exam unless they participate in the second partial exam.

Computer Science topics using Linear Algebra

The Association for Computing Machinery (ACM) has developed the 2012 ACM Computing Classification System for the research topics in the field of Computer Science (`https://www.acm.org`) under the form of a multi-level tree. We mention some higher level branches of this tree in which Linear Algebra has important applications.

Networks

- Network architectures
  - Network design principles
- Network types
  - Public Internet

Theory of Computation

- Models of computation
  - Quantum computation theory
- Computational complexity and cryptography
  - Cryptographic protocols
- Randomness, geometry and discrete structures
  - Error-correcting codes
- Theory and algorithms for application domains
  - Machine learning theory

Mathematics of Computing

- Information theory
  - Coding theory
- Mathematical analysis
  - Mathematical optimization

Information Systems

- World Wide Web
  - Web searching and information discovery
- Information retrieval
  - Retrieval models and ranking

Security and Privacy

- Cryptography
  - Symmetric cryptography and hash functions
- Network security
  - Security protocols

Computing Methodologies

- Machine learning
  - Machine learning approaches
- Computer graphics
  - Image manipulation

Applied Computing

- Electronic commerce
  - Online banking
- Operations research
  - Decision analysis

Chapter 1 PRELIMINARIES

1.1 Relations

Definition 1.1.1 A triple $r = (A, B, R)$, where A, B are sets and

$$R \subseteq A \times B = \{(a, b) \mid a \in A, b \in B\},$$

is called a *(binary) relation*.
The set A is called the *domain*, the set B is called the *codomain* and the set R is called the *graph* of the relation r.
If $A = B$, then the relation r is called *homogeneous*.
If $(a, b) \in R$, then we sometimes write $a\,r\,b$ and we say that *a has the relation r to b* or *a and b are related with respect to the relation r*.

Definition 1.1.2 Let $r = (A, B, R)$ be a relation and let $X \subseteq A$. Then the set

$$r(X) = \{b \in B \mid \exists x \in X : x\,r\,b\}$$

is called the *relation class of X with respect to r*. If $x \in X$, then we denote

$$r < x >= r(\{x\}) = \{b \in B \mid x\,r\,b\}.$$

Remark 1.1.3 (1) Let $r = (A, B, R)$ be a relation and let $X \subseteq A$. Notice that

$$r(X) = \bigcup_{x \in X} r < x > .$$

(2) As in the case of functions, if $A, B \subseteq \mathbb{R}$, then the graph of a relation $r = (A, B, R)$ may be represented as a subset of points of the real plane $\mathbb{R} \times \mathbb{R}$, whereas if A, B are any finite sets, then $r = (A, B, R)$ may be represented by a diagram consisting of two sets with elements and connecting arrows. For instance, let $r = (A, B, R)$, where $A = \{1, 2, 3\}$, $B = \{1, 2\}$ and

$$R = \{(1, 1), (1, 2), (3, 1)\}.$$

One may draw the two sets A and B, and arrows between the elements related by R, namely arrows from 1 to 1, from 1 to 2 and from 3 to 1.

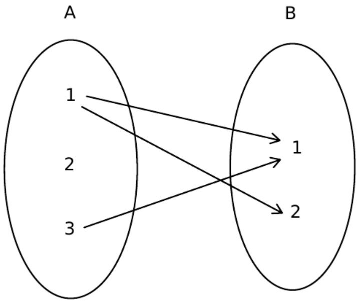


Figure 1.1: Diagram of a relation.

Also note that $r < 1 >= \{1, 2\} = r(A)$.

Example 1.1.4 (a) Let C be the set of all children and let P be the set of all parents. Then we may define the relation $r = (C, P, R)$, where

$$R = \{(c, p) \in C \times P \mid c \text{ is a child of } p\}.$$

(b) The triple $r = (\mathbb{R}, \mathbb{R}, R)$, where

$$R = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \leq y\}$$

is a homogeneous relation, called the *inequality relation* on $\mathbb{R}$. We have

$$r < 1 >= [1, \infty) = r([1, 2]).$$

(c) There are several examples from Number Theory, such as divisibility on $\mathbb{N}$ or on $\mathbb{Z}$, and Geometry, such as parallelism of lines, perpendicularity of lines, congruence of triangles, similarity of triangles.

(d) Let A and B be two sets. Then the triples

$$o = (A, B, \emptyset), \quad u = (A, B, A \times B)$$

are relations, called the *void relation* and the *universal relation* respectively.

(e) Let A be a set. Then the triple $\delta_A = (A, A, \Delta_A)$, where

$$\Delta_A = \{(a, a) \mid a \in A\}$$

is a relation called the *equality relation* on A.

(f) Every function is a relation. Indeed, a function $f : A \to B$ is determined by its domain A, its codomain B and its graph

$$G_f = \{(x, y) \in A \times B \mid y = f(x)\}.$$

Then the triple (A, B, G_f) is a relation.

(g) Every directed graph is a relation. Indeed, a directed graph (V, E) consists of a set V of vertices and a set E of directed edges ("arrows") between vertices. We may identify each directed edge with a pair in $V \times V$, where the first and the second component are respectively the starting and the ending vertex of that directed edge. Denote by P the set of those pairs. Then the triple (V, V, P) is a relation. For instance, the directed graph

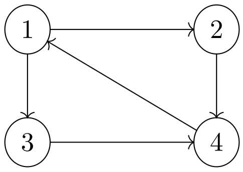


Figure 1.2: Directed graph.

may be seen as a relation (A, A, R), where $A = \{1, 2, 3, 4\}$ and

$$R = \{(1, 2), (1, 3), (2, 4), (3, 4), (4, 1)\}.$$

1.2 Functions

Definition 1.2.1 A relation $r = (A, B, R)$ is called a *function* if

$$\forall a \in A, \quad |r < a > | = 1,$$

that is, the relation class with respect to r of every $a \in A$ consists of exactly one element.

In other words, a relation r is a function if and only if every element of the domain has the relation r to exactly one element of the codomain.

In what follows, if $f = (A, B, F)$ is a function, we will mainly use the classical notation for a function, namely $f : A \to B$ or sometimes $A \xrightarrow{f} B$. The unique element of the set $f < a >$ will be denoted by $f(a)$. Then we have

$$(a, b) \in F \Longleftrightarrow f(a) = b.$$

In particular, from Definition 1.1.1 for a relation, we get the following corresponding notions for a function.

Definition 1.2.2 Let $f : A \to B$ be a function. Then A is called the *domain*, B is called the *codomain* and

$$F = \{(a, f(a)) \mid a \in A\}$$

is called the *graph* of the function f.

Example 1.2.3 (a) Let A be a set. Then the equality relation (A, A, Δ_A) is a function called the *identity function (map) on* A, that is denoted by $1_A : A \to A$ and is defined by $1_A(a) = a$, $\forall a \in A$.

(b) Let B be a set and let $A \subseteq B$. Then the relation (A, B, Δ_A) is a function called the *inclusion function of* A *into* B, that is denoted by $i : A \to B$ and is defined by $i(a) = a$, $\forall a \in A$.

(c) Let $A = \{1, 2, 3\}$, $B = \{1, 2\}$ and let $r = (A, B, R)$, $s = (A, B, S)$, $t = (A, B, T)$ be the relations having the graphs

$$R = \{(1, 1), (2, 1), (3, 2)\},$$
$$S = \{(1, 2), (3, 1)\},$$
$$T = \{(1, 1), (1, 2), (2, 1), (3, 2)\}.$$

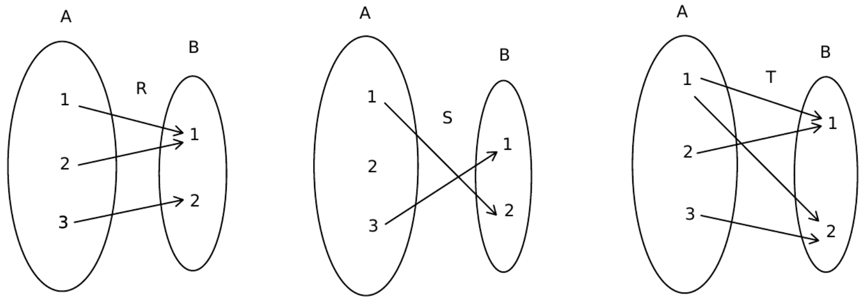


Figure 1.3: Diagrams of functions or relations.

Since $|r < a > | = 1$ for every $a \in A$, the relation r is a function. But s and t are not functions, because, for instance, we have $|s < 2 > | = 0$ and $|t < 1 > | = 2$.

Now we introduce a classical notation. Let A and B be two sets. Then we denote

$$B^A = \{f \mid f : A \to B \text{ is a function}\}.$$

If $|A| = n \in \mathbb{N}^*$, then the set B^A can be identified with the set $B^n = \underbrace{B \times \cdots \times B}_{n \text{ times}}$.

The notation is justified by the following nice result.

Theorem 1.2.4 *Let A and B be finite sets, say $|A| = n$ and $|B| = m$ $(m, n \in \mathbb{N}^*)$. Then*

$$|B^A| = m^n = |B|^{|A|}.$$

Definition 1.2.5 Let $f : A \to B$ be a function and let $X \subseteq A$.
We call the *image of* X *by* f the relation class of X with respect to f, that is,

$$f(X) = \{b \in B \mid \exists x \in X : x \, f \, b\} = \{f(x) \mid x \in X\}.$$

We denote $\mathrm{Im} f = f(A)$ and call it the *image of* f.

Homework: Recall the definitions and the properties of injective, surjective and bijective functions.

1.3 Equivalence relations and partitions

Recall that a relation $r = (A, B, R)$ is called *homogeneous* if $A = B$. Some special type of such relations is the subject of the present section.

Definition 1.3.1 A homogeneous relation $r = (A, A, R)$ on A is called:
(1) *reflexive* (r) if: $\forall x \in A,\ x\,r\,x$.
(2) *transitive* (t) if: $x, y, z \in A,\ x\,r\,y$ and $y\,r\,z \Longrightarrow x\,r\,z$.
(3) *symmetric* (s) if: $x, y \in A,\ x\,r\,y \Longrightarrow y\,r\,x$.
A homogeneous relation $r = (A, A, R)$ is called an *equivalence relation* if r has the properties (r), (t) and (s).

Example 1.3.2 (a) The equality relation δ_A on a set A has all 3 properties, hence δ_A is an equivalence relation on A.

(b) The similarity of triangles is an equivalence relations on the set of all triangles.

(c) The inequality relation " $\leq$ " on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$ has (r) and (t), but not (s). Hence it is not an equivalence relation on the corresponding set.

(d) Let $n \in \mathbb{N}$ and let ρ_n be the relation defined on $\mathbb{Z}$ by

$$x\,\rho_n\,y \Longleftrightarrow x \equiv y \pmod{n},$$

that is, $n|(x-y)$ or equivalently for $n \neq 0$, x and y give the same remainder when divided by n. Then ρ_n is called the *congruence modulo* n and it has the properties (r), (t) and (s), hence it is an equivalence relation.

For $n = 0$, we have $x\,\rho_0\,y \Longleftrightarrow 0|x-y \Longleftrightarrow x = y$, hence $\rho_0 = \delta_{\mathbb{Z}} = (\mathbb{Z}, \mathbb{Z}, \Delta_{\mathbb{Z}})$.

For $n = 1$, we have $x\,\rho_1\,y \Longleftrightarrow 1|x-y$, which is always true, and thus $\rho_1 = u = (\mathbb{Z}, \mathbb{Z}, \mathbb{Z} \times \mathbb{Z})$.

Definition 1.3.3 Let A be a non-empty set. Then a family $(A_i)_{i\in I}$ of non-empty subsets of A is called a *partition* of A if:
(i) The family $(A_i)_{i\in I}$ covers A, that is,

$$\bigcup_{i\in I} A_i = A.$$

(ii) The A_i's are pairwise disjoint, that is,

$$i, j \in I, i \neq j \Longrightarrow A_i \cap A_j = \emptyset.$$

Example 1.3.4 (a) Let $A = \{1, 2, 3, 4, 5\}$ and $A_1 = \{1, 2, 3\}$, $A_2 = \{4\}$, $A_3 = \{5\}$. Then $\{A_1, A_2, A_3\}$ is a partition of A.

(b) Let A be a set. Then $\{\{a\} \mid a \in A\}$ and $\{A\}$ are partitions of A.

(c) Let A_1 be the set of even integers and A_2 the set of odd integers. Then $\{A_1, A_2\}$ is a partition of $\mathbb{Z}$.

(d) Consider the intervals

$$A_n = [n, n+1)$$

for every $n \in \mathbb{Z}$. Then the family $(A_n)_{n\in\mathbb{Z}}$ is a partition of $\mathbb{R}$.

Denote by $E(A)$ the set of all equivalence relations and by $P(A)$ the set of all partitions on a set A.

Definition 1.3.5 Let $r \in E(A)$.
The relation class $r < x >$ of an element $x \in A$ with respect to r is called the *equivalence class of x with respect to r*, while the element x is called a *representative* of $r < x >$.
The set

$$A/r = \{r < x > \mid x \in A\},$$

which is the set of all equivalence classes of elements of A with respect to r, is called the *quotient set of A by r*.

Definition 1.3.6 Let $\pi = (A_i)_{i \in I} \in P(A)$ and define the relation r_π on A by

$$x \, r_\pi \, y \Longleftrightarrow \exists i \in I : x, y \in A_i \, .$$

Then r_π is called the *relation associated to the partition π*.

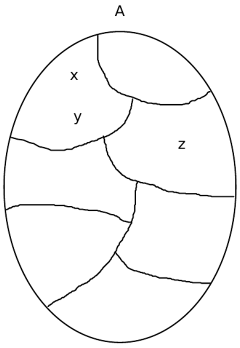


Figure 1.4: Relation associated to a partition.

The following theorem establishes the fundamental connection between equivalence relations and partitions.

Theorem 1.3.7 *(i) Let $r \in E(A)$. Then $A/r \in P(A)$.*
(ii) Let $\pi = (A_i)_{i \in I} \in P(A)$. Then $r_\pi \in E(A)$.
(iii) Let $F : E(A) \to P(A)$ be defined by

$$F(r) = A/r, \quad \forall r \in E(A).$$

Then F is a bijection, whose inverse is $G : P(A) \to E(A)$, defined by

$$G(\pi) = r_\pi, \quad \forall \pi \in P(A).$$

Example 1.3.8 (a) Consider the set A of all first-year students in Computer Science, and its partition, say $\pi = \{A_1, \dots, A_7\}$, where A_i denotes the set of all students in Group i for $i \in \{1, \dots, 7\}$. Then the equivalence relation r_π on A corresponding to the partition π is defined as follows: student $x \in A$ has relation r_π to student $y \in A$ if and only if students x and y are in the same group i.

(b) Let $A = \{1, 2, 3\}$ and let r and s be the homogeneous relations defined on A with the graphs

$$R = \{(1,1), (2,2), (3,3), (1,2), (2,1)\} \, ,$$
$$S = \{(1,1), (2,2), (3,3), (1,2), (2,3)\} \, .$$

Then r is an equivalence relation, but s is not. The partition corresponding to r is

$$A/r = \{\{1,2\}, \{3\}\} \, .$$

(c) Consider the following families of sets:

$$\pi = \{\{1\}, \{2, 3\}, \{4\}\},$$

$$\pi' = \{\{1, 2\}, \{2, 3\}, \{4\}\}.$$

Then π is a partition of $A = \{1, 2, 3, 4\}$, but π' is not. The equivalence relation corresponding to π has the graph

$$R_\pi = \{(1, 1), (2, 2), (2, 3), (3, 2), (3, 3), (4, 4)\}\,.$$

(d) The congruence relation modulo n is an equivalence relation on $\mathbb{Z}$ and its corresponding partition is

$$\mathbb{Z}/\rho_n = \{\rho_n < x >| \, x \in \mathbb{Z}\} = \{x + n\mathbb{Z} \mid x \in \mathbb{Z}\} = \{\widehat{x} \mid x \in \mathbb{Z}\}\,,$$

where an equivalence class is denoted by $\widehat{x}$. For $n \geq 2$, we denote

$$\mathbb{Z}_n = \mathbb{Z}/\rho_n = \{\widehat{0}, \widehat{1}, \dots, \widehat{n-1}\}\,.$$

For $n = 0$ and $n = 1$, we have seen in Example 1.3.2 that $\rho_0 = \delta_{\mathbb{Z}}$ and $\rho_1 = u$, and we get

$$\mathbb{Z}/\rho_0 = \{\{x\} \mid x \in \mathbb{Z}\} \quad \text{and} \quad \mathbb{Z}/\rho_1 = \{\mathbb{Z}\}\,,$$

that are the two extreme partitions of $\mathbb{Z}$.

Extra: Relational database

Binary relations may be naturally generalized as follows.

Definition 1.3.9 A (finite) tuple

$$r = (A_1, \dots, A_n, R),$$

where $A_1, \dots, A_n$ are sets and

$$R \subseteq A_1 \times \cdots \times A_n = \{(a_1, \dots, a_n) \mid a_1 \in A_1, \dots, a_n \in A_n\},$$

is called an *(n-ary) relation*. The sets $A_1, \dots, A_n$ are called the *domains* of r, and the set R is called the *graph* of r. The number n is called the *degree (arity)* of r. A *relational database* is a (finite) set of relations.

Example 1.3.10 Consider the relation

$$student = (Integer, String, String, Integer, Student),$$

where

$$Student \subseteq Integer \times String \times String \times Integer$$

is given by the following table:

ID (Integer)	**Surname** (String)	**Name** (String)	**Grade** (Integer)
7	Ionescu	Alina	9
11	Ardelean	Cristina	10
23	Ionescu	Dan	7

Remark 1.3.11 Some known relational database management systems are:

- Oracle and RDB – Oracle
- SQL Server and Access - Microsoft

Course 2

1.4 Operations

Definition 1.4.1 By an *operation* (or *composition law*) on a set A we understand a function

$$\varphi : A \times A \to A.$$

Usually, we denote operations by symbols like $\cdot$, $+$, $*$, so that $\varphi(x,y)$ is denoted by $x \cdot y$, $x+y$, $x*y$, $\forall (x,y) \in A \times A$. We denote by $(A,\cdot)$ the fact that " $\cdot$ " is an operation on a set A.

Example 1.4.2 The usual addition and multiplication are operations on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and the usual subtraction is an operation on $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, but not on $\mathbb{N}$. The usual division is not an operation on either of the five numerical sets, because of the element zero.

Definition 1.4.3 Let " $\cdot$ " be an operation on an arbitrary set A. Define the following laws:
(1) *Associative law*: $(x \cdot y) \cdot z = x \cdot (y \cdot z)\,, \quad \forall x,y,z \in A$.
(2) *Commutative law*: $x \cdot y = y \cdot x\,, \quad \forall x,y \in A$.
(3) *Identity law*: $\exists e \in A$ such that $\forall a \in A,\ \ a \cdot e = e \cdot a = a$. In this case, e is called an *identity element*.
(4) *Inverse law*: $\forall a \in A, \exists a' \in A$ such that $a \cdot a' = a' \cdot a = e$, where e is the identity element. In this case, a' is called an *inverse element for* a.

Lemma 1.4.4 *Let " $\cdot$ " be an operation on a set A.*
(i) If there exists an identity element in A, then it is unique.
(ii) Assume further that the operation "$\cdot$" is associative and has identity element e and let $a \in A$. If an inverse element for a does exist, then it is unique.

Let us now discuss some special subsets of sets endowed with an operation.

Definition 1.4.5 Consider an operation $\varphi : A \times A \to A$ on a set A and let $B \subseteq A$. Then B is called a *stable subset of A with respect to φ* (or *closed subset of A under the operation φ*) if

$$\forall x,y \in B\,, \quad \varphi(x,y) \in B\,.$$

In this case, we may consider the operation $\varphi' : B \times B \to B$ on B defined by

$$\varphi'(x,y) = \varphi(x,y)\,, \quad \forall (x,y) \in B \times B\,,$$

that is called the *operation induced by φ in the stable subset B of A*.
When using a symbol "$\cdot$" for φ, we simply say that B *is a stable subset of* $(A,\cdot)$.

Example 1.4.6 (a) The set $2\mathbb{Z} = \{2k \mid k \in \mathbb{Z}\}$ of even integers is stable in $(\mathbb{Z},+)$, but the set $2\mathbb{Z}+1 = \{2k+1 \mid k \in \mathbb{Z}\}$ of odd integers is not stable in $(\mathbb{Z},+)$.

(b) The interval $[0,1]$ is stable in $(\mathbb{R},\cdot)$, but the interval $[-1,0]$ is not stable in $(\mathbb{R},\cdot)$.

Remark 1.4.7 Notice that the associative, the commutative (and later on, the distributive laws) still hold in a stable subset (endowed with the induced operation), since they are true for every element in the initial set (only the universal quantifier $\forall$ appears in their definition). But the identity element and the inverse element do not transfer (their definition uses the existential quantifier $\exists$ as well).

1.5 Groups and rings

Definition 1.5.1 Let " $\cdot$ " be an operation on a set A. Then $(A, \cdot)$ is called a:
(1) *semigroup* if the associative law holds.
(2) *monoid* if it is a semigroup with identity element.
(3) *group* if it is a monoid in which every element has an inverse.

If the operation is commutative as well, then the structure is called *commutative*. A commutative group is also called an *abelian group* (after the name of N. H. Abel).

Remark 1.5.2 We denote by 1 the identity element of a group $(G, \cdot)$ and by x^{-1} the inverse of an element $x \in G$. In case of an additive group $(G, +)$, the identity element is denoted by 0, while the inverse of an element $x \in G$ is called the *symmetric* of x and is denoted by $-x$.

Definition 1.5.3 Let $(G, \cdot)$ be a semigroup, let $x \in G$ and let $n \in \mathbb{N}^*$. Then we may use the associative law and define

$$x^n = \underbrace{x \cdot x \cdot \ldots \cdot x}_{n \text{ times}}.$$

If $(G, \cdot)$ is a monoid, then we may also define $x^0 = 1$.
If $(G, \cdot)$ is a group, then we may also define $x^{-n} = (x^{-1})^n$.

Remark 1.5.4 If the operation is denoted by "+", then we replace the notation x^n by nx.

We may now give some standard properties of group computation.

Lemma 1.5.5 *Let $(G, \cdot)$ be a group, let $x \in G$ and let $m, n \in \mathbb{Z}$. Then:*
(i) $x^m \cdot x^n = x^{m+n}$.
(ii) $(x^m)^n = x^{mn}$.

Lemma 1.5.6 *Let $(G, \cdot)$ be a group and let $a, x, y \in G$. Then:*
(i) $a \cdot x = a \cdot y \Longrightarrow x = y$,
$x \cdot a = y \cdot a \Longrightarrow x = y$ *(cancellation laws).*
(ii) $(x^{-1})^{-1} = x$.
(iii) $(x \cdot y)^{-1} = y^{-1} \cdot x^{-1}$.

Remark 1.5.7 A finite group may be defined by its operation table, that specifies the result of any multiplication of two elements of the group. Using the cancellation laws, it is easy to see that the operation table of a group has the property that every element appears exactly once on each row and each column.

Example 1.5.8 (a) The operation "$-$" defined on $\mathbb{Z}$ is not associative.

(b) $(\mathbb{N}^*, +)$ is a semigroup, but not a monoid.

(c) $(\mathbb{N}, +)$, $(\mathbb{N}, \cdot)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$, $(\mathbb{C}, \cdot)$ are monoids, but not groups.

(d) $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$ and $(\mathbb{C}^*, \cdot)$ are groups.

(e) Let X be a non-empty set. By a *word on* X of length n we understand a string of n elements from X for some $n \in \mathbb{N}$. The word of length 0 is called the *void word* and is denoted by e. On the set X^* of words on X consider the operation "$\cdot$" given by concatenation. Then $(X^*, \cdot)$ is a monoid with identity element e, called the *free monoid* on the set X.

(f) Let $\{e\}$ be a single element set and let " $\cdot$ " be the only operation on $\{e\}$, defined by $e \cdot e = e$. Then $(\{e\}, \cdot)$ is an abelian group, called the *trivial group*.

(g) Let $n \in \mathbb{N}$, $n \geq 2$. Then $(\mathbb{Z}_n, +)$ is an abelian group, called the *group of residue classes modulo* n. The addition is defined by

$$\widehat{x} + \widehat{y} = \widehat{x+y}\,, \quad \forall \widehat{x}, \widehat{y} \in \mathbb{Z}_n\,.$$

(h) Let $n \in \mathbb{N}$ with $n \geq 2$. Denote by $M_{m,n}(\mathbb{R})$ the set of $m \times n$-matrices with entries in $\mathbb{R}$ and by $M_n(\mathbb{R})$ the set of $n \times n$-matrices with entries in $\mathbb{R}$. Then $(M_{m,n}(\mathbb{R}), +)$ is an abelian group and $(M_n(\mathbb{R}), \cdot)$ is a monoid.

Denote by $GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det(A) \neq 0\}$ the set of invertible $n \times n$-matrices with real entries. Then $(GL_n(\mathbb{R}), \cdot)$ is a group, called the *general linear group of rank* n.

(i) Let M be a set and let $S_M = \{f : M \to M \mid f \text{ is bijective}\}$. Then $(S_M, \circ)$ is a group, called the *symmetric group of* M. The identity element is the identity map 1_M and the inverse of an element f (which is a bijection) is the inverse function f^{-1}.

If $|M| = n$, then S_M is denoted by S_n, and the group $(S_n, \circ)$ is in fact the *permutation group* of n elements.

(j) Let $K = \{e, a, b, c\}$ and define an operation "$\cdot$" on K by the following table:

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Then $(K, \cdot)$ is an abelian group, called *Klein's group*. It comes from Geometry, and it may be viewed as the group of geometric transformations of a rectangle:

- e is the identical transformation,
- a is the symmetry with respect to the horizontal symmetry axis of the rectangle,
- b is the symmetry with respect to the vertical symmetry axis of the rectangle,
- c is the symmetry with respect to the center of the circumscribed circle of the rectangle.

The product $x \cdot y$ of two transformations x and y of K is defined by performing first y and then x.

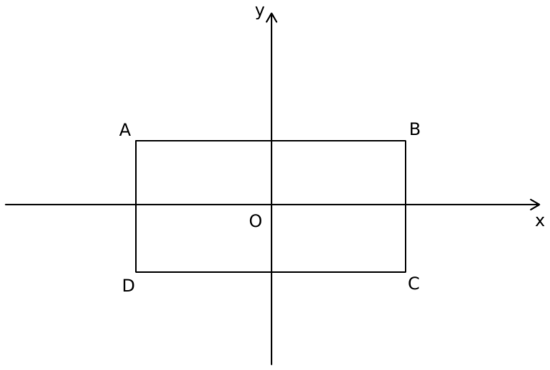


Figure 1.1: Klein's group.

Definition 1.5.9 Let R be a set. A structure with two operations $(R, +, \cdot)$ is called a:
(1) *ring* if $(R, +)$ is an abelian group, $(R, \cdot)$ is a semigroup and the *distributive laws* hold:

$$x \cdot (y+z) = x \cdot y + x \cdot z, \quad \forall x, y, z \in R,$$
$$(y+z) \cdot x = y \cdot x + z \cdot x, \quad \forall x, y, z \in R.$$

(2) *unitary ring* if $(R, +, \cdot)$ is a ring and there is an identity element with respect to "$\cdot$".
(3) *division ring* (or *skew field*) if $(R, +)$ is an abelian group, $(R^*, \cdot)$ is a group and the distributive laws hold.
(4) *field* if it is a commutative division ring.
The ring $(R, +, \cdot)$ is called *commutative* if the operation "$\cdot$" is commutative.

If $(R,+,\cdot)$ is a ring, then we denote the identity elements with respect to "$+$" and "$\cdot$" by 0 and 1 respectively. We also use the notation $R^* = R \setminus \{0\}$.

Remark 1.5.10 (1) A ring $(R,+,\cdot)$ is a division ring if and only if $|R| \geq 2$ and any $x \in R^*$ has an inverse $x^{-1} \in R^*$.

(2) If $(R,+,\cdot)$ is a ring, then $(R,+)$ is a group and $(R,\cdot)$ is a semigroup, so that we may talk about multiples and positive powers of elements of R.

Definition 1.5.11 Let $(R,+,\cdot)$ be a ring, let $x \in R$ and let $n \in \mathbb{N}^*$. Then we define

$$\begin{aligned} n\cdot x &= \underbrace{x+x+\cdots+x}_{n \text{ times}}\,,\\ 0\cdot x &= 0\,,\\ (-n)\cdot x &= -n\cdot x\,,\\ x^n &= \underbrace{x\cdot x\cdot\ldots\cdot x}_{n \text{ times}}\,. \end{aligned}$$

If R is a unitary ring, then we may also consider $x^0 = 1$.
If R is a division ring, then we may also define negative powers of x, namely $x^{-n} = (x^{-1})^n$.

Remark 1.5.12 Notice that in the definition $0\cdot x = 0$, the first 0 is the integer zero and the second 0 is the zero element of the ring R, that is, the identity element of the group $(R,+)$.

Clearly, the first computational properties of a ring $(R,+,\cdot)$ are the properties of the group $(R,+)$ and of the semigroup $(R,\cdot)$. Some relationship properties between the two operations are given in the following result, in which all zeros are the zero element of the ring R.

Lemma 1.5.13 *Let $(R,+,\cdot)$ be a ring and let $x,y,z \in R$. Then:*
(i) $x\cdot(y-z) = x\cdot y - x\cdot z$.
$(y-z)\cdot x = y\cdot x - z\cdot x$.
(ii) $x\cdot 0 = 0\cdot x = 0$.
(iii) $x\cdot(-y) = (-x)\cdot y = -x\cdot y$.

Example 1.5.14 (a) $(\mathbb{Z},+,\cdot)$ is a unitary ring, but not a field.

(b) $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$ and $(\mathbb{C},+,\cdot)$ are fields.

(c) Let $\{e\}$ be a single element set and let both "$+$" and "$\cdot$" be the only operation on $\{e\}$, defined by $e+e=e$ and $e\cdot e=e$. Then $(\{e\},+,\cdot)$ is a commutative unitary ring, called the *trivial ring*.

(d) Let $n \in \mathbb{N}$, $n \geq 2$. Then $(\mathbb{Z}_n,+,\cdot)$ is a commutative unitary ring, called the *ring of residue classes modulo n*. The addition and the multiplication are defined by

$$\widehat{x}+\widehat{y} = \widehat{x+y}\,, \quad \widehat{x}\cdot\widehat{y} = \widehat{x\cdot y}\,, \quad \forall \widehat{x},\widehat{y} \in \mathbb{Z}_n.$$

Note that $(\mathbb{Z}_n,+,\cdot)$ is a field if and only if n is prime.

(e) Let $(R,+,\cdot)$ be a commutative unitary ring. Then $(R[X],+,\cdot)$ is a commutative unitary ring, called the *polynomial ring over R in the indeterminate X*, where the operations are the usual addition and multiplication of polynomials.

(f) Let $n \in \mathbb{N}$, $n \geq 2$ and let $(R,+,\cdot)$ be a ring. Then $(M_n(R),+,\cdot)$ is a ring, called the *ring of matrices $n\times n$ with entries in R*, where the operations are the usual addition and multiplication of matrices.

(g) Let M be a non-empty set and let $(R,+,\cdot)$ be a ring. Define on the set

$$R^M = \{f \mid f : M \to R\}$$

two operations by: $\forall f, g \in R^M$, we have $f+g : M \to R$, $f \cdot g : M \to R$, where

$$(f+g)(x) = f(x) + g(x)\,, \quad \forall x \in M\,,$$

$$(f \cdot g)(x) = f(x) \cdot g(x)\,, \quad \forall x \in M\,.$$

Then $(R^M, +, \cdot)$ is a ring, called the *ring of functions with a set as domain and a ring as codomain.* The zero element is $\theta : M \to R$, $\theta(x) = 0$, $\forall x \in M$. The symmetric of any $f : M \to R$ is $-f : M \to R$, $(-f)(x) = -f(x)$, $\forall x \in M$.

(h) A ring $(R, +, \cdot)$ is called *Boolean* (after the name of G. Boole) if $a^2 = a$ for every $a \in R$. If M is a set and $\mathcal{P}(M)$ is the power set of M (that is, the set of all subsets of M), then $(\mathcal{P}(M), \Delta, \cap)$ is a Boolean ring, where Δ is the *symmetric difference* operation defined by

$$A \Delta B = (A \setminus B) \cup (B \setminus A)$$

for every $A, B \in \mathcal{P}(M)$.

1.6 Subgroups and subrings

We turn now our attention to the study of a group or ring inside another group or ring respectively. Recall that the associative and the commutative laws transfer in a stable subset, while the identity element and an inverse element do not transfer in general.

Definition 1.6.1 Let $(G, \cdot)$ be a group and let $H \subseteq G$. Then H is called a *subgroup* of G if:
(i) H is a stable subset of $(G, \cdot)$.
(ii) $(H, \cdot)$ is a group.

We denote by $H \leq G$ the fact that H is a subgroup of a group G.

The next characterization theorem gives more efficient ways to check that a subset of a group is a subgroup.

Theorem 1.6.2 *Let $(G, \cdot)$ be a group and let $H \subseteq G$. Then*

$$H \leq G \iff \begin{cases} H \neq \emptyset \ (1 \in H) \\ \forall x, y \in H,\ x \cdot y \in H \\ \forall x \in H,\ x^{-1} \in H. \end{cases} \iff \begin{cases} H \neq \emptyset \ (1 \in H) \\ \forall x, y \in H,\ x \cdot y^{-1} \in H. \end{cases}$$

Remark 1.6.3 (1) In case of an additive group $(G, +)$, the last two conditions in the middle of Theorem 1.6.2 become:

- $\forall x, y \in H$, $x + y \in H$.
- $\forall x \in H$, $-x \in H$.

(2) In case of an additive group $(G, +)$, the last condition in the end of Theorem 1.6.2 becomes:

- $\forall x, y \in H$, $x - y \in H$.

Let us now see some examples of subgroups.

Example 1.6.4 (a) Every non-trivial group $(G, \cdot)$ has two subgroups, namely $\{1\}$ and G, called the *trivial subgroups.*

(b) $\mathbb{Z}$ is a subgroup of $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$, $\mathbb{Q}$ is a subgroup of $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$, $\mathbb{R}$ is a subgroup of $(\mathbb{C}, +)$.

(c) The set $n\mathbb{Z} = \{nx \mid x \in \mathbb{Z}\}$ is a subgroup of $(\mathbb{Z}, +)$ for every $n \in \mathbb{N}$.

(d) The set $H = \{z \in \mathbb{C} \mid |z| = 1\}$ is a subgroup of the group $(\mathbb{C}^*, \cdot)$, called the *circle group*. But it is not a subgroup of the group $(\mathbb{C}, +)$.

(e) The set $U_n = \{z \in \mathbb{C} \mid z^n = 1\} \quad (n \in \mathbb{N}^*)$ is a subgroup of the group $(\mathbb{C}^*, \cdot)$, called the *group of* n^{th} *roots of unity*. Its elements are the following:

$$\varepsilon_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \quad k \in \{0, \dots, n-1\}.$$

(f) Consider the general linear group $(GL_n(\mathbb{R}), \cdot)$ of rank n, where $GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det(A) \neq 0\}$ $(n \in \mathbb{N},\ n \geq 2)$ and denote

$$SL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det(A) = 1\}.$$

Then $SL_n(\mathbb{R})$ is a subgroup of $(GL_n(\mathbb{R}), \cdot)$, called the *special linear group of rank n*.

Definition 1.6.5 Let $(R, +, \cdot)$ be a ring and let $A \subseteq R$. Then A is called a *subring* of R if:
(i) A is a stable subset of $(R, +, \cdot)$.
(ii) $(A, +, \cdot)$ is a ring.

Definition 1.6.6 Let $(K, +, \cdot)$ be a field and let $A \subseteq K$. Then A is called a *subfield* of K if:
(i) A is a stable subset of $(K, +, \cdot)$.
(ii) $(A, +, \cdot)$ is a field.

We denote by $A \leq R$ $(A \leq K)$ the fact that A is a subring (subfield) of a ring R (field K).

In practice, one checks that a subset of a ring (field) is a subring (subfield) by using one of the next two characterization theorems.

Theorem 1.6.7 *Let $(R, +, \cdot)$ be a ring and let $A \subseteq R$. Then*

$$A \textit{ is a subring of } R \iff \begin{cases} A \neq \emptyset \ (0 \in A) \\ \forall x, y \in A,\ \ x - y \in A \\ \forall x, y \in A,\ \ x \cdot y \in A. \end{cases}$$

Theorem 1.6.8 *Let $(K, +, \cdot)$ be a field and let $A \subseteq K$. Then*

$$A \textit{ is a subfield of } K \iff \begin{cases} |A| \geq 2 \ (0, 1 \in A) \\ \forall x, y \in A,\ \ x - y \in A \\ \forall x, y \in A \textit{ with } y \neq 0,\ \ x \cdot y^{-1} \in A. \end{cases}$$

Let us now see some examples of subrings and subfields.

Example 1.6.9 (a) Every non-trivial ring $(R, +, \cdot)$ has two subrings, namely $\{0\}$ and R, called the *trivial subrings*.

(b) $\mathbb{Z}$ is a subring of $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$.

(c) $\mathbb{Q}$ is a subfield of $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$, while $\mathbb{R}$ is a subfield of $(\mathbb{C}, +, \cdot)$.

(d) The set $n\mathbb{Z} = \{nx \mid x \in \mathbb{Z}\}$ is a subring of $(\mathbb{Z}, +, \cdot)$ for every $n \in \mathbb{N}$. Note that $n\mathbb{Z}$ does not have identity for $n \geq 2$.

(e) The set $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$ is a subring of the field $(\mathbb{C}, +, \cdot)$, but not a subfield. It is called the *ring of Gauss integers*.

1.7 Group and ring homomorphisms

Let us now define some special maps between groups or rings. For the sake of simplicity, we denote by the same symbol operations in different arbitrary structures.

Definition 1.7.1 Let $(G,\cdot)$ and $(G',\cdot)$ be groups and let $f : G \to G'$. Then f is called a *group homomorphism* if

$$f(x \cdot y) = f(x) \cdot f(y)\,, \quad \forall x, y \in G\,.$$

Also, f is called a *group isomorphism* if it is a bijective group homomorphism.

We denote by $G \simeq G'$ the fact that two groups G and G' are isomorphic.

Usually, we denote by 1 and $1'$ the identity elements in G and G' respectively.

Example 1.7.2 (a) Let $(G,\cdot)$ and $(G',\cdot)$ be groups and let $f : G \to G'$ be defined by $f(x) = 1'$, $\forall x \in G$. Then f is a homomorphism, called the *trivial group homomorphism.*

(b) Let $(G,\cdot)$ be a group. Then the identity map $1_G : G \to G$ is a group isomorphism.

(c) Let $n \in \mathbb{N}$ and let $f : \mathbb{Z} \to \mathbb{Z}$ be defined by $f(x) = nx$. Then f is a group homomorphism from the group $(\mathbb{Z},+)$ to itself.

(d) Let $n \in \mathbb{N}$ with $n \geq 2$. The map $f : \mathbb{Z} \to \mathbb{Z}_n$ defined by $f(x) = \widehat{x}$ is a group homomorphism between the groups $(\mathbb{Z},+)$ and $(\mathbb{Z}_n,+)$.

(e) Let $f : \mathbb{C}^* \to \mathbb{R}^*$ be defined by $f(z) = |z|$. Then f is a group homomorphism between $(\mathbb{C}^*,\cdot)$ and $(\mathbb{R}^*,\cdot)$. But $f : \mathbb{C} \to \mathbb{R}$ defined by $f(z) = |z|$ is not a group homomorphism between the groups $(\mathbb{C},+)$ and $(\mathbb{R},+)$.

(f) Let $n \in \mathbb{N}$, $n \geq 2$ and let $f : GL_n(\mathbb{R}) \to \mathbb{R}^*$ be defined by $f(A) = \det(A)$. Then f is a group homomorphism between the groups $(GL_n(\mathbb{R}),\cdot)$ and $(\mathbb{R}^*,\cdot)$.

Theorem 1.7.3 *Let $f : G \to G'$ be a group homomorphism. Then:*
(i) $f(1) = 1'$.
(ii) $(f(x))^{-1} = f(x^{-1})$, $\forall x \in G$.

Definition 1.7.4 Let $(R,+,\cdot)$ and $(R',+,\cdot)$ be rings and $f : R \to R'$. Then f is called a *ring homomorphism* if $\forall x, y \in R$ we have

$$f(x+y) = f(x) + f(y)\,,$$
$$f(x \cdot y) = f(x) \cdot f(y)\,.$$

Also, f is called a *ring isomorphism* if it is a bijective ring homomorphism.

We denote by $R \simeq R'$ the fact that two rings R and R' are isomorphic.

Example 1.7.5 (a) Let $(R,+,\cdot)$ and $(R',+,\cdot)$ be rings and let $f : R \to R'$ be defined by $f(x) = 0'$, $\forall x \in R$. Then f is a ring homomorphism, called the *trivial ring homomorphism.*

(b) Let $(R,+,\cdot)$ be a ring. Then the identity map $1_R : R \to R$ is a ring isomorphism.

(c) Let $n \in \mathbb{N}$ with $n \geq 2$. The map $f : \mathbb{Z} \to \mathbb{Z}_n$ defined by $f(x) = \widehat{x}$ is a ring homomorphism between the rings $(\mathbb{Z},+,\cdot)$ and $(\mathbb{Z}_n,+,\cdot)$.

(d) The map $f : \mathbb{C} \to \mathbb{R}$ defined by $f(z) = |z|$ is not a ring (field) homomorphism between the fields $(\mathbb{C},+,\cdot)$ and $(\mathbb{R},+,\cdot)$.

(e) Let $n \in \mathbb{N}$, $n \geq 2$ and let $f : M_n(\mathbb{R}) \to \mathbb{R}$ be defined by $f(A) = \det(A)$. Then f is not a ring homomorphism between the rings $(M_n(\mathbb{R}),+,\cdot)$ and $(\mathbb{R},+,\cdot)$.

Remark 1.7.6 If $f : R \to R'$ is a ring homomorphism, then the first condition from its definition tells us that f is a group homomorphism between $(R,+)$ and $(R',+)$. Then f takes the identity element of $(R,+)$ to the identity element of $(R',+)$, that is, $f(0) = 0'$ and we also have $f(-x) = -f(x)$, $\forall x \in R$. But in general, even if R and R' have identities, denoted by 1 and $1'$ respectively, in general it does not follow that a ring homomorphism $f : R \to R'$ has the property that $f(1) = 1'$.

Definition 1.7.7 Let $(R, +, \cdot)$ and $(R', +, \cdot)$ be rings with identity elements 1 and $1'$ respectively, and let $f : R \to R'$ be a ring homomorphism. Then f is called *unitary* if $f(1) = 1'$.

Theorem 1.7.8 *Let $(R, +, \cdot)$ and $(R', +, \cdot)$ be unitary rings with identity elements 1 and $1'$ respectively, and let $f : R \to R'$ be a ring homomorphism.*
(i) If f is surjective, then f is unitary.
(ii) If f is a ring isomorphism, then f is unitary.
(iii) If f is unitary and $x \in R$ has an inverse element $x^{-1} \in R$, then $f(x)$ has an inverse and

$$(f(x))^{-1} = f(x^{-1})\,.$$

Homework: Recall the definitions and the properties of matrices and determinants.

Extra: Fast adding

We describe a method for fast adding large natural numbers, following [Lidl, Pilz].

Remark 1.7.9 If a and b are two natural numbers, then it makes no difference if we add them as natural numbers or as elements (that is, residue classes) of some group $(\mathbb{Z}_n, +)$ for some $n > a + b$.

Theorem 1.7.10 (Chinese Remainder Theorem) *If $n = p_1^{r_1} \cdots p_k^{r_k}$ for some distinct primes $p_1, \dots, p_k$, then there is an isomorphism of additive groups:*

$$\mathbb{Z}_n \simeq \mathbb{Z}_{p_1^{r_1}} \times \cdots \times \mathbb{Z}_{p_k^{r_k}}$$

given by $\varphi : \mathbb{Z}_n \to \mathbb{Z}_{p_1^{r_1}} \times \cdots \times \mathbb{Z}_{p_k^{r_k}}$, $\varphi([x]_n) = ([x]_{p_1^{r_1}}, \dots, [x]_{p_k^{r_k}})$, where $[x]_m$ denotes the residue class of x modulo $m \in \mathbb{N}$. If we denote $n_i = p_i^{r_i}$, $N_i = \frac{n}{n_i}$ and $K_i = [N_i^{-1}]_{n_i}$ for every $i \in \{1, \dots, k\}$, then the inverse of φ is given by

$$\psi : \mathbb{Z}_{p_1^{r_1}} \times \cdots \times \mathbb{Z}_{p_k^{r_k}} \to \mathbb{Z}_n, \quad \psi(a_1, \dots, a_k) = \left[\sum_{i=1}^{k} a_i N_i K_i\right]_n .$$

This allows one (the computer) to replace the addition of large natural numbers by parallel "small" simultaneous additions. This technique is used in the design of computer software in order to speed up calculations.

Example 1.7.11 Let $a = 37$, $b = 56$, and choose $n = 140 = 2^2 \cdot 5 \cdot 7$.

$$\begin{aligned} a = 37 &\to [37]_{140} \to ([37]_4, [37]_5, [37]_7) = ([1]_4, [2]_5, [2]_7) \quad + \\ b = 56 &\to [56]_{140} \to ([56]_4, [56]_5, [56]_7) = ([0]_4, [1]_5, [0]_7) \\ a + b & \qquad\qquad\qquad\qquad\qquad\qquad\quad = ([1]_4, [3]_5, [2]_7) \end{aligned}$$

Now one solves the following system by the *Chinese Remainder Theorem*:

$$\begin{cases} x = 1 \pmod 4 \\ x = 3 \pmod 5 \\ x = 2 \pmod 7 \end{cases} .$$

We have:

$$n_1 = 4, n_2 = 5, n_3 = 7, n = n_1 \cdot n_2 \cdot n_3 = 140,$$
$$N_1 = \frac{n}{n_1} = 35, N_2 = \frac{n}{n_2} = 28, N_3 = \frac{n}{n_3} = 20.$$

Note that $K_i = [N_i^{-1}]_{n_i}$ means that $N_i K_i = 1 \pmod{n_i}$. Hence we have:

$$\begin{aligned} K_1 &= N_1^{-1} \bmod n_1 = 35^{-1} \bmod 4 = 3^{-1} \bmod 4 = 3, \\ K_2 &= N_2^{-1} \bmod n_2 = 28^{-1} \bmod 5 = 3^{-1} \bmod 5 = 7, \\ K_3 &= N_3^{-1} \bmod n_3 = 20^{-1} \bmod 7 = 6^{-1} \bmod 7 = 6. \end{aligned}$$

Finally, we get the solution

$$x = a_1 N_1 K_1 + a_2 N_2 K_2 + a_3 N_3 K_3 = 93$$

(unique solution modulo n). Hence $a + b = 93$.

Course 3

Chapter 2 VECTOR SPACES

This chapter deals with vector spaces (also called linear spaces), that are algebraic structures having also an "external operation" beside a usual operation, as it was the case in the previous chapter. They are the bricks of Linear Algebra and have numerous applications in different branches of Mathematics, but also in Physics, Computer Science and in other fields of Natural Sciences. Some of their algebraic applications will be studied in the next chapter, dedicated to the study of matrices and linear systems of equations.

Throughout the present chapter K will always denote a field.

2.1 Basic properties

Definition 2.1.1 A *vector space over* K (or a K*-vector space*) is an abelian group $(V,+)$ together with a so-called *external operation* or *scalar multiplication*

$$\cdot : K \times V \to V\,, \quad (k,v) \mapsto k \cdot v \quad (\text{or simply } kv)\,,$$

satisfying the following axioms:
(L_1) $k \cdot (v_1 + v_2) = k \cdot v_1 + k \cdot v_2$;
(L_2) $(k_1 + k_2) \cdot v = k_1 \cdot v + k_2 \cdot v$;
(L_3) $(k_1 \cdot k_2) \cdot v = k_1 \cdot (k_2 \cdot v)$;
(L_4) $1 \cdot v = v$,
for every $k, k_1, k_2 \in K$ and every $v, v_1, v_2 \in V$.
In this context, the elements of K are called *scalars* and the elements of V are called *vectors*.
Sometimes a vector space is also called a *linear space*.

We usually denote a vector space V over K by ${}_KV$, which emphasizes the fact that vectors are multiplied by scalars on the left hand side. Sometimes, we also use the notation $(V, K, +, \cdot)$.

Remark 2.1.2 (1) Notice that in the definition of a vector space there are present four operations, two denoted by the same symbol " $+$ " and two denoted by the same symbol " $\cdot$ ". Of course, they are not the same, but as we have already done it several times before, we use the convention to denote them identically for the sake of simplicity of writing. There are 3 operations in the classical sense, namely the addition and the multiplication in the field K and the addition in the group V and, on the other hand, there is also an external operation of multiplication by scalars.

(2) The axioms (L_1) and (L_2) look like some distributive laws and the axiom (L_3) looks like an associative law, but they are not, since the involved elements are not taken from the same set.

(3) The definition we have just given is that of a *left vector space*. It is also possible to give the definition of a *right vector space* by considering an external operation

$$\cdot : V \times K \to V\,, \quad (v,k) \mapsto v \cdot k,$$

satisfying some similar axioms, but on the right hand side.

Since one can show that there is a bijection between the left and the right vector spaces of the field K, we are going to study only the left vector spaces and omit the adjective "left".

Let us now see several important examples of vector spaces.

Example 2.1.3 (a) Let V_2 be the set of all vectors (in the classical sense) in the plane with a fixed origin O. Then V_2 is a vector space over $\mathbb{R}$ (or a *real vector space*), where the addition is the usual addition of two vectors by the parallelogram rule and the external operation is the usual scalar multiplication of vectors by real scalars.

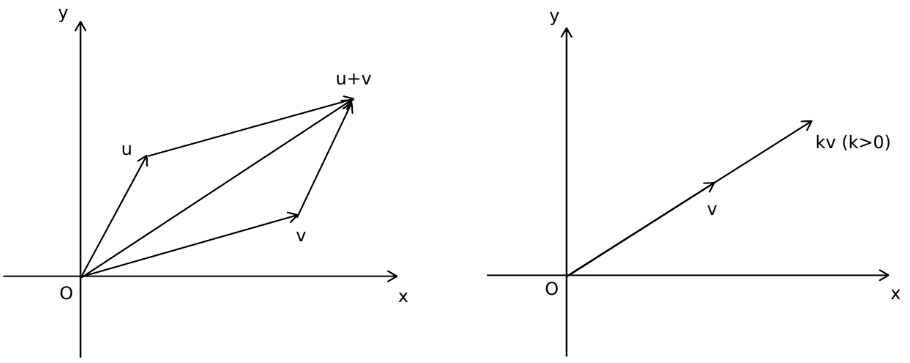


Figure 2.1: Vector addition and scalar multiplication.

If we consider two coordinate axes Ox and Oy in the plane, each vector in V_2 is perfectly determined by the coordinates of its ending point. Therefore, the addition of vectors and the scalar multiplication of vectors by real numbers become:

$$(x,y)+(x',y')=(x+x',y+y')\,,$$
$$k\cdot(x,y)=(k\cdot x,k\cdot y)\,,$$

$\forall k\in\mathbb{R}$ and $\forall(x,y),(x',y')\in\mathbb{R}\times\mathbb{R}$. Thus, $(\mathbb{R}^2,\mathbb{R},+,\cdot)$ is a vector space.

Similarly, one can consider the real vector space V_3 of all vectors in the space with a fixed origin. Moreover, a further, but more algebraical, generalization is possible, as we may see in the following example.

(b) Let $n\in\mathbb{N}^*$. Define

$$(x_1,\dots,x_n)+(y_1,\dots,y_n)=(x_1+y_1,\dots,x_n+y_n)\,,$$
$$k\cdot(x_1,\dots,x_n)=(kx_1,\dots,kx_n)\,,$$

$\forall(x_1,\dots,x_n),(y_1,\dots,y_n)\in K^n$ and $\forall k\in K$. Then $(K^n,K,+,\cdot)$ is a vector space, called the *canonical vector space* (or *standard vector space*) over K.

Let us discuss some particular cases. For $K=\mathbb{Z}_2$, $\mathbb{Z}_2^n$ is a vector space over $\mathbb{Z}_2$. For $n=1$, we get that ${}_KK$ is a vector space. Hence, as far as the classical numerical fields are concerned, ${}_{\mathbb{Q}}\mathbb{Q}$, ${}_{\mathbb{R}}\mathbb{R}$ and ${}_{\mathbb{C}}\mathbb{C}$ are vector spaces.

(c) If $V=\{e\}$ is a single element set, then we know that there is a unique structure of an abelian group for V, namely that one defined by $e+e=e$. Then we can define a unique scalar multiplication, namely $k\cdot e=e$, $\forall k\in K$. Thus, V is a vector space, called the *zero (null) vector space* and denoted by $\{0\}$.

(d) If A is a subfield of the field K, then K is a vector space over A, where the addition and the scalar multiplication are just the addition and the multiplication of elements in the field K.

In particular, ${}_{\mathbb{Q}}\mathbb{R}$, ${}_{\mathbb{Q}}\mathbb{C}$ and ${}_{\mathbb{R}}\mathbb{C}$ are vector spaces. Note that $\mathbb{R}$ may be viewed as a vector space over $\mathbb{Q}$ or $\mathbb{R}$, while $\mathbb{C}$ may be viewed as a vector space over any of the fields $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.

(e) $(K[X],K,+,\cdot)$ is a vector space, where the addition is the usual addition of polynomials and the scalar multiplication is defined as follows: $\forall f=a_0+a_1X+\cdots+a_nX^n\in K[X]$, $\forall k\in K$,

$$kf=(ka_0)+(ka_1)X+\cdots+(ka_n)X^n\,.$$

(f) Let $m,n\in\mathbb{N}$, $m,n\geq 2$. Then $(M_{m,n}(K),K,+,\cdot)$ is a vector space, where the operations are the usual addition and scalar multiplication of matrices.

(g) Let A be a non-empty set. Denote

$$K^A=\{f\mid f:A\to K\}\,.$$

Then $(K^A, K, +, \cdot)$ is a vector space, where the addition and the scalar multiplication are defined as follows: $\forall f, g \in K^A$, $\forall k \in K$, we have $f + g \in K^A$, $kf \in K^A$, where

$$\begin{gathered}(f+g)(x) = f(x) + g(x)\,, \\ (kf)(x) = kf(x)\end{gathered}$$

$\forall x \in A$. As a particular case, we obtain the vector space $(\mathbb{R}^{\mathbb{R}}, \mathbb{R}, +, \cdot)$ of real functions of a real variable.

(h) Let V and V' be K-vector spaces. Then the cartesian product $V \times V'$ is a K-vector space, called the *direct product* of V and V', where the addition and the scalar multiplication are defined as follows:

$$\begin{gathered}(v_1, v_1') + (v_2, v_2') = (v_1 + v_2, v_1' + v_2'), \\ k(v_1, v_1') = (kv_1, kv_1')\end{gathered}$$

$\forall (v_1, v_1'), (v_2, v_2') \in V \times V'$ and $\forall k \in K$.

(i) We have seen that $V = K \times K$ has a canonical structure of vector space over K. Let us now see what happens if we change the addition or the scalar multiplication.

Let us first define them as follows:

$$\begin{gathered}(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + 2y_2), \\ k \cdot (x_1, y_1) = (kx_1, ky_1)\end{gathered}$$

$\forall (x_1, y_1), (x_2, y_2) \in V$ and $\forall k \in K$. Then V is still a vector space over K, with a different structure of vector space than the canonical one.

Now let us define them as follows:

$$\begin{gathered}(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \\ k \cdot (x_1, y_1) = (kx_1, y_1)\end{gathered}$$

$\forall (x_1, y_1), (x_2, y_2) \in V$ and $\forall k \in K$. In general, they do not define a structure of vector space for V over K, because the axiom (L_2) does not hold. For instance, for $K = \mathbb{R}$, we have

$$(1+2) \cdot (3,4) = 3 \cdot (3,4) = (9,4) \neq (9,8) = (3,4) + (6,4) = 1 \cdot (3,4) + 2 \cdot (3,4).$$

Let us now state some computation rules in a vector space. Notice that we denote by 0 both the zero scalar and the zero vector.

Theorem 2.1.4 *Let V be a vector space over K. Then $\forall k, k' \in K$ and $\forall v, v' \in V$ we have:*
(i) $k \cdot 0 = 0 \cdot v = 0$.
(ii) $k(-v) = (-k)v = -kv$.
(iii) $k(v - v') = kv - kv'$.
(iv) $(k - k')v = kv - k'v$.

Proof. (i) We have:

$$\begin{gathered}k \cdot 0 + k \cdot v = k(0 + v) = kv \Longrightarrow k \cdot 0 = 0, \\ 0 \cdot v + k \cdot v = (0 + k)v = kv \Longrightarrow 0 \cdot v = 0.\end{gathered}$$

(ii) We have:

$$\begin{gathered}kv + k(-v) = k(v - v) = k \cdot 0 = 0 \Longrightarrow k(-v) = -kv, \\ kv + (-k)v = (k - k)v = 0 \cdot v = 0 \Longrightarrow (-k)v = -kv.\end{gathered}$$

(iii) We have:

$$k(v - v') + kv' = k(v - v' + v') = kv \Longrightarrow k(v - v') = kv - kv'.$$

(iv) We have:

$$(k - k')v + k'v = (k - k' + k')v = kv \Longrightarrow (k - k')v = kv - k'v.$$

Hence all the above properties are true. □

Theorem 2.1.5 *Let V be a vector space over K and let $k \in K$ and $v \in V$. Then*

$$kv = 0 \Longleftrightarrow k = 0 \text{ or } v = 0 \,.$$

Proof. $\boxed{\Longrightarrow}$ Assume that $kv = 0$. Suppose that $k \neq 0$. Then k is invertible in the field K and we have

$$kv = 0 \Longrightarrow kv = k \cdot 0 \Longrightarrow k^{-1}(kv) = k^{-1}(k \cdot 0) \Longrightarrow (k^{-1}k)v = (k^{-1}k) \cdot 0 \Longrightarrow v = 0 \,.$$

$\boxed{\Longleftarrow}$ This is Theorem 2.1.4 (i). □

2.2 Subspaces

Let us now discuss some special subsets of vector spaces, namely *subspaces*. We are going to define a subspace in the same general way as we did for subgroups or subrings.

Definition 2.2.1 Let V be a vector space over K and let $S \subseteq V$. Then S is a *subspace* of V if:
(i) $S \neq \emptyset$.
(ii) $\forall v_1, v_2 \in S$, $v_1 + v_2 \in S$.
(iii) $\forall k \in K$, $\forall v \in S$, $kv \in S$.

We usually denote by $S \leq_K V$, or simply by $S \leq V$, the fact that S is a subspace of the vector space V over K.

Remark 2.2.2 Notice that every subspace S of a vector space V over K is a subgroup of the additive group $(V, +)$, hence S must contain 0.

We have the following characterization theorem for subspaces.

Theorem 2.2.3 *Let V be a vector space over K and let $S \subseteq V$. Then*

$$S \leq V \Longleftrightarrow \begin{cases} S \neq \emptyset \quad (0 \in S) \\ \forall k_1, k_2 \in K \,, \ \forall v_1, v_2 \in S \,, \ k_1v_1 + k_2v_2 \in S \,. \end{cases}$$

Proof. $\boxed{\Longrightarrow}$ Taking $k = 0$ and $v_1 \in S \neq \emptyset$, we have $0 = 0 \cdot v_1 \in S$. Now let $k_1, k_2 \in K$ and $v_1, v_2 \in S$. Then we have $k_1v_1, k_2v_2 \in S$, and then $k_1v_1 + k_2v_2 \in S$.

$\boxed{\Longleftarrow}$ Choose $k_1 = k_2 = 1$ and then $k_2 = 0$ and use Definition 2.2.1. □

Example 2.2.4 (a) Every non-zero vector space V over K has two subspaces, namely $\{0\}$ and V. They are called the *trivial subspaces*.

(b) Let

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\} \,,$$
$$T = \{(x, y, z) \in \mathbb{R}^3 \mid x = y = z\} \,.$$

We have $S \neq \emptyset$, because $(0, 0, 0) \in S$. Now let $k_1, k_2 \in \mathbb{R}$ and $v_1, v_2 \in S$. Then $v_1 = (x_1, y_1, z_1)$ and $v_2 = (x_2, y_2, z_2)$ for some $x_1, y_1, z_1, x_2, y_2, z_2 \in \mathbb{R}$ such that $x_1 + y_1 + z_1 = 0$ and $x_2 + y_2 + z_2 = 0$. It follows that

$$k_1v_1 + k_2v_2 = (k_1x_1 + k_2x_2, k_1y_1 + k_2y_2, k_1z_1 + k_2z_2)$$

and we have

$$(k_1x_1 + k_2x_2) + (k_1y_1 + k_2y_2) + (k_1z_1 + k_2z_2) = k_1(x_1 + y_1 + z_1) + k_2(x_2 + y_2 + z_2) = 0.$$

Hence $k_1v_1 + k_2v_2 \in S$, and thus S is a subspace of the canonical real vector space $\mathbb{R}^3$. Note that S is a plane passing through the origin. For instance, the plane

$$\{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1\}$$

is not a subspace of $\mathbb{R}^3$ over $\mathbb{R}$.

We have $T \neq \emptyset$, because $(0, 0, 0) \in T$. Now let $k_1, k_2 \in \mathbb{R}$ and $v_1, v_2 \in T$. Then $v_1 = (x_1, x_1, x_1)$ and $v_2 = (x_2, x_2, x_2)$ for some $x_1, x_2 \in \mathbb{R}$. It follows that

$$k_1v_1 + k_2v_2 = (k_1x_1 + k_2x_2, k_1x_1 + k_2x_2, k_1x_1 + k_2x_2).$$

Hence $k_1v_1 + k_2v_2 \in T$, and thus T is a subspace of the canonical real vector space $\mathbb{R}^3$. Note that T is a line passing through the origin.

(c) More generally, the only subspaces of $\mathbb{R}^3$ are $\{(0, 0, 0)\}$, any line containing the origin, any plane containing the origin and $\mathbb{R}^3$.

(d) Let $n \in \mathbb{N}$ and let

$$K_n[X] = \{f \in K[X] \mid \text{degree}(f) \leq n\}\,.$$

Then $K_n[X]$ is a subspace of the polynomial vector space $K[X]$ over K. Note that the set $\{f \in K[X] \mid \text{degree}(f) = n\}$ is not a subspace of $K[X]$ over K.

(e) Let $I \subseteq \mathbb{R}$ be an interval. By Example 2.1.3,

$$\mathbb{R}^I = \{f \mid f : I \to \mathbb{R}\}$$

is a real vector space, where the addition and the scalar multiplication are defined as follows: $\forall f, g : I \to \mathbb{R}$, $\forall k \in K$, we have $f + g : I \to \mathbb{R}$, $kf : I \to \mathbb{R}$, where

$$\begin{aligned}(f + g)(x) &= f(x) + g(x)\,,\\ (kf)(x) &= kf(x)\,, \forall x \in I\,.\end{aligned}$$

The subsets

$$\begin{aligned}C(I, \mathbb{R}) &= \{f \in \mathbb{R}^I \mid f \text{ continuous on } I\},\\ D(I, \mathbb{R}) &= \{f \in \mathbb{R}^I \mid f \text{ derivable on } I\}\end{aligned}$$

are subspaces of $\mathbb{R}^I$, because they are nonempty and we have:

$$\forall k_1, k_2 \in \mathbb{R}, \forall f, g \in C(I, \mathbb{R}),\ k_1f + k_2g \in C(I, \mathbb{R}),$$

$$\forall k_1, k_2 \in \mathbb{R}, \forall f, g \in D(I, \mathbb{R}),\ k_1f + k_2g \in D(I, \mathbb{R}).$$

Extra: Vernam cipher

Following [Klein], we describe an easy, but secure cipher. Let $n \in \mathbb{N}^*$ and consider the canonical vector space $V = \mathbb{Z}_2^n$ over $\mathbb{Z}_2$. The vectors of V may be identified with n-bit binary strings. Suppose that Alice needs to send an n-bit plaintext $p \in \mathbb{Z}_2^n$ to Bob.

Vernam cipher:

1. (*Key establishment*) Alice and Bob randomly choose a vector $k \in \mathbb{Z}_2^n$ as a key.
2. (*Encryption*) Alice computes the ciphertext c according to the formula

$$c = p + k,$$

where the sum is a vector in $\mathbb{Z}_2^n$.

3. (*Decryption*) Bob computes the plaintext p according to the formula

$$p = c - k = c + k,$$

where the sum is a vector in $\mathbb{Z}_2^n$.

Remark 2.2.5 The system satisfies perfect secrecy, but the key k must be distributed in advance.

Example 2.2.6 Alice wants to send the message

$$p = (0,0,0,1,1,1,0,1,0,1) \in \mathbb{Z}_2^{10}$$

to Bob.
Alice and Bob agree on the vector

$$k = (0,1,1,0,1,0,0,0,0,1) \in \mathbb{Z}_2^{10}$$

as a key.
Alice encrypts the message by computing the ciphertext c as:

$$c = p + k = (0,0,0,1,1,1,0,1,0,1) + (0,1,1,0,1,0,0,0,0,1) = (0,1,1,1,0,1,0,1,0,0) \in \mathbb{Z}_2^{10}.$$

Bob decrypts the message by computing the plaintext p as:

$$p = c + k = (0,1,1,1,0,1,0,1,0,0) + (0,1,1,0,1,0,0,0,0,1) = (0,0,0,1,1,1,0,1,0,1) \in \mathbb{Z}_2^{10}.$$

Course 4

2.3 Generated subspace

For a vector space V over K, we denote by $S(V)$ the set of all subspaces of V. Sometimes, this set is denoted by $S_K(V)$ if we like to emphasize the field K.

Theorem 2.3.1 *Let V be a vector space over K and let $(S_i)_{i\in I}$ be a family of subspaces of V. Then $\bigcap_{i\in I} S_i \in S(V)$.*

Proof. For each $i \in I$, we have $S_i \in S(V)$, hence $0 \in S_i$. Then $0 \in \bigcap_{i\in I} S_i \neq \emptyset$. Now let $k_1, k_2 \in K$ and $x, y \in \bigcap_{i\in I} S_i$. Then $x, y \in S_i$, $\forall i \in I$. But $S_i \in S(V)$, $\forall i \in I$. It follows that $k_1x + k_2y \in S_i$, $\forall i \in I$, hence $k_1x + k_2y \in \bigcap_{i\in I} S_i$. Therefore, $\bigcap_{i\in I} S_i \in S(V)$. □

Remark 2.3.2 In general, the union of two subspaces of a vector space is not a subspace. For instance, $S = \{(x, 0) \mid x \in \mathbb{R}\}$ and $T = \{(0, y) \mid y \in \mathbb{R}\}$ are subspaces of the canonical real vector space $\mathbb{R}^2$, but $S\cup T$ is not a subspace of $\mathbb{R}^2$. Indeed, for instance, we have $(1, 0), (0, 1) \in S\cup T$, but $(1, 0)+(0, 1) = (1, 1) \notin S\cup T$.

Now we are interested in how to "complete" a given subset of a vector space to a subspace in a minimal way. This is the motivation for the following definition.

Definition 2.3.3 Let V be a vector space and let $X \subseteq V$. Then we denote

$$\langle X\rangle = \bigcap\{S \leq V \mid X \subseteq S\}$$

and we call it the *subspace generated by* X or the *subspace spanned by* X.
Here X is called the *generating set* of $\langle X\rangle$.
If $X = \{v_1, \dots, v_n\}$, we denote $\langle v_1, \dots, v_n\rangle = \langle\{v_1, \dots, v_n\}\rangle$.

Remark 2.3.4 (1) $\langle X\rangle$ is the "smallest" (with respect to inclusion) subspace of V containing X.

(2) $\langle\emptyset\rangle = \{0\}$.

(3) If $S \leq V$, then $\langle S\rangle = S$.

Definition 2.3.5 A vector space V over K is called *finitely generated* if $\exists v_1, \dots, v_n \in V$ $(n \in \mathbb{N})$ such that $V = \langle v_1, \dots, v_n\rangle$. Then the set $\{v_1, \dots, v_n\}$ is called a *system of generators for* V.

Definition 2.3.6 Let V be a vector space over K and $v_1, \dots, v_n \in V$ $(n \in \mathbb{N})$. A finite sum of the form

$$k_1v_1 + \cdots + k_nv_n\,,$$

where $k_i \in K$ $(i = 1, \dots, n)$, is called a (finite) *linear combination* of the vectors $v_1, \dots, v_n$.

Let us now determine how the elements of a generated subspace look like.

Theorem 2.3.7 *Let V be a vector space over K and let $\emptyset \neq X \subseteq V$. Then*

$$\langle X\rangle = \{k_1v_1 + \cdots + k_nv_n \mid k_i \in K\,,\ v_i \in X\,, i = 1, \dots, n\,,\ n \in \mathbb{N}^*\}\,,$$

that is, the set of all finite linear combinations of vectors of X.

Proof. We prove the result in 3 steps, by showing that

$$L = \{k_1v_1 + \cdots + k_nv_n \mid k_i \in K\,,\ v_i \in X\,, i = 1, \dots, n\,,\ n \in \mathbb{N}^*\}$$

is the smallest subspace of V containing X.

(i) Let $v \in X$. Then $v = 1 \cdot v \in L$, hence $L \neq \emptyset$. Now let $k, k' \in K$ and $v, v' \in L$. Then $v = \sum_{i=1}^{n} k_i v_i$ and $v' = \sum_{j=1}^{m} k'_j v'_j$ for some $k_1, \dots, k_n, k'_1, \dots, k'_m \in K$ and $v_1, \dots, v_n, v'_1, \dots, v'_m \in X$. Hence

$$kv + k'v' = k \sum_{i=1}^{n} k_i v_i + k' \sum_{j=1}^{m} k'_j v'_j = \sum_{i=1}^{n} (kk_i) v_i + \sum_{j=1}^{m} (k'k'_j) v'_j \in L\,,$$

because it is a finite linear combination of vectors of X. Hence we have $L \leq V$.

(ii) Choose $n = 1$ and $k_1 = 1$ in order to see that $X \subseteq L$.

(iii) Let $S \leq V$ be such that $X \subseteq S$. Let $k_1, \dots, k_n \in K$ and $v_1, \dots, v_n \in X$. Since $X \subseteq S$ and $S \leq V$, it follows that

$$k_1 v_1 + \dots + k_n v_n \in S.$$

Hence $L \subseteq S$.

Thus, we have $\langle X \rangle = L$ by the remark from the beginning of the proof. □

Corollary 2.3.8 *Let V be a vector space over K and let $x_1, \dots, x_n \in V$. Then*

$$\langle x_1, \dots, x_n \rangle = \{k_1 x_1 + \dots + k_n x_n \mid k_i \in K\,,\ x_i \in X\,, i = 1, \dots, n\}\,.$$

Example 2.3.9 (a) Consider the canonical real vector space $\mathbb{R}^3$. Then

$$\begin{aligned}\langle (1,0,0), (0,1,0), (0,0,1) \rangle &= \{k_1(1,0,0) + k_2(0,1,0) + k_3(0,0,1) \mid k_1, k_2, k_3 \in \mathbb{R}\} \\ &= \{(k_1,0,0) + (0,k_2,0) + (0,0,k_3) \mid k_1, k_2, k_3 \in \mathbb{R}\} \\ &= \{(k_1,k_2,k_3) \mid k_1, k_2, k_3 \in \mathbb{R}\} = \mathbb{R}^3\,.\end{aligned}$$

Hence $\mathbb{R}^3$ is generated by the three vectors $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$, and thus it is finitely generated.

(b) Consider the canonical vector space $\mathbb{Z}_2^3$ over $\mathbb{Z}_2$. Then

$$\begin{aligned}\langle (\widehat{1},\widehat{0},\widehat{0}), (\widehat{0},\widehat{1},\widehat{0}) \rangle &= \{k_1(\widehat{1},\widehat{0},\widehat{0}) + k_2(\widehat{0},\widehat{1},\widehat{0}) \mid k_1, k_2 \in \mathbb{Z}_2\} \\ &= \{(k_1,\widehat{0},\widehat{0}) + (\widehat{0},k_2,\widehat{0}) \mid k_1, k_2 \in \mathbb{Z}_2\} = \{(k_1,k_2,\widehat{0}) \mid k_1, k_2 \in \mathbb{Z}_2\} \neq \mathbb{Z}_2^3\,.\end{aligned}$$

Hence $\mathbb{Z}_2^3$ is not generated by the two vectors $(\widehat{1},\widehat{0},\widehat{0})$ and $(\widehat{0},\widehat{1},\widehat{0})$. But it is generated by $(\widehat{1},\widehat{0},\widehat{0})$, $(\widehat{0},\widehat{1},\widehat{0})$ and $(\widehat{0},\widehat{0},\widehat{1})$, hence it is finitely generated.

(c) Consider the subspace

$$S = \{(x,y,z) \in \mathbb{R}^3 \mid x - y - z = 0\}$$

of the canonical real vector space $\mathbb{R}^3$. Let us write it as a generated subspace. Expressing $x = y + z$, we have:

$$\begin{aligned}S &= \{(y+z,y,z) \mid y,z \in \mathbb{R}\} = \{(y,y,0) + (z,0,z) \mid y,z \in \mathbb{R}\} \\ &= \{y(1,1,0) + z(1,0,1) \mid y,z \in \mathbb{R}\} = \langle (1,1,0), (1,0,1) \rangle.\end{aligned}$$

Alternatively, one may express y or z by using the other two components and get other writings of S as a generated subspace, namely $S = \langle (1,1,0), (0,-1,1) \rangle = \langle (1,0,1), (0,1,-1) \rangle$. We see that S is finitely generated.

In what follows we shall be interested in "decomposing" a vector space into subspaces. This allows one to study the component subspaces and then deduce properties of the whole vector space.

Let us first define the sum and the direct sum of two subspaces of a vector space.

Definition 2.3.10 Let V be a vector space over K and let $S, T \leq V$.
We define the *sum* of the subspaces S and T as the set

$$S + T = \{s + t \mid s \in S\,,\ t \in T\}\,.$$

If $S \cap T = \{0\}$, then $S + T$ is denoted by $S \oplus T$ and is called the *direct sum* of the subspaces S and T.

Theorem 2.3.11 *Let V be a vector space over K and let $S, T \leq V$. Then*

$$S + T = \langle S \cup T \rangle\,.$$

Proof. We prove the equality by double inclusion.

First, let $v = s + t \in S + T$, for some $s \in S$ and $t \in T$. Then

$$v = 1 \cdot s + 1 \cdot t$$

is a linear combination of the vectors $s, t \in S \cup T$, hence $v \in \langle S \cup T \rangle$. Thus, $S + T \subseteq \langle S \cup T \rangle$.

Now let $v \in \langle S \cup T \rangle$. Then

$$v = \sum_{i=1}^{n} k_i v_i = \sum_{i \in I} k_i v_i + \sum_{j \in J} k_j v_j\,,$$

where $I = \{i \in \{1, \dots, n\} \mid v_i \in S\}$ and $J = \{j \in \{1, \dots, n\} \mid v_j \in T \setminus S\}$. But the first sum is a linear combination of vectors of S, hence it belongs to S, while the second sum is a linear combination of vectors of T, hence it belongs to T. Thus, $v \in S + T$ and consequently $\langle S \cup T \rangle \subseteq S + T$.

Therefore, $S + T = \langle S \cup T \rangle$. □

Corollary 2.3.12 *Let V be a vector space over K and let $S, T \leq V$. Then $S + T \leq V$.*

Proof. By Theorem 2.3.11. □

Theorem 2.3.13 *Let V be a vector space over K and let $S, T \leq V$. Then*

$$V = S \oplus T \Longleftrightarrow \forall v \in V\,, \exists! s \in S\,, t \in T\ : v = s + t\,.$$

Proof. $\boxed{\Longrightarrow}$ Assume that $V = S \oplus T$. Let $v \in V$. Then $\exists s \in S\,, t \in T$ such that $v = s + t$. Now suppose that $\exists s' \in S\,, t' \in T$ such that $v = s' + t'$. Then $s + t = s' + t'$, whence

$$s - s' = t' - t \in S \cap T = \{0\}\,.$$

Hence $s = s'$ and $t = t'$, that show the uniqueness.

$\boxed{\Longleftarrow}$ Assume that $\forall v \in V$, $\exists! s \in S\,, t \in T$ such that $v = s + t$. Then $V \subseteq S + T$. Clearly, we have $S + T \subseteq V$ and consequently $V = S + T$. Now suppose that $0 \neq v \in S \cap T$. Then

$$v = v + 0 = 0 + v.$$

But this is a contradiction, since we have the uniqueness of writing of v as a sum of an element of S and an element of T. Therefore, $S \cap T = \{0\}$ and thus, $V = S \oplus T$. □

Example 2.3.14 Consider the canonical real vector space $\mathbb{R}^2$. Then $\mathbb{R}^2 = S \oplus T$, where $S = \{(x, 0) \mid x \in \mathbb{R}\}$ and $T = \{(0, y) \mid y \in \mathbb{R}\}$.

2.4 Linear maps

Definition 2.4.1 Let V and V' be vector spaces over the same field K. A function $f : V \to V'$ is called:
(1) *(K-)linear map* (or *(vector space) homomorphism* or *linear transformation*) if

$$\begin{aligned} f(v_1+v_2) &= f(v_1)+f(v_2)\,, \quad \forall v_1,v_2 \in V, \\ f(kv) &= kf(v)\,, \quad \forall k \in K, \forall v \in V. \end{aligned}$$

(2) *isomorphism* if it is a bijective K-linear map.
(3) *endomorphism* if it is a K-linear map and $V = V'$.
(4) *automorphism* if it is a bijective K-linear map and $V = V'$.

Remark 2.4.2 If $f : V \to V'$ is a K-linear map, then the first condition from its definition tells us that f is a group homomorphism between the groups $(V,+)$ and $(V',+)$. Then we have $f(0) = 0'$ and $f(-v) = -f(v)$, $\forall v \in V$.

We denote by $V \simeq V'$ the fact that two vector spaces V and V' are isomorphic. We also denote

$$\begin{aligned} \mathrm{Hom}_K(V,V') &= \{f : V \to V' \mid f \text{ is } K\text{-linear}\}\,, \\ \mathrm{End}_K(V) &= \{f : V \to V \mid f \text{ is } K\text{-linear}\}\,, \\ \mathrm{Aut}_K(V) &= \{f : V \to V \mid f \text{ is bijective } K\text{-linear}\}\,. \end{aligned}$$

Let us now give a characterization theorem for linear maps.

Theorem 2.4.3 *Let V and V' be vector spaces over K and $f : V \to V'$. Then*

$$f \text{ is a } K\text{-linear map} \Longleftrightarrow f(k_1v_1+k_2v_2) = k_1f(v_1)+k_2f(v_2)\,, \forall k_1,k_2 \in K, \forall v_1,v_2 \in V.$$

Proof. $\boxed{\Longrightarrow}$ Let $k_1, k_2 \in K$ and $v_1, v_2 \in V$. Then

$$f(k_1v_1+k_2v_2) = f(k_1v_1)+f(k_2v_2) = k_1f(v_1)+k_2f(v_2)\,.$$

$\boxed{\Longleftarrow}$ Choose $k_1 = k_2 = 1$ and then $k_2 = 0$ to get the two conditions of a K-linear map. □

Example 2.4.4 (a) Let V and V' be vector spaces over K and let $f : V \to V'$ be defined by $f(v) = 0'$, $\forall v \in V$. Then f is a K-linear map, called the *trivial linear map*.

(b) Let V be a vector space over K. Then the identity map $1_V : V \to V$ is an automorphism of V.

(c) Let V be a vector space and $S \leq V$. Define $i : S \to V$ by $i(v) = v$, $\forall v \in S$. Then i is a K-linear map, called the *inclusion linear map*.

(d) Let V be a vector space over K and $a \in K$. Define $t_a : V \to V$ by $t_a(v) = av$, $\forall v \in V$. Then t_a is an endomorphism of V.

(e) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = x+y$. Then f is an $\mathbb{R}$-linear map, because we have

$$\begin{aligned} f(k_1(x_1,y_1)+k_2(x_2,y_2)) &= f(k_1x_1+k_2x_2, k_1y_1+k_2y_2) \\ &= (k_1x_1+k_2x_2)+(k_1y_1+k_2y_2) \\ &= k_1(x_1+y_1)+k_2(x_2+y_2) \\ &= k_1f(x_1,y_1)+k_2f(x_2,y_2) \end{aligned}$$

for every $k_1, k_2 \in K$ and for every $(x_1,y_1),(x_2,y_2) \in \mathbb{R}^2$.

On the other hand, $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x,y) = xy$ is not an $\mathbb{R}$-linear map, because, for instance, we have

$$f((1,0)+(0,1)) = f(1,1) = 1 \neq 0 = f(1,0)+f(0,1).$$

(f) Let $\theta \in \mathbb{R}$ and let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$f(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta),$$

which is the counterclockwise rotation of angle θ about the origin in the plane. Then f is an $\mathbb{R}$-linear map. In particular, for $\theta = \frac{\pi}{2}$, we have $f(x, y) = (-y, x)$.

(g) For an interval $I = [a, b] \subseteq \mathbb{R}$ we considered the real vector space

$$\mathbb{R}^I = \{f \mid f : I \to \mathbb{R}\}$$

and its subspaces

$$C(I, \mathbb{R}) = \{f \in \mathbb{R}^I \mid f \text{ continuous on } I\},$$
$$D(I, \mathbb{R}) = \{f \in \mathbb{R}^I \mid f \text{ derivable on } I\}.$$

Then

$$F : D(I, \mathbb{R}) \to \mathbb{R}^I, \quad F(f) = f',$$
$$G : C(I, \mathbb{R}) \to \mathbb{R}, \quad G(f) = \int_a^b f(t)\mathrm{d}t,$$

are $\mathbb{R}$-linear maps.

Theorem 2.4.5 *(i) Let $f : V \to V'$ be an isomorphism of vector spaces over K. Then $f^{-1} : V' \to V$ is again an isomorphism of vector spaces over K.*
(ii) Let $f : V \to V'$ and $g : V' \to V''$ be K-linear maps. Then $g \circ f : V \to V''$ is a K-linear map.

Proof. (i) Since f is an isomorphism of vector spaces over K, f is bijective, hence so is f^{-1}.

Now let $k_1, k_2 \in K$ and $v_1', v_2' \in V'$. We have to prove that

$$f^{-1}(k_1v_1' + k_2v_2') = k_1f^{-1}(v_1') + k_2f^{-1}(v_2').$$

Let us denote $v_1 = f^{-1}(v_1')$ and $v_2 = f^{-1}(v_2')$. Then $f(v_1) = v_1'$ and $f(v_2) = v_2'$, hence

$$k_1v_1' + k_2v_2' = k_1f(v_1) + k_2f(v_2) = f(k_1v_1 + k_2v_2).$$

Thus we have

$$f^{-1}(k_1v_1' + k_2v_2') = k_1v_1 + k_2v_2 = k_1f^{-1}(v_1') + k_2f^{-1}(v_2').$$

Hence f^{-1} is an isomorphism of vector spaces over K.

(ii) Let $k_1, k_2 \in K$ and $v_1, v_2 \in V$. We have:

$$\begin{aligned}(g \circ f)(k_1v_1 + k_2v_2) &= g(f(k_1v_1 + k_2v_2))\\ &= g(k_1f(v_1) + k_2f(v_2))\\ &= k_1g(f(v_1)) + k_2g(f(v_2))\\ &= k_1(g \circ f)(v_1) + k_2(g \circ f)(v_2).\end{aligned}$$

Hence $g \circ f$ is a K-linear map. □

Definition 2.4.6 Let $f : V \to V'$ be a K-linear map. Then the set

$$\operatorname{Ker} f = \{v \in V \mid f(v) = 0'\}$$

is called the *kernel* (or the *null space*) of the K-linear map f and the set

$$\operatorname{Im} f = \{f(v) \mid v \in V\}$$

is called the *image* (or the *range space*) of the K-linear map f.

Theorem 2.4.7 *Let $f : V \to V'$ be a K-linear map. Then*

$$\mathrm{Ker} f \leq V \text{ and } \mathrm{Im} f \leq V' .$$

Proof. First, note that $f(0) = 0'$, hence $0 \in \mathrm{Ker}\, f \neq \emptyset$. Let $k_1, k_2 \in K$ and $v_1, v_2 \in \mathrm{Ker}\, f$. We prove that $k_1v_1 + k_2v_2 \in \mathrm{Ker}\, f$. Indeed, we have:

$$f(k_1v_1 + k_2v_2) = k_1f(v_1) + k_2f(v_2) = k_1 \cdot 0' + k_2 \cdot 0' = 0',$$

and thus $k_1v_1 + k_2v_2 \in \mathrm{Ker}\, f$. Hence $\mathrm{Ker}\, f \leq V$.

Now note that $0' = f(0) \in \mathrm{Im}\, f \neq \emptyset$. Let $k_1, k_2 \in K$ and $v_1', v_2' \in \mathrm{Im}\, f$. We prove that $k_1v_1' + k_2v_2' \in \mathrm{Im}\, f$. We have $v_1' = f(v_1)$ and $v_2' = f(v_2)$ for some $v_1, v_2 \in V$. Then:

$$k_1v_1' + k_2v_2' = k_1f(v_1) + k_2f(v_2) = f(k_1v_1 + k_2v_2) \in \mathrm{Im}\, f.$$

Hence $\mathrm{Im}\, f \leq V'$. □

Theorem 2.4.8 *Let $f : V \to V'$ be a K-linear map. Then*

$$\mathrm{Ker}\, f = \{0\} \Longleftrightarrow f \text{ is injective} .$$

Proof. $\boxed{\Longrightarrow}$ Assume that $\mathrm{Ker}\, f = \{0\}$. Let $v_1, v_2 \in V$ be such that $f(v_1) = f(v_2)$. It follows that $f(v_1 - v_2) = 0$, hence $v_1 - v_2 \in \mathrm{Ker}\, f = \{0\}$, and thus $v_1 = v_2$. Therefore, f is injective.

$\boxed{\Longleftarrow}$ Assume that f is injective. Clearly, we have $\{0\} \subseteq \mathrm{Ker}\, f$. Now let $v \in \mathrm{Ker}\, f$. Then $f(v) = 0' = f(0)$. By the injectivity of f, we deduce that $v = 0$. Thus $\mathrm{Ker}\, f \subseteq \{0\}$, and consequently, $\mathrm{Ker}\, f = \{0\}$. □

Theorem 2.4.9 *Let $f : V \to V'$ be a K-linear map and let $X \subseteq V$. Then*

$$f(\langle X \rangle) = \langle f(X) \rangle .$$

Proof. If $X = \emptyset$, then we have:

$$f(\langle \emptyset \rangle) = f(\{0\}) = \{f(0)\} = \{0'\} = \langle \emptyset \rangle = \langle f(\emptyset) \rangle.$$

Now assume that $X \neq \emptyset$. By Theorem 2.3.7 we have

$$\langle X \rangle = \{k_1v_1 + \cdots + k_nv_n \mid k_i \in K\,,\ v_i \in X\,, i = 1, \ldots, n\,,\ n \in \mathbb{N}^*\} .$$

Since f is a K-linear map, it follows by Theorem 2.4.3 that

$$\begin{aligned} f(\langle X \rangle) &= \{f(k_1v_1 + \cdots + k_nv_n) \mid k_i \in K\,,\ v_i \in X\,, i = 1, \ldots, n\,,\ n \in \mathbb{N}^*\} \\ &= \{k_1f(v_1) + \cdots + k_nf(v_n) \mid k_i \in K\,,\ v_i \in X\,, i = 1, \ldots, n\,,\ n \in \mathbb{N}^*\} \\ &= \langle f(X) \rangle\,, \end{aligned}$$

which proves the result. □

Theorem 2.4.10 *Let V and V' be vector spaces over K. Consider on $\mathrm{Hom}_K(V, V')$ the operations: $\forall f, g \in \mathrm{Hom}_K(V, V')$ and $\forall k \in K$, $f + g, k \cdot f \in \mathrm{Hom}_K(V, V')$, where*

$$\begin{aligned} (f + g)(v) &= f(v) + g(v)\,, \\ (kf)(v) &= kf(v) \end{aligned}$$

$\forall v \in V$. Then $\mathrm{Hom}_K(V, V')$ is a vector space over K.

Proof. Let $k \in K$ and $f, g \in \mathrm{Hom}_K(V, V')$. Let us prove first that the operations are well-defined, that is, $f + g, kf \in \mathrm{Hom}_K(V, V')$. Let $k_1, k_2 \in K$ and $v_1, v_2 \in V$. Then:

$$\begin{aligned}(f+g)(k_1xv_1+k_2v_2) &= f(k_1v_1+k_2v_2)+g(k_1v_1+k_2v_2)\\ &= k_1f(v_1)+k_2f(v_2)+k_1g(v_1)+k_2g(v_2)\\ &= k_1(f(v_1)+g(v_1))+k_2(f(v_2)+g(v_2))\\ &= k_1(f+g)(v_1)+k_2(f+g)(v_2)\,.\end{aligned}$$

We also have:

$$\begin{aligned}(kf)(k_1v_1+k_2v_2) &= kf(k_1v_1+k_2v_2)\\ &= k(k_1f(v_1))+k(k_2f(v_2))\\ &= (kk_1)f(v_1)+(kk_2)f(v_2)\\ &= k_1(kf(v_1))+k_2(kf(v_2))\,.\end{aligned}$$

Therefore, $f + g, kf \in \mathrm{Hom}_K(V, V')$.

It is easy to check that $(\mathrm{Hom}_K(V, V'), +)$ is an abelian group, where the identity element is the trivial linear map

$$\theta : V \to V', \quad \theta(v) = 0', \quad \forall v \in V$$

and every element $f \in \mathrm{Hom}_K(V, V')$ has a symmetric

$$-f \in \mathrm{Hom}_K(V, V'), \quad (-f)(v) = -f(v), \quad \forall v \in V.$$

Checking the axioms of the vector space for $\mathrm{Hom}_K(V, V')$ reduces, by the definitions of operations, to the axioms for the vector space V'. □

Corollary 2.4.11 *Let V be a vector space over K. Then $\mathrm{End}_K(V)$ is a vector space over K.*

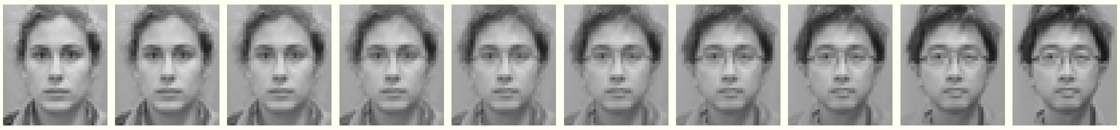
Proof. Take $V = V'$ in Theorem 2.4.10. □

Extra: Image crossfade

Following [Klein], we describe a way to achieve an image crossfade effect.
A black-and-white image of (say) $n = 1024 \times 768$ pixels can be viewed as a vector in the real canonical vector space $\mathbb{R}^n$, where each component of the vector is the intensity of the corresponding pixel.
Let us consider two vectors representing images:

$v_1 =$, $v_2 =$.

Now consider the following intermediate images:



The vectors corresponding to the above images are the following linear combinations of the vectors v_1 and v_2:

$$\begin{gathered} v_1, \quad \frac{8}{9}v_1+\frac{1}{9}v_2, \quad \frac{7}{9}v_1+\frac{2}{9}v_2, \quad \frac{6}{9}v_1+\frac{3}{9}v_2, \quad \frac{5}{9}v_1+\frac{4}{9}v_2,\\ \frac{4}{9}v_1+\frac{5}{9}v_2, \quad \frac{3}{9}v_1+\frac{6}{9}v_2, \quad \frac{2}{9}v_1+\frac{7}{9}v_2, \quad \frac{1}{9}v_1+\frac{8}{9}v_2, \quad v_2. \end{gathered}$$

One may use these images as frames in a video in order to get a crossfade effect.

Course 5

2.5 Linear independence

Definition 2.5.1 Let V be a vector space over K. We say that the vectors $v_1,\dots,v_n \in V$ are (or the set of vectors $\{v_1,\dots,v_n\}$ is):
(1) *linearly independent* in V if for every $k_1,\dots,k_n \in K$,

$$k_1v_1+\dots+k_nv_n=0 \Longrightarrow k_1=\dots=k_n=0\,.$$

(2) *linearly dependent* in V if they are not linearly independent, that is, $\exists k_1,\dots,k_n \in K$ not all zero such that

$$k_1v_1+\dots+k_nv_n=0\,.$$

Remark 2.5.2 (1) A set consisting of a single vector v is linearly dependent $\Longleftrightarrow v=0$.

(2) As an immediate consequence of the definition, we notice that if V is a vector space over K and $X,Y\subseteq V$ such that $X\subseteq Y$, then:
(i) If Y is linearly independent, then X is linearly independent.
(ii) If X is linearly dependent, then Y is linearly dependent. Thus, every set of vectors containing the zero vector is linearly dependent.

(3) More generally, an infinite set of vectors of V is called *linearly independent* if any finite subset is linearly independent, and *linearly dependent* if there exists a finite subset which is linearly dependent.

Theorem 2.5.3 *Let V be a vector space over K. Then the vectors $v_1,\dots,v_n\in V$ are linearly dependent if and only if one of the vectors is a linear combination of the others, that is, $\exists j\in\{1,\dots,n\}$ such that*

$$v_j=\sum_{\substack{i=1\\ i\neq j}}^{n}\alpha_iv_i$$

for some $\alpha_i\in K$, where $i\in\{1,\dots,n\}$ and $i\neq j$.

Proof. $\boxed{\Longrightarrow}$ Assume that $v_1,\dots,v_n\in V$ are linearly dependent. Then $\exists k_1,\dots,k_n\in K$ not all zero, say $k_j\neq 0$, such that $k_1v_1+\dots+k_nv_n=0$. But this implies

$$-k_jv_j=\sum_{\substack{i=1\\ i\neq j}}^{n}k_iv_i$$

and further,

$$v_j=\sum_{\substack{i=1\\ i\neq j}}^{n}(-k_j^{-1}k_i)v_i\,.$$

Now choose $\alpha_i=-k_j^{-1}k_i$ for each $i\neq j$ to get the conclusion.

$\boxed{\Longleftarrow}$ Assume that $\exists j\in\{1,\dots,n\}$ such that

$$v_j=\sum_{\substack{i=1\\ i\neq j}}^{n}\alpha_iv_i$$

for some $\alpha_i\in K$, where $i\in\{1,\dots,n\}$ and $i\neq j$. Then

$$(-1)v_j+\sum_{\substack{i=1\\ i\neq j}}^{n}\alpha_iv_i=0\,.$$

Since there exists such a linear combination equal to zero and the scalars are not all zero, the vectors $v_1, \dots, v_n$ are linearly dependent. □

Example 2.5.4 (*a*) Let V_2 be the real vector space of all vectors (in the classical sense) in the plane with a fixed origin O. Recall that the addition is the usual addition of two vectors by the parallelogram rule and the external operation is the usual scalar multiplication of vectors by real scalars. Then:
(i) one vector v is linearly dependent in $V_2 \Longleftrightarrow v = 0$;
(ii) two vectors are linearly dependent in $V_2 \Longleftrightarrow$ they are collinear;
(iii) three vectors (or more) are always linearly dependent in V_2.

Now let V_3 be the real vector space of all vectors (in the classical sense) in the space with a fixed origin O. Then:
(i) one vector v is linearly dependent in $V_3 \Longleftrightarrow v = 0$;
(ii) two vectors are linearly dependent in $V_3 \Longleftrightarrow$ they are collinear;
(iii) three vectors are linearly dependent in $V_3 \Longleftrightarrow$ they are coplanar;
(iv) four vectors (or more) are always linearly dependent in V_3.

(*b*) If K is a field and $n \in \mathbb{N}^*$, then the vectors $e_1 = (1, 0, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, $\dots$, $e_n = (0, 0, 0, \dots, 1) \in K^n$ are linearly independent in the canonical vector space K^n over K. In order to show that, let $k_1, \dots, k_n \in K$ be such that

$$k_1e_1 + k_2e_2 + \cdots + k_ne_n = 0 \in K^n.$$

Then we have

$$k_1(1, 0, 0, \dots, 0) + k_2(0, 1, 0, \dots, 0) + \cdots + k_n(0, 0, 0, \dots, 1) = (0, \dots, 0),$$

and furthermore

$$(k_1, \dots, k_n) = (0, \dots, 0).$$

This implies that $k_1 = \cdots = k_n = 0$, and thus the vectors $e_1, \dots, e_n$ are linearly independent in K^n.

(*c*) Let K be a field and $n \in \mathbb{N}$. Then the vectors $1, X, X^2, \dots, X^n$ are linearly independent in the vector space $K_n[X] = \{f \in K[X] \mid \text{degree}(f) \leq n\}$ over K. More generally, the vectors

$$1, X, X^2, \dots, X^n, \dots \quad (n \in \mathbb{N})$$

form an infinite linearly independent set in the vector space $K[X]$ over K.

Now let us give a very useful practical result on linear dependence.

Theorem 2.5.5 *Let $n \in \mathbb{N}$, $n \geq 2$.*
(i) Two vectors in the canonical vector space K^n are linearly dependent $\Longleftrightarrow$ their components are respectively proportional.
(ii) n vectors in the canonical vector space K^n are linearly dependent $\Longleftrightarrow$ the determinant consisting of their components is zero.

Proof. (*i*) Let $v = (x_1, \dots, x_n)$, $v' = (x'_1, \dots, x'_n) \in K^n$. By Theorem 2.5.3, the vectors v and v' are linearly dependent if and only if one of them is a linear combination of the other, say $v' = kv$ for some $k \in K$. That is, $x'_i = kx_i$ for each $i \in \{1, \dots, n\}$.

(*ii*) Let $v_1 = (x_{11}, x_{21}, \dots, x_{n1})$, $\dots$, $v_n = (x_{1n}, x_{2n}, \dots, x_{nn}) \in K^n$. The vectors $v_1, \dots, v_n$ are linearly dependent if and only if $\exists k_1, \dots, k_n \in K$ not all zero such that

$$k_1v_1 + \cdots + k_nv_n = 0\,.$$

But this is equivalent to

$$k_1(x_{11}, x_{21}, \dots, x_{n1}) + \cdots + k_n(x_{1n}, x_{2n}, \dots, x_{nn}) = (0, \dots, 0)\,,$$

and further to

$$\begin{cases} k_1x_{11} + k_2x_{12} + \cdots + k_nx_{1n} = 0 \\ k_1x_{21} + k_2x_{22} + \cdots + k_nx_{2n} = 0 \\ \dots\dots\dots\dots\dots\dots \\ k_1x_{n1} + k_2x_{n2} + \cdots + k_nx_{nn} = 0\,. \end{cases}$$

We are interested in the existence of a non-zero solution for this homogeneous linear system. We will see later on that such a solution does exist if and only if the determinant of the system is zero. □

2.6 Bases

We are going to define a key notion related to a vector space, namely that of a *basis*, which will perfectly determine a vector space. For the sake of simplicity and because of our limited needs, til the end of the chapter, *by a vector space we will understand a finitely generated vector space.*

Definition 2.6.1 Let V be a vector space over K. A list of vectors $B = (v_1, \dots, v_n) \in V^n$ is called a *basis* of V if:
(i) B is linearly independent in V;
(ii) B is a system of generators for V, that is, $\langle B\rangle = V$.

Theorem 2.6.2 *Every vector space has a basis.*

Proof. Let V be a vector space over K. If $V = \{0\}$, then it has the basis $\emptyset$.

Now let $V = \langle B\rangle \neq \{0\}$, where $B = (v_1, \dots, v_n)$. If B is linearly independent, then B is a basis and we are done. Suppose that the list B is linearly dependent. Then by Theorem 2.5.3, $\exists j_1 \in \{1, \dots, n\}$ such that

$$v_{j_1} = \sum_{\substack{i=1\\ i\neq j_1}}^{n} k_iv_i$$

for some $k_i \in K$. It follows that $V = \langle B \setminus \{v_{j_1}\}\rangle$, because every vector of V can be written as a linear combination of the vectors of $B \setminus \{v_{j_1}\}$. If $B \setminus \{v_{j_1}\}$ is linearly independent, it is a basis and we are done. Otherwise, $\exists j_2 \in \{1, \dots, n\} \setminus \{j_1\}$ such that

$$v_{j_2} = \sum_{\substack{i=1\\ i\neq j_1,j_2}}^{n} k'_iv_i$$

for some $k'_i \in K$. It follows that $V = \langle B \setminus \{v_{j_1}, v_{j_2}\}\rangle$, because every vector of V can be written as a linear combination of the vectors of $B \setminus \{v_{j_1}, v_{j_2}\}$. If $B \setminus \{v_{j_1}, v_{j_2}\}$ is linearly independent, then it is a basis and we are done. Otherwise, we continue the procedure. If all the previous intermediate subsets are linearly dependent, we get to the step

$$V = \langle B \setminus \{v_{j_1}, \dots, v_{j_{n-1}}\}\rangle = \langle v_{j_n}\rangle.$$

If v_{j_n} were linearly dependent, then $v_{j_n} = 0$, hence $V = \langle v_{j_n}\rangle = \{0\}$, contradiction. Hence v_{j_n} is linearly independent and thus forms a single element basis of V. □

Remark 2.6.3 We are going to see that a vector space may have more than one basis.

Let us give now a characterization theorem for a basis of a vector space.

Theorem 2.6.4 *Let V be a vector space over K. A list $B = (v_1, \dots, v_n)$ of vectors in V is a basis of V if and only if every vector $v \in V$ can be uniquely written as a linear combination of the vectors $v_1, \dots, v_n$, that is,*

$$v = k_1v_1 + \cdots + k_nv_n$$

for some unique $k_1, \dots, k_n \in K$.

Proof. $\boxed{\Longrightarrow}$ Assume that B is a basis of V. Hence B is linearly independent and $\langle B\rangle = V$. The second condition assures us that every vector $v \in V$ can be written as a linear combination of the vectors of B. Suppose now that $v = k_1v_1 + \cdots + k_nv_n$ and $v = k_1'v_1 + \cdots + k_n'v_n$ for some $k_1, \dots, k_n, k_1', \dots, k_n' \in K$. It follows that

$$(k_1 - k_1')v_1 + \cdots + (k_n - k_n')v_n = 0\,.$$

By the linear independence of B, we must have $k_i = k_i'$ for each $i \in \{1, \dots, n\}$. Thus, we have proved the uniqueness of writing.

$\boxed{\Longleftarrow}$ Assume that every vector $v \in V$ can be uniquely written as a linear combination of the vectors of B. Then clearly, $V = \langle B\rangle$. For $k_1, \dots, k_n \in K$, we have by the uniqueness of writing

$$k_1v_1 + \cdots + k_nv_n = 0 \Longrightarrow k_1v_1 + \cdots + k_nv_n = 0 \cdot v_1 + \cdots + 0 \cdot v_n \Longrightarrow$$

$$\Longrightarrow k_1 = \cdots = k_n = 0\,,$$

hence B is linearly independent. Consequently, B is a basis of V. □

Definition 2.6.5 Let V be a vector space over K, $B = (v_1, \dots, v_n)$ a basis of V and $v \in V$. Then the scalars $k_1, \dots, k_n \in K$ appearing in the unique writing of v as a linear combination

$$v = k_1v_1 + \cdots + k_nv_n$$

of the vectors of B are called the *coordinates of v in the basis B*.

Example 2.6.6 (a) If K is a field and $n \in \mathbb{N}^*$, then the list $E = (e_1, \dots, e_n)$ of vectors of K^n, where

$$\begin{cases} e_1 = (1, 0, 0, \dots, 0) \\ e_2 = (0, 1, 0, \dots, 0) \\ \dots\dots\dots \\ e_n = (0, 0, 0, \dots, 1) \end{cases}$$

is a basis of the canonical vector space K^n over K, called the *canonical basis* (or *standard basis*). Indeed, each vector $v = (x_1, \dots, x_n) \in K^n$ has a unique writing $v = x_1e_1 + \cdots + x_ne_n$ as a linear combination of the vectors of E, hence E is a basis of V by Theorem 2.6.4.

Notice that the coordinates of a vector in the canonical basis are just the components of that vector, fact that is not true in general.

In particular, the canonical vector space $\mathbb{Z}_2^n$ over $\mathbb{Z}_2$ has the above canonical basis $E = (e_1, \dots, e_n)$, where 0 and 1 are just the elements $\widehat{0}$ and $\widehat{1}$ of $\mathbb{Z}_2$.

Also, if $n = 1$, the set $\{1\}$ is a basis of the canonical vector space K over K. For instance, $\{1\}$ is a basis of the vector space $\mathbb{C}$ over $\mathbb{C}$.

(b) Consider the canonical real vector space $\mathbb{R}^2$. We already know a basis of $\mathbb{R}^2$, namely the canonical basis $((1, 0), (0, 1))$. But it is easy to show that the list $((1, 1), (0, 1))$ is also a basis of $\mathbb{R}^2$. Therefore, a vector space may have more than one basis.

Also, note that $\{e_1\}$ is linearly independent, but not a system of generators, while the list $(e_1, e_2, e_1 + e_2)$ is a system of generators, but not linearly independent. Hence none of the two lists is a basis of the canonical real vector space $\mathbb{R}^2$.

(c) Let V_3 be the real vector space of all vectors (in the classical sense) in the space with a fixed origin O. Then a basis of V_3 consists of the three pairwise orthogonal *unit vectors* $\overrightarrow{i}$, $\overrightarrow{j}$, $\overrightarrow{k}$.

(d) Let K be a field and $n \in \mathbb{N}$. Then the list

$$E = (1, X, X^2, \dots, X^n)$$

is a basis of the vector space $K_n[X] = \{f \in K[X] \mid \text{degree}(f) \leq n\}$ over K, because every vector (polynomial) $f \in K_n[X]$ can be uniquely written as a linear combination $a_0 \cdot 1 + a_1 \cdot X + \cdots + a_n \cdot X^n$ $(a_0, \dots, a_n \in K)$ of the vectors of E (see Theorem 2.6.4).

In this case, the coordinates of a vector $f \in K_n[X]$ in the basis B are just its coefficients as a polynomial.

(e) Consider the real vector space $\mathbb{R}_2[X] = \{f \in \mathbb{R}[X] \mid \text{degree}(f) \leq 2\}$. We have seen that the list $E = (1, X, X^2)$ is a basis of $\mathbb{R}_2[X]$. Let us show that the list

$$B = (1, X-1, (X-1)^2)$$

is also a basis of $\mathbb{R}_2[X]$. Let $g = a_0 + a_1X + a_2X^2 \in \mathbb{R}_2[X]$. We look for unique $k_1, k_2, k_3 \in \mathbb{R}$ such that

$$g = k_1 \cdot 1 + k_2 \cdot (X-1) + k_3 \cdot (X-1)^2.$$

The equality is equivalent to the linear system of equations

$$\begin{cases} k_1 - k_2 + k_3 & = a_0 \\ \qquad k_2 - 2k_3 & = a_1 \\ \qquad\qquad k_3 & = a_2 \end{cases}$$

which has the unique solution $k_1 = a_0 + a_1 + a_2$, $k_2 = a_1 + 2a_2$, $k_3 = a_2$. Hence B is a basis of $\mathbb{R}_2[X]$, and the coordinates of a vector $g = a_0 + a_1X + a_2X^2 \in \mathbb{R}_2[X]$ in the basis B are $a_0 + a_1 + a_2$, $a_1 + 2a_2$, a_2.

(f) Let K be a field. The list

$$E = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

is a basis of the vector space $M_2(K)$ over K.

More generally, let $m, n \in \mathbb{N}$, $m, n \geq 2$ and consider the matrices $E_{ij} = (a_{kl})$, where

$$a_{kl} = \begin{cases} 1 & \text{if } k = i \text{ and } l = j \\ 0 & \text{otherwise} \end{cases}.$$

Then the list consisting of all matrices E_{ij} is a basis of the vector space $M_{m,n}(K)$ over K.

In this case, the coordinates of a vector $A \in M_{m,n}(K)$ in the above basis are just the entries of that matrix.

(g) Consider the real vector space $M_2(\mathbb{R})$. We have seen that

$$E = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

is a basis of $M_2(\mathbb{R})$. Let us show that the list

$$B = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right)$$

is also a basis of $M_2(\mathbb{R})$. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$. We look for unique $k_1, k_2, k_3, k_4 \in \mathbb{R}$ such that

$$A = k_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + k_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + k_3 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + k_4 \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

The equality is equivalent to the linear system of equations

$$\begin{cases} k_1 + k_2 + k_3 & = a \\ \qquad\qquad\quad k_4 & = b \\ \qquad\quad k_3 & = c \\ \quad k_2 + \qquad k_4 & = d \end{cases}$$

which has the unique solution

$$\begin{cases} k_1 & = a - d + b - c \\ k_2 & = d - b \\ k_3 & = c \\ k_4 & = b \end{cases} .$$

Hence B is a basis of $M_2(\mathbb{R})$, and the coordinates of a vector $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ in the basis B are $a - d + b - c$, $d - b$, c, b.

(h) Since $\forall z \in \mathbb{C}$, $\exists! x, y \in \mathbb{R}$ such that $z = x \cdot 1 + y \cdot i$, the list $B = (1, i)$ is a basis of the vector space $\mathbb{C}$ over $\mathbb{R}$.

The coordinates of a vector $z \in \mathbb{C}$ in the basis B are just its real and its imaginary part.

Theorem 2.6.7 *Let $f : V \to V'$ be a K-linear map and let $B = (v_1, \ldots, v_n)$ be a basis of V. Then f is determined by its values on the vectors of the basis B.*

Proof. Let $v \in V$. Since B is a basis of V, $\exists! k_1, \ldots, k_n \in K$ such that $v = k_1 v_1 + \cdots + k_n v_n$. Then

$$f(v) = f(k_1 v_1 + \cdots + k_n v_n) = k_1 f(v_1) + \cdots + k_n f(v_n) \,,$$

that is, f is determined by $f(v_1)$, ..., $f(v_n)$. □

Corollary 2.6.8 *Let $f, g : V \to V'$ be K-linear maps and let $B = (v_1, \ldots, v_n)$ be a basis of V. If $f(v_i) = g(v_i)$, $\forall i \in \{1, \ldots, n\}$, then $f = g$.*

Proof. Let $v \in V$. Then $v = k_1 v_1 + \cdots + k_n v_n$ for some $k_1, \ldots, k_n \in K$, hence

$$f(v) = f(k_1 v_1 + \cdots + k_n v_n) = k_1 f(v_1) + \cdots + k_n f(v_n) = k_1 g(v_1) + \cdots + k_n g(v_n) = g(v) \,.$$

Therefore, $f = g$. □

Theorem 2.6.9 *Let $f : V \to V'$ be a K-linear map, and let $X = (v_1, \ldots, v_n)$ be a list of vectors in V.*
(i) If f is injective and X is linearly independent in V, then $f(X)$ is linearly independent in V'.
(ii) If f is surjective and X is a system of generators for V, then $f(X)$ is a system of generators for V'.
(iii) If f is bijective and X is a basis of V, then $f(X)$ is a basis of V'.

Proof. We have $f(X) = (f(v_1), \ldots, f(v_n))$.

(i) Let $k_1, \ldots, k_n \in K$ be such that

$$k_1 f(v_1) + \cdots + k_n f(v_n) = 0'.$$

Since f is a K-linear map, it follows that

$$f(k_1 v_1 + \cdots + k_n v_n) = f(0),$$

whence by the injectivity of f we get

$$k_1 v_1 + \cdots + k_n v_n = 0.$$

But since X is linearly independent in V, we have $k_1 = \cdots = k_n = 0$. Hence $f(X)$ is linearly independent in V'.

(ii) Since X is a system of generators for V, we have $\langle X \rangle = V$. By the surjectivity of f we have:

$$\langle f(X) \rangle = f(\langle X \rangle) = f(V) = V' \,,$$

that is, $f(X)$ is a system of generators for V'.

(iii) This follows by (i) and (ii). □

Extra: Lossy compression

Following [Klein], we present a way of achieving lossy compression of images.

Definition 2.6.10 Let $k, n \in \mathbb{N}^*$ be such that $k < n$, and let u be a vector of the canonical vector space K^n over K. Then the *closest k-sparse* vector associated to u is defined as the vector obtained from u by replacing all but its k largest magnitude components by zero.

Example 2.6.11 Consider an image consisting of a single row of four pixels with intensities 200, 50, 200 and 75 respectively. We know that such an image can be viewed as a vector $u = (200, 50, 200, 75)$ in the real canonical vector space $\mathbb{R}^4$. The closest 2-sparse vector associated to u is the vector $\tilde{u} = (200, 0, 200, 0)$.

Suppose that we need to store a grayscale image of (say) $n = 2000 \times 1000$ pixels more compactly. We can view it as a vector v in the real canonical vector space $\mathbb{R}^n$. If we just store its associated closest k-sparse vector, then the compressed image may be far from the original.

One may use the following *lossy compression algorithm*:

Step 1. *Consider a* suitable *basis $B = (v_1, \dots, v_n)$ of the real canonical vector space $\mathbb{R}^n$.*

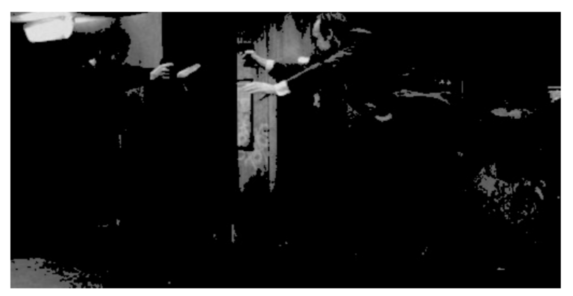
Step 2. *Determine the n-tuple u (which is desired to have as many zeros as possible) of the coordinates of v in the basis B.*

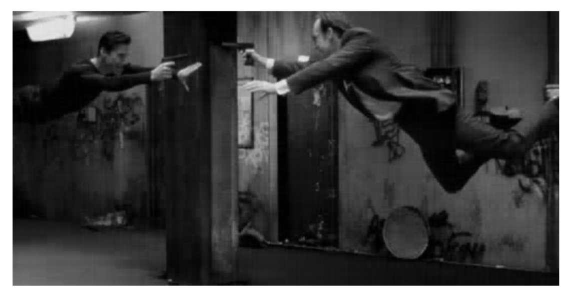
Step 3. *Replace u by the closest k-sparse n-tuple $\tilde{u}$ for a* suitable *k, and store $\tilde{u}$.*

Step 4. *In order to recover an image from $\tilde{u}$, compute the corresponding linear combination of the vectors of B with scalars the components of $\tilde{u}$.*

Consider the following image:

First, use the closest sparse vector which supresses all but 10% of the components of v, and secondly, use the lossy compression algorithm which supresses all but 10% of the components of u in order to get the following images respectively:





Course 6

2.7 Dimension

Recall that we consider only finitely generated vector spaces. Let us begin with a very useful lemma, that will be often implicitly used.

Lemma 2.7.1 *Let V be a vector space over K and let $Y = \langle y_1, \dots, y_n, z\rangle$. If $z \in \langle y_1, \dots, y_n\rangle$, then $Y = \langle y_1, \dots, y_n\rangle$.*

Proof. The generated subspace Y is the set of all linear combinations of the vectors $y_1, \dots, y_n, z$. Since $z \in \langle y_1, \dots, y_n\rangle$, z is a linear combination of the vectors $y_1, \dots, y_n$. It follows that every vector in Y can be written as a linear combination only of the vectors $y_1, \dots, y_n$. Consequently, $Y = \langle y_1, \dots, y_n\rangle$. □

The following result is a key theorem for proving that any two bases of a vector space have the same number of elements. But it is worth mentioning that it has a much broader importance in Linear Algebra.

Theorem 2.7.2 (Steinitz Theorem, Exchange Theorem) *Let V be a vector space over K, $X = (x_1, \dots, x_m)$ a linearly independent list of vectors of V and $Y = (y_1, \dots, y_n)$ a system of generators of V. Then:*
(i) $m \leq n$.
(ii) m vectors of Y can be replaced by the vectors of X obtaining again a system of generators for V.

Proof. We prove this result by induction on m.

The first step is to check it for $m = 1$. Then clearly $m \leq n$. Since Y is a system of generators for V, we have $x_1 = \sum_{i=1}^{n} k_i y_i$ for some $k_1, \dots, k_n \in K$. The list $X = \{x_1\}$ is linearly independent, hence $x_1 \neq 0$. It follows that $\exists j \in \{1, \dots, n\}$ such that $k_j \neq 0$. Then

$$y_j = k_j^{-1} x_1 - \sum_{\substack{i=1\\ i\neq j}}^{n} k_j^{-1} k_i y_i\,,$$

that is, y_j is a linear combination of the vectors $y_1, \dots, y_{j-1}, x_1, y_{j+1}, \dots, y_n$. Hence, in any linear combination of $y_1, \dots, y_n$, the vector y_j can be expressed as a linear combination of the other vectors and x_1. Therefore, we have

$$V = \langle y_1, \dots, y_n\rangle = \langle y_1, \dots, y_{j-1}, x_1, y_{j+1}, \dots, y_n\rangle\,.$$

Thus, we have obtained again a system of n generators for V containing x_1.

Let us now move on to the second step of the induction. We suppose the conclusion is true for $m-1$ and prove it for m. Let $X = (x_1, \dots, x_m)$ be a linearly independent list in V. Then $(x_1, \dots, x_{m-1})$ must be also linearly independent in V. By the induction hypothesis, we have $m - 1 \leq n$ and, after a renumbering,

$$V = \langle x_1, \dots, x_{m-1}, y_m, \dots, y_n\rangle\,.$$

If $m - 1 = n$, then $V = \langle x_1, \dots, x_{m-1}\rangle$, whence it follows that $x_m \in \langle x_1, \dots, x_{m-1}\rangle$, which contradicts the fact that X is linearly independent in V. Thus $m - 1 < n$, so that $m \leq n$.

We have $x_m \in V = \langle x_1, \dots, x_{m-1}, y_m, \dots, y_n\rangle$, whence

$$x_m = \sum_{i=1}^{m-1} k_i x_i + \sum_{i=m}^{n} k_i y_i$$

for some $k_1, \dots, k_n \in K$. The list X being linearly independent in V, it follows that $\exists m \leq j \leq n$ such that $k_j \neq 0$ (otherwise, $x_m = \sum_{i=1}^{m-1} k_i x_i$ and the list X would be linearly dependent in V). For simplicity of writing, assume that $j = m$. It follows that

$$y_m = k_m^{-1} x_m - \sum_{i=1}^{m-1} k_m^{-1} k_i x_i - \sum_{i=m+1}^{n} k_m^{-1} k_i y_i\,.$$

Thus, y_m is a linear combination of the vectors $x_1, \dots, x_m, y_{m+1}, \dots, y_n$, that is, we have $y_m \in \langle x_1, \dots, x_m, y_{m+1}, \dots, y_n \rangle$. Therefore, it follows that

$$V = \langle x_1, \dots, x_{m-1}, y_m, \dots, y_n \rangle = \langle x_1, \dots, x_m, y_{m+1}, \dots, y_n \rangle .$$

Thus, we have obtained again a system of generators for V, where m vectors of the list Y have been replaced by the vectors of the list X. This completes the proof. □

Remark 2.7.3 Let us point out that in Steinitz Theorem not necessarily the first m vectors of Y can be replaced by the m vectors of X.

Theorem 2.7.4 *Any two bases of a vector space have the same number of elements.*

Proof. Let V be a vector space over K and let $B = (v_1, \dots, v_m)$ and $B' = (v'_1, \dots, v'_n)$ be bases of V. Since B is linearly independent in V and B' is a system of generators for V, we have $m \leq n$ by Theorem 2.7.2. Since B is a system of generators for V and B' is linearly independent in V, we have $n \leq m$ by the same Theorem 2.7.2. Hence $m = n$. □

Definition 2.7.5 Let V be a vector space over K. Then the number of elements of any of its bases is called the *dimension of* V and is denoted by $\dim_K V$ or simply by $\dim V$.

Remark 2.7.6 If $V = \{0\}$, then V has the basis $\emptyset$ and $\dim V = 0$.

Example 2.7.7 Using the examples of bases given in the previous section, one can easily determine the dimension of each of those vector spaces.

(a) Let K be a field and $n \in \mathbb{N}^*$. Then $\dim_K K^n = n$.

(b) We have seen that the subspaces of $\mathbb{R}^3$ are $\{(0,0,0)\}$, any line containing the origin, any plane containing the origin and $\mathbb{R}^3$. Their dimensions are 0, 1, 2 and 3 respectively.

(c) Let K be a field and $n \in \mathbb{N}$. Then $\dim K_n[X] = n + 1$.

(d) Let K be a field. Then $\dim M_2(K) = 4$.
More generally, if $m, n \in \mathbb{N}$, $m, n \geq 2$, then $\dim M_{m,n}(K) = m \cdot n$.

(e) Consider the subspace

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x - y - z = 0\}$$

of the canonical real vector space $\mathbb{R}^3$. We have seen that $S = \langle (1,1,0), (1,0,1) \rangle$. Since the vectors $(1,1,0)$ and $(1,0,1)$ are linearly independent, it follows that $B = ((1,1,0), (1,0,1))$ is a basis of S. Hence $\dim S = 2$.

(f) We have $\dim_{\mathbb{C}}\mathbb{C} = 1$ and $\dim_{\mathbb{R}}\mathbb{C} = 2$.

Theorem 2.7.8 *Let V be a vector space over K. Then the following statements are equivalent:*
(i) $\dim V = n$.
(ii) The maximum number of linearly independent vectors in V is n.
(iii) The minimum number of generators for V is n.

Proof. $(i) \implies (ii)$ Assume that $\dim V = n$. Let $B = (v_1, \dots, v_n)$ be a basis of V. Then B is a list of n linearly independent vectors in V. Since B is a system of generators for V, any linearly independent list in V must have at most n elements by Theorem 2.7.2.

$(ii) \implies (i)$ Assume (ii). Let $B = (v_1, \dots, v_m)$ be a basis of V and let $(u_1, \dots, u_n)$ be a linearly independent list in V. Since B is linearly independent, we have $m \leq n$ by hypothesis. Since B is a system of generators for V, we have $n \leq m$ by Theorem 2.7.2. Hence $m = n$ and consequently $\dim V = n$.

$(i) \implies (iii)$ Assume that $\dim V = n$. Let $B = (v_1, \dots, v_n)$ be a basis of V. Then B is a system of n generators for V. Since B is a linearly independent list in V, any system of generators for V must have at least n elements by Theorem 2.7.2.

(iii) $\Longrightarrow$ *(i)* Assume *(iii)*. Let $B = (v_1, \dots, v_m)$ be a basis of V and let $(u_1, \dots, u_n)$ be a system of generators for V. Since B is a system of generators for V, we have $n \leq m$ by hypothesis. Since B is linearly independent, we have $m \leq n$ by Theorem 2.7.2. Hence $m = n$ and consequently $\dim V = n$. □

Theorem 2.7.9 *Let V be a vector space over K with $\dim V = n$ and $X = (u_1, \dots, u_n)$ a list of vectors in V. Then*

$$X \text{ is linearly independent in } V \Longleftrightarrow X \text{ is a system of generators for } V\,.$$

Proof. Let $B = (v_1, \dots, v_n)$ be a basis of V.

$\boxed{\Longrightarrow}$ Assume that X is linearly independent. Since B is a system of generators for V, we know by Theorem 2.7.2 that n vectors of B, that is, all the vectors of B, can be replaced by the vectors of X and we get another system of generators for V. Hence $\langle X \rangle = V$. Thus, X is a system of generators for V.

$\boxed{\Longleftarrow}$ Assume that X is a system of generators for V. Suppose that X is linearly dependent. Then $\exists j \in \{1, \dots, n\}$ such that

$$u_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} k_i u_i$$

for some $k_i \in K$. It follows that

$$V = \langle X \rangle = \langle u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n \rangle\,.$$

But the minimum number of generators for V is n by Theorem 2.7.8, which is a contradiction. Therefore, X is linearly independent. □

Corollary 2.7.10 *Let $n \in \mathbb{N}$, $n \geq 2$. Then n vectors in K^n form a basis of the canonical vector space K^n if and only if the determinant consisting of their components is non-zero.*

Proof. We have seen that n vectors in K^n are linearly independent if and only if the determinant consisting of their components is non-zero. But if this happens, then using the fact that $\dim_K K^n = n$ and Theorem 2.7.9, the vectors are also a system of generators, and thus a basis of K^n. □

Theorem 2.7.11 *Any linearly independent list of vectors in a vector space can be completed to a basis of the vector space.*

Proof. Let V be a vector space over K. Let $B = (v_1, \dots, v_n)$ be a basis of V and let $(u_1, \dots, u_m)$ be a linearly independent list in V. Since B is a system of generators for V, we know by Theorem 2.7.2 that $m \leq n$ and m vectors of B can be replaced by the vectors $(u_1, \dots, u_m)$ obtaining again a system of generators for V, say $(u_1, \dots, u_m, v_{m+1}, \dots, v_n)$. But by Theorem 2.7.9, this is also linearly independent in V and consequently a basis of V. □

Remark 2.7.12 The completion of a linearly independent list to a basis of the vector space is not unique.

Example 2.7.13 The list (e_1, e_2), where $e_1 = (1, 0, 0)$ and $e_2 = (0, 1, 0)$, is linearly independent in the canonical real vector space $\mathbb{R}^3$. It can be completed to the canonical basis of the space, namely (e_1, e_2, e_3), where $e_3 = (0, 0, 1)$. On the other hand, since $\dim_{\mathbb{R}} \mathbb{R}^3 = 3$, in order to obtain a basis of the space it is enough to add to our list a vector v_3 such that (e_1, e_2, v_3) is linearly independent (see Theorem 2.7.9). For instance, we may take $v_3 = (1, 1, 1)$, since the determinant consisting of the components of the three vectors is non-zero.

Corollary 2.7.14 *Let V be a vector space over K and $S \leq V$. Then:*
(i) Any basis of S is a part of a basis of V.
(ii) $\dim S \leq \dim V$.
(iii) $\dim S = \dim V \Longleftrightarrow S = V$.

Proof. (i) Let $(u_1,\dots,u_m)$ be a basis of S. Since the list is linearly independent, it can be completed to a basis $(u_1,\dots,u_m,v_{m+1},\dots,v_n)$ of V by Theorem 2.7.11.

(ii) It follows by (i).

(iii) Assume that $\dim S = \dim V = n$. Let $(u_1,\dots,u_n)$ be a basis of S. Then it is linearly independent in V, hence it is a basis of V by Theorem 2.7.9. Thus, if $v\in V$, then $v = k_1u_1+\dots+k_nu_n$ for some $k_1,\dots,k_n\in K$, hence $v\in S$. Therefore, $S=V$. □

Theorem 2.7.15 *Let V be a vector space over K and let $S\leq V$. Then there exists $\overline{S}\leq V$ such that $V = S\oplus\overline{S}$. In particular,*

$$\dim V = \dim S + \dim\overline{S}.$$

Proof. Let $(u_1,\dots,u_m)$ be a basis of S. Then by Corollary 2.7.14, it can be completed to a basis $B=(u_1,\dots,u_m,v_{m+1},\dots,v_n)$ of V. We consider

$$\overline{S} = \langle v_{m+1},\dots,v_n\rangle$$

and we prove that $V = S\oplus\overline{S}$.

Let $v\in V$. Then

$$v = \sum_{i=1}^{m} k_iu_i + \sum_{i=m+1}^{n} k_iv_i \in S+\overline{S}\,,$$

for some $k_1,\dots,k_n\in K$. Hence $V=S+\overline{S}$.

Now let $v\in S\cap\overline{S}$. Then

$$v = \sum_{i=1}^{m} k_iu_i = \sum_{i=m+1}^{n} k_iv_i\,,$$

for some $k_1,\dots,k_n\in K$. Hence

$$\sum_{i=1}^{m} k_iu_i - \sum_{i=m+1}^{n} k_iv_i = 0\,,$$

whence $k_i=0$, $\forall i\in\{1,\dots,n\}$, because B is a basis. Thus, $v=0$ and $S\cap\overline{S}=\{0\}$.

Therefore, $V=S\oplus\overline{S}$. □

Remark 2.7.16 This is an extremely important property of a vector space, that allows us to split it in "smaller" subspaces, that can be studied much easier and then to use that information to get information about the entire vector space.

Definition 2.7.17 Let V be a vector space over K and $S\leq V$. Then a subspace $\overline{S}$ of V such that

$$V = S\oplus\overline{S}$$

is called a *complement of S in V*.

Remark 2.7.18 A subspace of a vector space may have more than one complement (see also the remark following Theorem 2.7.11).

Example 2.7.19 Consider the subspace $S=\langle e_1,e_2\rangle$ of the canonical real vector space $\mathbb{R}^3$, where $e_1=(1,0,0)$, $e_2=(0,1,0)$. Then clearly (e_1,e_2) is a basis of S. Now by Example 2.7.13, it can be completed to a basis of $\mathbb{R}^3$, with the vector $e_3=(0,0,1)$ or with the vector $v_3=(1,1,1)$. Following the proof of Theorem 2.7.15, a complement in V of the subspace $S=\langle e_1,e_2\rangle$ is $\langle e_3\rangle$ or $\langle v_3\rangle$.

2.8 Dimension theorems

Theorem 2.8.1 *Let V and V' be vector spaces over K. Then*

$$V \simeq V' \Longleftrightarrow \dim V = \dim V' .$$

Proof. $\boxed{\Longrightarrow}$ Let $f : V \to V'$ be a K-isomorphism and let $B = (v_1, \dots, v_n)$ be a basis of V. Note that, since f is injective, we have $f(v_i) \neq f(v_j)$ for every $i, j \in \{1, \dots, n\}$ with $i \neq j$. Hence the list

$$B' = f(B) = (f(v_1), \dots, f(v_n))$$

has n elements. Then B' is a basis of V'. Now it follows that $\dim V = \dim V'$.

$\boxed{\Longleftarrow}$ Assume that $\dim V = \dim V' = n$. Let $B = (v_1, \dots, v_n)$ and $B' = (v'_1, \dots, v'_n)$ be bases of V and V' respectively. Define a function $f : V \to V'$ in the following way. For every $v = k_1v_1 + \dots + k_nv_n \in V$ (where $k_1, \dots, k_n \in K$ are uniquely determined), define

$$f(v) = k_1v'_1 + \dots + k_nv'_n.$$

Let us first prove that f is a K-linear map. Let $\alpha, \beta \in K$ and $v, w \in V$. Then $v = k_1v_1 + \dots + k_nv_n$ and $w = l_1v_1 + \dots + l_nv_n$ for some unique $k_1, \dots, k_n, l_1, \dots, l_n \in K$. It follows that

$$\begin{aligned} f(\alpha v + \beta w) &= f((\alpha k_1 + \beta l_1)v_1 + \dots + (\alpha k_n + \beta l_n)v_n) \\ &= (\alpha k_1 + \beta l_1)v'_1 + \dots + (\alpha k_n + \beta l_n)v'_n \\ &= \alpha(k_1v'_1 + \dots + k_nv'_n) + \beta(l_1v'_1 + \dots + l_nv'_n) \\ &= \alpha f(v) + \beta f(w). \end{aligned}$$

Hence f is a K-linear map. In particular, we have $f(v_i) = v'_i$ for every $i \in \{1, \dots, n\}$.

Now let us prove that f is bijective. Let $v' = k'_1v'_1 + \dots + k'_nv'_n \in V'$ (where $k'_1, \dots, k'_n \in K$ are uniquely determined). Using the fact that $f(v_i) = v'_i$ for every $i \in \{1, \dots, n\}$, it follows that

$$v' = k'_1f(v_1) + \dots + k'_nfv_n) = f(k'_1v_1 + \dots + k'_nv_n),$$

where the vector $k'_1v_1 + \dots + k'_nv_n \in V$ is uniquely determined. Hence f is bijective, and thus f is a K-isomorphism. □

We may immediately deduce the following result.

Theorem 2.8.2 *Any vector space V over K with $\dim V = n$ is isomorphic to the canonical vector space K^n over K.*

Remark 2.8.3 Theorem 2.8.2 is a very important structure theorem, saying that, up to an isomorphism, *any finite dimensional vector space over K is, in fact, the canonical vector space K^n over K.* For instance, we have the K-isomorphisms $K_n[X] \simeq K^{n+1}$ and $M_{m,n}(K) \simeq K^{mn}$. Now we have an explanation why we have used so often the canonical vector spaces: not only because the operations are very nice and easily defined, but they are, up to an isomorphism, the only types of finite dimensional vector spaces.

Definition 2.8.4 Let $f : V \to V'$ be a K-linear map. Then:
(1) $\dim(\mathrm{Ker} f)$ is called the *nullity* of f, and is denoted by $\mathrm{null}(f)$.
(2) $\dim(\mathrm{Im} f)$ is called the *rank* of f, and is denoted by $\mathrm{rank}(f)$.

Next we present an important theorem relating the nullity and the rank of a linear map.

Theorem 2.8.5 (First Dimension Theorem) *Let $f : V \to V'$ be a K-linear map. Then*

$$\dim V = \dim(\mathrm{Ker} f) + \dim(\mathrm{Im} f) .$$

In other words, $\dim V = \mathrm{null}(f) + \mathrm{rank}(f)$.

Theorem 2.8.6 (Second Dimension Theorem) *Let V be a vector space over K and let S, T be subspaces of V. Then*

$$\dim S + \dim T = \dim(S \cap T) + \dim(S + T)\,.$$

Corollary 2.8.7 *Let V be a vector space over K, and let S and T be subspaces of V such that $V = S \oplus T$. Then*

$$\dim V = \dim S + \dim T\,.$$

Extra: Checksum function

Following [Klein], we present a checksum function for detecting corrupted files.

Definition 2.8.8 Let $u = (x_1, \dots, x_n), v = (y_1, \dots, y_n) \in K^n$. Then the *dot-product* (or *scalar product*) of u and v is the scalar

$$u \cdot v = x_1 y_1 + \cdots + x_n y_n \in K.$$

Example 2.8.9 We give an example of a checksum function which may detect accidental random corruption of a file during transmission or storage.
Let $a_1, \dots, a_{64} \in \mathbb{Z}_2^n$ and let $f : \mathbb{Z}_2^n \to \mathbb{Z}_2^{64}$ be the $\mathbb{Z}_2$-linear map defined by

$$f(v) = (a_1 \cdot v, \dots, a_{64} \cdot v).$$

Suppose that v is a "file". We model corruption as the addition of a random vector $e \in \mathbb{Z}_2^n$ (the error), so the corrupted version of the file is $v + e$. We look for a formula for the probability that the corrupted file has the same checksum as the original file.
The checksum of the original file v is taken to be $f(v)$, hence the checksum of the corrupted file $v + e$ is $f(v + e)$. The original file and the corrupted file have the same checksum if and only if $f(v) = f(v + e)$ if and only if $f(e) = 0$ if and only if $e \in \operatorname{Ker} f$.
Every vector space V over the field $\mathbb{Z}_2$ with $\dim V = n$ is isomorphic to $\mathbb{Z}_2^n$, hence it has 2^n vectors. In particular, $\operatorname{Ker} f$ has 2^k vectors, where $k = \dim(\operatorname{Ker} f)$.
If the error is chosen according to the uniform distribution, the probability that $v + e$ has the same checksum as v is the following:

$$P = \frac{\text{number of vectors in } \operatorname{Ker} f}{\text{number of vectors in } \mathbb{Z}_2^n} = \frac{2^k}{2^n}.$$

One may show that $\dim(\operatorname{Im} f)$ is close to $\min(n, 64)$. So if we choose $n \geq 64$, we may assume that $\dim(\operatorname{Im} f) = 64$. By the First Dimension Theorem, we have

$$k = \dim(\operatorname{Ker} f) = \dim \mathbb{Z}_2^n - \dim(\operatorname{Im} f) = n - 64.$$

Hence

$$P = \frac{2^{n-64}}{2^n} = \frac{1}{2^{64}},$$

and thus there is only a very tiny chance that the change is undetected.