Topics in Linear Algebra and Geometry

Philippe Ellia

Alex Massarenti

Dipartimento di Matematica e Informatica, Università di Ferrara, Via Machiavelli 30, 44121 Ferrara, Italy

Email address: phe@unife.it

Dipartimento di Matematica e Informatica, Università di Ferrara, Via Machiavelli 30, 44121 Ferrara, Italy

Email address: msslxa@unife.it

2020 *Mathematics Subject Classification.* Primary 15-01; Secondary 51A05.

Contents

Preface

Starting from the principle that a well made brain is better than a well filled one, this book represents a selection of topics for a Linear Algebra course for Mathematics, Physics, and Engineering.

The first part is a sort of quick introduction to the basic notions of logic and algebra; in the second part, the fundamental concepts and results of linear algebra are introduced.

The third, fourth, and fifth parts, respectively on affine geometry, Euclidean geometry, and the Jordan form, can be seen as applications of the second part.

In the first parts, various informal connections and comparisons with projective geometry are made. These connections will be clarified and formalized in the sixth part in which some fundamental topics of projective geometry, heavily relying on notions of linear algebra covered in the previous parts, are developed.

The text contains 314 exercises, which are an essential part of the course. Indeed, we believe that learning theory without practical application is a sterile exercise.

Some texts, listed in the bibliography [**Hal74**], [**Lan87**], [**Ser00**], [**DF04**], were consulted during the preparation of this book and can provide interested readers with further insights into some of the covered topics.

Ferrara, May 2024

Philippe Ellia
Alex Massarenti

Part 1

Preliminaries

Among exact sciences *mathematics* is the only hypothetico-deductive science, physics and chemistry, for instance, are experimental sciences. Therefore, starting from statements which are assumed to be true (*axioms*) and proceeding by logical deductions (*proofs*), we obtain other true statements (*theorems*). Then we start again: using the axioms and theorems we prove new theorems.

Logical deductions are regulated by formal logic, axioms are not subject to any truth conditions, they only need to be logically compatible that is not contradictory. Therefore, there are different *axiomatic systems*, for instance mathematics with or without the *axiom of choice*. There is no need to go into details here, let us just say that the most used system is the *Zermelo-Fraenkel system* with the *uncountable axiom of choice*. This system, that will be called in what follows mathematics, responds well to our intuition and allows us to prove a great number of theorems.

However, we must recall that a famous result of *Gödel* states that *it can not be proved, with the tools of mathematics, that mathematics is not contradictory.* But do not worry, it has been going on like this for more than two thousand years, no one has found the contradiction, the theorems keep accumulating and the applications of mathematics are everywhere. Therefore we abandon the delicate questions relating to the foundations of mathematics and adopt the more down-to-earth insider point of view. From this perspective, a proof is exactly what the Ancient Greeks had already established: *a conscious effort to order the arguments in such a way that the transition from one step to the next leaves no doubt, so that a virtual interlocutor would be forced to agree.*

The formal presentation should not be deceiving. Mathematical activity consists of solving theoretical and practical problems, proving theorems and solving exercises. The various theories are not sterile formal constructions but serve to frame, in one fell swoop, a whole series of problems and therefore facilitate their understanding and resolution.

Furthermore, the proof of a theorem, or the resolution of an exercise, does not proceed in a rational, formal way, starting from the hypotheses and continuing automatically until reaching the thesis. Instead, it seems that three phases can be distinguished: two rational and one, the second and perhaps most important, non-rational. In the first phase (*analysis*) we try to understand the question, the problem (*thesis*). Once we have a good understanding of what we have to prove, we look at what we have available (*hypothesis*). At this point both the arrival point (*thesis*) and the starting point (*hypothesis*) are known. In the second phase, we will try to connect these two points. Often, but not always, our brain will proceed by analogy. Starting simultaneously from both the arrival point and the starting point, it will begin taking already known paths, like exercises already done and proofs already seen, until hopefully the connection is established between the thesis

and the hypothesis. At this stage we have an idea of proof. The third and final phase consists in verification and adjustment. We retrace the path in a rational way and, starting from the hypotheses and proceeding by deductions, perhaps fixing unclear passages, we reach the thesis. If a mistake comes up, do not be afraid to start over. The last phase is the drafting.

The final version must be concise and clear, without obscure passages. It is advisable to reread it with a critical eye and always remember that you write not just for yourself but also to communicate with others, and this should be done in such a way that a virtual interlocutor would be forced to agree. To achieve this skill, since this is something one learns, the student is recommended to practice solving the proposed exercises writing down complete solutions. Needless to say, it is more important to know how to solve the exercises than knowing the theory, also because it is often impossible to solve the exercises without knowing the theory.

In this first part we revise basic concepts of logic and set theory, and the main algebraic structure which will be frequently referred to in the rest of this book.

Chapter 1

Numbers

We denote the set of *natural numbers* by $\mathbb{N}$:

$$\mathbb{N} = \{0, 1, 2, \cdots, n, \cdots\}$$

and by $\mathbb{Z}$ the *integers*:

$$\mathbb{Z} = \{\cdots, -n, \cdots, -2, -1, 0, 1, 2, \cdots, -n, \cdots\}.$$

There is no easy way to define $\mathbb{N}$ (the existence of an infinite set is an axiom).

Rational numbers are fractions of integers:

$$\mathbb{Q} = \left\{ \frac{a}{b} \text{ such that } a, b \in \mathbb{Z} \text{ with } b \neq 0 \text{ and where } \frac{a}{b} = \frac{c}{d} \text{ if and only if } ad = bc \right\}.$$

For example, the fractions $\frac{1}{2}$ and $\frac{2}{4}$ are different but represent the same rational number 0.5. We will see later (equivalence relations) an algebraic way to construct $\mathbb{Q}$ from $\mathbb{Z}$.

There is no simple way to describe *real numbers* either. Morally, real numbers are all the points on the real line! (This is not a definition). There is a not-too-complicated process to construct $\mathbb{R}$ using $\mathbb{Q}$ (topological completion). This process clearly shows that $\mathbb{Q}$ is dense in $\mathbb{R}$ meaning that any non-empty real interval, no matter how small, contains a rational number (in fact, infinitely many rational numbers).

Complex (or *imaginary*) *numbers* were invented in the Middle Ages for solving cubic equations by Italian algebraists. Their problem was finding the square root of a negative number, which was seemingly impossible. They realized that if -1 were a square, it would solve the problem. So, they introduced the "imaginary quantity" i such that $i^2 = -1$. Now, $\sqrt{-5} = \sqrt{i^2 5} = \pm i\sqrt{5}$. The interesting part is that in their calculations, the imaginary quantity canceled out and did not appear in the

final formula. For them, the symbol i was just a trick to make the calculations work. Over time, imaginary numbers became complex numbers and are very useful, as we will see. We have

$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R},\, i^2 = -1\}.$$

Here too, there is a simple algebraic construction to build $\mathbb{C}$ from $\mathbb{R}$.

Summing up, we have

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

and if, as Kronecker said (look him up on Google), $\mathbb{N}$ is the work of God, once we have $\mathbb{N}$, it is not difficult, with mathematical constructions, to build all other numbers but the rest is the work of man.

We will also see other types of numbers (algebraic, transcendental, modular).

Chapter 2

Propositional logic

2.0.1. Logical Connectives. A proposition is a statement of the considered axiomatic theory. A proposition has a logical value that can be either T (true) or F (false). There are no other possibilities. Given a proposition $\mathcal{P}$, its negation, not $\mathcal{P}$ (or $\neg\mathcal{P}$), is false when $\mathcal{P}$ is true and true when $\mathcal{P}$ is false. This can be represented by the following truth table:

$\mathcal{P}$	$\neg\mathcal{P}$
T	F
F	T

Take two propositions $\mathcal{P}$ and $\mathcal{Q}$, using logical connectives it is possible to create others.

Definition 2.1 (*And connective*)**.** Given two propositions $\mathcal{P}$ and $\mathcal{Q}$, the proposition $\mathcal{P}$ and $\mathcal{Q}$ ($\mathcal{P} \wedge \mathcal{Q}$) is true if and only if both propositions are true.

Definition 2.2 (*Or connective*)**.** Given two propositions $\mathcal{P}$ and $\mathcal{Q}$, the proposition $\mathcal{P}$ or $\mathcal{Q}$ ($\mathcal{P} \vee \mathcal{Q}$) is true if and only if one of the two propositions is true.

Definition 2.3 (*Logical implication*)**.** Given two propositions $\mathcal{P}$ and $\mathcal{Q}$, the proposition $\neg\mathcal{P} \vee \mathcal{Q}$ (not $\mathcal{P}$ or $\mathcal{Q}$) is denoted by $\mathcal{P} \Rightarrow \mathcal{Q}$ ($\mathcal{P}$ implies $\mathcal{Q}$) and is called logical implication.

Definition 2.4. (*Logical equivalence*) Given two propositions $\mathcal{P}$ and $\mathcal{Q}$, the logical equivalence $\mathcal{P} \Leftrightarrow \mathcal{Q}$ is true if and only if $\mathcal{P}$ and $\mathcal{Q}$ have the same logical value.

One can verify (using truth tables) that $\mathcal{P} \Leftrightarrow \mathcal{Q}$ is the same as

$$(\mathcal{P} \Rightarrow \mathcal{Q}) \wedge (\mathcal{Q} \Rightarrow \mathcal{P}).$$

Proposition 2.5. *Given two propositions $\mathcal{P}$ and $\mathcal{Q}$, we have:*

(i) $\neg(\mathcal{P} \wedge \mathcal{Q}) \Leftrightarrow \neg\mathcal{P} \vee \neg\mathcal{Q}$;

(ii) $\neg(\mathcal{P} \vee \mathcal{Q}) \Leftrightarrow \neg\mathcal{P} \wedge \neg\mathcal{Q}$.

Proof. Let us prove the first statement. From the following truth table

$\mathcal{P}$	$\mathcal{Q}$	$\mathcal{P} \wedge \mathcal{Q}$	$\neg(\mathcal{P} \wedge \mathcal{Q})$	$\neg\mathcal{P}$	$\neg\mathcal{Q}$	$\neg\mathcal{P} \vee \neg\mathcal{Q}$
T	T	T	F	F	F	F
T	F	F	T	F	T	T
F	T	F	T	T	F	T
F	F	F	T	T	T	T

we see that $\neg(\mathcal{P} \wedge \mathcal{Q})$ and $\neg\mathcal{P} \vee \neg\mathcal{Q}$ are equivalent. The second claim can be proved similarly. □

2.0.2. Quantifiers. If $\mathcal{P}$ is a proposition that depends on x, it will be denoted by $\mathcal{P}(x)$ to highlight this dependence (x is a variable).

The symbol $\forall$ (*Universal quantifier*) means "for every" or "for all". So, "$\forall\, x \in \mathbb{R}, f(x) = x + 3$" means "for every x in $\mathbb{R}$, $f(x)$ equals $x + 3$".

The symbol $\exists$ (*Existential quantifier*) means "there exists". To show that the proposition $\exists\, x \mid \mathcal{P}(x)$ (there exists an x such that $\mathcal{P}(x)$ is true), one needs to find at least one x for which $\mathcal{P}(x)$ is true.

For example, to prove the proposition "$\exists\, x \mid 15$ divides x" one simply needs to exhibit a number which is divisible by 15, 15 itself for instance. Someone else could propose 30, and that would also be fine because 30 is divisible by 15.

Sometimes, $\exists!$ is used to indicate that the existing element is unique. For example, "$\exists!\, x \mid 1 < x < 5$ and 4 divides x". Indeed, 4 is the only integer between 1 and 5 that is divisible by 4.

The order in which quantifiers appear is very important. For example:

$$\exists\, e \mid \forall\, g,\ ge = eg = g$$

means there exists an e such that for all g, $ge = eg = g$. While

$$\forall\, g,\ \exists\, g^{-1} \mid gg^{-1} = g^{-1}g = e$$

means for every g, there exists a g^{-1} such that $gg^{-1} = g^{-1}g = e$. What is the difference? In the first case, e is the same for all g while in the second case g^{-1} depends on g.

The logical negation of $\forall\ x,\ \mathcal{P}(x)$ is $\exists\ x,\ \neg\mathcal{P}(x)$. Similarly, $\neg(\exists\ x,\ \mathcal{P}(x))$ is $\forall\ x,\ \neg\mathcal{P}(x)$. So, the logical negation of "all cars are red" is "there exists a non red car".

2.1. Axiomatic systems

Without going into too much detail, we can say that an axiomatic system consists of basic symbols, axioms, and proof rules.

Definition 2.6. A proposition $\mathcal{P}$ is proved (it is also said that $\mathcal{P}$ is true) if it can be included in a formalized text that starts with an axiom of the theory and develops according to the rules of the theory.

Definition 2.7. (*Decidability*) Suppose $\mathcal{P}$ is a proposition of an axiomatic theory $\mathfrak{T}$. If $\mathcal{P}$ is not true (that is $\mathcal{P}$ is not proved), it does not mean that $\mathcal{P}$ is false (which means that not $\mathcal{P}$ is true). It can happen that in $\mathfrak{T}$ it is not possible to prove that either $\mathcal{P}$ or not $\mathcal{P}$ is true. In this case, the proposition $\mathcal{P}$ is called *undecidable*.

Saying that a proposition in theory $\mathfrak{T}$ is undecidable means that the tools available in $\mathfrak{T}$ are insufficient to determine the logical value (true or false) of $\mathcal{P}$. If $\mathcal{P}$ is undecidable in $\mathfrak{T}$, we can consider the theory $\mathfrak{T}' = \mathfrak{T} + \mathcal{P}$, where the axioms are those of $\mathfrak{T}$ plus $\mathcal{P}$. The theory $\mathfrak{T}'$ is as consistent as $\mathfrak{T}$. In $\mathfrak{T}'$ the proposition $\mathcal{P}$ is true (it is an axiom). However, we can also consider the theory $\mathfrak{T}'' = T + \neg\mathcal{P}$, and $\mathfrak{T}''$ is as consistent as $\mathfrak{T}$, but in $\mathfrak{T}''$ the proposition $\mathcal{P}$ is false.

In set theory, there exist undecidable relations (such as the continuum hypothesis, which we will discuss later). In practice, we will not encounter undecidable relations, and if this were to happen, we would immediately make them into axioms.

Definition 2.8. (*Consistency*) If there exists a proposition $\mathcal{P}$ in a formalized theory $\mathfrak{T}$ such that both $\mathcal{P}$ and not $\mathcal{P}$ are true, $\mathfrak{T}$ is said to be contradictory (or inconsistent). In this case, every proposition in $\mathfrak{T}$ is both true and false, and the theory loses all significance.

Let us assume that both $\mathcal{P}$ and not $\mathcal{P}$ are true, which is a contradiction. Now, consider a proposition $\mathcal{Q}$. The proposition "not $\mathcal{P}$ implies ($\mathcal{Q}$ or not $\mathcal{P}$)" is true because not $\mathcal{P}$ is false (since $\mathcal{P}$ is true). Since not $\mathcal{P}$ is true, "$\mathcal{Q}$ or not $\mathcal{P}$" is true, which implies that $\mathcal{P}$ implies $\mathcal{Q}$ is true. Since $\mathcal{P}$ is true, $\mathcal{Q}$ is true. Similarly, by replacing $\mathcal{Q}$ with not $\mathcal{Q}$, we can show that not $\mathcal{Q}$ is true.

We are confident that mathematics is consistent, so we will reject any contradiction!

2.2. Methods of proof

Now, we will discuss four methods of proof: direct proof, indirect proof, proof by contradiction, and proof by induction.

2.2.1. Direct proof. Direct proof is based on the idea that if $\mathcal{P}$ is true and $\mathcal{P} \Rightarrow \mathcal{Q}$ is true, then $\mathcal{Q}$ is necessarily true.

Let us suppose we have our hypothesis $\mathcal{I}$ and we need to prove the thesis $\mathcal{T}$. Obviously, the hypothesis $\mathcal{I}$ is true. We must then prove that the logical implication $\mathcal{I} \Rightarrow \mathcal{T}$ is true. As observed, this will result in $\mathcal{T}$ being true. We can break the problem into easier problems:

$\mathcal{I} \Rightarrow \mathcal{P}_1$ (so $\mathcal{P}_1$ is true);
$\mathcal{P}_1 \Rightarrow \mathcal{P}_2$ (so $\mathcal{P}_2$ is true);
$\mathcal{P}_2 \Rightarrow \mathcal{P}_3$ (so $\mathcal{P}_3$ is true);
$\vdots$
$\mathcal{P}_k \Rightarrow \mathcal{T}$ (so $\mathcal{T}$ is true, and we have proved our theorem).

2.2.2. Indirect proof. First of all, we need to observe that $\mathcal{P} \Rightarrow \mathcal{Q}$ is logically equivalent to $\neg\mathcal{Q} \Rightarrow \neg\mathcal{P}$. This can be seen with truth tables.

Sometimes, one may want to prove a logical implication $\mathcal{P} \Rightarrow \mathcal{Q}$ that seems challenging to approach directly. In such cases, it is advisable to try to prove $\neg\mathcal{Q} \Rightarrow \neg\mathcal{P}$ instead.

Let us illustrate this with an example: Let $n \in \mathbb{N}$, and let $\mathcal{P}$ be "n^2 is even" and $\mathcal{Q}$ be "n is even". We want to prove $\mathcal{P} \Rightarrow \mathcal{Q}$. This is equivalent to prove that $\neg\mathcal{Q}$, that is n is odd, implies $\neg\mathcal{P}$, that is n^2 is odd. But this is straightforward. Let $n = 2t + 1$ be an odd number. Then $n^2 = 4t^2 + 1 + 4t = 2(2t^2 + 2t) + 1$ is also odd. Can anyone see a direct proof?

2.2.3. Proof by contradiction. We want to prove the proposition $\mathcal{P}$. In practice, the proof by contradiction proceeds as follows: Let us assume $\mathcal{P}$ is false (we are adding $\neg\mathcal{P}$ to the axioms of mathematics), we reason with this hypothesis until we identify a contradiction. Since we do not want any contradictions, we conclude by saying "But this is a contradiction, so $\mathcal{P}$ is true".

Here is an example: We want to prove that $\sqrt{2}$ is irrational. Suppose $\sqrt{2}$ is rational. Therefore, $\sqrt{2} = \frac{m}{n}$. We can assume that n and m are coprime that is their greatest common divisor is one. Indeed, if $m = km'$ and $n = kn'$ then $\frac{m}{n} = \frac{m'}{n'}$.

Squaring the equation we have $2n^2 = m^2$. So, m^2 is even. We have just seen that this implies m is even. Therefore, $m = 2t$ and $m^2 = 4t^2$. It follows that $2n^2 = 4t^2$ that is $n^2 = 2t^2$. Therefore, n^2 is even, so n is even. In conclusion, we have shown that 2 divides both m and n, but this is a contradiction because we had assumed that m and n are coprime. Therefore, $\sqrt{2}$ is irrational.

The fact that $\sqrt{2}$ is irrational was known in the time of Pythagoras. More generally, one can prove that if $d \in \mathbb{N}$ is not a perfect square then $\sqrt{d}$ is irrational.

Another famous example of a proof by contradiction is the proof that the set of prime numbers P is infinite. Recall that $p \in \mathbb{N}$ is a prime number if $p > 1$, and the only divisors of p are 1 and p that is $\mathrm{Div}(p) = \{1, p\}$, where $\mathrm{Div}(n)$ stands for the set of divisors of n. We have that

$$P = \{2, 3, 5, 7, 11, 13, 17, 19, \dots\}.$$

We observe that the divisors of n are symmetric with respect to $\sqrt{n}$. If $a|n$ (a divides n), then $n = ab$ (so b is also in $\mathrm{Div}(n)$), and if $a < b$, then $a < \sqrt{n}$, and $b > \sqrt{n}$. If $a < b \leq \sqrt{n}$, then $n = ab < n$, a contradiction. If $\sqrt{n} \geq a < b$, then $n = ab > n$. The only case in which equality can occur is when n is a perfect square, and $a = b = \sqrt{n}$. In particular, the number of elements (the "cardinality") of $\mathrm{Div}(n)$ is even unless n is a perfect square.

Lemma 2.9. *Let $n \in \mathbb{N}$, $n > 1$, and let $\mathrm{Div}(n) = \{1, d_2, d_3, \dots, d_r = n\}$ with $1 < d_2 < d_3 < \dots < d_r = n$. Then d_2 is a prime number. So, every $n > 1$ has a prime divisor.*

Proof. Let us assume, by contradiction, that d_2 is not prime. Then, there exist a and b such that $1 < a, b < d_2$ and $d_2 = ab$. But if a divides d_2, it also divides n which contradicts the fact that d_2 is the smallest divisor greater than 1 of n. Therefore, d_2 is prime. □

Theorem 2.10. *The set of prime numbers P is infinite.*

Proof. Let us assume, by contradiction, that P is finite: $P = \{p_1, p_2, \dots, p_k\}$. Let $N = p_1 p_2 \dots p_k + 1$. By Lemma 2.9, N has a prime divisor, which must be an element of P. Let us call this prime divisor p_i. We have:

$$\frac{N}{p_i} = \frac{p_1 p_2 \dots p_i \dots p_k}{p_i} + \frac{1}{p_i}.$$

In the first term, $\frac{p_1 p_2 \dots p_i \dots p_k}{p_i}$, p_i cancels out, and we get an integer (the product of all primes except p_i). By assumption, N/p_i is an integer. Therefore, $\frac{N}{p_i} - \frac{p_1 p_2 \dots p_i \dots p_k}{p_i}$ is an integer. But this is a contradiction since

$$\frac{N}{p_i} - \frac{p_1 p_2 \dots p_i \dots p_k}{p_i} = \frac{1}{p_i}$$

and $\frac{1}{p_i}$ is not an integer. □

This proof can be found in Euclid's elements. Do not confuse indirect proof with proof by contradiction. In both cases, the thesis is negated, but in a indirect proof there is no contradiction.

2.2.4. Proof by induction. The proof by contradiction is very powerful, but the proof by induction is the mathematician's atomic bomb!

Peano's axioms are an attempt to axiomatize the construction of natural numbers, the mathematician G. Peano formulated the following list of axioms. First, there are three primitive terms: number, zero, successor. The axioms are as follows:

(1) Every number has a successor.

(2) Zero is not the successor of any number.

(3) Two distinct numbers can not have the same successor.

(4) Let F be a set of numbers such that: Zero belongs to F, and if a number belongs to F, then its successor also belongs to F. Then, every number belongs to F.

If we represent the successor of x as $x+1$, axiom (4), known as the induction axiom or principle, can be reformulated as follows:

$$(0 \in F) \wedge (x \in F \Rightarrow x + 1 \in F) \Rightarrow F = \mathbb{N}.$$

Peano's axioms themselves do not define $\mathbb{N}$ as they use the term "set" without providing a formal definition of what a set is. However, the fourth axiom, known as the *Induction Principle* is one of the most powerful tools in mathematics (once one has successfully defined $\mathbb{N}$).

Let $\mathcal{P}(n)$ be a proposition that depends on the variable $n \in \mathbb{N}$. The Induction Principle allows one to prove $\mathcal{P}(n)$ by:

(1) Proving $\mathcal{P}(0)$.

(2) Proving that if $\mathcal{P}(n-1)$ is true then $\mathcal{P}(n)$ is true as well.

The first step is called the *initial case*. The second step is known as the *inductive step*.

Indeed, let us consider the set

$$X = \{m \in \mathbb{N} \mid \mathcal{P}(m) \text{ is true}\}.$$

According to (1), $0 \in X$. According to (2), if $x \in X$ then $x+1 \in X$. Using Peano's Axiom (4) (the induction principle), we conclude that $X = \mathbb{N}$. In other words, $\mathcal{P}(n)$ is true for every $n \in \mathbb{N}$, and the proposition is proved.

Let us make an example. We want to prove

$$\mathcal{P}(n) : \ 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}.$$

We have $\mathcal{P}(0) = 0$. This is true because $0 = \frac{0(0+1)}{2}$. Now, let us proceed with the induction step. We have:

$$1 + 2 + \cdots + n + (n+1) = [1 + 2 + \cdots + n] + (n+1) = \mathcal{P}(n) + (n+1).$$

By the induction hypothesis: $\mathcal{P}(n) = \frac{n(n+1)}{2}$. Therefore

$$\mathcal{P}(n+1) = \mathcal{P}(n) + (n+1) = \frac{n(n+1)}{2} + (n+1).$$

Since $\frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$ the claim is proved.

The induction principle can be illustrated in the following way. Imagine you have a ladder with rungs, and you want to make sure you step on all the rungs of the ladder. To achieve this, you just need to do two things:

(1) Place your foot on the first rung.

(2) Whenever you are standing on one rung you must know how to step onto the next rung.

The only problem is that our ladder is infinite, but that is another story.

In the induction step, $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$, we can assume not only that $\mathcal{P}(n)$ is true but also that $\mathcal{P}(k)$ is true for every k less than $n+1$.

It is not necessary to start from 0; starting from n_0, we can prove a proposition $P(n)$ for every $n \geq n_0$. Let us go through an example.

Example 2.11. We are wondering whether $2^n > n^2$ for all $n \in \mathbb{N}$. If $n = 0$: $2^0 = 1 > 0$, fine. If $n = 1$, $2^1 = 2 > 1$. Let us try the induction step. So, we want to prove: $2^{n+1} > (n+1)^2 = n^2 + 2n + 1$, knowing that $2^n > n^2$ (induction hypothesis). We have $2^{n+1} = 2 \cdot 2^n$. By the induction hypothesis, $2 \cdot 2^n > 2n^2 = n^2 + n^2$. So, if $n^2 > 2n + 1$, we are fine. We have $n^2 - 2n = n(n-2)$, and we see that if $n \geq 3$, $n(n-2) > 1$. However, the inequality $n(n-2) > 1$ is false for $n = 2$.

One might be tempted to say that the proposition is proved for $n \geq 3$. Be careful! In fact, $2^3 = 8$ while $3^2 = 9$. So, we can not start from 3 nor from 4 because $2^4 = 16 = 4^2$. But for $n = 5$, it works because $2^5 = 32 > 5^2 = 25$. In conclusion, we have proved: If $n \geq 5$, then $2^n > n^2$. We have also proved: If $n \geq 4$, then $2^n \geq n^2$.

The proof by induction is a powerful method, but there is a small issue. Think about the proof for $1 + 2 + \ldots + n = \frac{n(n+1)}{2}$. What is the problem?

The problem is that we need to know the formula. How do we find the formula? This is where you need to use your imagination and creativity.

First try: Consider a table with $n \times n$ cells, containing a total of n^2 elements. Let us focus on the elements that are on the diagonal and those below the diagonal:

$$\begin{pmatrix} \star & \dots & \dots & \dots & \dots \\ \star & \star & \dots & \dots & \dots \\ \star & \star & \star & \dots & \dots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \star & \star & \star & \cdots & \star \end{pmatrix}.$$

So, we are considering the elements marked with stars ($\star$). In the first row, there is one, in the second row two, in the third three, and so on, up to the n-th row where there are n stars. So, in total, we have $1+2+\cdots+n$ elements of $\star$ type. On the diagonal, we have $D = n$ elements (one per row or column). The number of elements strictly below the diagonal, denoted by s, is the same, due to symmetry, as the number, S, of elements strictly above the diagonal. Since $s + S + D = n^2$, and we know $s = S$ and $D = n$, it follows that $s = \frac{n^2-n}{2}$. Therefore, the number of $\star$ elements, denoted by $\mathcal{P}(n) = 1 + 2 + ... + n$, which is equal to $s + D$, is given by $\frac{n^2-n}{2} + n = \frac{n(n+1)}{2}$.

Second try: Let us write our numbers $1, 2, ..., n$, and then on the row below, rewrite them in reverse order:

$$\begin{array}{cccccc} 1 & 2 & \cdots & \cdots & n-1 & n \\ n & n-1 & \cdots & \cdots & 2 & 1 \end{array}$$

The sum of each column equals $n + 1$. There are n columns, so $2\mathcal{P}(n) = n(n + 1)$, which means $\mathcal{P}(n) = \frac{n(n+1)}{2}$.

This approach is inspired by the following anecdote: one day, the young Gauss's math teacher, wanting some time to read the newspaper in peace, asked the students to calculate the sum: $S = 1 + 2 + \ldots + 100$. The teacher had just managed to sit down and open the newspaper when he heard young Gauss saying "The answer is 5050". How had he done it? He paired the numbers so that the sum of each pair was 100: 0 with 100, 1 with 99, 2 with 98, and so on, up to 49 with 51. There are 50 pairs with a total sum of 100 each, so their total sum is $100 \times 50 = 5000$. To conclude, he added the single number 50, which was not paired with any other, making $S = 5050$!

Example 2.12. Sometimes intuition and rigor complement each other. Let us try to calculate $S(n) = 1^2+2^2+\ldots+n^2$. If we consider $\mathcal{P}(n) = 1+2+\ldots+n = \frac{n(n+1)}{2}$ we see that $\frac{n(n+1)}{2}$ is a degree two polynomial in n. One might think (by analogy?) that $S(n)$ is a degree three polynomial in n. Be careful! There is no mathematical or scientific justification for this, it is just an attempt!

So, let us assume $S(n) = an^3 + bn^2 + cn + d$. Since $S(0) = 0^2 = 0$ we can see that d should be 0. Similarly, we should have $S(1) = 1 = a + b + c$. From here you can continue on your own (Exercise 2.15). The idea is to determine the coefficients a, b, c and then verify, by induction, that the formula works (and it does).

Exercises

Exercise 2.13. Using the prime factorization theorem, provide a direct proof of the implication: n^2 is even $\Rightarrow n$ is even.

Exercise 2.14. For every $n \in \mathbb{N}$, the factorial $n!$ is defined as follows: $0! = 1$ (this is a convention), and if $n > 0$, $n! := 1 \cdot 2 \cdot \ldots \cdot n$.

(1) Compute 6!

(2) Show that if $d > 1$ divides $n! + 1$, then $d > n$.

Conclude that for every n, there exists a prime number p with $p > n$, and that the set of prime numbers is infinite.

Exercise 2.15. Find a formula for $S(n) = 1^2 + 2^2 + \ldots + n^2$.

Exercise 2.16. Prove by induction that for every $n \in \mathbb{N}$:

$$\forall x \in \mathbb{R} \setminus \{1\} : \sum_{i=0}^{n} x^i = \frac{1 - x^{n+1}}{1 - x} \quad (\star).$$

Prove $(\star)$ without using induction.

Exercise 2.17. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of elements in $\mathbb{Z}$ defined as $u_0 = 1, u_1 = 4$, and $u_{n+2} = 2u_{n+1} - u_n$ for all $n \in \mathbb{N}$. Compute the first terms and guess a formula for u_n that depends only on n (not on previous terms).

Prove the formula found by *strong induction*. Strong induction means that we assume a proposition $\mathcal{P}(m)$ to be true for all $m < n$ as opposed to the weak induction we have introduced where we assume $\mathcal{P}(m)$ to be true just for $m = n - 1$.

Chapter 3

Set theory and maps

The rigorous definition of set is beyond our scope, so we will settle for an intuitive definition: A set is a collection of objects sharing a certain properties. If X is the set, the objects are the elements of X. We write $x \in X$ to say that x is an element of X.

Given two sets X and Y, we can form their union and intersection:

- $X \cap Y = \{z \mid z \in X \wedge z \in Y\}$;
- $X \cup Y = \{z \mid z \in X \vee z \in Y\}$.

In addition, we have the notion of a subset: Y is a subset of X if every element in Y belongs to X: $Y \subset X \Leftrightarrow (y \in Y \Rightarrow y \in X)$.

With $Y \subset X$, we indicate an inclusion that can be either strict ($Y \neq X$) or non-strict ($Y = X$).

If $Y \subset X$, the complement, $X \setminus Y$, of Y in X is defined as:

$$X \setminus Y = \{z \in X \mid z \notin Y\}.$$

Two sets are equal if they have the same elements, more precisely:

$$X = Y \Leftrightarrow (X \subset Y) \wedge (Y \subset X).$$

This fact will be used repeatedly!

Proposition 3.1. (i) *There exists a set, denoted by $\emptyset$, such that: $\forall x, x \notin \emptyset$ (meaning $\emptyset$ has no elements).*

(ii) *For every set X we have $\emptyset \subset X$.*

(iii) *The set $\emptyset$ is called the empty set and is a unique set satisfying properties (i) and (ii).*

Proof. (i) Let X be any set. We have $X \subset X$ (meaning X is a subset of itself), and we can consider $\emptyset = X \setminus X$. If $x \in \emptyset$, then $x \in X \setminus X = \{x \mid x \in X \wedge x \notin X\}$. If $A(x)$ is the proposition $x \in X$, then $\neg A(x)$ is $x \notin X$. The proposition $A(x) \wedge \neg A(x)$ is always false. So, $x \notin \emptyset$.

(ii) We need to prove the implication: $\forall x, x \in \emptyset \Rightarrow x \in X$. Since $x \in \emptyset$ is always false, the implication is true.

(iii) Let Z be a set that satisfies (ii). We have $Z \subset \emptyset$. But from (ii), we also have $\emptyset \subset Z$. Therefore, $Z = \emptyset$. □

Definition 3.2. The *power set* of a set X is denoted as $\wp(X)$, and it consists of all the subsets of X.

Let us illustrate this with an example. If $X = \{1, 2\}$, the power set is $\wp(X) = \{\emptyset, 1, 2, X\}$.

The cardinality of a finite set, which is the number of its elements, is denoted as $\mathrm{card}(X)$ or $\sharp(X)$.

Lemma 3.3. *If X is a finite set with n elements ($n \in \mathbb{N}$) then $\sharp(\wp(X)) = 2^n$.*

Proof. We proceed by induction. If $n = 0$, then $X = \emptyset$ (by Proposition 3.1). We have $\wp(\emptyset) = \emptyset$, and since $2^0 = 1$, we are done (in fact, $2^0 = 1$ precisely because $\sharp(\wp(\emptyset)) = 1$). If you are not convinced, we can start the induction from 1: if X has just one element, then $\wp(X) = \{\emptyset, X\}$, which has $2^1 = 2$ elements.

Let us proceed with the induction step. We assume the lemma is true for $n-1$ and prove it for n. Let X be a set with $n > 1$ elements. Take $x_0 \in X$. The subsets of X fall into two categories: (a) those that contain x_0 and (b) those that do not contain x_0. Note that there is a perfect correspondence between subsets of type (a) and those of type (b). Indeed, if A is of type (a), then $A \setminus \{x_0\}$ is of type (b); if B is of type (b), then $B \cup \{x_0\}$ is of type (a). So, the number of subsets of type (a) is equal to the number of subsets of type (b). It follows that $\sharp\wp(X)$ is twice the number of subsets of type (b). But the subsets of type (b) are exactly the subsets of $Y := X \setminus \{x_0\}$. By the induction hypothesis, $\sharp\wp(Y) = 2^{n-1}$. Therefore, $\sharp\wp(X) = 2 \cdot 2^{n-1} = 2^n$ and the lemma is proved. □

Definition 3.4. Let X and Y be two sets. There exists a set Z whose elements are pairs of elements from X and Y:

$$Z = \{(x, y) \mid x \in X, y \in Y\}.$$

We denote this set as $Z = X \times Y$, it is called the *Cartesian product* of X and Y.

If $X = Y$, we denote the Cartesian product of X and Y as $X^2 = X \times X$. The Cartesian coordinate system: in the plane of elementary geometry, let Ox and Oy

be two axes. We choose units of length on these axes. Once this is done, every point in the plane has two coordinates (a, b). Therefore, we can view our plane as the Cartesian product $\mathbb{R} \times \mathbb{R}$, which is denoted as $\mathbb{R}^2$.

Similarly, we can define the Cartesian product of n sets, $E_1 \times E_2 \times \ldots \times E_n$. Note that $A \times B = \emptyset \Leftrightarrow (A = \emptyset) \vee (B = \emptyset)$.

Given two sets X and Y, a map $f : X \to Y$ is, morally speaking, a rule that assigns to each element x of X one and only one element y of Y. We write $y = f(x)$.

Of course, this is not a mathematical definition, but it conveys the idea well. A rigorous formulation is as follows. A *graph*, G, from X to Y is a subset of $X \times Y$ that satisfies the following properties:

- $\{x \in X \mid \exists\, y \in Y$ such that $(x, y) \in G\} = X$ (i.e. every element of X is in the graph).
- For every $x \in X$, the set $\{y \in Y \mid (x, y) \in G\}$ has a unique element.

Definition 3.5. A map f from X to Y is a triple (X, Y, G) where $G \subset X \times Y$ is a graph. For every $x \in X$, the unique element $y \in Y$ such that $(x, y) \in G$ is denoted by $f(x)$, and $y = f(x)$ is called the *image* of x through f.

In practice, we write $f : X \to Y$ and forget about the graph.

Definition 3.6. Let X and Y be two sets, and let $f : X \to Y$ be a map, X is called the domain of the map f, and Y is called the codomain (We will also say that X is the starting set, while Y is the target set).

Remark 3.7. There is no problem, a priori, in modifying the set Y (the codomain). For example, if $f : \mathbb{N} \to \mathbb{Z} : n \to 2n$, we can consider $f_1 : \mathbb{N} \to \mathbb{R} : n \to 2n$. What is the difference between f and f_1? A priori, none (they both assign $2n$ to n). However, the codomain of f_1 is different from that of f. Therefore, we must conclude that $f \neq f_1$. In fact, changing the codomain can change the properties of the map. A map $f : E \to F$ is a triple (E, F, f) (remember the graph).

Let us see some examples.

- *The identity map.* Let X be any set. In the Cartesian product $X \times X$, there is a distinguished subset, the diagonal, D, where $D = \{(x, x) \mid x \in X\}$. Clearly, D is a graph, and therefore, it corresponds to a map from X to X. This map is called the *identity* of X and is denoted by Id_X. So, for all $x \in X, Id_X(x) = x$. This map may seem trivial, but it is actually very important!
- *The constant map.* Let $y_0 \in Y$. The map $f : X \to Y$ such that for all $x \in X, f(x) = y_0$ is called the constant map with value y_0. For example, the zero map $f : \mathbb{R} \to \mathbb{R} : x \to 0$ is a constant map with value 0.

- *Restriction.* Let $f : X \to Y$ be a map and let $Z \subset X$. We can consider the map $Z \to Y : z \to f(z)$. This map is well-defined because $Z \subset X$ (so we know what $f(z)$ is since $z \in X$), and it is called the restriction of f to Z. It is denoted by $f|_Z$. Clearly, if $Z \neq X$, then $f|_Z \neq f$ (they do not have the same domain).
- *Composition of maps.* Let $f : X \to Y$ and $g : Y \to Z$ be two maps. From these two maps, we can create a third one, $g \circ f : X \to Z$, defined as $(g \circ f)(x) = g(f(x))$. In other words, we first apply f to obtain $f(x) \in Y$, and then we apply g to $f(x)$. (Pay attention to the order when writing $g \circ f$).

Lemma 3.8. *Let $f : X \to Y$, $g : Y \to Z$, $h : Z \to T$ be maps. Then*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Proof. We have $(h \circ (g \circ f))(x) = h(g \circ f)(x)) = h(g(f(x))$ and $((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x))$. Since this holds for every x, the two maps are equal. □

This is known as the associativity of the composition of maps.

Definition 3.9. Let $f : X \to Y$ be a map. The *image set* of f is defined as:

$$\mathrm{Im}(f) = \{y \in Y \mid \exists x \in X \text{ such that } y = f(x)\}.$$

In more concrete terms $\mathrm{Im}(f) = \{f(x) \mid x \in X\}$.

Definition 3.10. A map $f : X \to Y$ is *surjective* if $\mathrm{Im}(f) = Y$. In other terms f is surjective if

$$\forall y \in Y, \exists x \in X \text{ such that } y = f(x).$$

Definition 3.11. Let $f : X \to Y$ be a map. The map f is *injective* if different elements of X have different images:

$$x \neq x' \Rightarrow f(x) \neq f(x').$$

Said differently:

$$f(x) = f(x') \Rightarrow x = x'.$$

The typical example of a surjective map is a projection $p : X \times X \to X : (x, y) \to x$. We also write $p : X \twoheadrightarrow Y$ to highlight that p is surjective.

The typical example of an injective map is the inclusion: let $Y \subset X$, the inclusion map is defined by $i : Y \to X : y \to y$. We also write $i : Y \hookrightarrow X$ to highlight that i is injective.

Proposition 3.12. *Let $f : E \to F$ be a map. The following are equivalent:*

(1) *f is injective;*

(2) *there exists a map $h : F \to E$ such that $h \circ f = Id_E$.*

Proof. (1) $\Rightarrow$ (2). Let x_0 be any element of E. Let $y \in F$. If $y \in F \setminus f(E)$ we set $h(y) = x_0$. If instead $y \in f(E)$, then $y = f(x)$ and we set $h(y) = x$. This defines $h : F \to E$. In fact, if $y \in f(E)$ there is a unique x such that $y = f(x)$ because f is injective. By construction we have $h(f(x)) = x = Id_E(x)$.
(2) $\Rightarrow$ (1) If $f(x) = f(x')$ then $h(f(x)) = h(f(x'))$ and therefore $x = x'$ (since $h \circ f = Id_E$). □

We observe that the map h, in general, is not uniquely determined (if $f(E) \neq F$, i.e. if f is not surjective, we can choose the images of the elements of $F \setminus f(E)$ as we want).

The proof of the following proposition requires the use of the axiom of choice.

Axiom of choice: Let X, Y be two sets and let $s : X \to \wp(Y)$ be such that $\forall x \in X, s(x) \neq \emptyset$. Then there exists a map $r : X \to Y$ such that $r(x) \in s(x),\ \forall x \in X$.

The map r *chooses* an element in each of the subsets $s(x)$ as x varies in X. This choice is *evident* when X is a finite set but becomes more problematic when X is an infinite set, not to mention uncountably infinite (more on that later).

Proposition 3.13. *Let $f : E \to F$ be a map. The following are equivalent:*

(1) *f is surjective;*

(2) *there exists a map $h : F \to E$ such that $f \circ h = Id_F$.*

Proof. (1) $\Rightarrow$ (2). Intuitively it is clear what needs to be done. Let $y \in F$. Since f is surjective the set $\{x \in E \mid f(x) = y\}$ is non-empty. By the axiom above, we can choose an x in this set and set $h(y) = x$. Then $(f \circ h)(y) = f(h(y)) = f(x) = y$.
(2) $\Rightarrow$ (1). Let $y \in F$. By hypothesis $y = (f \circ h)(y)$, then $y = f(h(y))$, i.e. $y = f(x)$ with $x = h(y)$; therefore f is surjective. □

We observe that the map h, in general, is not uniquely determined (if there exists y such that $\{x \in E \mid f(x) = y\}$ has more than one element, that is, if f is not one-to-one, there are various choices of x in this set).

We now come to a very important class of maps:

Definition 3.14. A map $f : E \to F$ is *bijective* if it is injective and surjective.

Proposition 3.15. *Let $f : E \to F$ be a map. The following are equivalent:*

(1) *f is bijective;*

(2) *there are maps h, g such that: $h : F \to E$, $f \circ h = Id_F$, $g : F \to E$ and $g \circ f = Id_E$.*

Furthermore, if conditions (1) and (2) are verified then $g = h$ and g is the only map from F to E such that $g \circ f = Id_E$ and $f \circ g = Id_F$. The map g is called the inverse *map of f and is denote by $g = f^{-1}$.*

Proof. The existence of the maps h, g follows from what we saw before (Proposition 3.12, Proposition 3.13). Let (r, s) be a pair of maps from $F \to E$ such that $f \circ r = Id_F$, $s \circ f = Id_E$. We have $s \circ (f \circ r) = s \circ Id_F = s$. But we also have $s \circ (f \circ r) = (s \circ f) \circ r = Id_E \circ r = r$. So $s = r$. Since (r, s) was any, this shows the uniqueness of the inverse map: $f^{-1} = s = r$. □

Remark 3.16. In other words f is bijective if and only if there exists $f^{-1} : F \to E$ such that: $f \circ f^{-1} = Id_F$ and $f^{-1} \circ f = Id_E$.

So in practice we have two ways to prove that a map is bijective:
(a) show that f is injective and surjective;
(b) construct the inverse map f^{-1}.

Remark 3.17. If $f : E \to F$ is bijective then f establishes a perfect correspondence between the elements of E and the elements of F: to each element of E corresponds one and a single element of F. Somehow, f identifies the set E with the set F: everything that can be done in E can be 'transported' into F. From the point of view of set theory, E and F behave in the same way. We then say that they are *isomorphic* (from Greek: they have the same shape).

3.0.1. Counterimage, a sad story. Let $f : E \to F$ be a map and let $T \subset F$ be a subset. The set $f^{-1}(T) = \{x \in E \mid f(x) \in T\}$ is the counterimage of T via the map f.

This notation is a catastrophe. In fact it is used even if f is not bijective! This is one of the few cases in mathematics in which the same symbol is used for two different things: the counterimage via any (not necessarily bijective) map and the image of T via the inverse map of a bijective map f.

It is true that if f is bijective then the counterimage of T is the image of T via the inverse map f^{-1} but in general, if f is not bijective, the map f^{-1} does not exist!

Some authors use a different symbol to indicate the counter image. From a formal, logical point of view they are right; unfortunately the use of the notation $f^{-1}(T)$ to indicate the counter image is now rooted everywhere, so in order not to be misled when reading books and articles it is better to know (and use) this terrible notation. That's what we will do.

For example, the map $f : \mathbb{R} \to \mathbb{R} : x \to x^2$ is not bijective, but if $T = \{4\}$, we have $f^{-1}(T) = \{-2, 2\}$, while if $R = [-2, -1]$, $f^{-1}(R) = \emptyset$.

3.0.2. Infinities. Let us begin with the following question:

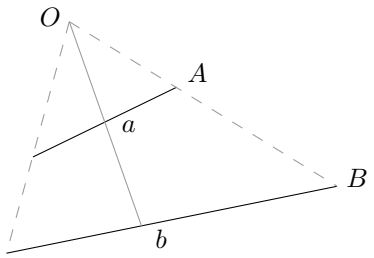
Do there exist more numbers or more even numbers?

Let $\mathcal{P} \subset \mathbb{N}$ be the set of even numbers. Since there are also odd numbers intuitively one would say that there exist more numbers than even numbers. Consider the map $f : \mathcal{P} \to \mathbb{N} : 2n \to n$. Note that the map $f^{-1} : \mathbb{N} \to \mathcal{P} : n \to 2n$ is the inverse of f. Therefore, there is a bijective correspondence between $\mathcal{P}$ and $\mathbb{N}$, and the answer to the previous question is:

There are just as many numbers as even numbers.

This fact is definitely very counter intuitive. Another example:

Are there more points in the segment A or in the segment B?



One would say that since B is longer than A there are more points in B than in A. However, the projection from O establishes a bijective correspondence between the points of A and those of B. Hence:

There are as many points in A as there are in B.

The problem we must face is the following: *How do we count the infinities?*

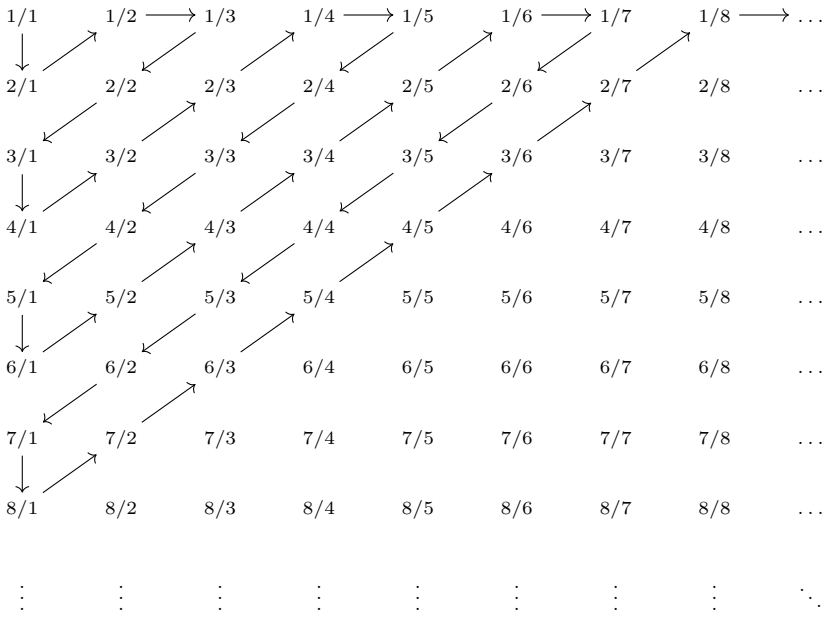
Definition 3.18. A set X is *infinite* if there exists a proper subset $Y \subsetneq X$ and a bijective correspondence $X \to Y$.

With this definition, taking into account the subset $\mathcal{P} \subsetneq \mathbb{N}$ of even numbers, we see that $\mathbb{N}$ is infinite. The bijection

$$f : \mathbb{N} \to \mathbb{Z} : \begin{cases} 2n & \to & n; \\ 2n+1 & \to & -n-1. \end{cases}$$

shows that natural numbers are as many as the integers.

An even more impressive fact is that rational numbers are as many as the natural numbers. This can be seen via Cantor's diagonal argument:



Following the path highlighted by the arrows and removing from it the fractions that are not reduced we see that we can label every positive rational number with an even integer, in exactly the same way we can label every negative rational number with an odd integer and thus we have a bijection from $\mathbb{N}$ to $\mathbb{Q}$.

Now, what about real numbers? Real numbers are "more infinite" than integers. Let us prove that there exists no bijective correspondence between $\mathbb{N}$ and the interval $J =]0, 1[\subset \mathbb{R}$ following another argument due to Cantor. Any $x \in J$ has a decimal expansion of the form $0.a_0a_1 \dots a_n \dots$ with $a_i \in \{0, 1, ..., 9\}$. Assume that there is a bijection between $\mathbb{N}$ and J, it gives us a table like this:

$\mathbb{N}$	J
0	$0.a_0a_1 \dots a_n \dots$
1	$0.b_0b_1 \dots b_n \dots$
2	$0.c_0c_1 \dots c_n \dots$
$\dots$	$\dots$
n	$0.d_0d_1 \dots d_n \dots$
$\dots$	$\dots$

Because of the bijection any $x \in J$ appears in the right column. Now, consider $z = 0.e_0e_1 \dots e_n \dots$ with $e_0 \neq a_0, e_0 \notin \{0, 9\}, e_1 \neq b_1, e_1 \notin \{0, 9\}, \dots, e_n \neq b_n, e_n \notin \{0, 9\}$. Note that:

- Since $z \neq 0 = 0.000 \dots$ and $z \neq 1 = 0.999 \dots$ we have that $z \in J$;

- z does not appear in the table since if z corresponded to n we would have $d_n = e_n$.

This proves that:

The infinity of $\mathbb{R}$ *is "greater" than the infinity of* $\mathbb{N}$.

Indeed, the map $J \to \mathbb{R} : x \to \frac{2x-1}{x-x^2}$ is bijective, and hence a bijection $\mathbb{R} \to \mathbb{N}$ would yield a bijection $J \to \mathbb{N}$.

Definition 3.19. Two sets X and Y have the same cardinality if there exists a bijective map $X \to Y$.

If X is a finite set, $\text{card}(X)$ (cardinality of X) is exactly the number of its elements. When X is infinite it is another thing, a new kind of number (transfinite cardinal).

Definition 3.20. Let X and Y be two sets. Then

- $\text{card}(X) \leq \text{card}(Y)$ if and only if there is an injective map $X \to Y$;
- $\text{card}(X) \geq \text{card}(Y)$ if and only if there is a surjective map $X \to Y$;
- $\text{card}(X) = \text{card}(Y)$ if and only if there is a bijective map $X \to Y$.

The cardinality of $\mathbb{N}$ is denote by $\aleph_0$ (*aleph-zero*). We will denote simply by c, which stands for *continuum*, the cardinality of $\mathbb{R}$.

The cardinal $\aleph_0$ is the smallest transfinite cardinal. An informal proof goes as follows. Let X be an infinite (hence non empty) set and take $x_0 \in X$. The set $X \setminus \{x_0\}$ is infinite, so we can choose $x_1 \in X \setminus \{x_0\}$. We go on this way: once we have $\{x_0, ..., x_n\}$ we take $x_{n+1} \in X \setminus \{x_0, ..., x_n\}$. We obtain an infinite sequence $\{x_i\}_{i \in \mathbb{N}}$ of distinct elements of X showing that $\text{card}(X) \geq \aleph_0$. This is just an informal sketch, to make it work properly we need the axiom of choice.

Definition 3.21. A set X is *countable* if it is finite or there exists a bijection $X \to \mathbb{N}$ that is if $\text{card}(X) = \aleph_0$.

Hence, $\mathbb{N}, \mathbb{Z}$ and $\mathbb{Q}$ are countable while $\mathbb{R}$ is not. We have that $\aleph_0 < c$ and we say that $\mathbb{R}$ is uncountable.

Note that if α, β are cardinals such that $\alpha \leq \beta$ and $\beta \leq \alpha$ it is not immediate that $\alpha = \beta$. For this we need a theorem.

Theorem 3.22 (Cantor-Bernstein)**.** *Let* X *and* Y *be two sets such that* $\text{card}(X) \leq \text{card}(Y)$ *and* $\text{card}(Y) \leq \text{card}(X)$ *then* $\text{card}(X) = \text{card}(Y)$.

This theorem says that if there exist $f : X \to Y$ injective and $g : X \to Y$ surjective then there exists $h : X \to Y$ bijective. This is not obvious at all since a priori there is no relation between f and g.

So far we have proved that $\aleph_0 < c$. Now, we look for an infinity greater than c. We have seen that $c = \text{card}(J)$ with $J =]0,1[$. The natural candidate is the set $K =]0,1[\times]0,1[\subset \mathbb{R}^2$.

We may associate to an element $(x,y) = (0.a_0a_1\dots a_n\dots, 0.b_0b_1\dots b_n\dots)$ the element $0.a_0b_0a_1b_1\dots a_nb_n\cdots \in J$. This yields an injective map $K \to J$. However, such map is not surjective: for instance $0.102010\dots1020\dots$ does not belong to the image since in this case it would be the image of $(0.1212\dots12\dots, 0.00\dots00\dots)$ which does not lie in K. Anyway, the map is injective and hence $\text{card}(K) \leq \text{card}(J) = c$.

On the other hand, since the map $J \to K : x \to (x, \frac{1}{2})$ is injective we get that $\text{card}(J) \leq \text{card}(K)$, and the Cantor-Bernstein's theorem yields that $\text{card}(K) = \text{card}(J) = c$. Therefore, in order to construct a set of cardinality bigger than c we need to come up with something more exotic. This will be the power set.

Theorem 3.23 (Cantor)**.** *Let X be a set (finite or infinite). There is no surjective map $X \to \mathcal{P}(X)$ from X to its power set $\mathcal{P}(X)$.*

Proof. Assume that there exists a surjective map $\varphi : X \to \mathcal{P}(X)$, and consider the subset

$$D = \{x \in X \mid x \notin \varphi(x)\} \subset X.$$

Since φ is surjective we have that $D = \varphi(z)$ for some $z \in X$. First, assume that $z \in D$. Then $z \notin \varphi(z) = D$, a contradiction. Now, assume that $z \notin D$. Then $z \in \varphi(z) = D$, again a contradiction. We conclude that D does not belong to the image of φ. □

Corollary 3.24. *For any set X we have that* $\text{card}(X) < \text{card}(\mathcal{P}(X))$.

Proof. The map $X \to \mathcal{P}(X) : x \to \{x\}$ is injective. Therefore, to conclude it is enough to apply Cantor's theorem. □

We got that $c = \text{card}(\mathbb{R}) < \text{card}(\mathcal{P}(\mathbb{R}))$, and more that this:

$$\aleph_0 < c < \text{card}(\mathcal{P}(\mathbb{R})) < \text{card}(\mathcal{P}(\mathcal{P}(\mathbb{R}))) < \text{card}(\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{R})))) < \dots$$

Remark 3.25. We have that $\text{card}(\mathcal{P}(\mathbb{N})) = c$ and we write $2^{\aleph_0} = c$.

We conclude with a crucial consequence of Cantor's results. Let us assume that there exists a set $\mathcal{I}$ of all sets, i.e. if x is a set, then $x \in \mathcal{I}$. Observe that $\mathcal{P}(\mathcal{I}) \subset \mathcal{I}$. Indeed every element, x, of $\mathcal{P}(\mathcal{I})$ is a subset of the set $\mathcal{I}$, hence x is a set. By this inclusion we get $\text{card}(\mathcal{I}) \geq \text{card}(\mathcal{P}(\mathcal{I}))$, in contradiction with Cantor's theorem. Cantor noticed this contradiction in set theory in 1895.

The following is only way to get out of this contradiction:

Theorem 3.26. *The collection of all sets is not a set.*

Another proof goes as follows. Assume $\mathcal{I}$ is a set and consider

$$X = \{F \in \mathcal{I} \mid F \notin F\}.$$

Since X is defined by a condition on the elements of a set, X is a set. Now, we ask the crucial question (compare with the proof of Cantor's theorem): does X belong to itself?

- If $X \in X$, by definition of X, $X \notin X$.
- If $X \notin X$, by definition of X, $X \in X$.

In both cases we get a contradiction.

Now if the collection $\mathcal{I}$ is not a set, we can no longer say that X is a set and wondering whether X belong to itself does not make sense.

This last proof is known as Russell's paradox and it is often formulated as follows:

In a village there is a barber who shaves only and all those who do not shave themselves. Who shaves the barber?

If the barber shaves himself then the barber can not shave himself. If the barber does not shave himself then he must shave himself.

As we have seen the only way to get out of this impasse is to decide that the collection of all sets is not a set, it is something else (a *proper collection*).

But then what is a set? There are certain "collections" that are sets and others that are not. How does one decide? It can be done thanks to mathematical logic but in general it is highly non-trivial.

We have seen that $\aleph_0 < c$. It is natural to wonder if there is a cardinal $\aleph_1$ in between $\aleph_0$ and c.

The *continuum hypothesis*, *CH* for short, states that there is no cardinal in between $\aleph_0$ and c.

In 1940 Gödel proved that it is not possible to prove that the continuum hypothesis is false. In other words it is logically consistent with the axioms of mathematics. But that does not mean it is true! In 1963 Cohen proved that it is not possible to prove that the continuum hypothesis is true either.

Therefore the continuum hypothesis is undecidable: one can construct a mathematical theory in which it is true and another one where it is false, both of them are equally valid!

Gödel himself proved later on a puzzling result: *It is not possible (with the tools of mathematics) to prove that mathematics is non-contradictory.*

So there is some uncertainty, someone might one day find a contradiction in mathematics? Since theorems have been proved for more than two millennia, and these theorems had, as Einstein said, an enormous impact on our understanding of nature, the working mathematician sleeps peacefully.

Exercises

Exercise 3.27. Let E, F be two sets. Show that $E \subset F$ if and only if $E \cap F = E$, and that $E \cap \emptyset = \emptyset$.

Exercise 3.28. List the elements of the set $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset)))$.

Exercise 3.29. Let X be a finite set and $f : X \to X$ a map. Show that the following are equivalent:
(i) f is injective;
(ii) f is surjective;
(iii) f is bijective.

Exercise 3.30. Let $f : X \to Y$ be a map, $X' \subset X$ a subset of X, and $f_{|X'} : X' \to Y$ the restriction of f to X'.
(i) Show that if f is injective then $f_{|X'}$ is injective.
(ii) Assume f surjective. Is $f_{|X'}$ necessarily surjective?
Now, consider the map $f' : X \to f(X) : x \to f(x)$. Show that f' is surjective, and that if f is injective then f' is bijective.

Exercise 3.31. Let $f : X \to Y$ be a map. For $y \in Y$ let

$$X_y := \{x \in X \mid f(x) = y\} \subset X$$

be the *fiber* of f over $y \in Y$. Show that if $y \neq y'$ then $X_y \cap X_{y'} = \emptyset$, and that

$$X = \bigcup_{y \in Y} X_y.$$

Exercise 3.32. Let E, F be two finite sets. Denote by $\mathcal{A}(E, F)$ is the set of maps from E into F. Show that $\text{card}(\mathcal{A}(E, F)) = f^e$ where $e = \text{card}(E), f = \text{card}(F)$ (Do two proofs, one by induction on e and the other without induction).

Exercise 3.33. Let $E, F_1, ..., F_n$ be non-empty sets. For i, $1 \leq i \leq n$ we have the map $p_i : F_1 \times \cdots \times F_n \to F_i : (x_1, ..., x_n) \to x_i$. This map is called the i-th projection.
(i) Show that p_i is surjective.
(ii) Let $f : E \to F_1 \times \cdots \times F_n$ be a map, and f_i the map $p_i \circ f : E \to F_i$. Show that the map $\Phi : \mathcal{A}(E, F_1 \times \cdots \times F_n) \to \mathcal{A}(E, F_1) \times \cdots \times \mathcal{A}(E, F_n) : f \to (f_1, ..., f_n)$ is bijective ($\mathcal{A}(X, Y)$ is the set of maps from X to Y). This means that to give a map of E in $F_1 \times \cdots \times F_n$ is equivalent to give n map $f_i : E \to F_i, 1 \leq i \leq n$.
(iii) Let $f : E \to F_1 \times \cdots \times F_n : x \to (f_1(x), ..., f_n(x))$. Show that if there exists i such that f_i is injective, then f is injective. Is it true that there exists i such that f_i is injective?

(iv) Is it true that f_i surjective $\forall i$ implies f surjective?
(v) State whether the following maps are injective, surjective:
(a) $f : \mathbb{R} \to \mathbb{R}^3 : x \to (1, 2x, x^2)$;
(b) $g : \mathbb{R} \to \mathbb{R}^3 : x \to (x - 1, x^2 - x, x^2)$.

Chapter 4

Algebraic structures

4.1. Equivalence relations

Intuitively a *binary relation*, R, on set E is a property concerning the pairs $(x, y) \in E \times E$. Some couples are related, some are not. For example, the relation R on the set of the Italian population $xRy \Leftrightarrow x$ and y are married.

Another example: on the set $\mathbb{R}$: $xRy \Leftrightarrow x \leq y$. Then $1\ R\ 2$ is verified while $3\ R\ 1$ is not.

A more formal definition is the following: a binary relation on the set E is a subset G of $E \times E$. We have $xRy \Leftrightarrow (x, y) \in G$ (be careful the order is important, it is possible to have $(x, y) \in G$ and $(y, x) \notin G$).

Definition 4.1. A binary relation R on the set E is called an *equivalence relation* if it satisfies:

(1) $\forall x \in E$, xRx (reflexivity);

(2) $\forall (x, y) \in E^2$: $xRy \Rightarrow yRx$ (symmetry);

(3) $\forall (x, y, z) \in E^3$: $(xRy) \wedge (yRz) \Rightarrow xRz$. (transitivity)

For example, let us consider the property $xRy \Leftrightarrow x$ and y have the same eye color. Assuming that there are no people with eyes of different colors we can partition the population according to various classes: Those who have blue eyes, those who have green eyes, those who have brown eyes.

On any set E the relation $xRy \Leftrightarrow x = y$ is an equivalence relation.

Let $f : E \to F$ be a map, the relation on E: $xRy \Leftrightarrow f(x) = f(y)$ is an equivalence relation.

On $\mathbb{R}$ the relation $xRy \Leftrightarrow x \geq y$ is not an equivalence relation. Why?

On $\mathbb{Z}$ the relation $nRm \Leftrightarrow n - m$ is divisible by a (where $a \in \mathbb{N}^*$) is an equivalence relation. Observe that being divisible by $-a$ is equivalent to being divisible by a: $n = k(-a) = (-k)a$. This relation can also be formulated by saying that n and m have the same remainder when dividing by a. The relation is noted $n \equiv m \pmod{a}$ and reads: n is congruent to m modulo.

Definition 4.2. Let R be an equivalence relation on the set E and let $x \in E$. We denote by $R(x)$ (or $\overline{x}$ if there is no ambiguity) the *equivalence class* of x:

$$R(x) := \{y \in E \mid xRy\}.$$

Observe that $\forall x \in E, R(x) \neq \emptyset$ because $x \in R(x)$. Furthermore, if $y, z \in R(x)$ then yRz since yRx and xRz imply yRz. In particular, $R(x) = R(y) = R(z)$.

Definition 4.3 (*Partition*)**.** Let E be a set. A family $(X_i)_{i \in I}$ of distinct, non-empty subsets of E is a *partition* of E if:

(1) $\bigcup_{i \in I} X_i = E$;
(2) $X_i \cap X_j \neq \emptyset \Rightarrow X_i = X_j$.

Lemma 4.4. *To give an equivalence relation on the set E is equivalent to give a partition of E.*

Proof. (1) Let $(X_i)_{i \in I}$ be a partition of E. We define a relation by: $xRy \Leftrightarrow \exists i$ such that $x \in X_i \wedge y \in X_i$. It is easily verified that R is an equivalence relation.
(2) Let R be an equivalence relation on E. We show that the set of equivalence classes $(R(x))_{x \in E}$ is a partition of E. The first condition of the Definition 4.3 is verified because $\forall x \in E, x \in R(x)$. Suppose $R(x) \cap R(y) \neq \emptyset$. Let $z \in R(x) \cap R(y)$. We show that any element $t \in R(y)$ belongs to $R(x)$. We have tRz and zRx. By transitivity tRx, i.e. $t \in R(x)$. This shows $R(y) \subset R(x)$. In the same way we prove $R(x) \subset R(y)$, so $R(x) = R(y)$. □

Definition 4.5 (*Quotient set*)**.** Let R be an equivalence relation on the set E. The quotient set is the set whose elements are the equivalence classes of R. It is denoted by E/R.

The map $p : E \to E/R : x \to R(x)$ is called the canonical map to the quotient.

The map p is always surjective ($\forall x, p(x) = R(x)$) but it is not injective except when R is the equality relation. Indeed p is injective $\Leftrightarrow \forall x, R(x) = \{x\}$.

Example 4.6. We denote $\mathbb{Z}^* = \{n \in \mathbb{Z} \mid n \neq 0\}$. On $E := \mathbb{Z} \times \mathbb{Z}^*$ we consider the relation $(a, b)\, R\, (c, d) \Leftrightarrow ad - bc = 0$. It can be verified that R is an equivalence

relation. The quotient set is nothing but $\mathbb{Q}$, the set of rational numbers. In fact $\mathbb{Z} \times \mathbb{Z}^*$ can be identified as the set of fractions: $(a, b) \leftrightarrow \frac{a}{b}$ and the relation R says that $\frac{a}{b} = \frac{c}{d} \Leftrightarrow ad = bc$. So a rational number is (the image in the quotient set) of an equivalence class of fractions.

Example 4.7. Let us consider the relation $n \equiv m \pmod 3$ on $\mathbb{Z}$. So two integers are related if they have the same remainder when divided by 3 (i.e. if $3 \mid n - m$). The quotient set is $\mathbb{Z}/R := \mathbb{Z}/3\mathbb{Z}$. How many elements does $\mathbb{Z}/3\mathbb{Z}$ have? The possible remainders in division by three are 0, 1, 2. So $\mathbb{Z}/3\mathbb{Z}$ has three elements. It is customary to write $\mathbb{Z}/3\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}\}$. Note however that $6 \equiv 0 \pmod 3$, $5 \equiv 2 \pmod 3$ and $4 \equiv 1 \pmod 3$, so we also have $\mathbb{Z}/3\mathbb{Z} = \{\overline{6}, \overline{4}, \overline{5}\}$. More generally $\mathbb{Z}/n\mathbb{Z}$ has n elements.

4.2. Groups

An internal composition law on the set E is essentially *an operation* which allows, starting from two elements of E, to obtain a third one (the result of the operation) for example: $2 + 4 = 6$. From a more formal point of view:

Definition 4.8. An *internal composition law* on the set E is a map:

$$f : E \times E \to E.$$

For psychological and practical reasons it is preferable to write: $x \star y$ instead of $f(x, y)$ (or $x \cdot y, x + y, xy, x \perp y, f \circ g$).

Definition 4.9 (*Group structure*)**.** A group structure on the set E is an internal composition law on E, $\star$, which satisfies the following conditions:

(1) $\forall (x, y, z) \in E^3$: $x \star (y \star z) = (x \star y) \star z$ (associativity);

(2) $\exists e \in E \mid \forall x \in E$: $x \star e = e \star x = x$ (neutral element);

(3) $\forall x \in E, \exists x^{-1} \in E \mid x \star x^{-1} = x^{-1} \star x = e$ (symmetric element).

In practice, when the law is specified by the context, it is common to say that E is a group, but it must be remembered that a group is a pair $(E, \star)$. The symmetric can be indicated in various ways: $x', x^{-1}, -x$. Note that the symmetric of the symmetric of x is x: $(x^{-1})^{-1} = x$.

Definition 4.10 (*Commutative group*)**.** Let $(G, \star)$ be a group. If

$$\forall (x, y) \in G^2 : x \star y = y \star x$$

we say that G is a *commutative* group (it is also called *abelian*).

The term *abelian* is in memory of the Norwegian mathematician Niels Henrik Abel (1802-1829).

Example 4.11. On $\mathbb{Z}$ the addition is an internal composition law, it is easily verified that $(\mathbb{Z}, +)$ is an abelian group. The neutral is 0 and the symmetric of n is $-n$.

The same goes for $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ (the symmetric of the complex number $z = x + iy$ is $-z = -x - iy$).

But $(\mathbb{N}, +)$ is not a group (symmetric elements are missing).

Example 4.12. The product is also an internal composition law on $\mathbb{Z}$, but $(\mathbb{Z}, \cdot)$ is not a group (symmetric elements are missing).

On $\mathbb{Q}$ the symmetric of $x = a/b$ is (if $x \neq 0$) $1/x = b/a$. But $(\mathbb{Q}, \cdot)$ is not a group. There is no way to have a symmetric for 0 (cannot divide by 0). However, if we consider $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$, we see that $(\mathbb{Q}^*, \cdot)$ is an abelian group (the neutral is 1, the symmetric of x is $1/x$).

The same goes for $\mathbb{R}^*$ and $\mathbb{C}^*$. The symmetric of the complex number $z = x+iy$ is $(x - iy)/(x^2 + y^2)$ which is also denoted by $\overline{z}/|z|^2$, $\overline{z} = x - iy$ is the *conjugate* of z and $|z| = \sqrt{x^2 + y^2}$ is the *modulus* of z.

Example 4.13. Let X be a set and let $\Sigma(X)$ be the set of bijective maps from X to X. On $\Sigma(X)$ we set the map composition as the internal composition law. Observe that the composition of two bijective maps is bijective. It is easily verified that $(\Sigma(X), \circ)$ is a group. The neutral element is Id_X. The symmetric of f is f^{-1}.

Example 4.14. We have defined $\mathbb{Z}/n\mathbb{Z}$ (Example 4.7), let us now try to define an addition on this set. The natural thing that comes to mind is to set: $\overline{x}+\overline{y} := \overline{x+y}$, that is, the sum of the two classes is the class of the sum of the two representatives. There is something to check. In fact, if $\overline{a} = \overline{x}$ (i.e. a and x are equivalent) and if $\overline{b} = \overline{y}$, then $\overline{x} + \overline{y} = \overline{a} + \overline{b}$ and therefore for this law to be well defined we must verify that $\overline{x+y} = \overline{a+b}$ (otherwise we would have two different results for our sum). But $\overline{x} = \overline{a}$ means that $n \mid a - x$, so $a - x = kn$ for some $k \in \mathbb{Z}$. Therefore $a = x + kn$. In the same way $b = y + tn$ for a suitable t. So $a + b = x + y + n(k + t)$ and $n \mid (a + b) - (x + y)$, that is, $\overline{a+b} = \overline{x+y}$. This shows that our addition (which we will denote by $+$) is well defined. It is easily verified that $(\mathbb{Z}/n\mathbb{Z}, +)$ is an abelian group.

Lemma 4.15. *In a group $(G, \star)$ the neutral element is unique.*

Proof. Let e be the neutral element of $(G, \star)$. Let e' be such that $e' \star x = x \star e' = x, \forall x \in G$. Then $e \star e' = e$ but $e \star e' = e'$ because e is the neutral element, so $e = e'$. □

Lemma 4.16. *Let $(G, \star)$ be a group, then $\forall x \in G$, $\exists! x^{-1}$ such that $x \star x^{-1} = x^{-1} \star x = e$.*

Proof. Let $z \in G$ be such that $z \star x = x \star z = e$. We want to show that $z = x^{-1}$. We have $(z \star x) \star x^{-1} = z \star (x \star x^{-1}) = z \star e = z$. On the other hand, $(z \star x) = e$ so $(z \star x) \star x^{-1} = e \star x^{-1} = x^{-1}$. We conclude that $z = x^{-1}$ and the lemma is proved. □

4.2.1. Subgroups. Let $(G, \star)$ be a group and let $H \subset G$ be a subset. We can simply compose elements of H and ask whether $(H, \star)$ is a group. A first condition is: $x, y \in H \Rightarrow x \star y \in H$, which is not obvious a priori. In fact $H \times H \subset G \times G$ and we can consider the map $\star$ restricted to $H \times H$ but a priori the image is contained in G, not in H. Another condition is $e \in H$. In fact the neutral element for $\star$ is uniquely determined. Finally if $x \in H$, due to uniqueness of the symmetric, we must have $x^{-1} \in H$.

Definition 4.17. Let $(G, \star)$ be a group and let $H \subset G$ be a subset. Let us assume that the following conditions are satisfied:

(1) $x, y \in H \Rightarrow x \star y \in H$;

(2) $e \in H$;

(3) $x \in H \Rightarrow x^{-1} \in H$.

Then the restriction of $\star$ to H defines a group structure on H; we say that H is a *subgroup* of G.

For example $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Q}, +)$.

Lemma 4.18. *Let $(G, \star)$ be a group and let $H \subset G$ be a subset. The following are equivalent:*

(1) *H is a subgroup of G;*

(2) *H is non-empty and $\forall x \in H$ and $\forall y \in H$ we have $x \star y^{-1} \in H$.*

Proof. (1) $\Rightarrow$ (2). We observe that a group is never empty, in fact it always contains at least one element: the neutral element. So if H is a subgroup it is non-empty, furthermore if $y \in H$, then $y^{-1} \in H$ (Definition 4.17, (3)). If $x \in H$ and $y^{-1} \in H$, then $x \star y^{-1} \in H$ (Definition 4.17 (1)).

(2) $\Rightarrow$ (1). Since H is non-empty, the exists $a \in H$. We apply (2) with $x = a$, $y = a$, $a \star a^{-1} = e \in H$ and the second condition of Definition 4.17 is verified. Let $x \in H$. We have $e \in H, x \in H$, so by our hypothesis $e \star x^{-1} = x^{-1} \in H$ and the third condition of Definition 4.17 is verified. Finally, for the first condition, let $x, y \in H$. We have seen that $y^{-1} \in H$. For our hypothesis $x \star (y^{-1})^{-1} = x \star y \in H$. □

4.2.2. Morphisms of groups. Let $(G, \star)$ and $(E, \circ)$ be two groups. Among the maps $f : G \to E$ there are those that respect the group structure. What does this mean? Taking two elements $x, y \in G$ we can compose them in G and take their image in E, we obtain $f(x \star y)$, or we can take the images in E and compose them into E, we get $f(x) \circ f(y)$. We want these two things to give the same result: $f(x \star y) = f(x) \circ f(y)$; such a map will be called a *group morphism*.

Definition 4.19. Let $(G, \star)$ and $(E, \circ)$ be two groups. A map $f : G \to E$ is a group morphism if:

$$\forall x, y \in G,\ f(x \star y) = f(x) \circ f(y).$$

Proposition 4.20. *Let $(G, \star)$ and $(E, \circ)$ be two groups with neutral elements $e \in G, \varepsilon \in E$. Let $f : G \to E$ be a group morphism, then:*

(1) $f(e) = \varepsilon$ *(the image of the neutral is the neutral);*

(2) $\forall x \in G,\ f(x^{-1}) = f(x)^{-1}$ *(the image of the symmetric is the symmetric of the image).*

Proof. (1) We have $f(x \star e) = f(x) \circ f(e)$. But $x \star e = x$ so $f(x) = f(x) \circ f(e)$. Composing on the left with the symmetric (in E) of $f(x)$: $f(x)^{-1} \circ f(x) = f(x)^{-1} \circ (f(x) \circ f(e))$. By associativity $f(x)^{-1} \circ (f(x) \circ f(e)) = (f(x)^{-1} \circ f(x)) \circ f(e)) = \varepsilon \circ f(e) = f(e)$. Since $f(x)^{-1} \circ f(x) = \varepsilon$ we conclude that $f(e) = \varepsilon$.

(2) We have $f(x \star x^{-1}) = f(x) \circ f(x^{-1})$. But $x \star x^{-1} = e$ and $f(e) = \varepsilon$, so $\varepsilon = f(x) \circ f(x^{-1})$. In a similar way we prove $f(x^{-1}) \circ f(x) = \varepsilon$. By uniqueness of the symmetric we conclude that $f(x^{-1}) = f(x)^{-1}$. □

Lemma 4.21. *Let $(G, \star)$ and $(E, \circ)$ be two groups and let $f : G \to E$ be a morphism of groups, then $\mathrm{Im}(f) = \{f(x) \mid x \in G\} \subset E$ is a subgroup of E.*

Proof. We have (Proposition 4.20) $f(e) = \varepsilon$ therefore $\varepsilon \in \mathrm{Im}(f)$. Let $z = f(x), t = f(y)$. We need to see $z \circ t^{-1} \in \mathrm{Im}(f)$ (Lemma 4.18). We have $z \circ t^{-1} = f(x) \circ f(y)^{-1}$. We know that $f(y)^{-1} = f(y^{-1})$ (Proposition 4.20) so we have $z \circ t^{-1} = f(x) \circ f(y^{-1}) = f(x \star y^{-1}) \in \mathrm{Im}(f)$. Therefore, $\mathrm{Im}(f)$ is a subgroup of E. □

The kernel of a group morphism $f : G \to E$, denote by $\mathrm{Ker}(f)$, is the set of elements of G whose image is the neutral of E.

Definition 4.22. Let $f : G \to E$ be a group morphism. We have

$$\mathrm{Ker}(f) := \{x \in G \mid f(x) = \varepsilon\}$$

where ε is the neutral of E.

The kernel of a morphism is a very important object.

Lemma 4.23. *Let $f : G \to E$ be a group morphism.*

(1) $\mathrm{Ker}(f) \subset G$ *is a subgroup of* G*;*

(2) f *is injective if and only if* $\mathrm{Ker}(f) = \{e\}$ *where* e *is the neutral of* G*.*

Proof. (1) We know that $e \in \mathrm{Ker}(f)$ (Proposition 4.20). Let $x, y \in \mathrm{Ker}(f)$, we must show $x \star y^{-1} \in \mathrm{Ker}(f)$, i.e. $f(x \star y^{-1}) = \varepsilon$. We have $f(x \star y^{-1}) = f(x) \circ f(y^{-1}) = f(x) \circ f(y)^{-1}$. By hypothesis $f(x) = f(y) = \varepsilon$ and the thesis follows.

(2) Suppose f is injective. Let $x \in \mathrm{Ker}(f)$. We have $f(x) = \varepsilon = f(e)$, by injectivity $x = e$, therefore $\mathrm{Ker}(f) = \{e\}$.

Conversely, we assume $\mathrm{Ker}(f) = \{e\}$ and show that f is one-to-one. Suppose $f(x) = f(y)$. We have $f(x) \circ f(y)^{-1} = \varepsilon$. So $\varepsilon = f(x) \circ f(y)^{-1} = f(x) \circ f(y^{-1}) = f(x \star y^{-1})$. It follows that $x \star y^{-1} \in \mathrm{Ker}(f) = \{e\}$. So $x \star y^{-1} = e$ and composing on the right with y we get $x = y$. □

The two subgroups $\mathrm{Im}(f), \mathrm{Ker}(f)$ tell us when f is surjective/injective; f is surjective when $\mathrm{Im}(f)$ is as large as possible ($\mathrm{Im}(f) = E$) and f is injective when $\mathrm{Ker}(f)$ is as small as possible ($Ker(f) = \{e\}$).

4.2.3. Complements. Among the axioms of group structure, associativity may seem of little influence. Meanwhile, we note that we have used it several times in proofs. For example if $x \star g = y \star g$, then $x = y$. In fact, composing with g^{-1} we have: $(x \star g) \star g^{-1} = (y \star g) \star g^{-1}$. By associativity this implies $x \star (g \star g^{-1}) = y \star (g \star g^{-1})$ and therefore $x = y$.

In particular this proves the following.

Lemma 4.24. *Let $(G, \star)$ be a group and $g \in G$, then the map $m_g : G \to G : x \to x \star g$ is injective (just like the map $m^g : G \to G : x \to g \star x$). In particular if G is finite, m_g and m^g are bijective.*

In fact, if X is a finite set and if $f : X \to X$ then f injective $\Leftrightarrow$ f is surjective $\Leftrightarrow$ f bijective.

Another interest of associativity is that it allows us to define $x_1 \star x_2 \star \cdots \star x_n$. To simplify the writing we take a multiplicative notation i.e. we denote $\star$ by $\cdot$, so we are in $(G, \cdot)$. There is no ambiguity in defining $x_1 \cdot x_2$ because we know how to compose two elements. However, the writing $x_1 \cdot x_2 \cdot x_3$ makes no sense a priori because we do not know how to compose three elements. However, we can consider $(x_1 \cdot x_2) \cdot x_3$ but also $x_1 \cdot (x_2 \cdot x_3)$. We are lucky because thanks to the associativity $(x_1 \cdot x_2) \cdot x_3 = x_1 \cdot (x_2 \cdot x_3)$, therefore we can set $x_1 \cdot x_2 \cdot x_3 := (x_1 \cdot x_2) \cdot x_3 = x_1 \cdot (x_2 \cdot x_3)$. We define by induction $x_1 \cdot \cdots \cdot x_n := (x_1 \cdot \cdots \cdot x_{n-1}) \cdot x_n$.

Let $x_1 \cdot \dots \cdot (x_i \cdot x_{i+1} \cdot \dots \cdot x_k) \cdot \dots \cdot x_n$ an expression containing parentheses in any order. We prove by induction that this expression is equal to $x_1 \cdot \dots \cdot x_n$. The case $n = 3$ has been done. Suppose the result is true for all $m < n$. Let us set $y = x_i \cdot \dots \cdot x_k$. It is this is well defined because there are less than n terms. Our expression becomes $x_1 \cdot \dots \cdot y \cdot \dots \cdot x_n$ and perhaps contains more parentheses but has fewer than n terms, so by induction hypothesis it is equal to $x_1 \cdot \dots \cdot y \cdot \dots x_n := (x_1 \cdot \dots y \cdot \dots x_{n-1}) \cdot x_n$. Entering the value of y gives $(x_1 \cdot \dots (x_i \cdot \dots \cdot x_k) \cdot \dots \cdot x_{n-1}) \cdot x_n$. The first parenthesis collects less than n terms, so by induction it equals $(x_1 \cdot \dots \cdot x_{n-1})$. This shows that our initial expression is equal to $(x_1 \cdot \dots \cdot x_{n-1}) \cdot x_n =: x_1 \cdot \dots \cdot x_n$, i.e. we have defined the product of n elements (*and for this only associativity is needed*). In particular, if the x_i are all equal to x, we have $x^n = x \cdot \dots \cdot x$. In additive notation: $nx := x + \dots + x$.

Let $(G, \cdot)$ be a finite group. Is there a way to describe it? Classify it? One possibility is to try to make a table of its law. For instance, let G be a group with two elements: $G = \{x, y\}$. One of them is the neutral let us say $x = e$, the table looks like this:

	e	y
e	$e \cdot e$	$e \cdot y$
y	$y \cdot e$	$y \cdot y$

Since e is the neutral: $e \cdot e = e, e \cdot y = y \cdot e = y$. Now, Lemma 4.24 tells us that all the elements of the group must appear in each row (and each column). So looking for example at the second line we see that $y \cdot y = e$. So the table is:

	e	y
e	e	y
y	y	e

There is only one possible table. In other words there is essentially a single group with two elements. However, we know at least two: $(\Sigma_2, \circ)$ and $(\mathbb{Z}/2\mathbb{Z}, +)$. The fact is that these two groups are *isomorphic* i.e. there exists $f : \Sigma_2 \to \mathbb{Z}/2\mathbb{Z}$ morphism of groups, bijective. In the same way we can see that there is a single table for groups with 3 elements.

4.3. Rings and fields

In this section we introduce more sophisticated algebraic structures, namely rings and fields.

Definition 4.25. Let A be a set with two internal composition laws $+$ and $\cdot$. The triple $(A, +, \cdot)$ is a ring structure on A if:

(1) $(A, +)$ is an abelian group;

(2) there is a neutral element 1 for $\cdot$: $\forall x \in A, 1 \cdot x = x \cdot 1 = x$;

(3) $\cdot$ is associative: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$;

(4) $\cdot$ is distributive with respect to $+$: $\forall (x, y, z) \in A^3$: $x \cdot (y + z) = x \cdot y + x \cdot z$, and $(x + y) \cdot z = x \cdot z + y \cdot z$.

Definition 4.26. If $\forall x, y \in A$ we have $x \cdot y = y \cdot x$, then A is a commutative ring.

Remark 4.27. Let G be a group. If the law of G is commutative, by convention, it is denoted $+$, while if the law is not commutative it is common to indicate it with $\cdot$. The symmetric of x for $+$ is $-x$ while the symmetric of x for $\cdot$ is denoted by x^{-1}. However, these are just notations and everyone is free to choose their own notations.

Example 4.28. $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ are commutative rings. But $(\mathbb{N}, +, \cdot)$ is not a ring. In fact $(\mathbb{N}, +)$ is not an abelian group.

In a ring A we have two neutral elements: 0 the neutral for $+$, and 1 the neutral for the multiplication $\cdot$. If $1 = 0$ then every element of A is equal to 0.

Lemma 4.29. *Let $(A, +, \cdot)$ be a ring, and 0 the neutral of the abelian group $(A, +)$. Then $\forall x \in A$, $x \cdot 0 = 0 \cdot x = 0$.*

Proof. We have $x = x \cdot 1 = x \cdot (1 + 0) = x \cdot 1 + x \cdot 0 = x + x \cdot 0$. So $-x + x = -x + (x + x \cdot 0)$. By associativity $-x + (x + x \cdot 0) = (-x + x) + x \cdot 0 = x \cdot 0$. Therefore $0 = x \cdot 0$. In the same way we show $0 = 0 \cdot x$. □

Lemma 4.30. *Let A be a ring. If $0 = 1$ then $x = 0$ $\forall x \in A$.*

Proof. We have $x = 1 \cdot x = 0 \cdot x = 0$ where the last equality follows from the previous lemma. □

The situation $1 = 0$ is certainly not very interesting and therefore in what follows we will always assume $1 \neq 0$.

Example 4.31. Consider $\mathbb{Z}$ be with the usual addition and multiplication operations: $(\mathbb{Z}, +, \cdot)$. Then it is easily verified that $(\mathbb{Z}, +, \cdot)$ is a commutative ring. Observe that the only elements that have a symmetric for $\cdot$ are 1 and -1.

Example 4.32. We have seen (Example 4.14) how to define an addition in $\mathbb{Z}/n\mathbb{Z}$. In the same way we can define a multiplication: $\overline{x} \cdot \overline{y} = \overline{xy}$. To see that this multiplication is well defined we need to show that it does not depend on the choice of representatives: if $\overline{a} = \overline{x}$ and $\overline{b} = \overline{y}$, we must see that $\overline{ab} = \overline{xy}$. So we need to show that $ab - xy$ is a multiple of n. By hypothesis $a - x = kn$ and $b - y = tn$, therefore $ab = (x + kn)(y + tn) = xy + n(xt + ky + ntk)$, and $ab - xy$ is a multiple of n. We then verify that $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ is a commutative ring.

So far all the examples given are commutative rings, but we will see important examples of non-commutative rings later.

In a ring not all elements have an inverse for $\cdot$. For example in $(\mathbb{Z}, +, \cdot)$, the only invertible elements for $\cdot$ are ± 1. However, there are rings where each non-zero element has an inverse for multiplication. For example $(\mathbb{Q}, +, \cdot)$ is such a ring.

Definition 4.33. A *field* K is a set with two laws, $+, \cdot$, of internal composition such that:

(1) $(K, +)$ is an abelian group;

(2) $(K^*, \cdot)$ is an abelian group, where $K^* = K \setminus \{0\}$;

(3) $\cdot$ is distributive with respect to $+$: $\forall (x, y, z) \in K^3$, $x \cdot (y + z) = x \cdot y + x \cdot z$ and $(x \cdot y) \cdot z = x \cdot z + y \cdot z$.

Observe that (2) implies $1 \neq 0$ (0 the neutral for $+$; 1 the neutral for $\cdot$).

Example 4.34. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ are fields, while $(\mathbb{N}, +, \cdot)$, $(\mathbb{Z}, +, \cdot)$ are not fields.

A more exotic example of a field is $(\mathbb{Z}/2\mathbb{Z}, +, \cdot)$, we will see others of this type. In fact, it can be shown that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime (Exercise 4.49).

Observe that the field structure reflects all the calculation rules that we are used to on $\mathbb{R}$, for this reason we will consider the elements of a field as *numbers* like real numbers.

Definition 4.35. Let A and B be two rings, a map $f : A \to B$ is a *ring morphism* if:

(1) f is a morphism of abelian groups: $f : (A, +) \to (B, +)$, i.e. $\forall (x, y) \in A^2 : f(x + y) = f(x) + f(y)$;

(2) $\forall (x, y) \in A^2 : f(xy) = f(x) \cdot f(y)$;

(3) $f(1_A) = 1_B$.

We observe that the third condition does not follow from the second as happened in the case of groups. In fact the procedure was: $f(1_A \cdot x) = f(x) = f(1_A) \cdot f(x)$, now composing with $f(x)^{-1}$ we get $f(1_A) = 1_B$. The problem is that a priori there is no reason why there should be $x \in A$ such that $f(x)$ is invertible with respect to the multiplication in B.

Let us see some calculation rules in a field.

Lemma 4.36. *Let K be a field and let $x, y \in K$. If $x \cdot y = 0$, then $x = 0$ or $y = 0$.*

Proof. Let $x \cdot y = 0$. If $x = 0$ we are done. If $x \neq 0$ these exists x^{-1} and we have: $x^{-1} \cdot (x \cdot y) = x^{-1} \cdot 0 = 0$. But $x^{-1} \cdot (x \cdot y) = (x^{-1} \cdot x) \cdot y = 1 \cdot y = y$. In conclusion $y = 0$. □

Remark 4.37. The previous lemma is not necessarily true in a ring $(A, +, \cdot)$. For example, let us consider $(\mathbb{Z}/4\mathbb{Z}, +, \cdot)$: we have $\overline{2} \cdot \overline{2} = \overline{4} = \overline{0}$ and $\overline{2} \neq \overline{0}$. Instead in the ring $(\mathbb{Z}, +, \cdot)$: $n \cdot m = 0 \Rightarrow n = 0$ or $m = 0$.

Definition 4.38. A ring that verifies the property of the Lemma 4.36 is called an *integral* ring.

The binomial coefficient $\binom{n}{k}$ is defined, for $0 \leq k \leq n$, by

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

Remember that $n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n$ with the convention: $0! = 1$.

Let K be a field and $x \in K$. We have seen that the writing $x^n = x \cdot \cdots \cdot x$ is well defined thanks to the associativity of the product.

Proposition 4.39 (*Binomial formula*)**.** *Let K be a field and $x, y \in K$, then for every integer $n \geq 1$ we have:*

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

This formula is proved by induction on n. Here we would like to draw the attention to the fact that binomial coefficients are integers. In fact we have:

Lemma 4.40. *For each $0 \leq k \leq n$, $\binom{n}{k}$ is the number of subsets with k elements of a set with n elements. In particular $\binom{n}{k}$ is an integer.*

Proof. If we denote $[n; k]$ the number of subsets with k elements of a set with n elements we obviously have $[n; 0] = 1$, the empty set is the only subset with 0 elements); $[n; 1] = n$ and $[n; n] = 1$. Clearly we also have $[n, k] = [n; n-k]$. In fact, every time one chooses a subset, Z, with k elements of a set X with n elements, one gets a subset with $n - k$ elements: the complementary $X \setminus Z$, and vice versa. Having said this it is not difficult to show that if $k \geq 1$:

$$[n; k] = [n-1; k] + [n-1; k-1]$$

In fact, let $X = \{1, 2, \ldots, n\}$. The subsets of X with k elements are of two types: (a) those that contain 1 and (b) those that do not contain 1. If Z is a subset as in the case (a), then $Z \setminus \{1\} \subset X \setminus \{1\}$ and we see that the number of such subsets is equal to $[n-1; k-1]$. If Y is a subset as in case (b) then $Y \subset X \setminus \{1\}$

and we see that the number of such subsets is equal to $[n-1;k]$. In conclusion $[n;k] = [n-1;k] + [n-1;k-1]$. Now with this formula and the cases $k = 0, 1, n$ it is easily proved, by double induction, that

$$[n;k] = \binom{n}{k}.$$

A result that still follows from the proof by induction of the binomial formula, but here too we need to know the formula. □

There is a more direct way to connect the two. What does it mean to calculate $(x+y)^n$? It means to compute the product:

$$(x+y)(x+y)\dots(x+y)$$

where there are n factors. To carry out this product we proceed as follows: in the first factor $(x+y)$ we choose either x or y, in the second factor we choose either x or y and so on till the last factor and calculate the product of the chosen terms (if we have chosen k times x we obtain $x^k y^{n-k}$). Then the operation is repeated in all possible ways. For example, to compute $(x+y)^3 = (x+y)(x+y)(x+y)$, we compute the product of the first two: $x^2 + xy + yx + y^2$, and then we compute $(x^2 + xy + yx + y^2) \cdot (x+y)$.

At this point we look at the coefficient of $x^k y^{n-k}$, it is exactly the number we have chosen x exactly k times and therefore $n-k$ times y. If we number the n factors $(x+y)$ with $1, 2, \dots, n$, to choose k times x means to choose a subset of $\{1, 2, \dots, n\}$ with k elements. Therefore the coefficient of $x^k y^{n-k}$ is $[n;k]$ the number of subsets with k elements of $\{1, 2, \dots, n\}$.

Exercises

Exercise 4.41. Let $f : X \to Y$ be a map, X_y the fiber of f over $y \in Y$, and consider the partition

$$X = \bigcup_{y \in Y} X_y$$

in Exercise 3.31. What equivalence relation does this partition represent?

Exercise 4.42. Let G be a group. Show that the map $f : G \to G : x \to x^2 = x \cdot x$ is a group morphism if and only if G is abelian.

Exercise 4.43. Let G be a group. Show that the map $g : G \to G : x \to x^{-1}$ is a group morphism if and only if G is abelian.

Exercise 4.44. Let G, H be two groups and $f : G \to H$ a bijective group morphism. Show that $f^{-1} : H \to G$ is a group morphism.

Exercise 4.45. Let A be an integral, commutative ring. The set A is assumed to be finite. Show that A is a field (Hint: if $x \in A, x \neq 0$ consider $m_x : A \to A : a \to ax$).

Exercise 4.46. Prove that any subgroup of $(\mathbb{Z}, +)$ is of the form

$$d\mathbb{Z} = \{dk \mid k \in \mathbb{Z}\}$$

with $d \in \mathbb{N} \setminus \{0\}$.

Exercise 4.47. Let K, K' be two fields and $f : K \to K'$ be a morphism of rings.
(i) Show that f is injective.
(ii) Note that $f(K)$, endowed with the restrictions of the operations of K', is a field.
(iii) Show that there is no morphism of rings $f : \mathbb{Z}/2\mathbb{Z} \to \mathbb{R}$.

Exercise 4.48 (Characteristic of a field)**.** Let K be a field and consider the map $f : \mathbb{Z} \to K : n \to n \cdot 1_K$, where $n \cdot 1_K = 1_K + \cdots + 1_K$ (n times). Show that f is a morphism of rings. Note that $\mathrm{Ker}(f)$ is a subgroup of $\mathbb{Z}$ and hence of the form $d\mathbb{Z}$ for some $d \in \mathbb{N} \setminus \{0\}$ (Exercise 4.46).

Let $\mathrm{char}(K)$ be the smallest non negative element of $\mathrm{Ker}(f)$, $\mathrm{char}(K)$ is called the *characteristic* of K.

Exercise 4.49. In this exercise we will use the following result:

Lemma 4.50 (Euclid)**.** *If p is prime and if $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Note that Lemma 4.50 is intuitively clear (p being prime can not be broken into two factors, one dividing a and the other dividing b).

Let K be a field. Show that if $\mathrm{char}(K) \neq 0$ then $\mathrm{char}(K)$ is a prime number. Using Lemma 4.50 and Exercise 4.45, prove that $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ is a field if and only if n is prime. If $p \in \mathbb{N}$ is a prime number what is the characteristic of $\mathbb{Z}/p\mathbb{Z}$?

Part 2

Linear Algebra

Linear algebra is the branch of mathematics concerning linear equations and the flat objects defined by them such as lines and planes.

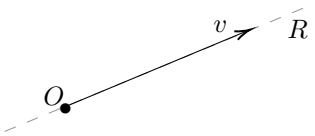
It is crucial to several areas of mathematics. For instance, in modern presentations of geometry, and in functional analysis, a branch of mathematical analysis, in which linear algebra is applied to function spaces.

Linear algebra is also very useful in most sciences and fields of engineering. In fact it allows modeling many natural phenomena, and computing efficiently with such models. For nonlinear systems, which cannot be modeled with linear algebra, it is often used for dealing with first-order approximations, using the fact that the differential of a multivariate function at a point is the linear map that best approximates the function near that point.

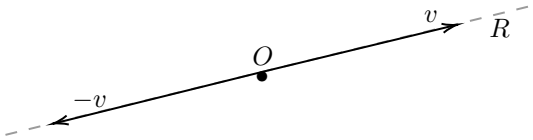
Chapter 5

Vector spaces

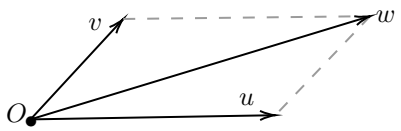
Let us do some physics. In the plane, a force applied to the point O is represented by a vector v:



The line R, identified by v, together with the path going from O to the end of the arrow is the *direction* of the vector v. The vector with the opposite direction is denoted by $-v$, it has the same *magnitude* (length) but opposite direction:

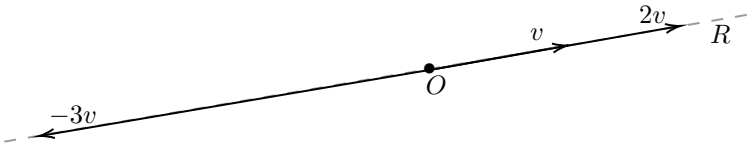


Given two forces applied at O, their action is the same as that of the *resultant force* obtained with the parallelogram rule:



The resulting force is $w = u + v$. We have therefore defined an internal composition law on the set, $\mathcal{F}$, of the forces applied in O. The zero force, 0, which consists in doing nothing, verifies $v + 0 = 0 + v = v$ and is the neutral element for this law. The symmetric of v is $-v$. It is easily verified that $(\mathcal{F}, +)$ is an *abelian group*.

Now, when we have a force v we can increase its magnitude, for example double it, and we have $2v$. For each $\alpha \in \mathbb{R}$ the vector αv is the vector that has the same direction as v if $\alpha > 0$, opposite direction if $\alpha < 0$, and magnitude $|\alpha|\cdot$ length of v. We have therefore defined a map: $\mathcal{F} \times \mathbb{R} \to \mathcal{F} : (v, \alpha) \to \alpha \cdot v$, let us call it an *external multiplication*:



It is not too difficult to convince yourself that the *external multiplication* verifies the following properties:

(1) $\forall v \in \mathcal{F}$: $1 \cdot v = v$;

(2) $\forall (\alpha, \beta) \in \mathbb{R}^2, \forall v \in \mathcal{F}$: $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$;

(3) $\forall \alpha \in \mathbb{R}, \forall (v, w) \in \mathcal{F}^2$: $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$;

(4) $\forall (\alpha, \beta) \in \mathbb{R}^2, \forall v \in \mathcal{F}$: $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$.

We then say that $\mathcal{F}$, with these laws, is an $\mathbb{R}$-*vector space*. More generally:

Definition 5.1. Let $(E, +)$ be an abelian group and let $(K, +, \cdot)$ be a field. We will say that E is a K-vector space if there exists a map $K \times E \to E$ (called *external multiplication*)

$$K \times E \to E : (\lambda, v) \to \lambda \cdot v$$

such that:

(1) $\forall v \in E$: $1_K \cdot v = v$;

(2) $\forall (\alpha, \beta) \in K^2, \forall v \in E$: $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$;

(3) $\forall \alpha \in K, \forall (v, w) \in E^2$: $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$;

(4) $\forall (\alpha, \beta) \in K^2, \forall v \in E$: $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$.

To be more rigorous one should say that, given a field K, a structure of K-vector space on the set E is a pair $((E, +), \cdot)$ where $(E, +)$ is an abelian group structure on E and where $\cdot$ is an external multiplication. We will say E is a K-vector space (or a vector space on K), but we must always keep in mind that there can be different vector space structures on the same set.

Usually, we will denote the elements of K with Greek letters and the elements of the abelian group $(E, +)$ with letters from our alphabet, for instance u, v, w, x, y. The elements of E are *vectors* while those of K are numbers, also called *scalars*.

We have two abelian group structures: the one on E and the one on K, each has its own composition law which we will however indicate with the same symbol:

$+$. It would be more correct to write $+_E$, $+_K$ but we won't do that because the context will always indicate which law it is. Both abelian groups have a neutral element: $0_E, 0_K$. The vector 0_E (which is an important character in this story) is called the *null vector*. We will denote the two neutrals with the same symbol 0, once again the context will tell us whether it is the null vector or the null scalar. We will also write 1 instead of 1_K (there can be no ambiguity).

We have two multiplications: the internal one of K and the external one; the first is denoted by $\alpha\beta$, the second by $\alpha \cdot v$, but there is no problem in indicating the two operations in the same way, and we will write αv to indicate the external multiplication as well.

Let $(G, +)$ be an abelian group written in additive notation. The symmetric of x is denoted as $-x$. Instead of writing $y + (-x)$ we write $y - x$. Since $(K, +), (E, +)$ are abelian groups we will write $\alpha - \beta$ and $u - v$.

Example 5.2. The most basic example is the following: on $E = \mathbb{R}^2$ we define an addition as follows: $(x, y) + (z, t) = (x + z, y + t)$ (coordinate-by-coordinate addition). This defines a $+$ law on $\mathbb{R}^2 = E$. It is easily verified that with this law $(E = \mathbb{R}^2, +)$ is a commutative group. The neutral is $(0, 0)$ (the origin); the symmetric of (x, y) is $(-x, -y)$.

Then we define an external multiplication as follows: $\mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2 : (\lambda, (x, y)) \to (\lambda x, \lambda y)$.

It is verified that $E = \mathbb{R}^2$, with these operations, is an $\mathbb{R}$-vector space. Actually, if you think about it, this is nothing more than the example seen before of the forces applied at a point. In fact, the construction of the parallelogram that gives the resultant is nothing but the addition that we have defined: if $u = (x, y)$ and $v = (x', y')$, the resulting vector $u + v$ corresponds to $(x + x', y + y')$.

Example 5.3. The previous example can be generalized further. Let $E := \mathbb{R}^n$. We define an addition on E as follows:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

It is verified that $(E, +)$ is an abelian group. Then we define the external multiplication as: $\mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n : (\lambda, (x_1, \ldots, x_n)) \to (\lambda x_1, \ldots, \lambda x_n)$ and it is easily verified that with these operations $\mathbb{R}^n$ is an $\mathbb{R}$-vector space.

In the previous example $\mathbb{R}$ is nothing special, the vector space conditions check works because $\mathbb{R}$ is a field; so we can generalize further.

Example 5.4. Let K be a field, on K^n we define an addition as follows:

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

Since K is a field, it is easy to verify that $(K^n, +)$ is an abelian group. Then we define an external multiplication via $K \times K^n \to K^n : (\lambda, (\alpha_1, \dots, \alpha_n)) \to (\lambda\alpha_1, \dots, \lambda\alpha_n)$. Once again it is easily verified that with these operations K^n is a K-vector space. You need to understand this example well because it is quite fundamental. In particular, K is a K-vector space (this is the case $n = 1$).

Example 5.5. A field K is a vector space over each of its subfields K. We have that $(K, +)$ is an abelian group because K is a field. The external multiplication is defined by $K \times K : (\alpha, u) \to \alpha u$.

For example $\mathbb{Q}$ is a subfield of $\mathbb{R}$, so $\mathbb{R}$ is a $\mathbb{Q}$-vector space. We have that $(\mathbb{R}, +)$ is an abelian group, furthermore $\mathbb{Q} \times \mathbb{R} \to \mathbb{R} : (\alpha, x) \to \alpha x$ verifies all the axioms of Definition 5.1.

In the same way $\mathbb{R}$ is an $\mathbb{R}$-vector space, we will see that these two structures are completely different.

Example 5.6. Let X be a set and K a field. Consider $\mathcal{F} = \{f \mid f : X \to K \mid f$ is a map$\}$. On the set $\mathcal{F}$ we define an internal composition law $+$ by: if $f, g \in \mathcal{F}$, then $f + g \in \mathcal{F}$ is defined by $f + g : x \to f(x) + g(x)$. The internal composition law is well defined, i.e. $f + g \in \mathcal{F}$. The neutral element is the constant map equal to 0, i.e. $0 : X \to K : x \to 0$. The symmetric of the map f is $-f$ where $-f : X \to X : x \to -f(x)$. With these definitions we derive that $(\mathcal{F}, +)$ is an abelian group. Now consider the following external multiplication: $K \times \mathcal{F} \to \mathcal{F} : (\lambda, f) \to \lambda f$, where $(\lambda f)(x) = \lambda f(x)$. We observe that this definition makes perfect sense, in fact $f(x) \in K$ and $\lambda \in K$, so the product $\lambda f(x)$ is an element of K. It turns out that with these definitions $\mathcal{F}$ is a K-vector space.

Example 5.7. Let $E_1, \dots, E_n$ be K-vector spaces. Let $E = E_1 \times \cdots \times E_n$ be the Cartesian product of the sets $E_1, \dots, E_n$. On E an addition is defined by $(v_1, \dots, v_n) + (u_1, \dots, u_n) = (v_1 + u_1, \dots, v_n + u_n)$. Note that the $+$ sign is used to indicate *different* operations, the addition in E_1 is not that in $E_i, i > 1$. It is verified that $(0_1, \dots, 0_n) := 0$ is the neutral element for this addition (here 0_i indicates the neutral for the addition of E_i; in the following we will omit the index).

We now define an external multiplication by: $K \times E \to E : (\lambda, v) \to \lambda v$, where if $v = (v_1, \dots, v_n)$, we have $\lambda v = (\lambda v_1, \dots, \lambda v_n)$. Observe that for each i, there is a specific law that tells us who λv_i is; that is, for every i we have that λv_i is an internal operation on E_i and therefore we can handle it.

Example 5.8 (Polynomials in one variable)**.** Let K be a field. A polynomial, $P(x)$, in the variable x, with coefficients in K, is an expression of the form: $P(x) = a_n x^n + \cdots + a_1 x + a_0$, with $a_i \in K, \forall i$. More rigorously, a polynomial is a sequence $(a_0, a_1, \dots, a_n, \dots)$ of elements of K, all zero except at most a finite

number. With this writing, the variable x disappears (a_i is the coefficient of x^i). Using the summation symbol we also write $P(x) = \sum_{i=0}^{n} a_i x^i$.

If $P(x) = (a_0, a_1, \dots, a_k, \dots)$ is a non-zero polynomial, its degree is $\deg(P) = \max\{i \mid a_i \neq 0\}$ (since the non-zero a_i are finite in number, the max is well defined). The null polynomial ($a_i = 0, \forall i$) has no degree (or sometimes, by convention, it has degree $-\infty$).

The set of polynomials in one variable with coefficients in K is usually denoted by $K[x]$.

On $K[x]$ we define an addition as follows: if $P(x) = \sum_{i\geq 0} a_i x^i$ and $Q(x) = \sum_{i\geq 0} b_i x^i$, then $P(x) + Q(x)$ is the polynomial $\sum_{i\geq 0}(a_i + b_i)x^i$. It is verified that $(K[x], +)$ is an abelian group, the neutral is the null polynomial and the symmetric of $P(x) = (a_0, a_1, \dots, a_k, \dots)$ is $-P(x) = (-a_0, -a_1, \dots, -a_k, \dots)$.

Finally, we define an external multiplication: $K \times K[x] \to K[x]$ as follows: $\lambda P(x) = \sum_{i\geq 0}(\lambda a_i)x^i$. With these operations, $K[x]$ is a K-vector space.

The last example (for now):

Example 5.9 (Matrices)**.** An $n \times m$ matrix with coefficients in K is a table of elements of K with n rows and m columns.

For everything concerning matrices we will always put the index of the rows first and the index of the columns after. So $a_{i,j}$ indicates the element that is at the intersection of the row i with the column j.

Let $M_{n,m}(K)$ be the set of matrices with n rows, m columns, and coefficients in K. If

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \in M_{n,m}(K),$$

we also write $A = (a_{ij}), 1 \leq i \leq n, 1 \leq j \leq m$. Let $B = (b_{ij}), 1 \leq i \leq n, 1 \leq j \leq m$ be another element of $M_{n,m}(K)$, then $A + B$ is defined by adding element by element:

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1m} + b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{nm} + b_{nm} \end{pmatrix}.$$

It is verified that $(M_{nm}(K), +)$ is an abelian group (the neutral is the zero matrix: $a_{ij} = 0, \forall i, j$), the symmetric of the matrix A is the matrix $-A = (-a_{ij})$.

An external multiplication $K \times M_{n,m}(K) \to M_{n,m}(K)) : (\lambda, A) \to \lambda A$ is then defined as follows:

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1m} \\ \vdots & \ddots & \vdots \\ \lambda a_{n1} & \cdots & \lambda a_{nm} \end{pmatrix}.$$

We observe that after having chosen an order (for example we write the elements of the first row first, then those of the second, an so on) we can define a map:

$$f : M_{n,m}(K) \to K^{nm} : A \to (a_{11}, \ldots, a_{1m}, \ldots, a_{n1}, \ldots, a_{nm}).$$

This map is obviously bijective (it is indeed something more because, as we will see, it respects the structure of K-vector space) and we see that the addition and external multiplication that we have defined on $M_{n,m}(K)$ are nothing but the ones we defined on K^{nm}.

5.1. Some calculation rules in a vector space

In the following E will indicate a K-vector space and we will denote the external multiplication simply by $\lambda u, \lambda \in K, u \in E$.

Proposition 5.10. *For every $u, v \in E$ and for every $\lambda, \mu \in K$ we have:*

(1) $\lambda(u - v) = \lambda u - \lambda v$;

(2) $\lambda 0_E = 0_E$;

(3) $\lambda(-v) = -(\lambda v)$;

(4) $(\lambda - \mu)u = \lambda u - \mu u$;

(5) $0_K u = 0_E$;

(6) $\lambda(-v) = -(\lambda v) = (-\lambda)v$.

Proof. (1) Using Definition 5.1 we have: $\lambda(u - v) + \lambda v = \lambda[(u - v) + v] = \lambda[u + (v - v)] = \lambda u$. Summing $-\lambda v$ to the right gives the result. (2) Set $u = v$ in (1). (3) Set $u = 0_E$ in (1) and use (2). (4) Compute $(\lambda - \mu)u + \mu u = [(\lambda - \mu) + \mu]u$ (use Definition 5.1). We have $(\lambda - \mu)u + \mu u = \lambda u$. (5) Set $\lambda = \mu$ in (4). (6) The first equality was proved in (3). If we set $\lambda = 0, \mu = \lambda$ in (4), we obtain the second equality. □

Remark 5.11. From (2) and (5) we see that there is no problem in denoting both 0_E and 0_K with 0. Furthermore, taking into account Definition 5.1, the calculations in expressions containing scalars and vectors are done with the usual rules (which we use with real numbers), as long as a vector is not added with a scalar and two vectors are not multiplied (these operations are not defined).

Lemma 5.12. *Let E be a K vector space, then $\forall \lambda \in K$ and $\forall v \in E$ we have*

$$\lambda v = 0 \Leftrightarrow \lambda = 0 \text{ or } v = 0.$$

Proof. We just show ($\Rightarrow$), the other implication follows from (2) and (5) of Proposition 5.10. If $\lambda = 0$ we are done. Otherwise, there exists λ^{-1} and we have: $\lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0$. So $0 = \lambda^{-1}(\lambda v) = (\lambda^{-1}\lambda)v = 1v = v$. □

5.2. Subvector spaces

Let E be a K-vector space and let $F \subset E$ be a subset. As usual, we would like to know when the restrictions on F of the laws of E, K endow F with a K-vector space structure.

A first condition is certainly that $(F, +)$ is a subgroup of $(E, +)$, so F is non-empty and $u, v \in F \Rightarrow u - v \in F$. The other condition is that F is stable for external multiplication: $v \in F \Rightarrow \lambda v \in F, \forall \lambda \in K$. We can summarize these three conditions as follows:

Proposition 5.13. *Let E be a K-vector space and let $F \subset E$ be a subset. The following are equivalent:*

(1) *F is a subvector space of E;*

(2) *the null vector belongs to F ($0 \in F$) and $\forall(\lambda, \mu) \in K^2$, $\forall(u, v) \in F^2$: $\lambda u + \mu v \in F$.*

Proof. (1) $\Rightarrow$ (2): clear. (2) $\Rightarrow$ (1): Since $0 \in F$, F is non-empty. By setting $\lambda = 1, \mu = -1$ in (2), we obtain that $(F, +)$ is a subgroup of $(E, +)$. By setting $\mu = 0$, we see that F is stable for external multiplication. □

Example 5.14. Let $E = K^n$ (Example 5.4) and let $F \subset K^n$ be defined by $F = \{(x_1, \ldots, x_n) \in K^n \mid x_1 = 0\}$. We show that F is a subvector space of E. We have $0 = (0, \ldots, 0) \in F$. Let $u = (0, x_2, \ldots, x_n)$, $v = (0, y_2, \ldots, y_n)$ be two elements of F and $\lambda, \mu \in K$. We have $\lambda u + \mu v = \lambda(0, x_2, \ldots, x_n) + \mu(0, y_2, \ldots, y_n) = (\lambda 0 + \mu 0, \lambda x_2 + \ mGy_2, \ldots, \lambda x_n + \mu y_n) = (0, \lambda x_2 + \mu y_2, \ldots, \lambda x_n + \mu y_n) \in F$. So F is a subvector space of E.

Observe that $G = \{(x_1, \ldots, x_n) \in K^n \mid x_1 = 1\}$ is not a subvector space of E. Why?

Example 5.15. Let $E = K[x]$ (Example 5.8). Let $F \subset E$ be defined by $F = \{P(x) \mid \text{ degree}(P(x)) \leq 3\}$. We show that F is a subvector space of E (we will use the convention that the degree of the null polynomial is $-\infty$). We therefore have $0 \in F$. Let $P(x) = \sum_{i=0}^{3} a_i x^i$, $Q(x) = \sum_{i=0}^{3} b_i x^i$ be two elements of F and let

$\lambda, \mu \in K$. Then $\lambda P(x) + \mu Q(x) = \sum_{i=0}^{3} (\lambda a_i + \mu b_i) x^i$ is still an element of F. So F is a subvector space of $K[x]$.

It is clear that 3 has nothing special, in fact if $d \in \mathbb{N}$ the set of polynomials of degree at most d is a subvector space of $K[x]$.

Observe that $G = \{P(x) \mid \text{degree } (P(x)) = 3\} \cup \{0\}$ is not a subvector space of $K[x]$. Why?

5.3. Linear combinations

Let E be a K-vector space and $v_1, \ldots, v_n \in E$. The vector $w \in E$ is a *linear combination* of the vectors v_i if there exist scalars $\lambda_i \in K$ such that $w = \lambda_1 v_1 + \cdots + \lambda_n v_n$.

More generally, let $A \subset E$ be a (non-empty) set, a linear combination of elements of A is a vector of the form $\lambda_1 v_1 + \cdots + \lambda_t v_t$ where $v_i \in A, \forall i$, $\langle A \rangle$ denotes the set of all linear combinations of elements of A.

Remark 5.16. Proposition 5.13 can be reformulated as follows: $F \subset E$ is a subvector space if $0 \in F$ and if every linear combination of two elements of F is still an element of F.

Remark 5.17. We have $A \subset \langle A \rangle$ since if $a \in A$ then $a = 1a \in \langle A \rangle$.

Remark 5.18. More generally $F \subset E$ is a subvector space if (a) $0 \in F$ and (b) if $v_1, \ldots, v_n \in F$ then every linear combination of the v_i belongs to F. In fact, if (b) is verified, every linear combination of two elements of F is an element of F and, taking (a) into account, F is a subvector space. Conversely, we show by induction on n that if F is a subspace, then every linear combination of n vectors of F is still a vector of F. The case $n = 1$ is clear: if $v \in F$, $\lambda v \in F$ because F is a subvector space. Let us assume by induction that the statement is true for $n - 1$. Set $w = \lambda_1 v_1 + \cdots + \lambda_n v_n$. We have $w = \lambda_1 v_1 + (\lambda_2 v_2 + \cdots + \lambda_n v_n)$. By induction hypothesis $u = \lambda_2 v_2 + \cdots + \lambda_n v_n \in F$. We have $\lambda_1 v_1 \in F$, so $w = \lambda_1 v_1 + u \in F$.

Proposition 5.19. *Let E be a K-vector space and let $A \subset E$ be a non-empty subset. Then*

(1) *$\langle A \rangle$ is a subvector space of E;*

(2) *$\langle A \rangle$ is the smallest subvector space of E containing A. That is, if $F \subset E$ is a subvector space of E such that $A \subset F$, then $\langle A \rangle \subset F$.*

Proof. (1) Using the Proposition 5.13 we must verify two things: (a) $0 \in \langle A \rangle$, (b) if $u, v \in \langle A \rangle$ and if $\lambda, \mu \in K$, then $\lambda u + \mu v \in \langle A \rangle$.

(a) Since $A \neq \emptyset$, $\exists a \in A$ and $0a$ is a linear combination of elements of A, then $0a = 0 \in \langle A \rangle$.

(b) Let $u, v \in \langle A \rangle$, for every $\lambda, \mu \in K$, $\lambda u + \mu v$ is still a linear combination of elements of A (a linear combination of linear combinations of elements of A is still a linear combination of elements of A).

(2) Let $w = \lambda_1 v_1 + \cdots + \lambda_n v_n$ with $v_i \in A$ (so $w \in \langle A \rangle$). If $A \subset F$ then $v_i \in F, \forall i$. Since F is a subvector space, by Remark 5.18, $w \in F$. Since w is any element of $\langle A \rangle$, this shows $\langle A \rangle \subset F$. □

Definition 5.20. The subvector space $\langle A \rangle$ is called the subvector space *generated* by the set A.

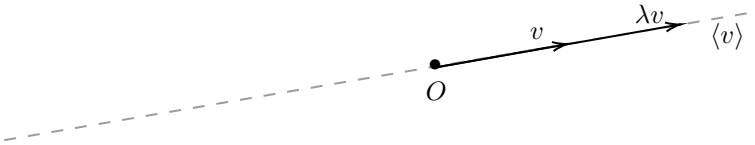
Saying that $\langle A \rangle$ is the smallest subspace containing A means that $\langle A \rangle$ is the intersection of all subvector spaces of E containing A.

Definition 5.21. If $A = \{v_1, \ldots, v_n\}$ the subspace $F := \langle A \rangle = \langle v_1, \ldots, v_n \rangle$ is the subspace generated by the v_i, and we say that $v_1, \ldots, v_n$ are *generators* of F.

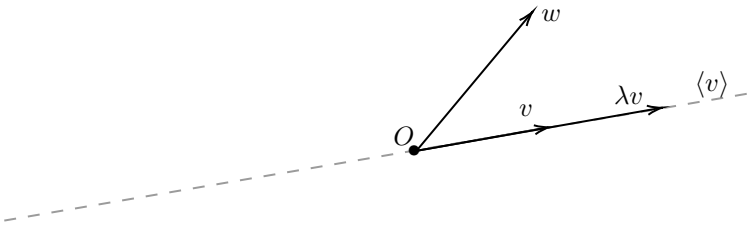
5.4. Subvector spaces of $\mathbb{R}^2$ and $\mathbb{R}^3$

First we observe that a K-vector space, E, always contains two trivial subspaces: the null space $\{0\}$ (there is only the null vector) and the space E.

Let us now consider $E = \mathbb{R}^2$ with its natural structure of $\mathbb{R}$-vector space and try to understand what its subvector spaces are. Let $F \subset \mathbb{R}^2$ be a non-zero subspace. So there exists $v \in F$, $v \neq 0$. Clearly, F contains all multiples ($\lambda v, \lambda \in \mathbb{R}$) of v, i.e. F contains the subspace $\langle v \rangle$, generated by v. Graphically:

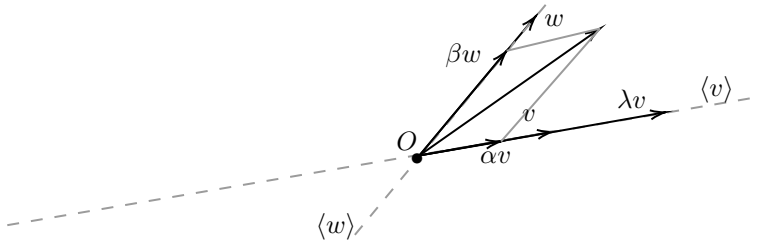


If F does not contain any other vectors then $F = \langle v \rangle$ which is actually a subvector space of $\mathbb{R}^2$. Now, suppose that F contains a vector $w \notin \langle v \rangle$:



So F contains $\langle v, w \rangle = \{\alpha v + \beta w \mid \alpha, \beta \in K\}$. We see that $F = \mathbb{R}^2$. In fact, let $u \in \mathbb{R}^2$ be any vector and consider its projection on the line $\langle v \rangle$: we obtain a vector

αv. Now let us consider its projection on the line $\langle w\rangle$: we obtain a vector βw. By the parallelogram rule $u = \alpha v + \beta w$, i.e. every vector of $\mathbb{R}^2$ is a linear combination of v, w, therefore $\langle v, w\rangle = \mathbb{R}^2$.



Finally, we have proved:

Lemma 5.22. *The subvector spaces of $\mathbb{R}^2$ are: $\{0\}$, $\mathbb{R}^2$ and the lines that pass through the origin.*

Similarly we have:

Lemma 5.23. *The subvector spaces of $\mathbb{R}^3$ are: $\{0\}$, $\mathbb{R}^3$, the lines passing through the origin ($\langle v\rangle$, $v \neq 0$) and the planes that pass through the origin ($\langle v, w\rangle$ with $v \neq 0$ and $w \notin \langle v\rangle$).*

Proof. We reason as in the case of $\mathbb{R}^2$ the only thing to observe is that if a subvector space F contains three vectors e_1, e_2, e_3 such that $e_1 \notin \langle e_2\rangle$ and $e_3 \notin \langle e_1, e_2\rangle$, then $F = \mathbb{R}^3$. This can be seen as before, taking any vector u and projecting it onto the plane $\langle e_1, e_2\rangle$ thus obtaining a vector v linear combination of e_1, e_2. Then u is projected onto the line $\langle e_3\rangle$ obtaining a vector λe_3 such that $u = v + \lambda e_3$. □

5.5. Intersections, unions and sums

Let E be a K-vector space and let $F, G \subset E$ be two subvector spaces. We ask whether $F \cap G$ is a subvector space of E.

Proposition 5.24. *Let E be a K-vector space and $F, G \subset E$ be subvector spaces. Then $F \cap G := H$ is still a subvector space of E.*

Proof. We must verify (Proposition 5.13): (a) $0 \in H$ and (b) given any two vectors in H, every linear combination of them is still an element of $H = F \cap G$. Regarding (a) since $(0 \in F) \wedge (0 \in G)$, we have $0 \in F \cap G$.

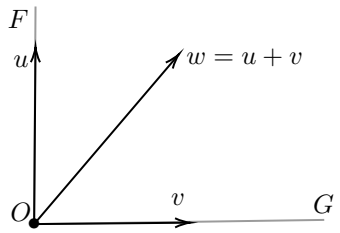
Let $u, v \in H$ and $\alpha, \beta \in K$. We have $u, v \in F$ so $\alpha u + \beta v \in F$, since F is a K-vector space. In the same way $\alpha u + \beta v \in G$. So $\alpha u + \beta v \in H = F \cap G$. □

More generally, with a similar proof, we have:

Lemma 5.25. *Let E be a K-vector space and let $(F_i)_{i\in I}$ be a family of subvector spaces of E. Then $\bigcap_{i\in I} F_i$ is a subvector space of E.*

What can we say about the union of two subvector spaces $F, G \subset E$? Is it still a subvector space? The bad news is that, in general, $F \cup G$, is not a subvector space of E.

Let us take a simple example. In $E = \mathbb{R}^2$ let $F = \{(x, y) \in \mathbb{R}^2 \mid x = 0\}$ and $G = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$. Both are subvector spaces of $\mathbb{R}^2$, but $F \cup G$ is not a subvector space of $\mathbb{R}^2$. In fact, let $u = (0, 1) \in F$ and $v = (1, 0) \in G$. Then $w := u + v = (1, 1) \notin F \cup G$ since $w \notin F$ and $w \notin G$. This can be understood very well with a picture:



This is a bit of a catastrophe because it would be nice if the fundamental operations on sets (intersection, union) respected the vector space structure. We would therefore like to have an analogue of the union that produces a subvector space. We can consider the subvector space generated by the union i.e. $\langle F \cup G\rangle$, the smallest subvector space containing $F \cup G$.

Definition 5.26. Let E be a K-vector space and $F, G \subset E$ two subvector spaces. The subspace generated by $F \cup G$ is denoted by $F + G$ and is called the *sum* of F and G.

Proposition 5.27. *Let E be a K-vector space and $F, G \subset E$ two subvector spaces. We have*

$$\langle F \cup G\rangle = F + G = \{f + g \mid f \in F \wedge g \in G\}.$$

Proof. A vector $w \in F+G := \langle F\cup G\rangle$ is a linear combination of vectors in $F\cup G$ i.e. $w = \lambda_1 v_1+\cdots+\lambda_n v_n$ with $v_i \in F$ or $v_i \in G$, since F, G are subvector spaces if $v_i \in F$ then also $\lambda_i v_i \in F$ (same thing for G). So we can rewrite $w = u_1 + \cdots + u_n$ with $u_i \in F \cup G$. We arbitrarily decide to treat a vector of $F \cap G$ as a vector of F. Once this is done, each vector is a vector of F or a vector of G. We can therefore assume (reordering the indexes if needed) that $u_1, \ldots, u_t \in F$ while $u_{t+1}, \ldots, u_n \in G$. We have $w = (u_1 + \cdots + u_t) + (u_{t+1} + \cdots + u_n) = f + g$ with $f = u_1 + \cdots + u_t$, $g = u_{t+1} + \cdots + u_n$ and $f \in F, g \in G$, since F, G are subvector spaces. □

There is a particularly important situation when considering the sum of two subspaces.

Definition 5.28. Let E be a K-vector space and F, G two subvector spaces of E. If $F \cap G = \{0\}$ we say that the sum $F + G$ is *direct* and we write $F \oplus G$ instead of $F + G$.

The following proposition explains the importance of direct sum.

Proposition 5.29. *Let E be a K-vector space and F, G two subvector spaces of E. The following are equivalent:*

(1) $F \cap G = \{0\}$*;*

(2) *each vector of $F + G$ is written uniquely as the sum of a vector of F and a vector of G;*

(3) *if $u \in F$ and $v \in G$ then $u + v = 0 \Leftrightarrow u = 0 \wedge v = 0$.*

Proof. (1) $\Rightarrow$ (2). Suppose $v = f + g = f' + g'$ with $f, f' \in F; g, g' \in G$. We have $f - f' = g' - g =: w$. The vector $w \in F$ since $w = f - f'$ and F is a subvector space. We also have $w \in G$ since $w = g' - g$ and G is a subvector space. It follows that $w \in F \cap G$, so $w = 0$. Therefore $f = f'$ and $g = g'$.

(2) $\Rightarrow$ (3). If $u + v = 0$ then $u + v = 0 + 0$ with $0 \in F, 0 \in G$. By (2) $u = 0$ and $v = 0$.

(3) $\Rightarrow$ (1). Let $v \in F \cap G$. So $v \in F$ and $v \in G$. We have $-v \in G$ since G is a subvector space. Now, $v + (-v) = 0$ with $v \in F, -v \in G$. From (3) it follows that $v = 0$. □

Definition 5.30. Let E be a K-vector space and F, G two subvector spaces of E. We say that F, G are *supplementary* spaces if $F \oplus G = E$.

The notion of direct sum can be extended to the case of more that two subspaces.

Proposition 5.31. *Let E be a K-vector space and $F_1, \ldots, F_t$ subvector spaces of E. The following are equivalent:*

(1) $F_i \cap \left(\sum_{j \neq i} F_j \right) = \{0\}$*, for every i;*

(2) *each vector, w, of $F_1 + \cdots + F_n$ is written uniquely in the form $w = v_1 + \cdots + v_n$ with $v_i \in F_i$;*

(3) *if $v_i \in F_i$, $\forall i$, then $v_1 + \cdots + v_n = 0 \Leftrightarrow v_i = 0, \forall i$.*

Proof. The proof is similar to that of Proposition 5.29. □

Exercises

Exercise 5.32. State whether E is a subvector space of $\mathbb{R}^n$ ($\mathbb{R}^n$ with the natural structure of $\mathbb{R}$-vector space):
(a) $E = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - 3z = 0\}$;
(b) $E = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - 3z = 0, \text{ e } 2x - y + z = 0\}$;
(c) $E = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - z = 1\}$;
(d) $E = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0\}$;
(e) $E = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y = 0\}$.

Exercise 5.33. State whether A is a subvector space of $M_2(\mathbb{R})$:
(a) $A = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$;
(b) $A = \left\{ \begin{pmatrix} a & a+b \\ b & 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$;
(c) $A = \left\{ \begin{pmatrix} a & ab \\ b & 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$;
(d) $A = \left\{ \begin{pmatrix} a & 1 \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$.

Exercise 5.34. Let E be a $\mathbb{C}$-vector space, and a, b two elements of E. We set $u = a + b$, $v = a - b$.
(i) Prove that the subspace $\langle a, b \rangle$ generated by a and b is equal to the subspace $\langle u, v \rangle$ generated by u and v.
(ii) Does (i) still hold if E is a $\mathbb{Z}/2\mathbb{Z}$-vector space?

Exercise 5.35. Let E be a K-vector space and F, G two subvector spaces of E. Give a necessary and sufficient condition for $F \cup G$ to be a subvector space of E.

Exercise 5.36. In $\mathbb{R}^3$ consider three pairwise non-proportional vectors v_1, v_2, v_3. Are the subspaces $\langle v_1 \rangle$, $\langle v_1 \rangle$, $\langle v_1 \rangle$ necessarily in direct sum? (Draw a picture).

Exercise 5.37. Let E be a K-vector space and F, G two subvector spaces of E. Give a necessary and sufficient condition (on F and G), so that $F \cup G$ is a subvector space of E.

Exercise 5.38. (i) Let $(G, \cdot)$ be a group and F, H two proper subgroups of G (i.e. $F \neq G$, $H \neq G$). Show that $F \cup H \neq G$.
(ii) Deduce from (i) that a K-vector space can not be written as the union of two proper sub vector spaces.
(iii) Let $K = \mathbb{F}_2$ be the field with two elements. Consider $E = K^2$ with its natural

structure of K-vector space. Show that E can be written as the union of three proper subvector spaces.

Chapter 6

Linear Maps

As in the case of groups, rings and fields, we consider maps that respect the algebraic structure we are interested in.

Definition 6.1. Let E, F be two K-vector spaces and let $f : E \to F$ be a map. We say that f is a *K-linear map* (we also say that f is a *K-linear morphism*) if:

(1) f is a morphism of groups, from the group $(E, +)$ into the group $(F, +)$, that is $f(u + v) = f(u) + f(v), \forall u, v \in E$;

(2) f respects external multiplication that is $f(\lambda v) = \lambda f(v), \forall \lambda \in K, \forall v \in E$.

Remark 6.2. (i) These are also called K-linear operator. If K is clear from the context we will call them linear maps.

(ii) Note that if f is linear, then $f(\lambda_1 v_+ \cdots + \lambda_n v_n) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n)$.

Lemma 6.3 (*Linearity criterion*)**.** *Let E, F be two K-vector spaces. A map $f : E \to F$ is linear if and only if:*

$$\forall u, v \in E, \forall \alpha, \beta \in K : \ f(\alpha u + \beta v) = \alpha f(u) + \beta f(v).$$

Proof. If f satisfies the lemma condition by setting $\alpha = \beta = 1$ we see that f is a morphism of groups and by setting $\beta = 0$ we see that it respects external multiplication.

Conversely, if f is a linear morphism, we have $f(\alpha u + \beta v) = f(\alpha u) + f(\beta v)$ (group morphism) and this is equal to $\alpha f(u) + \beta f(v)$ because f respects external multiplication. □

So f is linear if the image of a linear combination is the linear combination of the images.

Example 6.4. (1) If f is linear, since it is a morphism of groups we have $f(0) = 0$.

(2) Let $f : \mathbb{R} \to \mathbb{R}$ be an $\mathbb{R}$-linear map (here $\mathbb{R}$ is considered as a vector space over itself). Since f respects the external multiplication (which in this case is the usual multiplication in $\mathbb{R}$), we have $f(xy) = xf(y) = yf(x), \forall x, y \in \mathbb{R}$. In particular $f(x) = f(x1) = xf(1)$. We see that f is completely determined by $f(1)$. So if $f(1) = a$, we have: $f : \mathbb{R} \to \mathbb{R} : x \to ax$ (the graph of this function is a line passing through the origin).

(3) More generally if E is a K-vector space every linear map $f : K \to E$ is completely determined by its value on a non-zero element $\alpha \in K$. In fact $f(\lambda) = f(\lambda\alpha^{-1}\alpha) = \lambda\alpha^{-1}f(\alpha)$. Clearly we usually take $\alpha = 1$.

(4) The map $f : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to (ax + by, cx + dy)$ is a linear map.

Lemma 6.5 (*Composition of linear maps*)**.** *Let E, F, G be three K-vector spaces and let $f : E \to F$, $g : F \to G$ be two linear maps, then $g \circ f : E \to G$ is linear.*

Proof. It is a simple check, try. □

If $f : E \to F$ is a linear map, f is in particular a morphism of groups and we have its kernel $\mathrm{Ker}(f) = \{u \in E \mid f(u) = 0\}$.

Lemma 6.6. *Let E, F be two K-vector spaces and $f : E \to F$ a linear map.*

(1) $\mathrm{Ker}(f) \subset E$ *is a subvector space;*

(2) $\mathrm{Im}(f) \subset F$ *is a subvector space;*

(3) *f is injective* $\Leftrightarrow \mathrm{Ker}(f) = \{0\}$*;*

(4) *f is surjective* $\Leftrightarrow \mathrm{Im}(f) = F$*.*

Proof. (1) Since f is a morphism of groups $f(0) = 0$, then $0 \in \mathrm{Ker}(f)$. It remains to show that if $u, v \in \mathrm{Ker}(f)$ then $\forall \alpha, \beta \in K$, $\alpha u + \beta v \in \mathrm{Ker}(f)$. By linearity $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) = 0$ (since $f(u) = f(v) = 0$ by hypothesis).

(2) Since $f(0) = 0$, $0 \in \mathrm{Im}(f)$. Let $x = f(u), y = f(v)$, we must see that $\forall \alpha, \beta \in K$, $\alpha x + \beta y \in \mathrm{Im}(f)$. We have $\alpha x + \beta y = \alpha f(u) + \beta f(v) = f(\alpha u + \beta v)$ (the last step follows from the linearity of f). In conclusion $\alpha x + \beta y \in \mathrm{Im}(f)$.

(3) Follows from the fact that f is a morphism of groups.

(4) It is the definition. □

Proposition 6.7. *Let E, F be two K-vector spaces and $f : E \to F$ a K-linear morphism. If the map f is bijective, then $f^{-1} : F \to E$ is also a K-linear morphism.*

Proof. We need to show $f^{-1}(\alpha x + \beta y) = \alpha f^{-1}(x) + \beta f^{-1}(y), \forall x, y \in F, \forall \alpha, \beta \in K$. Since f is bijective $\exists u, v \in E$ such that $x = f(u), y = f(v)$. So $f^{-1}(\alpha x + \beta y) =$

$f^{-1}(\alpha f(u) + \beta f(v))$, by linearity of f this is equal to $f^{-1}(f(\alpha u + \beta v))$, which is $\alpha u + \beta v$. Since $u = f^{-1}(x), v = f^{-1}(y)$, the proposition is proved. □

A bijective linear map is called a *linear isomorphism*. In the situation of Proposition 6.7 we say that E and F are *isomorphic* (as K-vector spaces); these two spaces have the same properties, a linear algebra theorem proven for E translates, via f, into a theorem for F. Conversely, a theorem in F is transposed, via f^{-1}, to E. As far as linear algebra is concerned these two spaces can be identified.

We conclude this section with an important result.

Definition 6.8. A *homogeneous degree one* polynomial in the variables $x_1, \ldots, x_n$, with coefficients in the field K, is an expression of the form $P(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n$, with $a_i \in K, \forall i$. In other words $P(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i$.

It is called homogeneous because all the terms have the same degree.

Proposition 6.9. *Every linear map $f : K^n \to K^m$ is of the form*

$$K^n \to K^m : (x_1, \ldots, x_n) \to (P_1(x_1, \ldots, x_n), \ldots, P_m(x_1, \ldots, x_n))$$

where the $P_i(x_1, \ldots, x_n)$ are homogeneous polynomials of degree one in the variables $x_1, \ldots, x_n$.

Proof. Let us start with the case $m = 1$: $f : K^n \to K$. Let $u = (x_1, \ldots, x_n) \in K^n$. We can write $u = x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \cdots + x_n(0, \ldots, 0, 1)$. That is, setting $e_i = (0, \ldots, 1, \ldots, 0)$ (all coordinates are zero except the i-th which is equal to 1), $u = x_1 e_1 + \cdots + x_n . e_n = \sum_{i=1}^{n} x_i e_i$. By linearity:

$$f\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i f(e_i).$$

Let $f(e_i) = a_i \in K$ (note that, by linearity, f is completely determined by the scalars $f(e_i)$). We set $P(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n$, then we have $f(x_1, \ldots, x_n) = P(x_1, \ldots, x_n)$ and the proposition is proved in the case $m = 1$.

We observe that the projection map $p_i : K^m \to K : (y_1, \ldots, y_i, \ldots, y_m) \to y_i$ is a linear map.

As we know (Exercise 3.33), to give a map $f : K^n \to K^m$ is equivalent to give the m maps $f_i : K^n \to K$, where $f_i = p_i \circ f$. As seen above, these maps are linear and therefore are homogeneous polynomials of degree one in the x_i: $f_i(x_1, \ldots, x_n) = P_i(x_1, \ldots, x_n)$. This proves the statement in the general case. □

Exercises

Exercise 6.10. State which of the following maps are linear:
(a) $f : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to (2x - y, x + 3y)$;
(b) $g : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to (x + y - 1, 2x - y)$;
(c) $h : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to (x + 3y, -x^2 + y)$.

Exercise 6.11. Let $f : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to (3x - y, x + y)$. Show that f is linear and determine the kernel and image of f. Tell whether f is injective, surjective, bijective.

Exercise 6.12. Let E be a K-vector space, F a set, and $f : E \to F$ a bijection. For every $x, y \in F$ we set $x + y := f(f^{-1}(x) + f^{-1}(y))$. Then we define $m : F \times K \to F$ by $m(x, \alpha) = f(\alpha f^{-1}(x))$.

Prove that these operations define on F a structure of K-vector space. Also show that with respect to this structure f is a linear isomorphism between E and F.

Exercise 6.13. Give an example of a non linear map $f : \mathbb{R}^2 \to \mathbb{R}$ such that $f(\lambda v) = \lambda f(v)$ for any $\lambda \in \mathbb{R}, v \in \mathbb{R}^2$.

Chapter 7

Finitely generated spaces

Let us now see a first approach to the notion of dimension of a K-vector space.

Definition 7.1. Let E be a K-vector space and $A \subset E$ a subset. The elements of A constitute a *generator system* of E if each vector of E can be written as a linear combination of elements of A. In other words, the elements of A are a system of generators of E if $\langle A \rangle = E$.

We will also say that A *generates* E.

Example 7.2. Let $A \subset \mathbb{R}^2$, $A = \{u, v\}$, with $u = (1, 0), v = (a, b)$ with $b \neq 0$ then A is a system of generators of $\mathbb{R}^2$ (since $b \neq 0$, $v \notin \langle u \rangle$).

Definition 7.3. Let E be a K-vector space. If there exists a finite subset, $A \subset E$, which generates E, we say that E has *finite dimension* (or that E is *finitely generated*). If, however, no *finite* system of generators exists, we say that E has *infinite dimension* and we write $\dim_K E = +\infty$.

Remark 7.4. Every vector space E has a system of generators since obviously $E = \langle E \rangle$.

In the rest of this book we will develop the theory essentially for finite dimensional spaces. Almost all (but not all) of the results we will see are also valid for spaces of infinite dimension but the proofs are technically more complicated.

Example 7.5. However, it is good to know that there are vector spaces of infinite dimension which are natural mathematical objects.

(1) The vector space $K[x]$ has infinite dimension. Why?

(2) The vector space of maps (or continuous maps, differentiable maps) from $\mathbb{R}$ to $\mathbb{R}$ has infinite dimension. Why?

(3) A more exotic example. Since $\mathbb{Q} \subset \mathbb{R}$ we can consider $\mathbb{R}$ as a $\mathbb{Q}$-vector space. We have $\dim_{\mathbb{Q}} \mathbb{R} = +\infty$. This follows from the fact that $\mathbb{Q}$ is *countable*, while $\mathbb{R}$ is not.

Obviously $\mathbb{R}$ is also an $\mathbb{R}$-vector space and since $\langle 1 \rangle = \mathbb{R}$, as $\mathbb{R}$-vector space, $\mathbb{R}$ has finite dimension.

We have $\mathbb{R} \subset \mathbb{C}$, so we can consider $\mathbb{C}$ as an $\mathbb{R}$-vector space. We have that, as $\mathbb{R}$-vector space, $\mathbb{C}$ has finite dimension. Why?

Exercises

Exercise 7.6. Prove that the vector space $\mathbb{R}[x]$ is not finitely generated (consider the degree). Deduce from this that the real vector space of maps (resp. of continuous maps, resp. of differentiable maps) from $\mathbb{R}$ in $\mathbb{R}$ is not finitely generated.

Exercise 7.7. Let $K = \mathbb{Z}/2\mathbb{Z}$. State whether the K-vector space $K[x]$ is finitely generated. Let $\mathcal{A}$ be the set of maps from K into K. Determine whether the K-vector space $\mathcal{A}$ is finitely generated (Exercise 3.32). Do you see anything strange?

Exercise 7.8. Let E, F be two K-vector spaces and $f : E \to F$ a surjective linear map. Prove that if E is finitely generated, then F is also finitely generated.

Chapter 8

Linear independence and bases

Sometimes to explain what something is it is preferable to say what it is not.

Definition 8.1. Let E be a K-vector space and $v_1, \ldots, v_n$ vectors of E. The vectors v_i are *linearly dependent* if there exist scalars $\alpha_1, \ldots, \alpha_n$ *not all zero* such that

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \quad (\star)$$

The relation $(\star)$ is a linear dependence relation between the vectors v_i.

The meaning of this definition is the following: since the α_i are not all zero $\exists j$ such that $\alpha_j \neq 0$. To simplify the writing we can assume $j = 1$, i.e. $\alpha_1 \neq 0$. So we can divide by α_1 and write

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \cdots - \frac{\alpha_n}{\alpha_1} v_n.$$

We see that given $v_2, \ldots, v_n$ we can obtain v_1, i.e. v_1 *depends* on $v_2, \ldots, v_n$ and this dependence is linear because v_1 is obtained as a linear combination of $v_2, \ldots, v_n$.

In conclusion, if $v_1, \ldots, v_n$ are linearly dependent then one of them is a linear combination of the others, i.e. the vectors are (linearly) *related* to each other.

The vectors $v_1, \ldots, v_n$ are *linearly independent* if they are not linearly dependent.

Definition 8.2. The vectors $v_1, \ldots, v_n$ are *linearly independent* if and only if:

$$\forall (\alpha_1, \ldots, \alpha_n) \in K^n : \alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \Leftrightarrow \alpha_i = 0, \forall i = 1, \ldots, n.$$

It is clear that $0v_1 + \cdots + 0v_n = 0$ (whatever the vectors v_i are); the definition says that the vectors v_i are linearly independent if this (trivial) linear combination of the v_i is the only one that gives the zero vector. That is, it is not possible to find a linear combination $\alpha_1 v_1 + \cdots + \alpha_n v_n$ equal to 0 with the α_i not all zero.

Example 8.3. Let us take in $\mathbb{R}^2$, $u = (1, 0)$ and $v = (a, b)$ with $b \neq 0$. If $\alpha u + \beta v = 0$, then $(\alpha + \beta a, \beta b) = 0 = (0, 0)$, so $\beta b = 0$. Since $b \neq 0$ this implies $\beta = 0$ and looking at the first coordinate we have $\alpha = 0$. We conclude that u and v are linearly independent.

Example 8.4. Let us take in $\mathbb{R}^2$ two proportional vectors u and $v = \lambda u$. We have $-\lambda u + 1v = 0$, since $1 \neq 0$ we conclude that u and v are linearly dependent.

Note that in any vector space any two proportional vectors are always linearly dependent.

Example 8.5. Let us see some examples of linear dependence.

(1) Let E be a K-vector space and let $u \in E$. The vector u is linearly independent if and only if $u \neq 0$. In fact $\lambda u = 0 \Leftrightarrow \lambda = 0$ or $u = 0$.

(2) Let $v_1 = 0$ and $v_2, \ldots, v_n \in E$ any vectors. Then $v_1, v_2, \ldots, v_n$ are linearly dependent. In fact $1 \cdot 0 + 0v_2 + \cdots + 0v_n = 0$ is a linear dependence relation between $v_1 = 0, v_2, \ldots, v_n$ because $1 \neq 0$.

(3) Let $v_1 = v = v_2$ and $v_3, \ldots, v_n$ any $n - 2$ vectors of E. Then $v_1, v_2, v_3, \ldots, v_n$ are linearly dependent. In fact $1v_1 - 1v_2 + 0v_3 + \cdots + 0v_n = 0$ is a linear dependence relation because $1 \neq 0$ (also $-1 \neq 0$ but a non-zero coefficient is enough).

The above example shows that if $v_1, \ldots, v_n$ are linearly independent then they are distinct and none of the v_i is zero. Obviously, these are necessary but not sufficient conditions.

The following two lemmas will be useful later.

Lemma 8.6. *Let E be a K-vector space and $v_1, \ldots, v_n$ vectors of E. The vectors v_i are linearly dependent if and only if: $n = 1$ and $v_1 = 0$; $n > 1$ and there exists j, $1 \leq j \leq n$ such that v_j is a linear combination of $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$.*

Proof. We have already proved a part of this lemma. It remains to prove that if v_j is a linear combination of the other vectors, then the vectors are dependent. If $v_j = \sum_{i \neq j} \lambda_i v_i$, then $1v_j - \sum_{i \neq j} \lambda_i v_i = 0$ is a linear dependence relation since $1 \neq 0$. □

Lemma 8.7. *Let E be a K-vector space and $v_1, \ldots, v_n$ linearly independent vectors of E. Let $w \in E$ be a vector such that $v_1, \ldots, v_n, w$ are dependent, then w is a linear combination of the v_i.*

Proof. If $v_1, \ldots, v_n, w$ are dependent, there exist scalars which are not all zero $\alpha_1, \ldots, \alpha_n, \beta$ such that $\alpha_1 v_1 + \cdots + \alpha_n v_n + \beta w = 0$. If $\beta = 0$ then $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$. Since the vectors v_i are independent this implies $\alpha_i = 0, \forall i$, contradicting the hypothesis. So $\beta \neq 0$. This means we can divide by β and express w as a linear combination of $v_1, \ldots, v_n$. □

One last observation to conclude (for now) on the topic.

Remark 8.8. Let E be a K-vector space and $v_1, \ldots, v_n \in E$.

(1) If $v_1, \ldots, v_n$ are dependent and if $w \in E$, are the vectors $v_1, \ldots, v_n, w$ dependent or independent? Think about it for a moment.

The vectors $v_1, \ldots, v_n, w$ are dependent. In fact, by hypothesis we have the existence of a linear dependence relation: $\sum \lambda_i v_i = 0$ with the λ_i not all zero. From this relation we deduce that: $\sum \lambda_i v_i + o.w = 0$, i.e. a linear dependence relation for the vectors $v_1, \ldots, v_n, w$. So if we add a vector to some dependent vectors, we get dependent vectors.

(2) If $v_1, \ldots, v_n$ are dependent and if we remove one of the vectors, say v_j, are the remaining vectors $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$ dependent or independent?

It depends, they can be dependent or independent (give examples of both cases).

(3) If $v_1, \ldots, v_n$ are independent and if $w \in E$, are the vectors $v_1, \ldots, v_n, w$ dependent or independent?

Again it depends, they can be dependent or independent (give examples for both cases).

(4) If $v_1, \ldots, v_n$ are independent and if we remove one of the vectors, say v_j, are the remaining vectors $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$ dependent or independent?

The remaining vectors are independent. In fact, if $\sum_{i \neq j} \lambda_i v_i = 0$, then $\sum_{i \neq j} \lambda_i v_i + 0 v_j = 0$. Since $v_1, \ldots, v_n$ are independent, this implies $\lambda_i = 0, \forall i$.

8.1. Bases

The following is a fundamental notion.

Definition 8.9. Let E be a K-vector space and $v_1, \ldots, v_n \in E$ be such that:

(1) $v_1, \ldots, v_n$ are linearly independent

(2) $v_1, \ldots, v_n$ generate E (i.e. $\langle v_1, \ldots, v_n \rangle = E$).

Then $\mathcal{B} = (e_1, \ldots, e_n)$ is a basis of E.

The importance of this notion arises from the following fact.

Lemma 8.10. *If $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis of E, then every $w \in E$ can be written uniquely as a linear combination of the v_i.*

Proof. Since $\langle v_1, \ldots, v_n \rangle = E$, each $w \in E$ is written as a linear combination of the v_i. Suppose we have $w = \lambda_1 v_1 + \cdots + \lambda_n v_n = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Then $(\lambda_1 - \alpha_1)v_1 + \cdots + (\lambda_n - \alpha_n)v_n = 0$. Since the v_i are independent, this implies $\lambda_i - \alpha_i = 0, \forall i$ and therefore the writing is unique. □

Remark 8.11.

(1) From Definition 8.9 it follows that E has finite dimension. There is a notion of basis even for spaces of infinite dimension.

(2) A basis $\mathcal{B} = (v_1, \ldots, v_n)$ is an n-tuple of vectors (there is an order on the set $\{v_1, \ldots, v_n\}$) . If $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ is a basis then $\mathcal{B}' = (v_2, v_1, \ldots, v_n)$ is also a basis (the fact of being independent generators do not depend on the order of the vectors); however $\mathcal{B}$ and $\mathcal{B}'$ are two *different* bases.

This distinction is explained as follows. If $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis of E, then each vector $w \in E$ is uniquely written as $w = \lambda_1 v_1 + \cdots + \lambda_n v_n$. We can therefore identify the vector w with the coefficients $(\lambda_1, \ldots, \lambda_n) \in K^n$. Conversely, to the element $(\alpha_1, \ldots, \alpha_n) \in K^n$ we can associate to the vector $u = \alpha_1 v_1 + \cdots + \alpha_n v_n$. We say that $(\lambda_1, \ldots, \lambda_n)$ are the *coordinates* of w in the basis $\mathcal{B}$.

Observe that the coordinates of w in the basis $\mathcal{B}'$ are $(\lambda_2, \lambda_1, \ldots, \lambda_n)$. This is why we say that $\mathcal{B}$ and $\mathcal{B}'$ are two different bases.

(3) As already said, for vectors v_i the fact of being linearly independent generators of E does not depend on the order of the vectors. So if we do not want to specify an order we will say that the vectors v_i *form* (in some order) a basis.

We now show some characteristic properties of a basis, but first we need a definition.

Definition 8.12. Let E be a K-vector space and $e_1, \ldots, e_n \in E$.

(1) The vectors (e_i) form a *minimal system of generators* of E if $\langle e_1, \ldots, e_n \rangle = E$ and if $\forall j, 1 \leq j \leq n$, the vectors $e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n$ no longer generate all the space E.

(2) The vectors (e_i) form a *maximal system of independent vectors*, if they are independent and if $\forall w \in E$, the vectors $e_1, \ldots, e_n, w$ are linearly dependent.

Proposition 8.13. *Let $E \neq \{0\}$ be a K-vector space and $e_1, \ldots, e_n$ vectors of E. The following are equivalent:*

(1) *$e_1, \ldots, e_n$ form a basis of E;*

(2) $G = \{e_1, \ldots, e_n\}$ *is a minimal system of generators;*

(3) G *is a maximal system of independent vectors.*

Proof. (1) $\Rightarrow$ (2). By hypothesis $\langle G\rangle = E$. If G is not a minimal system of generators, there exists j such that $G \setminus \{e_j\}$ generates E; in particular e_j is a linear combination of the other e_i. This implies (Lemma 8.6) that $e_1, \ldots, e_n$ are dependent, a contradiction since the vectors e_i form a basis.

(2) $\Rightarrow$ (3). We show that the vectors e_i are independent. If $n = 1$, since $E \neq \{0\}$, $e_1 \neq 0$ and e_1 is independent. Suppose $n > 1$. If the vectors e_i are dependent one of them, e_j, is a linear combination of the others (Lemma 8.6). In this case we have $\langle e_1, \ldots, e_n\rangle = \langle e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n\rangle$ and G is not a minimal system of generators, a contradiction. So the vectors e_i are independent. Let $w \in E$, since $\langle G\rangle = E$ by hypothesis, w is a linear combination of the e_i, therefore (Lemma 8.6) $e_1, ..., e_n, w$ are dependent. This shows that G is a maximal system of independent vectors.

(3) $\Rightarrow$ (1). The vectors e_i are independent by hypothesis, we show that they generate E. Let $w \in E$. Since G is a maximal system of independent vectors, $e_1, \ldots, e_n, w$ are dependent. This implies (Lemma 8.7) that w is a linear combination of the e_i, therefore $\langle G\rangle = E$ and $e_1, \ldots, e_n$ form a basis of E. □

Remark 8.14. If $E = \{0\}$, then 0 is a minimal system of generators of E but 0 is dependent. We will see later how to fix this situation.

Our aim now is to show that every finitely generated K-vector space admits a basis. This is a *fundamental* result.

Theorem 8.15. *Let E be a K-vector space and A, A' two non-empty finite subsets of E such that $A \subset A'$. Suppose that the vectors of A are linearly independent and that the vectors of A' generate E. Then there exists B, $A \subset B \subset A'$ such that the vectors of B form a basis of E.*

Proof. Let $\mathcal{F} := \{X \mid A \subset X \subset A'$ and such that the vectors of X are linearly independent$\}$. We observe that $\mathcal{F}$ is non-empty ($A \in \mathcal{F}$), furthermore $\mathcal{F}$ is a finite set (because it is contained in the power set A' and A' is finite). Let $B \in \mathcal{F}$ be such that: $\forall X \in \mathcal{F}, \mathrm{card}(X) \leq \mathrm{card}(B)$ (B exists, perhaps not unique, because $\mathcal{F}$ is finite). By construction the vectors of B are linearly independent. We show that they generate E. Let $v \in A'$, $v \notin B$, the vectors of $B' = B \cup \{v\}$ are linearly dependent (otherwise we will have $B' \in \mathcal{F}$ with $\mathrm{card}(B') > \mathrm{card}(B)$). Therefore (Lemma 8.7) v is a linear combination of the vectors of B. This shows that $\langle A'\rangle = \langle B\rangle$. Since $\langle A'\rangle = E$ by hypothesis, the vectors of B form a basis of E. □

Corollary 8.16. *Let E be a finite-dimensional K-vector space. If $E \neq \{0\}$ then E admits a basis.*

Proof. By definition E is generated by a finite, non-empty set, A', of vectors (E is *finitely generated*). Since $E \neq \{0\}$, there exists $v \in A'$ with $v \neq 0$, so v is independent. We conclude by applying Theorem 8.15 with $A = \{v\} \subset A'$. □

Remark 8.17. Let $E = \{0\}$. The space E is finitely generated (from the vector 0) but has no non-empty basis because the unique vector in E is dependent. By convention the empty set is said to be a basis of E. With this convention every finite-dimensional K-vector space admits a basis.

It can be shown that every vector space of infinite dimension admits a basis. In this case a basis is an infinite set of vectors $(e_i)_{i\in I}$ such that every finite number of these vectors are linearly independent and such that every vector in the space is a linear combination of a finite number of these vectors. For example the polynomials x^n with $n \in \mathbb{N}$ form a basis of $K[x]$.

A reformulation of Theorem 8.15:

Corollary 8.18. *Let E be a K-vector space. If $v_1, \ldots, v_k$ generate E, then the set $\{v_1, \ldots, v_k\}$ contains a basis of E.*

An *important* consequence of Theorem 8.15:

Corollary 8.19 (Incomplete basis theorem)**.** *Let E be a finitely generated (i.e. finite-dimensional) K-vector space and $v_1, \ldots, v_p$ linearly independent vectors of E. Then there exists a basis of E containing $v_1, \ldots, v_p$.*

Proof. Let us set $A = \{v_1, \ldots, v_p\}$. Since E is finitely generated there is a finite set, G, which generates E. We set $A' = G \cup A$. Clearly A' generates E and $A \subset A'$. We conclude by Theorem 8.15. □

8.2. Infinite-dimensional vector spaces

We have seen in Corollary 8.16 that all finite-dimensional vector spaces admit a basis. This is true also for vector spaces of infinite dimension. The proof of this fact requires Zorn's lemma which is equivalent to the axiom of choice.

Zorn's lemma states that a partially ordered set containing upper bounds for every chain necessarily contains at least one maximal element. The lemma was proved, assuming the axiom of choice, by Kazimierz Kuratowski in 1922 and independently by Max Zorn in 1935. This is why it is also known as the Kuratowski–Zorn lemma.

Zorn's lemma is equivalent to the well-ordering principle and to the axiom of choice. Within the Zermelo–Fraenkel set theory any one of the three is sufficient to prove the other two.

The *well-ordering principle* states that every non-empty set of positive integers contains a least element that is an element smaller than every other of its elements.

The *axiom of choice* states that for any family $(S_i)_{i\in I}$ of non-empty sets there exists a set $(s_i)_{i\in I}$ such that $s_i \in S_i$ for all $i \in I$. Informally, this says that given any collection of sets, even infinite, each containing at least one element it is possible to construct a new set by arbitrarily choosing one element from each set.

We refer to [**Hal74**, Sections 15, 16, 17] for a comprehensive treatment of these topics. In order to state Zorn's lemma we need to introduce some preliminary notions.

8.2.1. Partially ordered sets. A partially ordered set is a set P equipped with a binary relation $\leq$ that is reflexive ($x \leq x$ for all $x \in P$), antisymmetric (if $x \leq y$ and $y \leq x$ then $x = y$), and transitive (if $x \leq y$ and $y \leq z$ then $x \leq z$).

In a partially ordered set there might be pairs of elements that are not comparable, the word partial is meant to highlight this fact. An ordered set in which every pair of elements is comparable is called a totally ordered set.

A subset P' of a partially order set P is itself a partially ordered set whit respect to the order relation inherited from P. It is called a *chain* of P if it is totally ordered with respect to the inherited order.

An element $m \in P$ is *maximal* if for any $x \in P$ with $x \neq m$ we have $x \leq m$. A partially ordered set may have any number of maximal elements. However, a totally ordered set can have at most one maximal element.

If P' is a subset of P an element $b \in P$ is an *upper bound* of P' if $x \leq b$ for all $x \in P'$. In particular, any element of P' is required to be comparable with b. Note that b is not required to belong to P'.

Lemma 8.20 (Zorn's Lemma)**.** *Let P be a partially ordered set such that any chain of P has an upper bound in P. Then P contains at least one maximal element.*

Proof. See [**Hal74**, Section 16]. □

Now, we are ready to prove that any vector space, also of infinite dimension, has a basis.

Theorem 8.21. *Let E be a K-vector space. If $E \neq \{0\}$ then E admits a basis.*

Proof. Let $\mathcal{F}$ be the set of all linearly independent subsets of E. Note that $\mathcal{F}$ is partially ordered by the inclusion $\subseteq$. Let Y be a subset of $\mathcal{F}$ that is totally ordered

and consider

$$B_Y = \bigcup_{S \in Y} S.$$

Let F be a finite subset of B_Y. Since Y is totally ordered we have that $F \subseteq S$ for some $S \in Y$ which is a linearly independent subset of E. Hence, $B_Y \in \mathcal{F}$. Summing-up B_Y is an element of $\mathcal{F}$ that contains any element of Y that is B_Y is an upper bound for Y.

By Lemma 8.20 there exists and element $M \in \mathcal{F}$ such that if $M \subset S$ for some $S \in \mathcal{F}$ then $M = S$. We now prove that M is a basis of E.

Since $M \in \mathcal{F}$ we already know that M is linearly independent. Assume that there exists a vector $v \in E$ such that $v \notin \langle M \rangle$. Then $v \notin M$. Set $M_v = M \cup \{v\}$. Since $v \notin \langle M \rangle$ we have that $M_v \notin \mathcal{F}$. On the other hand, $M \subseteq M_v$ and $M \neq M_v$ since $v \notin M$, contradicting the maximality of M. We then conclude that M spans E. □

Vector spaces of infinite dimension arise naturally in mathematics. We have already seen the K-vector space $K[x]$ of polynomials in x with coefficients in K a basis of which is given by $\{1, x, x^2, \ldots, x^d, \ldots\}$. Similarly, we can consider the K-vector space $K[x_1, \ldots, x_n]$ of polynomials in several variables with coefficients in K.

Another example is given by the $\mathbb{R}$-vector space $C^h(\mathbb{R})$, where $h \in \mathbb{N}$ is a fixed natural number, of functions $f : \mathbb{R} \to \mathbb{R}$ whose derivatives of order up to h are continuous.

For any natural number $p \geq 1$ the space

$$L^p(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{R} \text{ such that } \left(\int_{-\infty}^{+\infty} |f(x)|^p dx \right)^{1/p} < \infty \right\}$$

is an $\mathbb{R}$-vector space of infinite dimension. These are examples of Banach spaces. We will see later the notion of inner product. One can prove that $L^p(\mathbb{R})$ admits an inner product if and only if $p = 2$, and hence $L^2(\mathbb{R})$ furnishes an example of Hilbert space. Banach and Hilbert spaces are the fundamental objects of study of functional analysis.

Exercises

Exercise 8.22. Tell whether the following vectors of $\mathbb{R}^3$ are linearly independent.
(a) $v_1 = (1, -1, 0)$, $v_2 = (2, 0, 1)$, $v_3 = (1, 1, 1)$.
(b) $v_1 = (1, -1, 0)$, $w = (3, 0, 1)$, $v_3 = (1, 1, 1)$.

Exercise 8.23. Let E be a K-vector space and $w, v_1, \ldots, v_k \in E$, where $v_1, \ldots, v_k$ are linearly independent. Show that $w, v_1, \ldots, v_k$ are linearly independent if and only if $w \notin \langle v_1, \ldots, v_k \rangle$.

Exercise 8.24. In $\mathbb{R}^3$ consider the vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (2, 1, 0)$. Show that these vectors are two by two linearly independent. Show that e_1, e_2, e_3 are not linearly independent (Exercise 5.36). Let $v = (4, 3, 0)$. Write v as a linear combination of e_1, e_2, e_3 in two different ways.

Exercise 8.25. Let E be a finitely generated K-vector space and e_i vectors such that $\langle e_1, \ldots, e_k \rangle = E$. We can assume $e_i \neq 0$, $\forall i$. Let us take e_1 and consider the two vectors e_1, e_2. If e_1 and e_2 are dependent then $\langle e_1, e_2, e_3, \ldots, e_k \rangle = \langle e_1, e_3, \ldots, e_k \rangle$ and we can discard e_2 (and keep e_1). If the two vectors are independent we keep e_1 and e_2. Continuing the procedure, prove the existence of a basis contained in $\{e_1, \ldots, e_k\}$.

Exercise 8.26. Let E be a finite-dimensional K-vector space. We assume $E = F \oplus G$ where F, G are two subvector spaces of E. Let $x_1 = f_1 + g_1$, $\ldots$, $x_n = f_n + g_n$ be vectors of E where $f_i \in F, g_i \in G$. Show that:
1) If $f_1, \ldots, f_n$ are linearly independent, then $x_1, \ldots, x_n$ are linearly independent.
2) Give examples where $x_1, \ldots, x_n$ are linearly independent, $f_1, \ldots, f_n$ are linearly dependent and where $g_1, \ldots, g_n$ are

(a) linearly dependent;

(b) linearly independent.

3) If $x_1, \ldots, x_n$ are linearly independent then $f_1, \ldots, f_n$ are linearly independent $\Leftrightarrow \langle f_1, \ldots, f_n \rangle \cap G = \{0\}$. Is this equivalence still true if $x_1, \ldots, x_n$ are not linearly independent?

Exercise 8.27. Let E be a K-vector space and F, G two subvector spaces of E. Give a necessary and sufficient condition (on F, G) such that $F \cup G$ is a subvector space of E.

Exercise 8.28. (i) Let G be a group and F, H two proper subgroups of G. Show that $F \cup H \neq G$.
(ii) Deduce from (i) that a K-vector space can not be written as the union of two

proper subvector spaces.
(iii) Let $K = F_2$ be the field with two elements and consider $E = K^2$ with its natural structure of K-vector space. Show that E can be written as the union of three proper subvector spaces.

Exercise 8.29. Let E be a finite-dimensional K-vector space. We assume $E = E' \oplus E''$. Let

$$x_1 = x'_1 + x''_1, \ldots, x_n = x'_n + x''_n$$

be vectors of E, with $x'_i \in E'$, $x''_i \in E''$.
(i) Prove that if the x'_i are independent then also the x_i are independent.
(ii) Give examples where $x_1, \ldots, x_n$ are independent, $x'_1, \ldots, x'_n$ are dependent and $x''_1, \ldots, x''_n$ are

(a) linearly dependent;

(b) linearly independent.

(iii) Prove that if $x'_1, \ldots, x'_n$ and $x''_1, \ldots, x''_n$ are dependent then $x_1, \ldots, x_n$ are not necessarily dependent.

Exercise 8.30. Let E be a finite-dimensional K-vector space. Assume $E = E' \oplus E''$. Let $x_1 = x'_1 + x''_1, \ldots, x_n = x'_n + x''_n$ be vectors of E ($x'_i \in E'$, $x''_i \in E''$).

(1) Show that if $x'_1, \ldots, x'_n$ are independent then $x_1, \ldots, x_n$ are also independent.

(2) Give examples where $x_1, \ldots, x_n$ are linearly independent, $x'_1, \ldots, x'_n$ are linearly dependent and where $x''_1, \ldots, x''_n$ are
   (a) linearly dependent;
   (b) linearly independent.

(3) Show that if $x'_1, \ldots, x'_n$ and $x''_1, \ldots, x''_n$ are linearly dependent, then $x_1, \ldots, x_n$ are not necessarily dependent.

Exercise 8.31. Let E be a finite-dimensional K-vector space, and assume that $E = E' \bigoplus E''$. Let $x_1 = x'_1 + x''_1, \ldots, x_n = x'_n + x''_n$ be vectors of E ($x'_i \in E'$, $x''_i \in E''$). It is assumed that $x_1, \ldots, x_n$ are linearly independent, show that: $x'_1, \ldots, x'_n$ are linearly independent $\Leftrightarrow \langle x_1, \ldots, x_n \rangle \cap E'' = \{0\}$.

Chapter 9

Bases, dimension and linear maps

9.1. Dimension

So far we have shown that every finite-dimensional K-vector space, E, admits a basis $\mathcal{B}$. If $\mathcal{B} = (e_1, \ldots, e_n)$ each vector can be written via its coordinates in the basis $\mathcal{B}$. This creates a bijection between E and K^n. What if there was now a basis of E, $\mathcal{C} = (v_1, \ldots, v_t)$ with $t \neq n$? Could a vector be identified with n coordinates or with t coordinates?

Fortunately, as we will see, the problem does not arise. Let us highlight a result that we have in a certain way already used.

Lemma 9.1. *Let E be a K-vector space and $g_1, \ldots, g_t, u$ vectors of E.*

(1) *If $u \in \langle g_1, \ldots, g_t \rangle$ and if g_1 is a linear combination of $u, g_2, \ldots, g_t$ (i.e. $g_1 \in \langle u, g_2, \ldots, g_t \rangle$), then $\langle g_1, g_2, \ldots, g_t \rangle = \langle u, g_2, \ldots, g_t \rangle$.*

(2) *In particular if $E = \langle g_1, \ldots, g_t \rangle$ and $g_1 \in \langle u, g_2, \ldots, g_t \rangle$ then $E = \langle u, g_2, \ldots, g_t \rangle$.*

Proof. 1) We show $A := \langle u, g_2, \ldots, g_t \rangle \subset \langle g_1, g_2, \ldots, g_t \rangle =: B$. A vector $w \in A$ is a linear combination of $u, g_2, \ldots, g_n$. Since u is a linear combination of the g_i, w is a linear combination of the g_i, and hence $A \subset B$.

Take now $v \in B$. Then v is a linear combination of the g_i. Since g_1 is a linear combination of $u, g_2, \ldots, g_n$, we can express v as a linear combination of $u, g_2, \ldots, g_n$.

Finally 2) follows from 1). □

Proposition 9.2. *Let E be a K-vector space. Let $g_1, \dots, g_t$ be vectors of E such that $\langle g_1, \dots, g_t \rangle = E$ (the vectors g_i generate E) and let $e_1, \dots, e_n$ be vectors of E such that $e_1, \dots, e_n$ are linearly independent. Then $n \leq t$.*

Proof. Since $\langle g_1, \dots, g_t \rangle = E$, e_1 is a linear combination of the vectors g_i: $e_1 = \sum \alpha_i g_i$. One of the α_i is non-zero otherwise we would have $e_1 = 0$ and this is not possible because the vectors $e_1, \dots, e_n$ are independent. By rearranging the indexes if needed we can assume $\alpha_1 \neq 0$. Then we have $g_1 = e_1 - \sum_{i>1}(\alpha_i/\alpha_1)g_i$. That is, $g_1 \in \langle e_1, g_2, \dots, g_t \rangle$. By Lemma 9.1, $E = \langle e_1, g_2, \dots, g_t \rangle$, that is, we replaced one of the generators g_j with one of the e_i. We can repeat this operation until we have used all the e_i.

In fact, suppose we have $E = \langle e_1, \dots, e_{i-1}, g_i, \dots, g_t \rangle$. Then $e_i = \sum_{k=1}^{i-1} \lambda_k e_k + \sum_{k=i}^{t} \mu_k g_k$. We observe that one of the μ_k is non-zero. In fact, otherwise we would have that e_i is a linear combination of $e_1, \dots, e_{i-1}$ but this is not possible because the e_i are linearly independent. By rearranging the indexes if needed we can assume $\mu_i \neq 0$. This implies that g_i is a linear combination of $e_1, \dots, e_i, g_{i+1}, \dots, g_t$. By Lemma 9.1 we have $E = \langle e_1, \dots, e_i, g_{i+1}, \dots, g_t \rangle$.

If $n > t$ we get $E = \langle e_1, \dots, e_t \rangle$ with e_{t+1} linear combination of $e_1, \dots, e_t$. This is not possible because the e_i are independent. So $n \leq t$. □

Theorem 9.3. *Let E be a K-vector space and let $(e_1, \dots, e_n)$, $(v_1, \dots, v_p)$ be two bases of E. Then $n = p$. So all bases of E have the same number of elements.*

Proof. Since the e_i generate E and since the v_j are independent, we have (Proposition 9.2) $n \geq p$. But we also have that the v_j generate E and the e_i are independent, so $p \geq n$. □

This allows us to define the notion of dimension.

Definition 9.4. Let E be a finitely generated K-vector space. The *dimension* of E (over K) is the number of vectors of a basis of E.

Remark 9.5. This definition makes perfect sense because every K-vector space of finite dimension admits a basis (Corollary 8.16, Observation 8.17) and because two bases of E have the same number of elements (Theorem 9.3).

The dimension of E (over K) is denoted by $\dim_K(E)$ (or simply by $\dim E$).

The unique basis of $E = \{0\}$ is the empty set that has zero elements so $\dim\{0\} = 0$ and this reflects our intuition that a point has dimension zero (while a line has dimension one, a plane has dimension two, and so on).

Example 9.6 (Canonical basis of K^n)**.** In K^n consider the following vectors: $e_1 = (1, 0, \dots, 0), \dots, e_i = (0, \dots, 0, 1, 0, \dots 0), \dots, e_n = (0, \dots, 0, 1)$. These vectors are

independent because $\sum \lambda_i e_i = 0 \Leftrightarrow (\lambda_1, \ldots, \lambda_n) = (0, \ldots, 0) \Leftrightarrow \lambda_i = 0, \forall i$. Furthermore, since $(\lambda_1, \ldots, \lambda_n) = \sum \lambda_i e_i$, the vectors e_i generate K^n. So $\mathcal{B} = (e_1, \ldots, e_n)$ is a basis of K^n. In particular, $\dim K^n = n$.

The basis $\mathcal{B} = (e_1, \ldots, e_n)$ is called the *canonical* basis of K^n. This comes from the fact that if we take a vector $(\lambda_1, \ldots, \lambda_n) \in K^n$ we just need to look at it to get its coordinates in the basis $\mathcal{B}$.

The following result is very useful in practice.

Proposition 9.7. *Let E be a K-vector space of dimension n.*

(1) *If $p > n$, p vectors of E are always linearly dependent;*

(2) *n vectors of E form a basis of E if and only if they are linearly independent.*

Proof. (1) Let $p > n$ and assume $v_1, \ldots, v_p$ linearly independent. By the Incomplete Basis Theorem (Theorem 8.19) there exists a basis of E containing the v_i. This is a contradiction because all bases of E have n elements. So the v_i are dependent.

(2) If $e_1, \ldots, e_n$ are independent, then $\forall w \in E$ the vectors $e_1, \ldots, e_n, w$ are dependent (by (1)). Therefore (Lemma 8.7) w is a linear combination of the e_i, i.e. $\langle e_1, \ldots, e_n \rangle = E$. □

The following is another very useful result.

Proposition 9.8. *Let E be a K-vector space of dimension n and $F \subset E$ a subvector space of E. Then $\dim F$ is finite and $\dim F \leq \dim E$, with equality if and only if $F = E$.*

Proof. If $F = \{0\}$ the result is trivial. Then consider the case $F \neq \{0\}$. If $f_1, \ldots, f_p$ are p vectors of F with $p > n$, then $f_1, \ldots, f_p$ are linearly dependent. Let r be the maximum number of independent vectors of F, if $v_1, \ldots, v_r$ are r independent vectors of F, then these vectors are a maximal system of independent vectors of F hence a basis of F. This shows $\dim F \leq n$.

If $\dim F = n$ a basis of F is also a basis of E (Proposition 9.7) then $F = E$ (every $u \in E$ is a linear combination of the basis vectors and therefore it belongs to F). □

Remark 9.9. We therefore have two ways to show that two subvector spaces $F, G \subset E$ are equal:

(1) Show that $F \subset G$ and $G \subset F$;

(2) Show that $F \subset G$ and $\dim F = \dim G$.

9.2. Bases and linear maps

Regarding linear maps there is only one thing to know.

Theorem 9.10. *Let E, F be two K-vector spaces, $\mathcal{B} = (e_1, \ldots, e_n)$ a basis of E and $f_1, \ldots, f_n$ some vectors of F. Then*

(1) *there exists one and only one linear map $\varphi : E \to F$ such that $\varphi(e_i) = f_i, \forall i, 1 \leq i \leq n$;*

(2) *the map φ is injective if and only if the vectors f_i are independent;*

(3) *the map φ is surjective if and only if $\langle f_1, \ldots, f_n \rangle = F$ (the vectors f_i generate F);*

(4) *the map φ is bijective if and only if $(f_1, \ldots, f_n)$ is a basis of F.*

Proof. (1) Let $v \in E$, then v can be written in the basis $\mathcal{B}$, that is: $v = \sum_{i=1}^{n} \alpha_i e_i$. Since we want φ to be linear we set $\varphi(v) = \sum_i \alpha_i \varphi(e_i) = \sum_i \alpha_i f_i$. In conclusion we define $\varphi : E \to F$ as follows if $v = \sum \alpha_i e_i$, then $\varphi(v) = \sum \alpha_i f_i$.

This is well defined since each $v \in E$ is written uniquely as a linear combination of the e_i. The application is clearly linear.

Uniqueness of φ: let $\psi : E \to F$ be a linear map such that $\psi(e_i) = f_i$. If $v \in E, v = \sum \lambda_i e_i$, $\psi(v) = \psi(\sum \lambda_i e_i)$, but by linearity of ψ we have $\sum_i \lambda_i \psi(e_i)$. Since $\psi(e_i) = f_i$, we get $\psi(v) = \sum \lambda_i f_i$. But this last expression is nothing but $\varphi(v)$. We conclude that $\psi = \varphi$.

(2) Suppose φ is injective and show that the vectors f_i are linearly independent. Let $\sum_i \alpha_i f_i = 0$. This can be rewritten as follows: $\sum_i \alpha_i \varphi(e_i) = 0$. By linearity of φ: $\sum_i \alpha_i \varphi(e_i) = \varphi(\sum_i \alpha_i e_i)$. So the vector $u := \sum_i \alpha_i e_i$ belongs to $\mathrm{Ker}(\varphi)$. But since φ is injective, $\mathrm{Ker}(\varphi) = \{0\}$. So $u := \sum_i \alpha_i e_i = 0$. Since the vectors e_i are independent, this implies $\alpha_i = 0, \forall i$. In conclusion $\sum_i \alpha_i f_i = 0$ implies $\alpha_i = 0, \forall i$, this shows that the vectors f_i are linearly independent.

Now we assume the vectors f_i to be linearly independent. Let $v \in \mathrm{Ker}(\varphi)$, then $v = \sum \alpha_i e_i$ and $\varphi(v) = \sum \alpha_i f_i = 0$. Since the vectors f_i are independent this implies $\alpha_i = 0, \forall i$. So $v = 0$, i.e. $\mathrm{Ker}(\varphi) = \{0\}$ and φ is injective.

(3) We show that φ surjective implies $\langle f_1, \ldots, f_n \rangle = F$. Let $f \in F$. Since φ is surjective, there exists $v \in E$ such that $\varphi(v) = f$. If $v = \sum \alpha_i e_i$, $\varphi(v) = \sum \alpha_i f_i$. So $f = \sum \alpha_i f_i$. This shows that every vector of F is linear combination of $f_1, \ldots, f_n$.

Conversely we show that $\langle f_1, \ldots, f_n \rangle = F$ implies that φ is surjective. Let $f \in F$, by hypothesis $f = \sum \alpha_i f_i$. Since $f_i = \varphi(e_i)$ we can write $f = \sum \alpha_i \varphi(e_i)$. By linearity of φ: $f = \varphi(\sum \alpha_i e_i)$. This shows that every vector of F belongs to $\mathrm{Im}(\varphi)$, so φ is surjective.

(4) is a consequence of (2) and (3). □

Remark 9.11. Point (1) of the Theorem 9.10 can be reformulated as follows: a linear map is completely determined by its values on the elements of a basis.

Remember that an isomorphism $f : E \to F$ is a bijective linear map (we know that f^{-1} is linear).

Corollary 9.12. (1) *Two K-vector spaces are isomorphic if and only if they have the same dimension.*

(2) *Every K-vector space of dimension n is isomorphic to K^n.*

Proof. (1) Suppose E and F are isomorphic. If $E = \{0\}$ then $F = \{0\}$ and $\dim E = \dim F$. If $E \neq \{0\}$ let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E. Then if $f : E \to F$ is an isomorphism, by Theorem 9.10 (4), $(f(e_1), \ldots, f(e_n))$ is a basis of F. So $n = \dim E = \dim F$.

Suppose $\dim E = \dim F = n > 0$. Let $(e_1, \ldots, e_n)$ be a basis of E, $(f_1, \ldots, f_n)$ a basis of F and define $\varphi : E \to F$ by $\varphi(e_i) = f_i, \forall i$. From Theorem 9.10 it follows that φ is linear and bijective, so E and F are isomorphic.

(2) Follows from (1). □

The Corollary 9.12 completes the *classification* of the K-vector spaces of finite dimension on a field K, that is, it describes, up to isomorphisms, all K-vector spaces.

Completing the classification of a class of mathematical objects is always a great achievement for mathematics.

However, bear in mind that the classification of finite groups has not been completed yet!

9.3. Rank theorem and Grassmann relation

Needless to say, the two main results of this section are fundamental. Let us start with a simple lemma.

Lemma 9.13. *Let E, F be two K-vector spaces and $f : E \to F$ a linear map. Let $A \subset E$ such that $\langle A \rangle = E$, then $f(A)$ generates* $\mathrm{Im}(f)$.

Proof. Let $v = f(u) \in \mathrm{Im}(f)$. We have $u = \sum \alpha_i w_i$ where $w_i \in A$, $\forall i$ because, by hypothesis, every vector of E is a linear combination of vectors of A, so $v = f(u) = \sum \alpha_i f(w_i)$ and each element of $\mathrm{Im}(f)$ is a linear combination of elements of $f(A)$. □

Theorem 9.14 (Rank theorem)**.** *Let $f : E \to F$ be a linear map between two K-vector spaces, E is assumed to be of finite dimension. Then* $\mathrm{Im}(f)$ *has finite dimension and:*

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

Proof. If $(e_1, \ldots, e_n)$ is a basis of E, by Lemma 9.13, the vectors $f(e_1), \ldots, f(e_n)$ generate $\mathrm{Im}(f)$, so $\mathrm{Im}(f)$ has finite dimension.

Let $(v_1, \ldots, v_p)$ be a basis of $\mathrm{Ker}(f)$ (the basis is empty if f is injective). By the Incomplete Basis Theorem there exists a basis $\mathcal{B}$ containing the vectors $v_1, \ldots, v_p$: $\mathcal{B} = (v_1, \ldots, v_p, v_{p+1}, \ldots, v_n)$. We have $\langle f(v_i) \rangle = \langle f(v_{p+1}, \ldots, f(v_n) \rangle = \mathrm{Im}(f)$ (because $f(v_i) = 0$ if $1 \leq i \leq p$). We show that $f(v_{p+1}), \ldots, f(v_n)$ form a basis of $\mathrm{Im}(f)$. This will prove the theorem because $\dim E = n$, $\dim \mathrm{Ker}(f) = p$ and $\dim \mathrm{Im}(f) = n - p$.

To show that $f(v_{p+1}), \ldots, f(v_n)$ form a basis of $\mathrm{Im}(f)$ it is enough to show that they are linearly independent because we already know that they generate $\mathrm{Im}(f)$. Let $\alpha_{p+1} f(v_{p+1} + \cdots + \alpha_n f(v_n) = 0$ be a linear combination equal to the zero vector, we must show $\alpha_i = 0, \forall i$. By linearity of f we have $0 = \alpha_{p+1} f(v_{p+1} + \cdots \alpha_n f(v_n) = f(\alpha_{p+1} v_{p+1} + \cdots + \alpha_n v_n)$. Therefore, $w = \alpha_{p+1} v_{p+1} + \cdots + \alpha_n v_n \in Ker(f)$. Since $(v_1, \ldots, v_p)$ is a basis of $\mathrm{Ker}(f)$, we have $w = \alpha_1 v_1 + \cdots + \alpha_p v_p$. Putting it all together $\alpha_{p+1} v_{p+1} + \cdots + \alpha_n v_n = w = \alpha_1 v_1 + \cdots + \alpha_p v_p$. In conclusion $\alpha_1 v_1 + \cdots + \alpha_p v_p - (\alpha_{p+1} v_{p+1} + \cdots + \alpha_n v_n) = 0$. Since the vectors $v_1, \ldots, v_n$ are independent this implies $\alpha_1 = \cdots = \alpha_n = 0$. In particular, $\alpha_{p+1} = \cdots = \alpha_n = 0$ and the theorem is proved. □

Remark 9.15. In Theorem 9.14 F is not assumed to be of finite dimension. The rank theorem is also called the rank-nullity theorem.

Corollary 9.16. *Let $f : E \to F$ be a linear map between two vector spaces with* $\dim E = \dim F$*. The following are equivalent:*

(1) *f is injective;*

(2) *f is surjective;*

(3) *f is bijective.*

Proof. (1) $\Rightarrow$ (2). If f is injective then $\mathrm{Ker}(f) = \{0\}$ and by Theorem 9.14 $\dim F = \dim E = \dim \mathrm{Im}(f)$. So $\mathrm{Im}(f) \subset F$ and $\dim F = \dim \mathrm{Im}(f)$. We conclude (Remark 9.9) $F = \mathrm{Im}(f)$, so f is surjective.

(2) $\Rightarrow$ (1) If f is surjective $\dim \mathrm{Im}(f) = \dim F = \dim E$ and from Theorem 9.14 it follows that $\dim \mathrm{Ker}(f) = 0$, so f is injective.

So (1) and (2) are equivalent and it is clear that (1) and (2) $\Leftrightarrow$ (3). □

Remark 9.17. Actually you do not need the Rank Theorem to prove this corollary. Just use Theorem 9.10. Suppose f is injective. Let $(e_1, \ldots, e_n)$ be a basis of E. Since f is injective, the vectors $f(e_1), \ldots, f(e_n)$ are independent (Theorem 9.10). We therefore have n independent vectors in a space of dimension n, these vectors form a basis of F. Therefore f is bijective (Theorem 9.10).

Suppose f is surjective. The vectors $f(e_1), \ldots, f(e_n)$ generate F. We have n vectors that generate a space of dimension n, these vectors are independent. In fact, otherwise one of them would be a linear combination of the others and F would be generated by less than n vectors, but this is impossible (because every system of generators contains a basis). So $(f(e_1), \ldots, f(e_n))$ is a basis of F and f is bijective (Theorem 9.10).

For the next application we need a lemma that is nevertheless interesting in itself.

Lemma 9.18. *Let E, F be two K-vector spaces of finite dimension. We have*

$$\dim E \times F = \dim E + \dim F.$$

Proof. Let $(e_1, \ldots, e_n)$ be a basis of E and let $(v_1, \ldots, v_p)$ be a basis of F. It is easily verified that the vectors $(e_1, 0), \ldots, (e_n, 0), (0, v_1), \ldots, (0, v_p)$ form a basis of $E \times F$. □

Theorem 9.19 (Grassmann relation)**.** *Let E be a K-vector space of finite dimension and $F, G \subset E$ two subvector spaces of E. Then*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Proof. Consider the map $\varphi : F \times G \to E : (f, g) \to f + g$. It is easily verified that φ is linear. Clearly $\mathrm{Im}(\varphi) = F + G$. If $(f, g) \in \mathrm{Ker}(\varphi)$, then $f = -g$. So $f \in F \cap G$. In conclusion $\mathrm{Ker}(\varphi) = \{(f, -f) \mid f \in F \cap G\}$ (it is clear that if $f \in F \cap G$, $(f, -f) \in \mathrm{Ker}(\varphi)$). It follows that $\dim \mathrm{Ker}(\varphi) = \dim(F \cap G)$ (a basis of $F \cap G$ gives a basis of $\mathrm{Ker}(\varphi)$). By Theorem 9.14 $\dim(F \times G) = \dim(F + G) + \dim(F \cap G)$. The result follows from Lemma 9.18. □

Theorem 9.19 can be proved in another way (without using the Rank Theorem).

Alternative proof of the Theorem 9.19. Let $e_1, \ldots, e_p$ be a basis of $F \cap G$. The vectors $e_1, \ldots, e_p$ are p independent vectors of F and therefore, by the Incomplete Basis Theorem, they can be completed to a basis $(e_1, \ldots, e_p, f_{p+1}, \ldots, f_n)$ of F. We observe that $f_i \notin F \cap G$. If this were not the case, we would have $f_i = \sum \lambda_j e_j$ and the vectors $e_1, \ldots, e_p, f_{p+1}, \ldots, f_n$ would not be independent.

In the same way the vectors $e_1, \ldots, e_p$ are p independent vectors of G and therefore, by the Incomplete Basis Theorem, they can be completed to a basis of G: $(e_1, ..., e_p, g_{p+1}, \ldots, g_m)$.

As before $g_j \notin F \cap G$. The vectors $e_1, \ldots, e_p, f_{p+1}, \ldots, f_n, g_{p+1}, \ldots, g_m$ are $p + (n-p) + (m-p)) = (n+m) - p = \dim F + \dim G - \dim(F \cap G)$. It remains to show that $\mathcal{B} = (e_1, \ldots, e_p, f_{p+1}, \ldots, f_n, g_{p+1}, \ldots, g_m)$ is a basis of $F + G$.

It is clear that $\langle e_1, \ldots, e_p, f_{p+1}, \ldots, f_n, g_{p+1}, \ldots, g_m \rangle = F + G$ because $F + G = \{f + g \mid f \in F, g \in G\}$. So f is a linear combination of the basis vectors of F and g is a linear combination of the basis vectors of G, putting everything together we get a vector of $\langle e_1, \ldots, e_p, f_{p+1}, \ldots, f_n, g_{p+1}, \ldots, g_m \rangle$. Vice versa it is clear that a vector of this space is a vector of $F + G$. It remains to show that the vectors are independent.

Let $\sum_i \lambda_i e_i + \sum_j \mu_j f_j + \sum_k \alpha_k g_k = 0$. We have $w := \sum_i \lambda_i e_i + \sum_j \mu_j f_j = -\sum_k \alpha_k g_k$. Since $-\sum_k \alpha_k g_k \in G$ and $\sum_i \lambda_i e_i + \sum_j \mu_j f_j \in F$, we see that $w \in F \cap G$, so w is linear combination of the vectors e_i and we have: $\sum_i \lambda_i e_i + \sum_j \mu_j f_j = \sum_i \beta_i e_i$. So $\sum_i (\lambda_i - \beta_i) e_i + \sum_j \mu_j f_j = 0$. Since the vectors (e_i, f_j) are independent (they form a basis of F) we obtain $\alpha_i = \beta_i$ and $\mu_j = 0$, $\forall j$. Going back to the initial relation we have $\sum_i \lambda_i e_i + \sum_k \alpha_k g_k = 0$. Since (e_i, g_k) is a basis of G we get $\lambda_i = 0$, $\forall i$ and $\alpha_k = 0$, $\forall k$ and the theorem is proved. □

9.4. Ring of endomorphisms, linear group

Let us start with a definition.

Definition 9.20. If E, F are two K-vector spaces we denote by $\mathcal{L}_K(E, F)$ (or $\mathcal{L}(E, F)$ if there is no risk of confusion) the set of linear maps from E to F:

$$\mathcal{L}_K(E, F) := \{f : E \to F \mid f \text{ is } K - \text{linear}\}.$$

Now, we define an addition on $\mathcal{L}(E, F)$ via $f + g : E \to F : v \to f(v) + g(v)$ and an external multiplication with $\lambda f : E \to F : v \to \lambda f(v)$.

Remark 9.21. It can be verified that with these two operations $\mathcal{L}(E, F)$ is a K-vector space.

If E and F have finite dimension, it is legitimate to ask whether $\mathcal{L}(E, F)$ also has finite dimension.

As we have seen a linear map is completely determined by its values on the elements of a basis. Therefore, let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E. A linear map $f \in \mathcal{L}(E, F)$ is therefore completely determined by the vectors $f(e_1), \ldots, f(e_n) \in F$.

A vector of F is completely determined by its coordinates in a basis of F. Let $\mathcal{C} = (v_1, \ldots, v_p)$ be a basis of F and write the vectors $f(e_i)$ in this basis:

$$\begin{aligned} f(e_1) = &\quad \alpha_{11}v_1 + \alpha_{21}v_2 + \cdots + \alpha_{p1}v_p; \\ f(e_2) = &\quad \alpha_{12}v_1 + \alpha_{22}v_2 + \cdots + \alpha_{p2}v_p; \\ &\vdots \\ f(e_n) = &\quad \alpha_{1n}v_1 + \alpha_{2n}v_2 + \cdots + \alpha_{pn}v_p. \end{aligned}$$

We conclude that the map f is completely determined by the np scalar α_{ij}. We can arrange these scalars in a matrix like this:

$$M = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pn} \end{pmatrix}.$$

The j-th column of the matrix contains the coordinates of $f(e_j)$ in the basis $\mathcal{C}$.

The matrix M is the matrix associated to f with respect to the bases $\mathcal{B}$ of E and $\mathcal{C}$ of F.

In conclusion, the matrix M completely determines the linear application f. Conversely, a matrix N with p rows and n columns determines a linear application $g : E \to F$. If $N = (\beta_{ij})$, then g is defined as $g(e_j) = \sum_{i=1}^{p} \beta_{ij} v_i$. We have therefore defined a map:

$$\mathrm{Mat}(-; \mathcal{B}, \mathcal{C}) : \mathcal{L}(E, F) \to M_{p,n}(K) : f \to M.$$

Since a linear map is uniquely determined by its values on the vectors of a basis and since a vector is uniquely determined by its coordinates with respect to a basis, we conclude that the map $\mathrm{Mat}(-; \mathcal{B}, \mathcal{C})$ is bijective. Furthermore, it is easy to verify that the application is linear. Taking into account that $M_{p,n}(K)$ is isomorphic to K^{pn} and therefore has dimension np, we have proved:

Proposition 9.22. *With the previous notations $\mathcal{L}_K(E, F)$ is isomorphic to $M_{p,n}(K)$. In particular*

$$\dim \mathcal{L}(E, F) = \dim E \times \dim F.$$

Let us now move on to the special case $E = F$.

Definition 9.23. A linear map $f : E \to E$ is also called an *endomorphism* of E and $\mathcal{L}(E, E)$ is denoted by $\mathrm{End}(E)$.

This case is interesting because on $\mathrm{End}(E)$ there is a further operation: $\circ$, the composition of maps. In fact, if $f, g \in \mathrm{End}(E)$, $f \circ g \in \mathrm{End}(E)$. The neutral element for this operation is the identity, the composition of maps is associative

and distributive with respect to the addition: $f \circ (g + h) = f \circ g + f \circ h$ and $(f + g) \circ h = f \circ h + g \circ h$. This tells us that $(\mathrm{End}(E), +, \circ)$ is a ring.

In general this ring is not commutative ($\exists f, g \in \mathrm{End}(E)$ with $f \circ g \neq g \circ f$) and it is not integral ($\exists f \neq 0, g \neq 0$ with $f \circ g = 0$).

So we have two algebraic structures on $\mathrm{End}(E)$:

(1) $(End(E), +, \cdot_\lambda)$ is a K-vector space (here $\cdot_\lambda$ denotes external multiplication).

(2) $(End(E), +, \circ)$ is a ring.

Furthermore, the three operations are compatible with each other $(\lambda(f \circ g) = (\lambda f) \circ g = f \circ (\lambda g)$. This gives to $\mathrm{End}(E)$ a structure of K-algebra.

Among the endomorphisms of E we can distinguish the bijective ones, such endomorphisms are called *automorphisms*.

Definition 9.24. An *automorphism* of the K-vector space E is a bijective endomorphism of E, i.e. a bijective linear map $f : E \to E$.

The set of automorphisms of E is denoted by $\mathrm{Aut}(E)$ or also $Gl(E)$.

The composition of maps is an internal composition law on $\mathrm{Aut}(E) = Gl(E)$. It is verified that $(Gl(E), \circ)$ is a group; this group is called the *General linear group of* E.

Finally, the last interesting special case is when $F = K$. That is, we are considering $\mathcal{L}(E, K)$. This space is called the *dual* of E and is denoted by E^*. We will talk about it in the next chapter.

Exercises

Exercise 9.25. Let E be a K-vector space and $F \subset E$ subvector space. Show that F admits a supplementary i.e. there exists a subvector space $G \subset E$ such that $E = F \oplus G$ (Hint: take a basis of F and use the Incomplete Basis Theorem).

In the real plane $E = \mathbb{R}^2$ describe the supplementary spaces of the line $F = \{(x, y) \mid x = 0\}$.

Exercise 9.26. We denote by $\mathbb{F}_2$ the field $\mathbb{Z}/2\mathbb{Z}$. Let $E = \mathbb{F}_2 \times \mathbb{F}_2$. In this exercise E is considered as a $\mathbb{F}_2$-vector space.

(1) What is the dimension of E?

(2) What is the cardinality of E?

(3) Let v, w be any two non-zero vectors of E. Show that v and w are linearly independent.

(4) List all subvector spaces of E.

(5) Let $u \in E$, $u = (1, 0)$ and let $F = \langle u \rangle$. Make a list of all the supplementary of F.

Exercise 9.27. In $\mathbb{R}^3$ let $F = \{(x, y, z) \mid x + y + z = 0 \text{ and } 2x - z = 0\}$.
1) Prove that F is a subvector space of $\mathbb{R}^3$. Determine a basis of F (and therefore the dimension of F).
2) Find a supplementary of F (let $\mathcal{C}$ be the canonical basis and $\mathcal{B}$ the basis of F from the previous point, then $\mathcal{C} \cup \mathcal{B}$ generates $\mathbb{R}^3$. Find a basis of $\mathbb{R}^3$ containing $\mathcal{B}$).

Exercise 9.28. Let E be a K-vector space. We assume $E = F \oplus G$ where $F, G \subset E$ are two subvector spaces of E. Let $(f_1, \ldots, f_r)$ be a basis of F and $(g_1, \ldots, g_t)$ a basis of G. Show that $(f_1, \ldots, f_r, g_1, \ldots, g_t)$ is a basis of E.

Exercise 9.29. Let $K[x]_{\leq n}$ be the subvector space of $K[x]$ of polynomials of degree at most n (by convention the zero polynomial has degree $-\infty$).
1) Determine the dimension of $K[x]_{\leq n}$.
2) Let $V \subset K[x]_{\leq n}$, $V = \{a_0 + a_1 x + \cdots + a_n x^n \in K[x]_{\leq n} \mid \sum_{0 \leq i \leq n} a_i = 0\}$. Show that V is a subvector space of $K[x]_{\leq n}$ and determine its dimension (begin with the case $n = 2$ and then generalize).

Exercise 9.30. Does there exist a linear map $g : \mathbb{R}^2 \to \mathbb{R}^3$ such that $g((1, 1)) = (1, 0, 3)$, $g((0, -2)) = (4, -1, 0)$ and $g((2, 6)) = (-6, 3, 6)$?

Exercise 9.31. Let $f : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \to (x + 2y + z, -x + y, y - z)$.
1) Show that f is linear.

2) State whether f is surjective (consider $f(e_i)$ where $\mathcal{C} = (e_i)$ is the canonical basis of $\mathbb{R}^3$).

Exercise 9.32. In $\mathbb{R}^3$ let $F = \{(x, y, z) \mid 2x - y + z = 0\}$.
1) Consider $f : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to 2x - y + z$ and conclude that f is a subvector space of $\mathbb{R}^3$.
2) Prove that there exists a linear map $g : \mathbb{R}^3 \to \mathbb{R}^3$ such that $\mathrm{Ker}(g) = F$ and $g((1, 0, 0)) = (1, 1, 1)$. Determine $\dim(\mathrm{Im}(g))$.

Exercise 9.33. Redo point 2) of Exercise 9.29 using the Rank Theorem.

Exercise 9.34. Let $M_n(K)$ be the vector space of $n \times n$ matrices with coefficients in K. Let $\mathcal{T} = \{A \in M_n(K) \mid A = (a_{ij} \text{ and } Tr(A) := a_{11} + \cdots + a_{nn} = 0\}$. So $\mathcal{T}$ is the set of matrices whose sum of the elements on the diagonal is zero. The sum of the elements on the diagonal of A is called the *trace* and is denoted by $Tr(A)$. Prove that $\mathcal{T}$ is a subvector space of $M_n(K)$ and determine its dimension (use the Rank Theorem).

Exercise 9.35. Redo Exercise 9.31 using the Rank Theorem.

Exercise 9.36. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$, $g : \mathbb{R}^2 \to \mathbb{R}^3$ be two linear maps. Show that $g \circ f : \mathbb{R}^3 \to \mathbb{R}^3$ is neither injective nor surjective. Can $f \circ g : \mathbb{R}^2 \to \mathbb{R}^2$ be injective, surjective, bijective?

Exercise 9.37. Let E be a K-vector space and $F \subset E$ a subvector space. Let $H \subset E$ be a hyperplane (i.e. $\dim(H) = \dim(E) - 1$). Show that if $F \subset H$ then $\dim(H \cap F) = \dim(F)$, while if $F \not\subset H$, $\dim(H \cap F) = \dim(F) - 1$ (use the Grassmann relation).

Exercise 9.38. Let E be a finite-dimensional K-vector space of dimension at least two. Show that there exists $f \in \mathrm{End}(E)$ such that $f \neq 0$ and $f^2 := f \circ f = 0$ (use 9.10 Theorem). This shows that the ring $\mathrm{End}(E)$ is not integral.

Exercise 9.39. Let E be a K-vector space of dimension at least two. Show that the ring $(\mathrm{End}(E), +, \circ)$ is not commutative.

Exercise 9.40. Let E be a K-vector space. Show that $(Gl(E), +)$ is not a group.

Exercise 9.41. Let $\varphi : E \to F$ be a linear map between K-vector spaces. Let E_1, E_2 be two subvector spaces of E, such that the sum $E_1 + E_2$ is direct.

(1) Show that $(E_1 \oplus E_2) \cap \mathrm{Ker}(\varphi) = \{0\} \Rightarrow$ the sum $\varphi(E_1) + \varphi(E_2)$ is direct.

(2) Is the vice versa true? That is, if E_1, E_2 and $\varphi(E_1), \varphi(E_2)$ are in direct sum, then $(E_1 \oplus E_2) \cap \mathrm{Ker}(\varphi) = \{0\}$?

Exercise 9.42. Let f, g be two endomorphisms of the K-vector space E. Show that: $\mathrm{rk}(f \circ g) \geq \mathrm{rk}(f) + \mathrm{rk}(g) - \dim(E)$, where $\mathrm{rk}(f)$ is the dimension of $\mathrm{Im}(f)$ (consider the restriction of f to the image of g).

Exercise 9.43. Let E be a K-vector space, $f \in \mathrm{End}(E)$ an endomorphism of E and $V \subset E$ a subvector space.
(i) Show that $f^{-1}(V)$ is a subvector space of E.
(ii) Show that $\dim(f^{-1}(V)) \leq \dim(\mathrm{Ker}(f)) + \dim(V)$.
(iii) Let h, g be two endomorphisms of E. Describe $\mathrm{Ker}(g \circ h)$ and using (ii) show that

$$\dim(\mathrm{Ker}(g \circ h)) \leq \dim(\mathrm{Ker}(g)) + \dim(\mathrm{Ker}(h)).$$

Conclude that $\mathrm{rk}(g \circ h) \geq \mathrm{rk}(g) + \mathrm{rk}(h) - \dim(E)$.

Exercise 9.44. (i) Let E be a K-vector space, $f \in \mathrm{End}(E)$ an endomorphism of E and $V \subset E$ a subvector space.
Show that $\dim(f^{-1}(V)) = \dim(\mathrm{Ker}(f)) + \dim(\Im(f) \cap V)$. Find the inequality from Exercise 9.43.
(ii) Let h, g be two endomorphisms of E. Show that:

$$\mathrm{rk}(g \circ h) = \mathrm{rk}(h) - \dim(\mathrm{Ker}(g) \cap \Im(h)).$$

Find the inequality from Exercise 9.43.

Exercise 9.45. Let E be a K-vector space of dimension n and $f \in \mathrm{End}(E)$.
(i) The following are equivalent:

(1) $\mathrm{Im}(f) = \mathrm{Ker}(f)$;

(2) $f^2 = 0$ and $\dim \mathrm{Im}(f) = \frac{n}{2}$ (in particular n is even).

(ii) Infer from (i) that the following are equivalent:

(1) $\mathrm{Im}(f) = \mathrm{Ker}(f)$;

(2) there exists a basis $\mathcal{B}$ of E such that

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix}$$

where A is an invertible matrix of order $\frac{n}{2}$ and where the zeros indicate null matrices of order $\frac{n}{2}$ (in particular n is even).

Exercise 9.46. Let E be a K-vector space and f an endomorphism of E. Show that:

$$\mathrm{rk}(f) = \mathrm{rk}(f^2) \Leftrightarrow \mathrm{Im}(f) \oplus \mathrm{Ker}(f) = E.$$

Exercise 9.47. Let E be a K-vector space and $f \in \mathrm{End}(E)$. For every integer $m \geq 0$ we set $J_m = \mathrm{Im}(f^m)$, $K_m = \mathrm{Ker}(f^m)$ (by convention $f^0 = Id$).

(1) Show that $J_p \subset J_{p-1}$, $\forall p \geq 1$. Observe that the sequence $(J_m)_{m\geq 0}$ is stationary (i.e. there exists K such that $J_m = J_k$ if $m \geq k$). More precisely, let $s := \min\{p \mid J_p = J_{p+1}\}$, show that $J_p = J_s$, $\forall p \geq s$. Observe that $s \leq n$.

(2) Show that $K_{m-1} \subset K_m$, $\forall m \geq 1$ and that the sequence $(K_m)_{m\geq 0}$ is stationary.

(3) Show that $K_s = K_{s+1}$ (s as in (1)); furthermore $K_{r-1} \neq K_r$ if $s \geq r$.

(4) Show that $E = K_s \oplus J_s$.

(5) Show that $K_m \subsetneqq K_{m+1} \Rightarrow K_m \cap J_m \neq \{0\}$. Conclude that $s = \min\{m \mid K_m \oplus J_m = E\}$. Deduce that $\mathrm{rk}(f) = \mathrm{rk}(f^2) \Leftrightarrow \mathrm{Im}(f) \oplus \mathrm{Ker}(f) = E$.

Chapter 10

Duality

In mathematics *duality* is synonymous with *we pay one and we get two*, that is, we prove a theorem and we have, by duality, two theorems. Let us take an example.

(1) One and only one straight line passes through two distinct points.

(2) In the plane two distinct lines meet at a point.

While statement (1) is true, we know that (2) is not always true (parallel lines). Let us pretend for a while that (2) is true. In this scenario we see that by exchanging 'pass' with 'meet' and 'point' with 'line', the two statements are equivalent. Therefore every theorem proven on points/lines becomes a theorem on lines/points. That is, if we have proved a theorem on 'points/lines', by duality, we have another on 'lines/points'.

Nice, unfortunately it is not true that two lines meet (always) in one point (consider two parallel lines).

Well it is not true in *Euclidean geometry*, but it is true in *projective geometry*. Projective geometry was born from the effort of painters to represent on a canvas (therefore a plane, of dimension two) the reality of space (of dimension three).

There are various ways of representing a 3-dimensional object on a sheet of paper. The most realistic and most used by painters of the Middle Ages (and not only) is perspective. In this technique, parallel lines in reality are represented as lines passing through a point (*vanishing point*). The classic example is when you draw a road (or tracks) the edges of the road (or tracks) appear to be approaching in the distance. The width of the road becomes smaller and smaller.

A classic example: The Last Supper by Leonardo da Vinci (central perspective, the vanishing point is the head of Christ). Today, especially in comics, it is common to make drawings with more than one vanishing point.

Painters found the rules of perspective empirically, by trial and error, even building mechanical machinery to project the 3-dimensional space onto a plane. The mathematical theory capable of explaining the constructions of perspective arrived well after the first paintings that used this technique. In fact, it was only in 1639 that Desargues published the first treatise on projective geometry (Desargues was an architect). This theory was not understood by contemporaries (with the exception of Pascal). The theory was taken up more successfully by Poncelet (1822). Let us quickly see what this mathematical model (the projective plane) consists of.

Let us try to explain, in almost mathematical language, the projective plane model. Imagine having our eye positioned at the origin, O, of space. Our eye is represented by a sphere centered on the origin. Taking a point m in space, the line, R, passing through the origin and the point m intersects the sphere at a point p, its projection (from O) onto the sphere. But all the points of the line R are projected into p. For our eye (the sphere) all the points of the line R identify with the point p (in particular the line has 'lost a dimension').

Similarly, a plane passing through the origin intersects the sphere in a *great circle* (i.e. a circumference with the origin as its center, therefore having a diameter equal to that of the sphere). Now, two planes passing through the origin intersect in a line passing through the origin and in fact two great circles on the sphere always intersect (in the 'projective' plane two straight lines always meet at a point, there is no is more parallelism, and the principle of duality is always valid).

Ah! But the intersection is made by two points!? Yes, but they are *antipodal* points and therefore they are identified in the projective plane. In fact, from a set-theoretic point of view, the real projective plane is in bijection with the set of lines of $\mathbb{R}^3$ that pass through the origin. Therefore a point of the real projective plane, denoted by $\mathbb{P}^2_{\mathbb{R}}$ or simply $\mathbb{P}^2$, corresponds to a line passing through the origin. How can we represent this?

Let us take our sphere centered at the origin. Every line through the origin intersects the sphere at two antipodal points; we only keep the one in the upper hemisphere. This works except for the lines of the equatorial plane which have their two points of intersection in the upper (and also lower) hemisphere. In conclusion $\mathbb{P}^2$, the projective plane, is the upper semi-sphere with the antipodal points of the base circle identified.

What kind of geometric entity is this? Well it is a bit complicated to visualize (one can represent it as a subset of $\mathbb{R}^3$, the Boy's surface).

From an algebraic point of view the projective plane has a simple description as a quotient set. On $(\mathbb{R}^3)^* = \mathbb{R}^3 \setminus \{O\}$ ($\mathbb{R}^3$ minus the origin) we consider the relation $(x, y, z) \sim (x', y', z') \Leftrightarrow \exists\lambda$ such $(\lambda x, \lambda y, \lambda z) = (x', y', z')$ or $u = (x, y, z) \sim v = (x', y', z')$ if and only if $u \in \langle v\rangle$ (the three points u, v and O are collinear. Then $\sim$ is an equivalence relation and the quotient set is nothing other than the projective plane $\mathbb{P}^2$.

This construction generalizes to any K-vector space E. On E^* the relation $u \sim v \Leftrightarrow u \in \langle v\rangle$ is an equivalence relation and the quotient set is $\mathbb{P}(E)$ the projective space associated with E. At this point we see that projective geometry 'derives' from linear algebra. We will see that the same happens for affine geometry and Euclidean geometry (a bit of multilinear algebra is also needed here).

To conclude, let us say that our perception of the world is closer to the 'projective space' than to the 'Euclidean space'. In fact, if we look at the tracks in any station our visual impression will be that these tracks tend to get closer and closer in the distance. If we stand under a very tall building and look up, the walls of the structure will seem closer and closer. We actually live in a projective world! Where there are no parallels.

Therefore, in the projective plane the principle of duality applies: we can exchange point with line and line with point. This principle of projective duality is a translation of vector duality which states that given a K-vector space E, of finite dimension, there exists a *canonical* isomorphism $E \simeq E^{**}$, i.e. E is *canonically* isomorphic to its bi-dual (the dual of the dual).

10.1. Dual space

We all know that a line in the plane passing through the origin has an equation of the form $y = ax$. Said like this, it is not entirely correct, we are missing the y axis which has the equation $x = 0$ (it is not a graph). For this reason we prefer to say that every line in the plane passing through the origin has an equation of the form $ax + by = 0$. If $b \neq 0$ we can rewrite $y = cx$ with $c = -a/b$; if $b = 0$ (and $a \neq 0$) we find the equation $x = 0$. Actually, a line through O has infinite equations: if $ax + by = 0$ is an equation of the line R, then also $\alpha(ax + by) = 0$ with $\alpha \neq 0$, is an equation of R. We also know that $ax + by = 0$ and $a'x + b'y = 0$ define the same line if and only if there exists λ such that $a = \lambda a', b = \lambda b'$.

Now, we will see that the same thing happens for planes in $\mathbb{R}^3$ and more generally for *hyperplanes* in any vector space E, where:

Definition 10.1. Let E be a K-vector space. A *hyperplane* of E is a subvector space of E of dimension $\dim E - 1$.

First we need to get acquainted with dual spaces.

Definition 10.2. A linear map $\varphi : E \to K$ is called a *linear form on* E. The set of all linear forms on E is called the *dual space of* E and is denoted by E^*. In other words

$$E^* = \{\varphi : E \to K \mid \varphi \text{ is linear}\}.$$

Given a K-vector space E, the bi-dual of E is the space $E^{**} := (E^*)^*$.

Remark 10.3. We have $E^* = \mathcal{L}(E, K)$ so $\dim E^* = \dim E \times \dim K = \dim E$.

Proposition 10.4. *Let E be a finite-dimensional K-vector space with basis $\mathcal{B} = (e_1, \cdots, en)$. Then*

(1) *there exists a basis $(e_1^*, \cdots ., e_n^*)$ of E^*, such that $e_i^*(e_j) = \delta_{i,j}$ (here δ_{ij} is the Kronecker symbol which is 1 if $i = j$ and 0 if $i \neq j$);*

(2) $\dim E = \dim E^* = \dim E^{**}$.

Proof. (1). Let $\mathcal{B} = (e_1, \cdots, e_n)$ be a basis of E. We define $e_i^* : E \to K$ via the images of the elements of the basis $\mathcal{B}$ (Theorem 9.10). More specifically we set $e_i^*(e_j) = \delta_{ij}$.

We show that $e_1^*, \cdots, e_n^*$ are independent. Let $\sum \lambda_i e_i^* = 0$. Then $\sum \lambda_i e_i^*(e_k) = 0$. Since $e_i^*(e_k) = \delta_{ik}$, we have $\sum \lambda_i e_i^*(e_k) = \lambda_k = 0$. Since this is true for every K, we conclude that the vectors e_i^* are linearly independent. Since $\dim E^* = n$ we can already conclude that (e_i^*) is a basis of E^*. We nevertheless show that the vectors (e_i^*) are a system of generators. Let $\varphi \in E^*$. Let $\varphi(e_i) = \lambda_i$. Then $\varphi = \sum_i \lambda_i e_i^*$. In fact $(\sum_i \lambda_i e_i^*)(e_k) = \sum_i [(\lambda_i e_i^*)(e_k)] = \lambda_k = \varphi(e_k)$. So $\sum_i \lambda_i e_i^*$ and φ coincide on the basis vectors e_i and therefore are equal (because a linear map is completely determined by its values on the elements of a basis; if these values are the same, the maps are the same).

(2) Follows from (1). □

Definition 10.5. With the notations of the Proposition 10.4, the basis (e_i^*) is called the *dual basis* of (e_i).

Remark 10.6. The linear form e_i^* is nothing other than the i-th coordinate function with respect to the basis $\mathcal{B} = (e_i)$: if $(\lambda_1, \cdots, \lambda_i, \cdots, \lambda_n)$ are the coordinates of v in the basis $\mathcal{B}$, then $e_i^*(v) = \lambda_i$.

Lemma 10.7. *Let $\varphi \in E^*$, $\varphi \neq 0$. Then φ is surjective and* $\mathrm{Ker}(\varphi)$ *is a hyperplane of E.*

Proof. We have $\varphi : E \to K$. Since $\varphi \neq 0$ there exists $v \in E$ such that $\varphi(v) \neq 0$. This implies $\dim \mathrm{Im}(\varphi) \geq 1$ ($\mathrm{Im}(\varphi)$ contains a non-zero vector). Since $\mathrm{Im}(\varphi) \subset K$

and $\dim K = 1$, we have $\mathrm{Im}(\varphi) = K$. By the rank theorem $\dim \mathrm{Ker}(\varphi) = \dim E - 1$. That is, $\mathrm{Ker}(\varphi)$ is a hyperplane. □

Proposition 10.8. *Let E be a K-vector space of dimension n.*

(1) *If $H \subset E$ is a hyperplane of E then there exists a linear form $\varphi \in E^*$ such that $\mathrm{Ker}(\varphi) = H$.*

(2) *Two linear forms φ, ψ are proportional ($\varphi = \lambda\psi$) if and only if $\mathrm{Ker}(\varphi) = \mathrm{Ker}(\psi)$.*

Proof. (1) Let $(e_1, \cdots, e_{n-1})$ be a basis of H. By the Incomplete Basis Theorem we can complete it to a basis of E: $\mathcal{B} = (e_1, \cdots, e_{n-1}, e_n)$. We define $\varphi : E \to K$ by its values on the elements of $\mathcal{B}$ as follows: $\varphi(e_i) = 0$ for $1 \leq i \leq n-1$ and $\varphi(e_n) = 1$. Then $\varphi \neq 0$ and $H \subset \mathrm{Ker}(\varphi)$. Since $\dim H = \dim \mathrm{Ker}(\varphi)$ we have $H = \mathrm{Ker}(\varphi)$.

(2) It is clear that if $\varphi = \lambda\psi$, then $\mathrm{Ker}(\varphi) = \mathrm{Ker}(\psi)$. Conversely, suppose $\mathrm{Ker}(\varphi) = \mathrm{Ker}(\psi) =: H$. Let $(e_1, \cdots, e_{n-1})$ be a basis of H and complete it (Incomplete Basis Theorem) to a basis of E: $\mathcal{B} = (e_1, \cdots, e_{n-1}, e_n)$. Let $\varphi(e_n) = \alpha$ and $\psi(e_n) = \beta$. We have $\alpha\beta \neq 0$ because the two forms are non-zero. Let $\lambda = \alpha/\beta$, then $\varphi(e_n) = \lambda\psi(e_n)$. Therefore $\varphi(e_i) = \lambda\psi(e_i)$, $\forall i, 1 \leq i \leq n$ ($\varphi(e_i) = \psi(e_i) = 0$ if $i < n$) and hence $\varphi = \lambda\psi$, the two forms are proportional. □

Let $\mathcal{B} = (e_1, \cdots, e_n)$ be a basis of E and let $\psi : E \to K$ be a linear form. We set $\psi(e_i) = a_i$. If $v = x_1e_1 + \cdots + x_ne_n$ is a vector of E, $\psi(v) = a_1x_1 + \cdots + a_nx_n$ and we see that ψ is a homogeneous polynomial of degree one in $x_1, \cdots, x_n$. The equation $\psi(v) = 0$ becomes $a_1x_1 + \cdots + a_nx_n = 0$ and defines $\mathrm{Ker}(\psi)$.

We have seen that every hyperplane has an equation of the form $\psi(v) = 0$, that every equation $\psi(v) = 0$ defines a hyperplane and that two linear forms define the same hyperplane if and only if they are proportional.

This can be formulated as follows:

Corollary 10.9. *Let $Gr(1, E^*)$ be the set of (vector) lines of the dual E^* and $Gr(n-1, E)$ the set of hyperplanes of E. The map*

$$d : Gr(1, E^*) \to Gr(n-1, E) : \langle\varphi\rangle \to \mathrm{Ker}(\varphi)$$

is a bijection.

Proof. The application is injective because if $d(\langle\psi\rangle) = d(\langle\varphi\rangle)$ then $\mathrm{Ker}(\psi) = \mathrm{Ker}(\varphi)$ and this implies ((2) of Proposition 10.8) that ψ and φ are proportional i.e. $\langle\psi\rangle = \langle\varphi\rangle$. The application is surjective due to (1) of Proposition 10.8. □

Now, we want to do something similar for subvector spaces $F \subset E$ of any dimension and this is where the duality theorem comes into play.

10.2. The duality theorem

We have seen that if E is a finite-dimensional K-vector space then $\dim E = \dim E^* = \dim E^{**}$, so these three spaces are *isomorphic*. To establish an isomorphism between E and its dual E^* we can do this: take a basis of E, $\mathcal{B} = (e_i)$ and consider the dual basis $\mathcal{B}^* = (e_i^*)$ then $f : E \to E^* : e_i \to e_i^*$ is an isomorphism. In the same way we can construct an isomorphism between E^* and E^{**} and then, by composing we obtain an isomorphism between E and its bi-dual E^{**}. However, there is a *canonical* isomorphism, 'better than the others' between E and E^{**}, the canonical isomorphism is more natural because it can be defined without using bases, i.e. it does not depend on the choice of bases.

An element, Θ, of E^{**} is a linear map $\Theta : E^* \to K : \varphi \to \Theta(\varphi)$. Let $u \in E$, if $\varphi \in E^*$ we can consider the value of φ in u: $\varphi(u) \in K$. If we let φ vary in E^* we obtain an *evaluation* map in u: $\nu_u : E^* \to K : \varphi \to \varphi(u)$. It is easily verified that the map ν_u is linear. So $\nu_u \in E^{**}$.

This allows us to define a map: $\nu : E \to E^{**} : u \to \nu_u$. We will verify that this map is linear.

Theorem 10.10 (Duality theorem)**.** *Let E be a finite-dimensional K-vector space. The map*

$$\nu : E \to E^{**} : u \to \nu_u$$

is a linear isomorphism.

Proof. Since $\dim E = \dim E^{**}$ it suffices to show that ν is injective (Corollary 9.16). Suppose $\nu_u = 0$. Then $\varphi(u) = 0, \forall \varphi \in E^*$. Geometrically this means that the vector u belongs to every hyperplane of E. Let $\mathcal{B} = (e_1, \cdots, e_n)$ be a basis of E and let $u = \alpha_1 e_1 + \cdots + \alpha_n e_n$ be the writing of u on this basis. We have $e_i^*(u) = \alpha_i = 0$, $\forall i$, so $u = 0$. So $\mathrm{Ker}(\nu) = \{0\}$ and ν is injective and therefore bijective. □

Remark 10.11.

(1) As promised the ν isomorphism was defined without using bases.

(2) The theorem is false in infinite dimension. In that case the map ν is always injective but never surjective.

(3) Another way of formulating the theorem: every linear map $\Theta : E^* \to K$ is evaluation map in u for some vector $u \in E$ (i.e. there exists $u \in E$ such that $\Theta = \nu_u$).

(4) Theorem 10.10 is the principle of duality which can be stated like this: if we have proved a result for (E, E^*) then we have this result for (E^*, E^{**}). But since $E \simeq E^{**}$ *canonically*, we have the result for (E^*, E).

Definition 10.12 (Orthogonal space)**.** Let $F \subset E$ be a subvector space of the K-vector space E. We define the *orthogonal* (or conjugate) of F as the subvector space, $F^\circ \subset E^*$, defined by

$$F^\circ = \{\psi \in E^* \mid \psi|_F = 0\} = \{\psi \in E^* \mid \psi(v) = 0,\, \forall v \in F\}.$$

Remark 10.13. Stated differently $F^\circ = \{\psi \in E^* \mid F \subset \mathrm{Ker}(\psi)\}$, i.e. F° is the space of equations of hyperplanes containing F.

We have $F^\circ \subset E^*$ and we can consider $F^{\circ\circ} := (F^\circ)^\circ \subset E^{**}$. Now since $E^{**} \simeq E$ i.e. since E^{**} canonically identifies with E, we can consider $F^{\circ\circ}$ as a subspace of E.

If $\Theta \in F^{\circ\circ}$, then $\Theta(\psi) = 0, \forall \psi \in F^\circ$. But $\Theta = \nu_u$ ((3) of Remark 10.11). So $\nu_u(\psi) = 0$, $\forall \psi \in F^\circ$. Since $\nu_u(\psi) = \psi(u)$, we see that

$$F^{\circ\circ} \simeq \{u \in E \mid \psi(u) = 0, \forall \psi \in F^\circ\}.$$

In other words $F^{\circ\circ}$ is the intersection of all hyperplanes containing F. Clearly $F \subset F^{\circ\circ}$.

Theorem 10.14. *Let E be a finite-dimensional K-vector space and $F \subset E$ a subvector space.*

(1) *We have* $\dim F + \dim F^\circ = \dim E$*;*

(2) $F = F^{\circ\circ}$.

Proof. (1) Let $(e_1, \cdots, e_p)$ be a basis of F (so $\dim F = p$). by the Incomplete Basis Theorem we can complete it to a basis of E: $\mathcal{B} = (e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$. Let us consider the dual basis $\mathcal{B}^* = (e_i^*)$. We show that $(e_{p+1}^*, \ldots, e_n^*)$ is a basis of F°. The vectors $e_{p+1}^*, \ldots, e_n^*$ are obviously independent (because $\mathcal{B}^*$ is a basis), so it is enough to show that they generate F°. Let $\psi \in F^\circ$, then ψ is written on the basis $\mathcal{B}^*$: $\psi = \alpha_1 e_1^* + \cdots + \alpha_n e_n^*$. We have $\psi(e_i) = 0$ if $1 \leq i \leq p$ because $\psi \in F^\circ$. We have $\psi(e_i) = (\alpha_1 e_1^* + \cdots + \alpha_n e_n^*)(e_i) = \alpha_i$. So $\alpha_i = 0$ if $1 \leq i \leq p$, therefore $\psi = \alpha_{p+1} e_{p+1}^* + \cdots + \alpha_n e_n^*$. This shows $F^\circ = \langle e_{p+1}^*, \cdots, e_n^* \rangle$ and $\dim F^\circ = n - p$.

(2) Applying (1) to $F^\circ \subset E^*$, we have $\dim F^\circ + \dim F^{\circ\circ} = \dim E^* = \dim E$. It turns out that $\dim F = \dim F^{\circ\circ}$. Since $F \subset F^{\circ\circ}$ (after canonically identifying $F^{\circ\circ}$ with a subspace of E), we conclude that $F = F^{\circ\circ}$. □

We have therefore proved the following: A subvector space is equal to the intersection of all hyperplanes containing it.

Definition 10.15. Let E be a K-vector space of dimension n. For every r, $1 \leq r \leq n$ the set of subvector spaces of dimension r of E is called the *Grassmannian*

of the subspaces of dimension r and is denoted by $Gr(r, E)$. So

$$Gr(r, E) = \{F \subset E \text{ such that } F \text{ is a vector subspace of dimension } r\}.$$

Remark 10.16. We have $\mathbb{P}(E) = Gr(1, E)$; the projective space is a special case of Grassmannian. We will study Grassmannians in more detail in Chapter 34.

Theorem 10.17 (Duality)**.** *Let E be a K-vector space of dimension n. For every r, $1 \leq r \leq n$ there exists a natural bijection given by:*

$$d : Gr(r, E) \to Gr(n - r, E^*) : F \to F^\circ$$

Proof. The inverse application is given by: $d : Gr(n - r, E^*) \to Gr(r, E) : X \to X^\circ$, after canonically identifying E^{**} with E. In fact $F^{\circ\circ} = F$. □

10.3. Cartesian equations of subvector spaces

Let E be a K-vector space of dimension n and $F \subset E$ a subvector space of dimension p. The space $F^\circ \subset E^*$ has dimension $n - p$. Let $(\psi_1, \cdots, \psi_{n-p})$ be a basis of F°. We have $F^{\circ\circ} = \{v \in E \mid \psi_i(v) = 0, \forall i, 1 \leq i \leq n - p\}$. But, as we have just seen, $F^{\circ\circ} = F$, we conclude that $F = \{v \in E \mid \psi_i(v) = 0, \forall i, \leq i \leq n - p\}$. In other words F is the solution set of the linear system:

$$\begin{cases} \psi_1(v) & = 0; \\ \vdots & \\ \psi_{n-p}(v) & = 0. \end{cases}$$

We conclude that the subspace F, of dimension p, can be defined by $n - p$ linearly independent equations.

We observe that F can not be defined by less than $n - p$ equations, in fact in this case we would have $F^\circ = \langle \varphi_1, \ldots, \varphi_r \rangle$ with $r < n - p$ and therefore $F^{\circ\circ}$ would have dimension $n - r > p$.

On the other hand, if $\psi_1(v) = 0, \ldots, \psi_{n-r}(v) = 0$ define F then also $\psi_1(v) = 0, \ldots, \psi_{n-r}(v) = 0, \sum \lambda_i \psi_i(v) = 0$ define F, in fact the last equation is a combination of the previous ones.

Proposition 10.18. *Let E be a K-vector space of dimension n, and $F \subset E$ a subvector space with* $\dim F = p$.

(1) *There exist $n - p$ linearly independent linear forms ψ_i, $1 \leq i \leq n - p$, such that $F = \{v \in E \mid \psi_i(v) = 0, 1 \leq i \leq n - p\}$;*

(2) *If $r < n - p$ and $\varphi_1, \ldots, \varphi_r$ are r linear forms, the solution set, G, of the linear system:*

$$\begin{cases} \varphi_1(v) & = 0; \\ \vdots & \\ \varphi_r(v) & = 0; \end{cases}$$

is such that $G \neq F$.

Proof. (1) As seen above $\dim F^\circ = n - p$ and if $(\psi_1, \ldots, \psi_{n-p})$ is a basis of F°, then $F = \{v \in E \mid \psi_i(v) = 0, 1 \leq i \leq n - p\}$.

(2) If $r < n - p$ the solution set G is $\langle \varphi_1, \ldots, \varphi_r \rangle^\circ$ which has dimension $n - s$ with $s \leq r$. So $n - s \geq n - r > p$. Therefore $G \neq F$. □

Remark 10.19. The equations $\psi_i(v) = 0$ that define a subvector space F are not uniquely determined, in fact the subspace F° has many different bases.

To conclude (at least for now) with duality we would say that there is only one thing that you absolutely need to know and that is: In a vector space of dimension n a subspace of dimension p is defined by $n-p$ linear, linearly independent, equations.

10.4. Homogeneous linear systems and duality

A linear system of type

$$\mathcal{S} = \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0; \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = 0; \end{cases}$$

with $a_{ij} \in K$ is called a *homogeneous linear system.* A solution of the system $\mathcal{S}$ is an element $b = (b_1, \cdots, b_n) \in K^n$ that verifies all the equations. For example, $0 = (0, \cdots, 0)$ is always a solution of a homogeneous linear system. Let Σ be the set of solutions of $\mathcal{S}$, we have:

Lemma 10.20. *The set $\Sigma \subset K^n$ is a subvector space.*

Proof. It is clear that $0 \in \Sigma$ and it is easily verified that if b, c are two elements of Σ then every linear combination of these two vectors is still an element of Σ. □

Let us now see a more interesting proof of the previous lemma.

Proof. Let us consider the maps $\psi_i : K^n \to K : (x_1, \ldots, x_n) \to a_{i1}x_1 + \cdots + a_{in}x_n$, we know that $\psi_1, ..., \psi_p$ are linear forms. Let $Y = \langle \psi_1, \ldots, \psi_p \rangle \subset (K^n)^*$. Then $\Sigma = Y^\circ$ and therefore $\mathcal{S}$ is a subvector space. Another way of saying this is as follows: $f : K^n \to K^p : (x_1, \ldots, x_n) \to (\psi_1(x_1, \ldots, x_n), \ldots, \psi_p(x_1, \ldots, x_n))$, then

f is linear because it is given by homogeneous polynomials of the degree one. The solution set is given by $\Sigma = \mathrm{Ker}(f)$ and therefore is a subvector space of K^n. □

Definition 10.21. With the previous notations $r := \dim Y$ is called the *rank* of the system $\mathcal{S}$.

Proposition 10.22. *The solution set Σ is a subvector space of dimension $n - r$.*

Proof. Since $\Sigma = Y^\circ$ this follows from Theorem 10.14. □

Remark 10.23. We observe that $r = \dim Y \leq p$, where p is the number of equations. So if $n > p$ (i.e. if the number of unknowns is strictly greater than the number of equations, then $\dim \Sigma > 0$, i.e. there exists a non-trivial solution. In fact, since $p \geq r$, we have $n - r > 0$.

We also observe that $\dim \Sigma \leq n$ ($\Sigma \subset K^n$) and the equality holds if and only if $r = 0$ i.e. if and only if all the equations are zero, therefore for a non-trivial system $\dim \Sigma \leq n - 1$.

Exercises

Exercise 10.24. State whether the linear forms φ, ψ on $\mathbb{R}^3$ are linearly independent or not where:
$\varphi : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to x + 2y - z$;
$\psi : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to -x + y + 2z$.

Exercise 10.25. In $\mathbb{R}^2$ consider the vectors $u = (1, -1), v = (1, 1)$.
1) Show that $\mathcal{B} = (u, v)$ is a basis of $\mathbb{R}^2$.
2) Determine the dual basis $\mathcal{B}^*$.

Exercise 10.26. Let E, F be K-vector spaces and $f : E \to F$ a linear map.
1) Prove that f is not surjective $\Leftrightarrow$ there exists a non-zero linear form $\phi \in F^*$ such that $\phi \circ f = 0$.
2) More precisely if $\dim(\mathrm{Im}(f)) = \dim(F) - r$, there exist r linear, linearly independent forms, $\phi_i \in F^*$, such that $\phi_i \circ f = 0$, $\forall i$.

Exercise 10.27. In $\mathbb{R}^3$ let $v = (1, -1, 2)$. Find Cartesian equations of the line $\langle v \rangle$. Let $w = (1, 0, 1)$. What is the dimension of $V = \langle v, w \rangle$? Give a Cartesian representation of V (i.e. give Cartesian equations of V).

Exercise 10.28. Let $V \subset \mathbb{R}^4$ be defined by the equations

$$\begin{cases} x + y - z + t = 0; \\ 5x - y + z - t = 0; \\ 2x - y + z - t = 0. \end{cases}$$

Determine $\dim(V)$.

Exercise 10.29. Show that duality reverses inclusions, i.e. if $V, W \subset E$ are subvector spaces with $V \subset W$, then $W^\circ \subset V^\circ$.

Exercise 10.30. Let E be a K-vector space of dimension n, and $H \subset E$ a hyperplane. Show that there are $n - 1$ hyperplanes $H_2, \cdots, H_n$ such that $\dim(H \cap H_2 \cap \cdots \cap H_n) = 0$.

Exercise 10.31. Let E be a K-vector space of dimension n. Show that two distinct hyperplanes intersect in a vector subspace of dimension $n - 2$ (Exercise 9.37). In particular, in $\mathbb{R}^3$ two distinct planes through the origin intersect in a line.

Exercise 10.32. In the $\mathbb{R}$-vector space $\mathbb{R}[x]$ consider the vectors

$$\begin{aligned} P_1(x) &= \pi x^3 - 3x^2 + x + 7; \\ P_2(x) &= x^3 - 7x + 1; \\ P_3(x) &= \sqrt{2}x^3 - 3x^2 + 11x - 2; \\ P_4(x) &= \pi\sqrt{2}x^3 + 2x + 4; \\ P_5(x) &= x^3 - 3x^2 + 27x - 11. \end{aligned}$$

State whether these vectors are linearly independent.

Exercise 10.33. For each $\alpha \in \mathbb{R}$ let S_α be the subvector space of $\mathbb{R}^3$ defined by the linear system

$$\begin{cases} 3\alpha x - 4y + \alpha z = 0; \\ x - 2\alpha y + 3\alpha z = 0. \end{cases}$$

Determine $\dim(S_\alpha)$ for any $\alpha \in \mathbb{R}$.

Chapter 11

Linear maps and matrices

11.1. Matrix writing of linear maps

Let $f : E \to F$ be a linear map between two K-vector spaces. Let $\mathcal{B}$ be a basis of E and $\mathcal{C}$ a basis of F. The basic principle is that f is completely determined by the vectors $f(e_i)$ ($\mathcal{B} = (e_1, \ldots, e_n)$) which in turn are completely determined by their coordinates in the basis $\mathcal{C} = (v_1, \ldots, v_p)$. So if $f(e_i) = a_{1i}v_1 + \cdots + a_{pi}v_p$, the associated matrix, M, is the matrix with p rows and n columns whose i -th column is formed by the coordinates of $f(e_i)$ in the basis $\mathcal{C}$:

$$M = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix}.$$

We insist on the fact that the i-th column vector of M is given by the coordinates in the base $\mathcal{C}$ of the vector $f(e_i)$.

Definition 11.1. With the previous notations the matrix M is the matrix associated to the map f with respect to the bases $\mathcal{B}, \mathcal{C}$.

In this way we get a map

$$\mathrm{Mat}(-; \mathcal{B}, \mathcal{C}) : \mathcal{L}(E, F) \to M_{pn}(K) : f \to M.$$

So $\mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) = M$. The map $\mathrm{Mat}(-; \mathcal{B}, \mathcal{C})$ is a linear isomorphism.

Let $f : E \to F$ be a linear map between the two K-vector spaces E, F, $\mathcal{B}$, $\mathcal{C}$ bases of E, F, $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$ and $u \in E$, $u = \alpha_1 e_1 + \cdots + \alpha_n e_n$, where $\mathcal{B} = (e_1, \ldots, e_n)$.

We have

$$f(u) = f(\sum \alpha_i e_i) = \sum \alpha_i f(e_i) = \sum_i \alpha_i (\sum_k a_{ki} v_k) = \sum_k (\sum_i \alpha_i a_{ki}) v_k.$$

In other words, the j-th coordinate of $f(u)$ in the basis $\mathcal{C}$ is

$$a_{j1}\alpha_1 + \cdots + a_{jn}\alpha_n.$$

Let us introduce a definition.

Definition 11.2. Let $x = (x_1, \ldots, x_n) \in K^n$, $y = (y_1, \ldots, y_n) \in K^n$. The *scalar product*, $(x \mid y)$, is equal, by definition, to $x_1y_1 + \cdots + x_ny_n$.

We will indicate this scalar product as follows:

$$(x \mid y) = (x_1, \ldots, x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := x_1y_1 + \cdots + x_ny_n.$$

The j-th coordinate of $f(u)$ in the basis $\mathcal{C}$ is the scalar product of the j-th row of $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$ with $(\alpha_1, \ldots, \alpha_n)$, the coordinate vector of u in the basis $\mathcal{B}$.

So if $f(u) = b_1v_1 + \cdots + b_pv_p$ we have:

$$b_j = (a_{j1}, \ldots, a_{jn}) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} := a_{j1}\alpha_1 + \cdots + a_{jn}\alpha_n. \tag{11.1}$$

As we will now see, this is a special case of matrix product.

Definition 11.3. Let A be a matrix of type (m, n) (m rows, n columns) with coefficients in K and let B be a matrix of type (t, p) with coefficients in K.

(1) The *row-by-column* product AB is defined if and only if $n = t$.

(2) If $n = t$ the product AB is a matrix D of type (m, p).

(3) If $AB = D = (d_{ij})$ then d_{ij} is equal to the scalar product of the i-th row of A with the j-th column of B:

$$d_{ij} = (R_i(A) \mid C_j(B)).$$

The matrix, M, of f is a matrix of type (p, n). If $u = \alpha_1e_1 + \cdots + \alpha_ne_n)$ and $f(u) = b_1v_1 + \cdots + b_pv_p$ then

$$M \cdot u = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}.$$

Example 11.4. Let A be the following matrix of type $(3,2)$:

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

and B be the following matrix $(2,2)$:

$$B = \begin{pmatrix} 1 & 2 \\ 0 & -2 \end{pmatrix}.$$

The product AB is the following matrix of type $(3,2)$:

$$AB = \begin{pmatrix} 1 = 1\cdot 1 + 0\cdot 0 & 2 = 1\cdot 2 + 0\cdot(-2) \\ 2 = 2\cdot 1 + -1\cdot 0 & 6 = 2\cdot 2 + (-1)\cdot(-2) \\ 0 = 0\cdot 1 + 3\cdot 0 & -6 = 0\cdot 2 + 3\cdot(-2) \end{pmatrix}.$$

The product BA however is not defined.

This row-by-column product definition does not fall from the clouds.

Proposition 11.5. *Let E, F, G be three K-vector spaces and $f : E \to F$, $g : F \to G$ two linear maps. Set $h = g \circ f : E \to G$.*

Let $\mathcal{B} = (e_1, ..., e_n)$, $\mathcal{C} = (v_1, ..., v_p)$ and $\mathcal{D} = (w_1, ..., w_t)$ be bases of E, F, G. Finally, let $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$, $B = \mathrm{Mat}(g; \mathcal{C}, \mathcal{D})$. Then

$$\mathrm{Mat}(g \circ f; \mathcal{B}, \mathcal{D}) = BA.$$

Proof. Let $C = \mathrm{Mat}(g \circ f; \mathcal{B}, \mathcal{D})$, $C = (c_{ij})$. Note that c_{ij} is the i-th coordinate in the basis $\mathcal{D}$ of $(g \circ f)(e_j) = g(f(e_j))$. But $f(e_j)$ is the vector of F whose coordinates in the basis $\mathcal{C}$ are given by the j-th column of the matrix A. So $g(f(e_j))$ is the vector of G whose i-th coordinate is given by the scalar product of the i-th row of $B = \mathrm{Mat}(g; \mathcal{C}, \mathcal{D})$ with the coordinate vector of $f(e_j)$ in the basis $\mathcal{C}$, i.e. $c_{ij} = (R_i(B) \mid C_j(A))$. This shows that $C = BA$. □

Remark 11.6. Pay attention to the order, it is BA and not AB, this is because in $g \circ f$ the first map that is applied is f. If $u \in E$, $u = \alpha_1 e_1 + \cdots + \alpha_n e_n$ then the coordinates of $g(f(u))$ are obtained by the matrix product $BAu := B(Au)$.

Remark 11.7. The matrix product has all the good properties of the composition of linear maps for instance $A(B + C) = AB + AC$.

11.2. Matrices associated with endomorphisms

Let E be a K-vector space of dimension n and $f \in \mathrm{End}(E)$. If $\mathcal{B}$, $\mathcal{C}$ are two bases of E, $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) \in M_n(K)$ is a square matrix of type (n, n).

If $g \in \mathrm{End}(E)$ and $B = \mathrm{Mat}(g; \mathcal{B}, \mathcal{C})$, then it is not true that $\mathrm{Mat}(g \circ f; \mathcal{B}, \mathcal{C}) = BA$. In fact we are in the following situation:

$$E_{\mathcal{B}} \xrightarrow{f} E_{\mathcal{C}}, E_{\mathcal{B}} \xrightarrow{g} E_{\mathcal{C}}.$$

Here $E_{\mathcal{B}}$ means E with the basis $\mathcal{B}$. The problem is that in the middle space Proposition 11.5 does not apply. To be able to apply the Proposition 11.5 we must have $\mathcal{B} = \mathcal{C}$.

For this reason, when studying the endomorphisms of a space E we will tend to take the same basis in the domain and in the codomain, in fact if $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ and if $B = \mathrm{Mat}(g; \mathcal{B}, \mathcal{B})$ then $\mathrm{Mat}(g \circ f; \mathcal{B}, \mathcal{B}) = BA$. Thus the isomorphism of K-vector spaces

$$\mathrm{Mat}(-; \mathcal{B}, \mathcal{B}) : End(E) \to M_n(K)$$

is an isomorphism of K-algebras. That is, it is also a ring isomorphism for the ring structures $(\mathrm{End}(E), +, \circ)$ and $(M_n(K), +, \cdot)$ where $\cdot$ denotes the row-by-column product of the matrices. Indeed $(M_n(K), +, \cdot)$ is a ring.

Lemma 11.8. *Let E be a K-vector space of dimension n. For every basis $\mathcal{B}$ of E we have*

$$\mathrm{Mat}(Id_E; \mathcal{B}, \mathcal{B}) = I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

That is, $I_n = (a_{ij})$, with $a_{ij} = \delta_{ij}$, where δ_{ij} is the Kronecker symbol (equal to 1 if $i = j$, equal to 0 otherwise).

Proof. It follows from $\mathcal{B} = (e_i)$, $Id_E(e_i) = e_i$. □

Remark 11.9. The previous lemma shows that the identity element for the row-by-column product is the matrix I_n. The fact that $\mathrm{Mat}(-; \mathcal{B}, \mathcal{B})$ is a ring isomorphism explains why for *every* basis $\mathcal{B}$ the image of Id_E (the neutral) is always the same matrix I_n (the neutral of $M_n(K)$ for the row-by-column product).

A first application of what we have seen so far:

Proposition 11.10. *Let $M \in M_n(K)$ be a square matrix. The following are equivalent:*

(1) *M is invertible*

(2) *There exists $N \in M_n(K)$ such that $MN = I_n$;*

(3) *There exists $Q \in M_n(K)$ such that $QM = I_n$.*

If these conditions are verified then $N = Q = M^{-1}$.

Proof. (1) $\Rightarrow$ (2). Just take $N = M^{-1}$.

(2) $\Rightarrow$ (3). Let $\mathcal{B}$ be a basis of E. We have the isomorphism $\mathrm{Mat}(-;\mathcal{B},\mathcal{B}) : \mathrm{End}(E) \to M_n(K)$. Let f, g be the endomorphisms corresponding to M, N. The matrix relation becomes $f \circ g = Id_E$. This implies that f is surjective. Since f is an endomorphism of a finite-dimensional vector space, if f is surjective, then it is also bijective. So there exists f^{-1} such that $f \circ f^{-1} = f^{-1} \circ f = Id_E$. Translating into matrix terms via the isomorphism $\mathrm{Mat}(-;\mathcal{B},\mathcal{B})$ we obtain the existence of Q (and we also see that $Q = N$).

(3) $\Rightarrow$ (1). The proof is similar to the previous one. This time we get h such that $h \circ f = Id_E$. We conclude that f is injective and therefore an isomorphism. $\square$

Definition 11.11. The invertible square matrices of type (n, n) form a group for the matrix product, isomorphic to the group $\mathrm{Aut}_K(E)$. This group is called *the general linear group* and is noted as $Gl_n(K)$.

Let us see some notable matrices.

- **Transpose matrix.** Let $A \in M_{p,n}(K)$, $A = (a_{ij})$. The transpose of A is denoted by ${}^tA$. Is is the matrix of type (n, p) whose rows are the columns of A (and whose columns are the rows of A). If ${}^tA = (b_{kl})$, then $b_{kl} = a_{lk}$. The transposition $M_{n,p}(K) \to M_{p,n}(K) : A \to {}^tA$ is a linear isomorphism.

- **Symmetric matrices.** A *square* matrix $M \in M_n(K)$ is *symmetric* if $M = {}^tM$; that is, if $M = (a_{ij})$, M is symmetric if and only if $a_{ij} = a_{ji}$ (symmetry with respect to the diagonal).

- **Antisymmetric matrices.** A square matrix $M \in M_n(K)$ is *anti-symmetric* if $M = {}^tM$, i.e. $a_{ij} = -a_{ji}$, where as usual $M = (a_{ij})$. This is a notion that depends on the field K we are working on. In fact $(i = j)$ we have $a_{ii} = -a_{ii} \Leftrightarrow 2a_{ii} = 0$. If $2 \neq 0$ in K then we can conclude that $a_{ii} = 0$, otherwise we can not conclude anything. This happens for example if $K = \mathbb{Z}/2\mathbb{Z}$ where $2 = 0$.

- **Triangular matrices.** A square matrix is upper triangular if all elements under the diagonal are zero. That is, $a_{ij} = 0$ if $i > j$ $(M = (a_{ij}))$. Similarly we have the notion of lower triangular matrix.

- **Diagonal matrices.** They are the most interesting matrices for us. A matrix $M = (a_{ij})$ is diagonal if $a_{ij} = 0$ when $i \neq j$: outside the diagonal all coefficients are zero (this does not prevent coefficients on the diagonal from being zero).

11.3. Base changes

Let E be a K-vector space and $\mathcal{B} = (e_1, ..., e_n)$, $\mathcal{C} = (v_1, ..., v_n)$ two bases of E. Consider the map $Id_E : E \to E$ and try to associate to it a matrix taking $\mathcal{B}$ as the basis in the domain and $\mathcal{C}$ as the basis in the codomain: $P = \mathrm{Mat}(Id_E; \mathcal{B}, \mathcal{C})$. So the columns of P are the components of $Id_E(e_j) = e_j$ in the basis $\mathcal{C}$.

If $Id_E(e_j) = e_j = \beta_{1j}v_1 + \cdots + \beta_{nj}v_n$, we have $P = (\beta_{ij})$. We observe that this matrix is very different from the I_n matrix (Lemma 11.8).

Example 11.12. Let $E = \mathbb{R}^2$, $\mathcal{B} = (e_1, e_2)$ with $e_1 = (1, 1), e_2 = (0, 1)$; $\mathcal{C} = (v_1, v_2)$, with $v_1 = (1, -1), v_2 = (1, 0)$. Check that $\mathcal{B}, \mathcal{C}$ are indeed bases. We have $e_1 = -v_1 + 2v_2$, $e_2 = -v_1 + v_2$. So in this case the matrix P is equal to

$$\begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Let $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$. To get the coordinates of v in the basis $\mathcal{C} = (v_i)$ just do the row-by-column product of the matrix P with the matrix

$$X = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

The result is a matrix of type $(n, 1)$ giving the coordinates in the basis $\mathcal{C}$ of the vector $Id_E(v) = v$. Note that

$$P \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \tag{11.2}$$

where $v = \gamma_1 v_1 + \cdots + \gamma_n v_n$. In other words $\gamma_i = (R_i(P) \mid X) = \sum_j \beta_{ij}\alpha_j$. This formula expresses the coordinates of v in the basis $\mathcal{C}$ as a function of its coordinates in the basis $\mathcal{B}$. For this reason the matrix P is called the *change of basis matrix* from the basis $\mathcal{B}$ to the basis $\mathcal{C}$.

We observe that the matrix P is invertible, in fact its inverse is $\mathrm{Mat}(Id_E; \mathcal{C}, \mathcal{B})$. This is perhaps clearer by looking at the diagram:

$$E_{\mathcal{B}} \overset{Id_E}{\to} E_{\mathcal{C}} \overset{Id_E}{\to} E_{\mathcal{B}}.$$

The composition is $E_{\mathcal{B}} \overset{Id_E}{\to} E_{\mathcal{B}}$ whose matrix is I_n. It follows from Proposition 11.10 that $P^{-1} = \mathrm{Mat}(Id_E; \mathcal{C}, \mathcal{B})$. Obviously P^{-1} does the reverse job: given the coordinates (γ_i) in the basis $\mathcal{C}$ it allows us to recover the coordinates (α_i) in the basis $\mathcal{B}$.

Let E, F be two K-vector spaces and $f : E \to F$ a linear map. Let $\mathcal{B}$ be a basis of E and $\mathcal{C}$ a basis of F. Then we have $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$. Suppose we change base

in both E and F. Let $\mathcal{B}', \mathcal{C}'$ be the new bases. We have $A' = \mathrm{Mat}(f; \mathcal{B}', \mathcal{C}')$. What is the connection between A and A'?

The answer is given by a simple diagram:

$$E_{\mathcal{B}'} \overset{Id_E}{\to} E_{\mathcal{B}} \overset{f}{\to} F_{\mathcal{C}} \overset{Id_F}{\to} F_{\mathcal{C}'}. \tag{11.3}$$

The composition is $E_{\mathcal{B}'} \overset{f}{\to} F_{\mathcal{C}'}$, whose matrix is A'. So if $P = \mathrm{Mat}(Id_E; \mathcal{B}', \mathcal{B})$ and if $Q = \mathrm{Mat}(Id_f; \mathcal{C}, \mathcal{C}')$, then:

$$QAP = A'. \tag{11.4}$$

Once again pay attention to the order. The first map that acts is $\mathrm{Mat}(Id_E; \mathcal{B}', \mathcal{B})$.

On the set $M_{p,n}(K)$ of matrices with p rows and n columns, we define the binary relation $M \sim N$ if and only if there exist square, invertible matrices, A, B, such that $N = AMB$.

Verify that $\sim$ is an *equivalence relation* on the set $M_{p,n}(K)$.

Definition 11.13. Two matrices $M, N \in M_{p,n}(K)$, are said to be *equivalent* if $M \sim N$.

What does it mean that two matrices are equivalent?

Proposition 11.14. *Two matrices $M, N \in M_{p,n}(K)$ are equivalent if and only if they represent the same linear map in different bases. More precisely $M \sim N \Leftrightarrow$ there exist two vector spaces E, F of dimension n, p, a linear map $f : E \to F$, bases $\mathcal{B}, \mathcal{B}'$, of E and $\mathcal{C}, \mathcal{C}'$ of F such that*

$$M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C}),\ N = \mathrm{Mat}(f; \mathcal{B}', \mathcal{C}').$$

Proof. (1) Suppose we have $N = AMB$, with $N, M \in M_{p,n}(K)$, $B \in M_n(K)$, $A \in M_p(K)$, A and B invertible. Let E, F be K-vector spaces of dimensions n, p. Let $\mathcal{B} = (e_i)$ be a basis of E and $\mathcal{C} = (v_k)$ a basis of F. Using the isomorphism $\mathrm{Mat}(-; \mathcal{B}, \mathcal{C}) : \mathcal{L}(E, F) \to M_{p,n}(K)$ we define f with $\mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) = M$. Let $e'_i \in E$ be the vectors whose coordinates in the basis $\mathcal{B}$ are given by the columns of the matrix B. The vectors $(e'_1, \ldots, e'_n)$ form a basis, $\mathcal{B}'$, of E because B is invertible (B is the matrix of a bijective linear map which, thus, transform a basis into a basis). We have $\mathrm{Mat}(Id_E; \mathcal{B}', \mathcal{B}) = B$. Now let $v'_1, \ldots, v'_p$ be the vectors whose coordinates in the basis $\mathcal{C}$ are the columns of A^{-1}. The vectors v'_k form a basis $\mathcal{C}'$ (because A^{-1} is invertible). We have $\mathrm{Mat}(Id_F; \mathcal{C}, \mathcal{C}') = A$, and (Proposition 11.5) $AMB = \mathrm{Mat}(f; \mathcal{B}', \mathcal{C}') = N$.

(2) Conversely, suppose we have $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$ and $N = \mathrm{Mat}(f; \mathcal{B}', \mathcal{C}')$. If $B = \mathrm{Mat}(Id_E; \mathcal{B}', \mathcal{B})$ and if $A = \mathrm{Mat}(Id_F; \mathcal{C}, \mathcal{C}')$, then $N = AMB$. □

In a similar way we define on $M_n(K)$ the binary relation $M \approx N \Leftrightarrow \exists P \in M_n(K)$ invertible such that $N = P^{-1}MP$. It turns out that the relation $\approx$ is an equivalence relation.

Definition 11.15. Two matrices $M, N \in M_n(K)$ are *similar* if $M \approx N$.

Proposition 11.16. *Two matrices $M, N \in M_n(K)$ are similar if and only if they represent the same endomorphism, with the same basis in the domain and in the codomain. More precisely, M and N are similar if there exists a K-vector space E, and bases $\mathcal{B}$, $\mathcal{C}$ of E and $f \in \mathrm{End}(E)$ such that*

$$M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) \text{ and } N = \mathrm{Mat}(f; \mathcal{C}, \mathcal{C}).$$

Proof. We reason as in the the proof of the previous proposition. □

In the next paragraphs we will give a complete description of the quotient set $M_{n,p}(K)/\sim$. The description of the quotient set $M_n(K)/\approx$ is out of our reach for the moment.

Let us take a trivial example in dimension one. Saying that $x \approx y$ means: $\exists z \neq 0$ such that $x = z^{-1}yz$ i.e. $x = y$. So, $\overline{x}$, the equivalence class of x is equal to $\{x\}$ and $K/\approx = K$.

Saying that x and y are equivalent means that there exist a, b such that $x = ayb$. If $y \neq 0$ (and therefore also $x \neq 0$) we take $b = 1/y$, $a = x$ and conclude $x \sim y$, that is, two non-zero elements are always equivalent i.e. $\forall x \neq 0, x \sim 1$. Instead, the only element equivalent to 0 is 0 because $0 = x(ab)$ implies $x = 0$ (because $ab \neq 0$). We thus see that the quotient set has two elements: $K/\sim = \{\overline{0}, \overline{1}\}$.

11.4. Rank of a matrix

We have already encountered the notion of rank (rank of a homogeneous linear system), it is an important notion in linear algebra.

Definition 11.17. Let E be a K-vector space and $v_1, \ldots, v_p$ vectors of E. The *rank* of the vectors v_i is the dimension of the subspace they generate: $\dim\langle v_1, \ldots, v_p\rangle$. Let $f : E \to F$ be a linear map. The *rank* of f is the dimension of $\mathrm{Im}(f)$.

We see that if $(e_1, \ldots, e_n)$ is a basis of E then $\mathrm{rk}(f)$, the rank of f, is equal to $\dim\langle f(e_1), \ldots, f(e_n)\rangle$. In fact we know that the vectors $f(e_i)$ generate $\mathrm{Im}(f)$. This explains why the rank theorem is called that: it gives the rank of f as a function of $\dim E$ and $\dim \mathrm{Ker}(f)$.

Definition 11.18. Let $M \in M_{p,n}(K)$. The rank of the matrix M is the rank of the column vectors (seen as vectors of K^p).

Remark 11.19. The column vectors of the matrix M are n vectors of K^p, so $\mathrm{rk}(M) \leq \min\{n, p\}$.

One might ask why column vectors and not row vectors? This choice is motivated by the following:

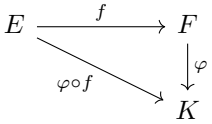
Lemma 11.20. *Let $f : E \to F$ be a linear map between two K-vector spaces E, F. Let $\mathcal{B}, \mathcal{C}$ be the bases of E, F and $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$. Under these conditions* $\mathrm{rk}(f) = \mathrm{rk}(M)$.

Proof. We have $\mathrm{rk}(f) = \dim\langle f(e_1), ..., f(e_n)\rangle = \dim\langle C_1, ..., C_n\rangle$, where C_i is the i-th column of the matrix M. It follows that $\mathrm{rk}(f) = \mathrm{rk}(M)$. □

We will now see that the *rank by columns* is equal to the *rank by rows*. For this we need some tools.

Definition 11.21. Let $f : E \to F$ be a linear map between two K-vector spaces. The *transposed* map, ${}^tf$, is defined by: ${}^tf : F^* \to E^* : \varphi \to \varphi \circ f$.

A diagram helps to understand:



An element $\varphi \in F^*$ is a linear map from F to K, composing with f we get $\varphi \circ f$, a linear map from E to K, i.e. an element of E^*. It turns out that ${}^tf$ is a linear map.

Proposition 11.22. *With the previous notations:*

(1) $\mathrm{Ker}({}^tf) = \mathrm{Im}(f)^\circ$;

(2) $\dim(\mathrm{Im}({}^tf) = \dim(\mathrm{Im}(f))$.

Proof. (1) We have

$$\mathrm{Ker}({}^tf) = \{\varphi \in F^* \mid \varphi \circ f = 0\} = \{\varphi \in F^* \; mid \varphi(f(E)) = \{0\}\}$$

that is

$$\mathrm{Ker}({}^tf) = \{\varphi \in F^* \mid \mathrm{Im}(f) \subset \mathrm{Ker}(\; fG)\} = \mathrm{Im}(f)^\circ.$$

(2) From the Rank Theorem it follows that $\dim F^* = \dim \mathrm{Ker}({}^tf) + \dim \mathrm{Im}({}^tf)$. So, for (1): $\dim F^* = \dim \mathrm{Im}(f)^\circ + \dim \mathrm{Im}({}^tf)$. Now (Theorem 10.14): $\dim \mathrm{Im}(f)^\circ + \dim \mathrm{Im}(f)^{\circ\circ} = \dim F^*$. We conclude by observing that $\dim \mathrm{Im}(f) = \dim \mathrm{Im}(f)^{\circ\circ}$, because $\mathrm{Im}(f) = \mathrm{Im}(f)^{\circ\circ}$ (after the canonical identification, Theorem 10.14). □

Let us go back to our matrices.

Lemma 11.23. *Let $f : E \to F$ be a linear map between two K-vector spaces, $\mathcal{B}, \mathcal{C}$ bases of E, F. Let $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$ and consider the dual bases $\mathcal{B}^*, \mathcal{C}^*$ of E^*, F^*. Then $\mathrm{Mat}({}^t f; \mathcal{C}^*, \mathcal{B}^*) = {}^t A$.*

Proof. Let $\mathcal{B} = (e_1, ..., e_n)$, $\mathcal{C} = (v_1, ..., v_p)$ and $M = \mathrm{Mat}({}^t f; \mathcal{C}^*, \mathcal{B}^*)$. The coefficient m_{ij} of M is, by definition of a matrix associated with a linear map, the i-th coordinate of ${}^t f(v_j^*)$ in the basis $\mathcal{B}^*$. We have ${}^t f(v_j^*) = \sum_k m_{kj} e_k^*$. So ${}^t f(v_j^*)(e_i) = \sum_k m_{kj} e_k^*(e_i) = m_{ij}$. On the other hand ${}^t f(v_j^*)(e_i) = v_j^*(f(e_i))$. By definition of a matrix associated with a linear map, $f(e_i)$ is the n-th column vector of the matrix A, i.e. $f(e_i) = \sum_t a_{ti} v_t$. Therefore $v_j^*(f(e_i)) = v_j^*(\sum_t a_{ti} v_t) = a_{ji}$. In conclusion $m_{ij} = a_{ji}$. This shows that $M = {}^t A$. □

Finally, we arrive at the desired result.

Corollary 11.24. *Let $M \in M_{p,n}(K)$. The rank of M is equal to the rank of ${}^t M$. That is, the column rank of M is equal to the row rank of M.*

Proof. We can see M, after choice of bases $\mathcal{B}, \mathcal{C}$ in E, F, as the matrix of a linear map $f : E \to F$. We therefore have $\mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) = M$. The rank of f is $\dim(\mathrm{Im}(f))$ and is the column rank of the matrix M.

If we consider ${}^t f : F^* \to E^*$, then $\mathrm{Mat}({}^t f; \mathcal{C}^*, \mathcal{B}^*) = {}^t M$ (Lemma 11.22) and its rank is $\dim(\mathrm{Im}({}^t f))$, the rank of the columns of ${}^t M$, i.e. the rank of the rows of M. But for the Proposition 11.22 $\dim(\mathrm{Im}(f)) = \dim(\mathrm{Im}({}^t f))$. So the column rank of M is equal to the row rank of M. □

The rank is an important invariant that allows us, among other things, to describe the quotient set $M_{p,n}(K)/\sim$.

Proposition 11.25. *Let $f : E \to F$ be a linear map between two K-vector spaces. Then $\mathrm{rk}(f) = r$ if and only if there exist bases $\mathcal{B}$, $\mathcal{C}$ of E, F such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) = M_r$ where:*

$$M_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Here I_r is the identity matrix $r \times r$ and the zeros indicate null matrices of the appropriate size.

Proof. (1) Suppose $\mathrm{rk}(f) = r$. From the Rank Theorem $\dim \mathrm{Ker}(F) = n - r$, where $n := \dim E$. Let $(e_{r+1}, \ldots, e_n)$ be a basis of $\mathrm{Ker}(f)$. By the incomplete basis theorem we can complete this basis to a basis of E: $\mathcal{B} = (e_1, \ldots, e_r, e_{r+1}, \ldots, e_n)$. We know that $f(e_1), \ldots, f(e_r)$ is a basis of $\mathrm{Im}(f)$ (see the proof of the Rank Theorem). These vectors can be completed to a basis of F: $\mathcal{C} = (f(e_1), \ldots, f(e_r), f_{r+1}, \ldots, f_p)$. At this point it is clear that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) = M_r$.

(2) If $\mathrm{Mat}(f;\mathcal{B},\mathcal{C}) = M_r$ then the rank of f is the column rank of M_r and this rank is clearly equal to r. □

We can now describe the quotient set $M_{p,n}(K)/\sim$.

Corollary 11.26. *Two matrices $M, N \in M_{p,n}(K)$ are equivalent if and only if they have the same rank.*

Proof. (1) If $M \sim N$ then M and N represent the same linear map $f : E \to F$ in different bases. Therefore $\mathrm{rk}(M) = \mathrm{rk}(f) = \mathrm{rk}(N)$ (Lemma 11.20).

(2) Conversely, suppose $\mathrm{rk}(M) = \mathrm{rk}(N) =: r$. In this case we have $M \sim M_r$ and $N \sim N_r$ (in fact we can associate to M, N linear maps f, g of rank r and just apply Proposition 11.25), therefore $M \sim N$ since $\sim$ is an equivalence relation. □

The rank of a matrix of $M_{p,n}(K)$ is always at most $m := min\{p, n\}$ and we have

$$M_{p,n}(K)/\sim = \{0, 1, ..., m\}. \tag{11.5}$$

We highlight another difference between equivalence and similarity of matrices. In $M_n(K)$ a matrix is equivalent to I_n if and only if it has rank n i.e. if and only if it is invertible, it is the generic case because n random vectors in K^n will in general be linearly independent (this will become clearer shortly). Instead, if $M \approx I_n$ then by definition $\exists P \in M_n(K)$ invertible such that $M = P^{-1}I_nP = I_n$, therefore the similarity class of I_n is $\{I_n\}$.

Exercises

Exercise 11.27. Let $f : \mathbb{R}^3 \to \mathbb{R}^2 : (x, y, z) \to (2x - z + y, y - z)$.
1) Write the matrix of f in the canonical bases.
2) Give a system of generators of $\mathrm{Im}(f)$.
3) In $\mathbb{R}^3$ consider the vectors $u = (1, 1, 0), v = (-1, 0, 0), w = (2, 0, 1)$. Show that $\mathcal{B} = (u, v, w)$ is a basis of $\mathbb{R}^3$. Determine $\mathrm{Mat}(\mathcal{B}, \mathcal{C})$ ($\mathcal{C}$ denotes the canonical basis of $\mathbb{R}^2$).

Exercise 11.28. Let $g : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to (x - y, y)$. Determine $M = \mathrm{Mat}(g, \mathcal{C}, \mathcal{C})$ ($\mathcal{C}$ is the canonical basis). State whether M is invertible and if the answer is affirmative, compute M^{-1}.

Exercise 11.29. Let E be a K-vector space of dimension n. Let $f \in \mathrm{End}(E)$ be such that there exist two linearly independent vectors $u, v \in E$ (hence $n \geq 2$) such that $f(u) = v$.
1) Show that there exists $g \in \mathrm{End}(E)$ such that $g(u) = g(v) = u$.
2) Show that f and g do not commute (i.e. $f \circ g \neq g \circ f$).
3) So if an endomorphism h commutes with all the other endomorphisms, we must have $h(u)$ and u linearly dependent $\forall u \in E$, i.e. $h(u) = \alpha_u u$.

Let $\mathcal{B} = (e_i)$ be a basis of E (assume $n \geq 2$). Considering the vector $u := e_1 + \cdots + e_n$, conclude that $h = \alpha \cdot Id$ (i.e. the proportionality factor between u and $h(u)$ does not depend on u). In conclusion, the only endomorphisms that commute with all the other endomorphisms are all and only those of the form $\alpha \cdot Id$.
4) Let $A \in M_n(K)$, conclude that $AM = MA$, $\forall M \in M_n(K) \Leftrightarrow A = \alpha I_n$ for some $\alpha \in K$.

Exercise 11.30 (Magic matrices)**.** A matrix $M = (a_{ij}) \in M_3(\mathbb{R})$ is said to be *magic* if the eight sums of the coefficients of the three rows, the three columns and the two diagonals are equal. So M is magic (of sum $s(M) = \alpha$) if: $\sum_{i=1}^{3} a_{ij} = \alpha$, $\sum_{j=1}^{3} a_{ij} = \alpha$, for each i, j, $\sum_{i=1}^{3} a_{ii} = \alpha$ and $a_{13} + a_{22} + a_{31} = \alpha$. We denote by $\mathcal{M}$ the set of magic matrices:

$$\mathcal{M} = \{M \in M_3(\mathbb{R}) \mid \exists \alpha \in \mathbb{R}, M \text{ is magic with } s(M) = \ aG\}.$$

1) Prove that $\mathcal{M}$ is a subvector space of $M_3(\mathbb{R})$.
2) If $M \in M_3(\mathbb{R})$

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

we set

$$\sigma(M) = \begin{pmatrix} a_{13} & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \end{pmatrix}$$

(the first and third columns have been permuted). Observe that if M is magic, then $\sigma(M)$ is also magic. We therefore have a map $\sigma : \mathcal{M} \to \mathcal{M}$. Tell whether this map is: a) linear, b) injective, c) surjective.

3) We denote by $\mathcal{M}_-$ the set of antisymmetric magic matrices ($a_{ij} = -a_{ji}$). If $M \in \mathcal{M}_-$, what is the value of $s(M)$?

4) Determine $\mathcal{M}_-$, show that it is a subvector space of $\mathcal{M}$ and determine its dimension.

5) Conclude that $\dim(\mathcal{M}) \geq 2$.

Exercise 11.31. Consider the following vectors of $\mathbb{R}^3$ (coordinates in the canonical basis $\mathcal{C}$): $v_1 = (1, 1, 0)$, $v_2 = (0, 1, 1)$, $v_3 = (-1, 0, 2)$, $u = (1, -1, -2)$.

(1) Show that $\mathcal{B} = (v_1, v_2, v_3)$ is a basis of $\mathbb{R}^3$.

(2) Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear map such that

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Determine the coordinates of $f(u)$ in the basis $\mathcal{B}$.

(3) Determine $\mathrm{Mat}(f; \mathcal{C}, \mathcal{C})$.

Exercise 11.32. In $\mathbb{R}^2$ let $v_1 = (1, 1)$, $v_2 = (0, -1)$ (coordinates in the canonical basis $\mathcal{C}$).

(1) Show that $\mathcal{B} = (v_1, v_2)$ is a basis of $\mathbb{R}^2$.

(2) Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map such that

$$\mathrm{Mat}(f; \mathcal{C}, \mathcal{B}) = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Show that f is bijective.

(3) Determine $\mathrm{Mat}(f^{-1}; \mathcal{B}, \mathcal{C})$, $\mathrm{Mat}(f^{-1}; \mathcal{C}, \mathcal{C})$.

Exercise 11.33. In the notations of Exercise 11.32 let $g : \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$\mathrm{Mat}(g; \mathcal{C}, \mathcal{B}) = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}.$$

Determine $\mathrm{Mat}(f \circ g; \mathcal{B}, \mathcal{C})$.

Exercise 11.34. For $\alpha \in \mathbb{R}$ consider the vectors of $\mathbb{R}^3$: $u_\alpha = (-1, 2\alpha, -1)$, $v_\alpha = (1, \alpha, 0)$, $w_\alpha = (0, 1, 2\alpha)$. For each $\alpha \in \mathbb{R}$ determine the rank of $(u_\alpha, v_\alpha, w_\alpha)$.

Exercise 11.35. Let $M \in M_{n,p}(K)$ be such that $rk(M) = r$ and

$$\mathcal{V} = \{B \in M_{p,m}(K) \mid MB = 0\}.$$

Prove that $\mathcal{V}$ is a subvector space of $M_{p,m}(K)$ and determine its dimension.

Exercise 11.36. Consider the following vectors of $\mathbb{R}^3$ (coordinates in the canonical basis C):

$$v_1 = (1,1,0), v_2 = (0,1,1), v_3 = (-1,0,2), u = (1,-1,-2).$$

(1) Show that $\mathcal{B} = (v_1, v_2, v_3)$ is a basis of $\mathbb{R}^3$.

(2) Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear map such that

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{C}) = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Determine the coordinates of $f(u)$ in the basis $\mathcal{B}$.

Exercise 11.37. (1) Does there exist a linear map $g : \mathbb{R}^2 \to \mathbb{R}^3$ such that: $g((1,1)) = (1,0,3)$, $g((0,-2)) = (4,-1,0)$ and $g((2,6)) = (-6,3,6)$?

(2) Show that $\forall A \in M_{2,3}(K), B \in M_{3,2}(K)$ there exists $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $X \neq 0$ such that $BAX = 0$.

Exercise 11.38. Let E be a K-vector space of dimension three and $V, W \subset E$ two subvector spaces of dimension two. Show that there exists a linear automorphism, f, of E such that $f(V) = W$ and $f(W) = V$.

Exercise 11.39. Let K be a field. In K^4, with coordinates (x,y,z,t) (in the canonical basis), let $W \subset K^4$ be the subvector space of equations $x = 0 = y$. So

$$W = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \\ t \end{pmatrix} \mid z, t \in K \right\}.$$

Let $V \subset K^4$ be a 2-dimensional subvector space such that $V \cap W = \{0\}$. We assume

$$\det \begin{pmatrix} x & z \\ y & t \end{pmatrix} = 0$$

for every $(x, y, z, t) \in V$. Show that V has a basis of the form

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ \lambda \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \lambda \end{pmatrix}$$

with $\lambda \in K$.

Exercise 11.40. Let E be a 2-dimensional K-vector space and $f \in \mathrm{End}(E)$. Let

$$C = \begin{pmatrix} a \\ b \end{pmatrix} \in M_{2,1}(K).$$

(1) Show that if f is not a homothety (i.e. f is not of the form λId_E for $\lambda \in K$) and if $b \neq 0$, then there exists a basis $\mathcal{B}$ of E such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ has C as its first column.
(2) Show that the assumption $b \neq 0$ is essential.

Chapter 12

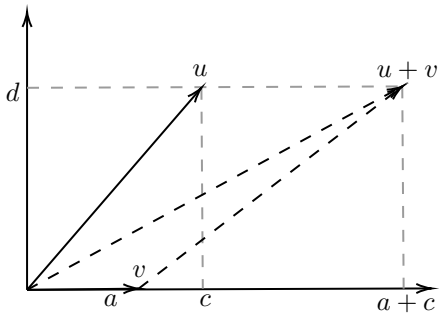
Determinants

We have already seen that in an n-dimensional vector space E, n vectors $v_1, \ldots, v_n$ form a basis if and only if they are linearly independent. We want to associate a number, D, to n vectors that tells us whether or not the vectors are independent. For example, the vectors are linearly independent if and only if $D \neq 0$.

Let us consider the case $n = 2$, $E = \mathbb{R}^2$. Let $v = (a, b)$, $u = (c, d)$. We know that v and u are linearly dependent if and only if they are proportional i.e. if and only if $\frac{a}{b} = \frac{c}{d}$. We are assuming $bd \neq 0$. To avoid having to distinguish the cases $b = 0, d = 0$ it is preferable to rewrite the relation in this form: $ad = bc$. We can conclude that the vectors $v = (a, b)$ and $u = (c, d)$ are linearly independent if and only if $D := ad - bc \neq 0$.

In these terms, however, it is not clear how to generalize to the case $n = 3$. We look for a geometric interpretation of our number $D = ad - bc$.

If u, v are independent, $|ac - bd|$ is the *area of the parallelogram* constructed on u, v. It is an exercise in analytic geometry in the plane that perhaps you did in high school. Here is a proof in the case $b = 0$:



Let $D = (0, d)$, $M = (a + c, 0)$, $S = (a + c, d)$, and $O = (0, 0)$. The area of the rectangle $OMSD$ is $d(a+c)$. To get the area of the parallelogram built on v, u you need to remove the area of the two right triangles ODu and vMS, they have the same area $\frac{dc}{2}$; so the searched area is ad.

Hence, we can visualize our number D, and in fact two vectors are dependent if and only if they are aligned and in this case the parallelogram is degenerated, its area is equal to zero.

Well this gives us the idea we should generalize. For three vectors in $\mathbb{R}^3$ we consider the volume of the parallelepiped constructed on the three vectors: the vectors are independent if and only if this volume is not zero. In fact, the three vectors are dependent if and only if they are contained in a plane and in this case the parallelepiped is degenerated, its volume is equal to zero.

We take this opportunity to underline the following fact: in a space E of dimension n, p vectors $v_1, \ldots, v_p$ are linearly independent if and only if $\dim F = p$, where $F = \langle v_1, \ldots, v_p \rangle$. This comes from the fact that every system of generators contains a basis. If $\dim F < p$ it is possible to extract from $v_1, \ldots, v_p$ a basis with less than p elements and one of the v_i will be a linear combination of the others: the vectors are dependent. If $\dim F = p$, no v_i is a linear combination of the others: the vectors are independent.

Now, we have an idea on how to associate a number D to n vectors of a space of dimension n that detects whether or not these vectors are independent. This number is called the *determinant* of the n vectors. However, we will not look for complicated analytic formulas, we will instead approach the question in a more abstract and elegant way.

12.1. Multilinear maps and forms

Let $E_1, \ldots, E_n, F$ be sets, $E := E_1 \times \cdots \times E_n$ their Cartesian product and $f : E \to F : (x_1, \ldots, x_n) \to f(x_1, \ldots, x_n)$ a map.

If we fix the variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ (all except x_i) for example $x_1 = a_1, \ldots, x_{i-1} = a_{i-1}, x_{i+1} = a_{i+1}, \ldots, x_n = a_n$, we obtain a map $f_i : E_i \to F : x_i \to f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$. It would have been more precise to write $f_{a_1,\ldots a_n}$ instead of f_i because the map depends on the a_i but for simplicity of notation we will settle for f_i.

Definition 12.1. The map f_i is *the partial map* from E_i to F relative to the values assigned to the variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$.

Now, suppose that $E_1, \ldots, E_n, F$ are K-vector spaces.

Definition 12.2. With the previous notations the application f is *multilinear* if for every i, $1 \leq i \leq n$, the partial map $f_i : E_i \to F$ is linear.

In the following we will be interested in the case $E_1 = \cdots = E_n$ and $F = K$. Then we will consider multilinear maps $f : E^n \to K$. Such a map is called a *n-linear form* (or multilinear form). If $n = 2$ it is called *bilinear form* (and *trilinear* if $n = 3$).

Remark 12.3. Be careful not to confuse linear applications and n-multilinear applications. If $f : E_1 \times E_2 \to F$ is bilinear then

(1) $f((x_1 + x_2, y_1 + y_2)) = f(x_1, y_1) + f(x_1, y_2) + f(x_2, y_1) + f((x_2, y_2)$;

(2) $f(\lambda(x, y)) = f(\lambda x, \lambda y) = \lambda^2 f(x, y)$.

If instead f is linear then

(1) $f((x_1 + x_2, y_1 + y_2) = f((x_1, y_1) + (x_2, y_2)) = f(x_1, y_1) + f(x_2, y_2)$;

(2) $f(\lambda(x, y)) = \lambda f(x, y)$.

Example 12.4. (1) The map $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x, y) \to xy$ is a bilinear map.

(2) The map $E \times E^* \to K : (v, \varphi) \to \varphi(v)$ is bilinear.

Proposition 12.5. *The set of multilinear maps from $E_1 \times \cdots \times E_n$ to F is a K-vector space for the usual addition and external multiplication operations.*

Proof. Verification is left to the reader. □

From now on we will only be interested in the case of multilinear forms $f : E^n = E \times \cdots \times E \to K : (u_1, \ldots, u_n) \to f(u_1, \ldots, u_n)$.

Definition 12.6. A multilinear form $f : E^n \to K$ is called *alternating* if $u_i = u_j$ implies $f(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) = 0$.

So if f is alternating and if two of the vectors $u_1, \ldots, u_n$ are equal, we have $f((u_1, \ldots, u_n) = 0$.

Lemma 12.7. *Let $f : E^n \to K$ be an alternating multilinear form, then:*

$$f(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) = -f(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n).$$

Proof. Consider $f(u_1, \ldots, u_i + u_j, \ldots, u_j + u_i, \ldots, u_n)$. Since the i-th coordinate $u_i + u_j$ is equal to the j-th $u_j + u_i$, we have $f(u_1, \ldots, u_i + u_j, \ldots, u_j + u_i, \ldots, u_n) = 0$. By multilinearity $f(u_1, \ldots, u_i + u_j, \ldots, u_j + u_i, \ldots, u_n) = f(u_1, \ldots, u_i, \ldots, u_j + u_i, \ldots, u_n) + f(u_1, \ldots, u_j, \ldots, u_j + u_i, \ldots, u_n)$. In fact, if we fix all the variables except the i-th one we are left with a partial linear map. Let us look at the first term from the member on the right.

By fixing all the variables except the j-th one we have a linear partial map. Note that $f(u_1, \ldots, u_i, \ldots, u_j + u_i, \ldots, u_n)$ is equal to

$$f(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) + f(u_1, \ldots, u_i, \ldots, u_i, \ldots, u_n)$$

which in turn is equal to $f(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n)$. In fact

$$f(u_1, \ldots, u_i, \ldots, u_i, \ldots, u_n) = 0$$

since f is alternating.

Now, $f(u_1, \ldots, u_j, \ldots, u_j + u_i, \ldots, u_n)$ equals to

$$f(u_1, \ldots, u_j, \ldots, u_j, \ldots, u_n) + f(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n)$$

which in turn is equal to $f(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n)$. Finally, $0 = f(u_1, \ldots, u_i + u_j, \ldots, u_j + u_i, \ldots, u_n) = f(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) + f(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n)$ and the lemma is proved. □

A multilinear form that verifies the lemma property is called *antisymmetric*.

Lemma 12.8. *Let $f : E^n \to K$ be an alternating multilinear form. If the vectors $u_1, \ldots, u_n$ are linearly dependent, then $f(u_1, \ldots, u_n) = 0$.*

Proof. If the vectors are dependent, one of them is written as a linear combination of the others. Reordering the indexes, if needed, we can assume $u_1 = \alpha_2 u_2 + \cdots + \alpha_n u_n$. So $f(u_1, \ldots, u_n) = f(\alpha_2 u_2 + \cdots + \alpha_n u_n, u_2, \ldots, u_n)$. By multilinearity $f(\alpha_2 u_2 + \cdots + \alpha_n u_n, u_2, \ldots, u_n) = \alpha_2 f(u_2, u_2, \ldots, u_n) + \cdots + \alpha_n f(u_n, u_2, \ldots, u_n)$. We see that in each term there are two equal variables therefore, since f is alternating, $\alpha_2 f(u_2, u_2, \ldots, u_n) + \cdots + \alpha_n f(u_n, u_2, \ldots, u_n) = 0$. □

Remark 12.9. So if there exists an alternating p-form f such that $f(v_1, \ldots, v_p) \neq 0$, the vectors $v_1, \ldots, v_p$ are linearly independent. This is pretty much what we are looking for.

Definition 12.10. We will denote by $\mathcal{A}^p(E)$ the set of alternating multilinear p-forms on E. It is easily verified that $\mathcal{A}^p(E)$ is a K-vector space.

As a consequence of Lemma 12.8 we have:

Corollary 12.11. *Let E be a K-vector space of dimension n. If $p > n$ then $\mathcal{A}^P(E) = \{0\}$.*

Proof. In fact p vectors of E are always dependent if $p > n$. □

We are looking for a criterion to test the independence of n vectors in a vector space of dimension n. Due to Remark 12.9 it seems natural to consider $\mathcal{A}^n(E)$ where $n = \dim E$.

We observe that $\mathcal{A}^1(E) = E^*$ and that, according to the corollary just seen, $\mathcal{A}^n(E)$ is the last space $\mathcal{A}^i(E)$ that perhaps is not trivial.

For these reasons we will not develop the complete theory of alternating forms in all generality but we will limit ourselves to the case $n = \dim E$.

12.2. Alternating 2-forms

Let E be a 2-dimensional K-vector space, we want to describe $\mathcal{A}^2(E)$. Let $\mathcal{B} = (e_1, e_2)$ be a basis of E and $f \in \mathcal{A}^2(E)$. Taking $u, v \in E$ how do we compute $f(u, v)$?

We write the vectors on the basis $\mathcal{B}$: $u = ae_1 + be_2$, $v = ce_1 + de_2$. Then $f(u, v) = f(ae_1+be_2, ce_1+de_2)$. If we fix the second variable we are left with a linear map, therefore: $f(u, v) = f(ae_1 + be_2, ce_1 + de_2) = af(e_1, ce_1 + de_2) + bf(e_2, ce_1 + de_2)$. Now, by fixing the first variable we are left with a linear map of the second variable $f(u, v) = acf(e_1, e_1)+adf(e_1, e_2)+bcf(e_2, e_1)+bdf(e_2, e_2)$. But $f(e_1, e_1) = f(e_2, e_2) = 0$ because f is alternating. So $f(u, v) = adf(e_1, e_2)+bcf(e_2, e_1)$. Finally, using the Lemma 12.7:

$$f(u, v) = (ad - bc)f(e_1, e_2).$$

Look who is back, $ad - bc$, our old determinant. The moral of this calculation is that f is completely determined by $f(e_1, e_2)$. Knowing $f(e_1, e_2)$ we can compute $f(u, v)$ for every pair (u, v) of vectors of E.

Proposition 12.12. *Let E be a 2-dimensional K-vector space. We have:*

(1) $\dim \mathcal{A}^2(E) = 1$.

(2) *If $\mathcal{B} = (e_1, e_2)$ is a basis of E there exists an unique alternating bilinear form, f, such that $f(e_1, e_2) = 1$. This form is denoted $\det_{\mathcal{B}}$ and is called the determinant in the basis $\mathcal{B}$.*

Proof. (1) Let $f, g \in \mathcal{A}^2(E)$ be two non zero elements. Set $f(e_1, e_2) = \alpha$, $g(e_1, e_2) = \beta$. Then $\alpha\beta \neq 0$ and $f(u, v) = \frac{\alpha}{\beta} g(u, v)$ for every $u, v \in E$, so $f = \frac{\alpha}{\beta} g$. This shows $\mathcal{A}^2(E) = \langle g \rangle$.

(2) It follows from the discussion in the beginning of this section. □

Proposition 12.13. *Let E be a K-vector space of dimension two.*

(1) *Two vectors $u, v \in E$ are linearly independent if and only if $f(u, v) \neq 0$, $\forall f \in \mathcal{A}^2(E), f \neq 0$.*

(2) *The vectors u, v are linearly independent if and only if there exists a basis $\mathcal{B}$ of E such $\det_{\mathcal{B}}(u, v) \neq 0$.*

Proof. (1) If u and v are independent they form a basis $\mathcal{C} = (u, v)$ and $\det_{\mathcal{C}}(u, v) = 1 \neq 0$. Now if $f \in \mathcal{A}^2(E), f \neq 0$, then $f = \lambda \det_{\mathcal{C}}$ $(\lambda \neq 0)$ and $f(u, v) = \lambda$.

Conversely, if $g(u, v) \neq 0$ for every $g \in \mathcal{A}^2(E), g \neq 0$ (actually one is enough), then u and v are linearly independent (Remark 12.9).

(2) It follows from (1) and the fact that $\dim \mathcal{A}^2(E) = 1$. □

We therefore have a way to see whether two vectors of E are independent: we take their coordinates in a basis $\mathcal{B} = (e_1, e_2)$: $u = ae_1 + be_2$, $v = ce_1 + de_2$ and calculate their determinant in the basis $\mathcal{B}$. Let us look at a way of representing all of this: Consider the matrix whose column vectors are the components of u, v:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

We have that $\det_{\mathcal{B}}(u, v)$ is the difference of the cross products. The determinant is represented in the form of a table but to distinguish it from matrices, slashes are used instead of brackets:

$$\det_{\mathcal{B}}(u, v) = \begin{vmatrix} a & c \\ b & d \end{vmatrix} := ad - bc.$$

It would be more correct to write

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix}_{\mathcal{B}}$$

because the determinant of the two vectors depends on the chosen basis but is not used in practice (the coordinates indicate the basis).

Well, now we have to repeat everything in the general case $\dim E = n$. We proceed exactly the same way, the math is just a little more complicated.

12.3. Alternating n-forms

Let E be a K-vector space of dimension n with a basis $\mathcal{B} = (e_1, \ldots, e_n)$ and $u_1, \ldots, u_n$ vectors of E. Let us consider the coordinates of these vectors in the basis $\mathcal{B}$:

$$\begin{aligned} u_1 &= \quad \xi_1^1 e_1 + \cdots + \xi_n^1 e_n; \\ &\vdots \\ u_n &= \quad \xi_1^n e_1 + \cdots + \xi_n^n e_n. \end{aligned}$$

Let $f : E^n \to K$ be an alternating n-form from E^n into K and let us compute $f(u_1, \ldots, u_n)$. We have $f(u_1, \ldots, u_n) = f(\xi_1^1 e_1 + \cdots + \xi_n^1 e_n, \ldots, \xi_1^n e_1 + \cdots + \xi_n^n e_n)$.

If we expand the right hand side, using the multilinearity of f, we obtain a sum of terms of the form:

$$\xi_{i_1}^1 \ldots \xi_{i_n}^n f(e_{i_1}, \ldots, e_{i_n}).$$

It is like in the expansion of the binomial, for the first vector $\xi_1^n e_1 + \cdots + \xi_n^n e_n$ we have to choose one of the terms, say $\xi_{i_1}^1 e_{i_1}$, for the second vector one of the terms of the sum $\xi_1^2 e_1 + \cdots + \xi_n^2 e_n$, say $\xi_{i_2}^2 e_{i_2}$ and so on, up to the last vector $\xi_1^n e_1 + \cdots + \xi_n^n e_n$ which gives us $\xi_{i_n}^n e_{i_n}$. So we have $f(\xi_{i_1}^1 e_{i_1}, \ldots, \xi_{i_n}^n e_{i_n}) = \xi_{i_1}^1 \ldots \xi_{i_n}^n . f(e_{i_1}, \ldots, e_{i_n})$. And we must do this in every possible way.

We observe that $i_1, \ldots, i_n$ correspond to an application from $X = \{1, 2, \ldots, n\}$ into itself: $X \to X : j \to i_j$. We have $\text{card}(A(X, X)) = n^n$, where $A(X, X)$ denotes the set of all maps from X to itself. The sum therefore contains n^n terms. In the case $n = 2$ we had four terms (corresponding to $f(e_1, e_1), f(e_1, e_2), f(e_2, e_1), f(e_2, e_2)$.

But since f is alternating if $i_k = i_t$, the value is zero and we can leave out this term. Said differently we must consider only the maps $X \to X$ which are injective. Since X is finite an injective map $X \to X$ is bijective. That is, we only have to look at the *permutations* of X. Therefore, let

$$\mathcal{S}_n = \{\sigma : X \to X \mid \sigma \text{ is bijective}\}$$

be the *symmetric group* on n elements. We have:

$$f(u_1, \ldots, u_n) = \sum_{\sigma \in \mathcal{S}_n} \xi_{\sigma(1)}^1 \ldots \xi_{\sigma(n)}^n . f(e_{\sigma(1)}, \ldots, e_{\sigma(n)}). \tag{12.1}$$

This sum contains $n!$ terms (in the case $n = 2$ they correspond to $f(e_1, e_2), f(e_2, e_1)$).

Now, the idea is to express everything as a function of $f(e_1, \ldots, e_n)$. In fact we know that $f(\ldots, e_i, \ldots, e_j, \ldots) = -f(\ldots, e_j, \ldots, e_i, \ldots)$ (Lemma 12.7); that is, when two variables exchange, the sign changes.

Starting from $f(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$, with subsequent exchanges we put the indexes back in place and if the number of exchanges is even, the sign does not changes, if it is odd the sign changes.

For example, if we have ($n = 3$) $f(e_2, e_3, e_1)$, we swap e_2 and e_3: $f(e_2, e_3, e_1) = -f(e_3, e_2, e_1)$. Now we exchange e_3 and e_1: $-f(e_3, e_2, e_1) = f(e_1, e_2, e_3)$ and conclude that $f(e_2, e_3, e_1) = f(e_1, e_2, e_3)$.

Observe that we could have done it differently: starting from $(2, 3, 1)$ (we simplify the writing) we can exchange 2 and 3: $(3, 2, 1)$, then 2 and 1: $(3, 1, 2)$, then 1 and 3: $(1, 3, 2)$ and finally 3 and 2: $(1, 2, 3)$. The number of exchanges is still even.

On the other hand, it can not be otherwise. If there exists an alternating non-zero f then $f(e_2, e_3, e_1)$ equals to $f(e_1, e_2, e_3)$ or $-f(e_1, e_2, e_3)$, but it can not be equal $f(e_2, e_3, e_1)$ at a time and $-f(e_2, e_3, e_1)$ at another.

The problem is that we do not yet know that there is a non-zero alternating n-form. Let us try to summarize what we have said so far.

Definition 12.14. A permutation $\tau \in \mathcal{S}_n$ is a *transposition* if it exchanges two elements while leaving the others unchanged. That is, $\tau(i) = k, \tau(k) = i$ and $\tau(m) = m$ if $m \notin \{i, k\}$.

Lemma 12.15. *Each permutation $\sigma \in \mathcal{S}_n$ is written as a product of at most $n-1$ transpositions.*

Proof. Let us write a permutation $\sigma \in \mathcal{S}_n$ as follows:

$$\begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix}.$$

Putting the indexes $(\sigma(1), \ldots, \sigma(n))$ back means applying σ^{-1}. We then show that σ^{-1} is a product of at most $n-1$ transpositions. Since every permutation is the inverse of a permutation ($\sigma = (\sigma^{-1})^{-1}$) this proves the lemma. If $\sigma(1) \neq 1$, we look for 1 and exchange it with $\sigma(1)$. Now, 1 is in place and we move to second index. If $\sigma(1) = 1$ we move to $\sigma(2)$. After at most $n-1$ such operations $1, 2, \ldots, n-1$ are in place and therefore also n is. □

Lemma 12.16. *Let $\sigma \in \mathcal{S}_n$ be a permutation. We have seen that σ is written as a product of transpositions. Let $\sigma = \tau_1 \circ \cdots \circ \tau_r$; and $\sigma = \theta_1 \circ \cdots \circ \theta_t$, where τ_i, θ_j are transpositions. Then $r \equiv t \pmod 2$, i.e. r and t have the same parity.*

Proof. This is an algebra result. Let us give an idea of the proof.

(1) If $f : E^n \to K$ is antisymmetric and if $\sigma = \tau_1 \circ \cdots \circ \tau_r$ is a decomposition of σ into a product of transpositions then we have

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = (-1)^r f(x_1, \ldots, x_n).$$

The proof is by induction on r. The initial case $r = 1$ is true because f is antisymmetric. Then we write $\sigma = \tau \circ \nu$, with $\tau = \tau_1$, $\nu = \tau_2 \circ \cdots \circ \tau_n$ and set $y_i = x_{\tau(i)}$. We have $x_{\sigma(i)} = y_{\nu(i)}$. Therefore $f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = f(y_{\nu(1)}, \ldots, y_{\nu(n)}) = (-1)^{r-1} f(y_1, \ldots, y_n)$, by induction hypothesis. Now $f(y_1, \ldots, y_n) = f(x_{\tau(1)}, \ldots, x_{\tau(n)})$. But since τ is a transposition and since f is antisymmetric $f(x_{\tau(1)}, \ldots, x_{\tau(n)}) = -f(x_1, ..., x_n)$ and we have the result.

(2) Then we show that

$$g : \mathbb{R}^n \to \mathbb{R} : (x_1, \ldots, x_n) \to \prod_{1 \leq i < j \leq n} (x_i - x_j)$$

is antisymmetric. It is a bit boring but not hard, left to the reader.

(3) Let $\sigma = \tau_1 \circ \cdots \circ \tau_t = \xi_1 \circ \cdots \circ \xi_t$ be two decompositions of σ as a product of transpositions. By (1) we have $g(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = (-1)^r g(x_1, \ldots, x_n) = (-1)^t g(x_1, \ldots, x_n)$. So r and t have the same parity

For details we refer to [**DF04**, Section 1.3]. □

Definition 12.17. Let $\sigma \in \mathcal{S}_n$ be a permutation. If $\sigma = \tau_1 \circ \cdots \circ \tau_r$, where $\tau_1, \ldots, \tau_r$ are transpositions, the *signature* of σ is $\varepsilon(\sigma) := (-1)^r$.

Remark 12.18. We observe that the signature is well defined due to Lemma 12.16.

Having said this we can rewrite (12.1) as follows:

$$f(u_1, \ldots, u_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \xi^1_{\sigma(1)} \ldots \xi^n_{\sigma(n)} . f(e_1, \ldots, e_n). \tag{12.2}$$

At this point we see that f is completely determined by $f(e_1, \ldots, e_n)$ as in the case $n = 2$. Therefore we have:

Proposition 12.19. *Let E be a K-vector space of dimension n. Then*

(1) *The vector space $\mathcal{A}^n(E)$ has dimension one.*

(2) *If $\mathcal{B} = (e_1, \ldots, e_n)$ is a basis of E, there exists an unique alternating multilinear form f such that $f(e_1, \ldots, e_n) = 1$; this form f is denoted by $\det_{\mathcal{B}}$. So $\det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$.*

Proof. The proof is analogous to that of Proposition 12.12. □

We therefore see that if $\mathcal{B} = (e_i)$ is a basis of E and if $u_j = \sum_i \xi^j_i e_i$ we have:

$$\det{}_{\mathcal{B}}(u_1, \ldots, u_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \xi^1_{\sigma(1)} \ldots \ xi^n_{\sigma(n)}. \tag{12.3}$$

The scalar $\det_{\mathcal{B}}(u_1, \ldots, u_n)$ is the determinant of the vectors $u_1, \ldots, u_n$ with respect to the basis $\mathcal{B}$. As in the case $n = 2$ we have:

Proposition 12.20. *Let E be a K-vector space of dimension n, and $u_1, \ldots, u_n$ vectors of E. Then $u_1, \ldots, u_n$ are linearly independent if and only if one of the following two equivalent conditions is satisfied:*

(1) *$\forall f \in \mathcal{A}^n(E)$, $f \neq 0$, $f(u_1, \ldots, u_n) \neq 0$;*

(2) *there exists a basis $\mathcal{B}$ such that $\det_{\mathcal{B}}(u_1, \ldots, u_n) \neq 0$.*

Proof. At this point the proof of this proposition is a simple generalization of that of Proposition 12.13. □

As in the case $n = 2$ it is customary to write the determinant as an $n \times n$ table where the j-th column is made up of the coordinates of the j-th vector, u_j:

$$\det_{\mathcal{B}}(u_1, \ldots, u_n) = \begin{vmatrix} \xi^1_1 & \cdots & \xi^n_1 \\ \vdots & \vdots & \vdots \\ \xi^n_1 & \cdots & \xi^n_n \end{vmatrix} := \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \xi^1_{\sigma(1)} \ldots \xi^n_{\sigma(n)}.$$

At this point we have proved for any n the results shown in the case $n = 2$. The only thing we do not know how to do is calculate an $n \times n$ determinant effectively. For this some preliminaries are needed.

12.4. The determinant of an endomorphism

The determinant of n vectors of E ($\dim E = n$) depends on a basis $\mathcal{B}$ of E. We now show that there exists, for an endomorphism $f : E \to E$, a determinant that does not depend on the choice of a basis of E.

Proposition 12.21. *Let E be a finite-dimensional K-vector space and $u \in \mathrm{End}(E)$. There exists a unique scalar $\det(u)$, such that for every $(x_1, \ldots, x_n) \in E^n$ and for every $f \in \mathcal{A}^n(E)$ we have:*

$$f(u(x_1), \ldots, u(x_n)) = \det(u) f(x_1, \ldots, x_n).$$

Proof. Let $f \in \mathcal{A}^n(E)$ and set $f_u : E^n \to K : (x_1, \ldots, x_n) \to f(u(x_1), \ldots, u(x_n))$. We have that f_u is multilinear alternating. In fact, by fixing all the variables except the i-th one, the corresponding partial application is $f_i \circ u$ which is linear as composed of two linear applications (here f_i is the partial application with respect to $u(x_1), \ldots u(x_n)$, without $u(x_i)$).

The map is alternating because f is alternating. Since $\dim \mathcal{A}^n(E) = 1$, f_u is proportional to f that is $f_u = \lambda f$. This will be true for $f \in \mathcal{A}^n(E)$. We show that the proportionality coefficient is constant (i.e. it does not depend on f). Let $g \in \mathcal{A}^n(E)$. Then $g = \xi f$ (because $\dim \mathcal{A}^n(E) = 1$). We have $g_u = (\xi f)_u = \xi f_u = \xi(\lambda f) = \lambda(\xi f) = \lambda g$. We conclude by setting $\det(u) = \lambda$. □

Remark 12.22. The scalar $\det u$ is called the *determinant* of the endomorphism u, $\det u$ does not depend on the choice of a basis.

Let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E. We use Proposition 12.21 with $f = \det_{\mathcal{B}}$, $x_i = e_i$ to get $\det_{\mathcal{B}}(u(e_1, \ldots, u(e_n)) = \det(u)$.

Let $M = \mathrm{Mat}(u; \mathcal{B}, \mathcal{B})$, $M = (a_{ij})$. We have that $\det(u)$ is the determinant of the column vectors of the matrix M, i.e

$$\det(u) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma).a_{\sigma(1)1}.a_{\sigma(2)2} \ldots a_{\ sG(n)n}.$$

This leads us to the following definition:

Definition 12.23 (Determinant of a matrix)**.** Let $M \in M_n(K)$, $M = (a_{ij})$. The *determinant* of M, $\det(M)$ is the scalar:

$$\det(M) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \ldots a_{\sigma(n)n}.$$

To summarize we have:

(1) The determinant of the vectors v_i in the basis $\mathcal{B}$: $\det_{\mathcal{B}}(v_1, \ldots, v_n)$;

(2) The determinant of $u \in \mathrm{End}(E)$;

(3) The determinant of $M \in M_n(K)$, $\det(M)$.

These three things are connected in the following way:

(i) The determinant $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ is defined by (12.3).

(ii) The determinant $\det(u)$ is defined in Proposition 12.21.

(iii) The $\det(M)$ is the determinant of the column vectors (seen as vectors of K^n) of the matrix M with respect to the canonical basis of K^n in the sense of (i).

If $M = \mathrm{Mat}(u; \mathcal{B}, \mathcal{B})$ then $\det(u) = \det(M)$ (with the basis $\mathcal{B}$ in the domain and in the codomain).

These various avatars of the determinant will allow us to prove some properties of determinants.

Proposition 12.24. *Let E be a K-vector space of dimension n.*

(1) *We have* $\det(Id_E) = 1$.

(2) *If* $u, v \in \mathrm{End}(E)$, $\det(u \circ v) = \det(u)\det(v)$.

(3) *An endomorphism $u \in \mathrm{End}(E)$ is invertible if and only if* $\det(u) \neq 0$. *If u is invertible then* $\det(u^{-1}) = \frac{1}{\det(u)}$.

Proof. (1) Let $\mathcal{B} = (e_i)$ be a basis of E, then $\det(Id_E) = \det_{\mathcal{B}}(Id_E(e_1), \ldots, Id_E(e_n)) = \det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$.

(2) Let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E. We have $\det_{\mathcal{B}}((u \circ v)(e_1), \ldots, (u \circ v)(e_n)) = \det(u \circ v)$ (Proposition 12.21). But, again by Proposition 12.21 $\det_{\mathcal{B}}((u \circ v)(e_1), \ldots, (u \circ v)(e_n)) = \det(v)\det(u)\det_{\mathcal{B}}(e_1, \ldots, e_n) = \det(v)\det(u)$.

(3) If u is invertible we have $u \circ u^{-1} = Id_E$ and from (1) and (2) it follows that $\det(u) \neq 0$ and $\det(u^{-1}) = \frac{1}{\det(u)}$.

It remains to be seen that $\det(u) \neq 0$ implies that u is invertible. Let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E. We have $\det(u) = det_{\mathcal{B}}(u(e_1), \ldots, u(e_n)) \neq 0$, so the vectors $u(e_1), \ldots, u(e_n)$ are independent, i.e. they form a basis of E. Then u transforms the basis $\mathcal{B}$ into a basis, and hence it is bijective. □

Now, we can translate these results in terms of matrices: if $A, B \in M_n(K)$, we take a vector space E of dimension n, a basis $\mathcal{B} = (e_i)$ of E and we have our isomorphism: $\mathrm{Mat}(-; \mathcal{B}, \mathcal{B}) : \mathrm{End}(E) \to M_n(K)$. Recall that $\mathrm{Mat}(-; \mathcal{B}, \mathcal{B})$ is also an isomorphism of rings, so the matrix product corresponds to the composition

of the maps. Our matrices correspond to endomorphisms u, v while $\det u, \det v$ correspond to $\det A, \det B$.

Corollary 12.25. (1) *We have* $\det I_n = 1$, *where* $I_n \in M_n(K)$ *is the identity matrix.*

(2) *If* $A, B \in M_n(K)$ *then* $\det(AB) = \det(A)\det(B)$.

(3) *A matrix* $A \in M_n(K)$ *is invertible if and only if* $\det(A) \neq 0$. *If* A *is invertible, then* $\det(A^{-1}) = 1/\det(A)$.

Proof. (1) It follows from the fact that $\det(Id_E) = 1$ and $\mathrm{Mat}(Id_E; \mathcal{B}, \mathcal{B}) = I_n$.

(2) Let $\mathrm{Mat}(u; \mathcal{B}, \mathcal{B}) = A$, $\mathrm{Mat}(v; \mathcal{B}, \mathcal{B}) = B$. We have $\mathrm{Mat}(u \circ v; \mathcal{B}, \mathcal{B}) = AB$ (Proposition 11.5). Now $\det(AB) = \det(u \circ v) = \det(u).\det(v) = \det(A)\det(B)$.

(3) Let $\mathrm{Mat}(u; \mathcal{B}, \mathcal{B}) = A$, then A is invertible if and only if u is, we conclude with (3) of Proposition 12.24. □

Remark 12.26. We obtained the above results without doing any calculations. Obviously there are other ways to proceed. For example, let us see how to prove point (1) using the determinant formula. So if $A = (a_{ij})$, we know that $\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$. Let us take the case of $A = I_n$, then $a_{ij} = \delta_{ij}$, where δ_{ij} is the Kronecker symbol. If $\sigma \in \mathcal{S}_n$ is not the identity then there exists $i \in \{1, \dots, n\}$ such that $\sigma(i) \neq i$ and therefore we have $a_{\sigma(i)i} = \delta_{\sigma(i)i} = 0$. This implies that in the summation we have a single term to consider: the one corresponding to $\sigma = Id$. So $\det(I_n) = \varepsilon(Id) a_{11} \dots a_{nn} = \varepsilon(Id)$. Now, it is clear that $\varepsilon(Id) = 1$, because if τ is any transposition $Id = \tau \circ \tau$. Therefore $\det(I_n) = 1$.

12.5. Computation of a determinant

Let $A \in M_n(K)$. From what we have seen so far, it turns out that the determinant is an alternating multilinear function of the columns of A. Therefore:

(1) Via a permutation σ of the columns of A, $\det(A)$ transforms into $\varepsilon(\sigma)\det(A)$. If two columns are exchanged the determinant of the matrix so obtained is $-\det(A)$.

(2) We have $\det(A) = 0$ if and only if the columns of A are linearly dependent.

(3) The determinant does not change if a linear combination of columns is added to a column.

(4) For every $\lambda \in K$, $\det(\lambda A) = \lambda^n \det(A)$. In fact $\lambda A = (\lambda C_1, \dots, \lambda C_n)$.

If we replace a column with a linear combination of all columns, the determinant changes:

$$\det(\sum_{i=1}^{n} \lambda_i C_i, C_2, \dots, C_n) = \lambda_1 \det(A).$$

Proposition 12.27. *If $A \in M_n(K)$ then* $\det(A) = \det({}^tA)$.

Proof. Let $A = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq n$. Then

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$$

and

$$\det({}^tA) = \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho) a_{1\rho(1)} \dots a_{n\rho(n)}.$$

In fact, if ${}^tA = (m_{ij})$, then $m_{ij} = a_{ji}$. We can rewrite the formula for $\det(A)$ as follows:

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}.$$

We set $\sigma(i) = t$ so that $i = \sigma^{-1}(t)$. Since σ is a bijection between $\{1, \dots, n\}$ and itself, while t varies between 1 and n also i varies between 1 and n, and so we can write:

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{t=1}^{n} a_{t\sigma^{-1}(t)}.$$

Let us now observe that for each permutation $\sigma \in \mathcal{S}_n$ we have $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$.

In fact if $\sigma = \tau_1 \circ \dots \circ \tau_r$. Then, since $(\tau_i)^{-1} = \tau_i$ we have $\sigma^{-1} = \tau_r \circ \dots \circ \tau_1$. So $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$ and

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma^{-1}) \prod_{t=1}^{n} a_{t\sigma^{-1}(t)}.$$

Since the map $\mathcal{S}_n \to \mathcal{S}_n : \sigma \to \sigma^{-1}$ is a bijection making the sum over σ or over σ^{-1} is the same thing, then:

$$\det(A) = \sum_{\sigma^{-1} \in \mathcal{S}_n} \varepsilon(\sigma^{-1}) \prod_{t=1}^{n} a_{t\sigma^{-1}(t)}.$$

Now, setting $\rho = \sigma^{-1}$ we get:

$$\det(A) = \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho) \prod_{t=1}^{n} a_{t\rho(t)}$$

and this shows that $\det(A) = \det({}^tA)$. □

Remark 12.28. The determinant $\det({}^tA)$ is a multilinear function of the columns of ${}^tA$, therefore of the rows of A, then this function is equal to $\det(A)$. So the properties $(1), \dots, (4)$ above are valid with the word row instead of the word column. That is, $\det(A)$ is an alternating multilinear function of the columns and rows of A.

If $n = 2$ and

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then as we well know

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

For $n = 3$ and $A = (a_{ij})$ we have

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

Of course (for $n = 3$ or worse $n > 3$) these are not formulas that a mathematician wants to remember.

Actually, as we will see now, there is a way to reduce the calculation of an $n \times n$ determinant to the calculation of multiple 2×2 determinants.

Lemma 12.29. *Let $M \in M_n(K)$ and write*

$$M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Then $\det(M) = a_{11} \det N$ where

$$N = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Proof. We have

$$\det(M) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots . a_{\sigma(n)n}.$$

If $\sigma(1) \neq 1$, $a_{\sigma(1)1} = 0$. So we need to consider only the permutations that verify $\sigma(1) = 1$. Let $\mathcal{U} = \{\sigma \in \mathcal{S}_n \mid \sigma(1) = 1\}$. We have

$$\det(M) = a_{11} \sum_{\sigma \in \mathcal{U}} \varepsilon(\sigma) a_{\sigma(2)2} \ldots a_{\sigma(n)n}.$$

Let $X = \{2, \ldots, n\}$. Then $\mathcal{U} \simeq \mathcal{S}(X) \simeq \mathcal{S}_{n-1}$, here $\mathcal{S}(X)$ indicates the symmetric group of X i.e. the group of bijective maps from X in X, clearly isomorphic to $\mathcal{S}_{n-1}$. The isomorphism is given by $\varphi : \mathcal{U} \to \mathcal{S}(X) : \sigma \to \sigma'$, where $\sigma'(i) = \sigma(i)$, $\forall i > 1$. We have $\varphi^{-1}(\sigma') = \sigma$ where $\sigma(1) = 1$ and $\sigma(i) = \sigma'(i)$ if $i > 1$. We observe that $\varepsilon(\sigma) = \varepsilon(\sigma')$. In fact if $\sigma' = \tau'_1 \circ \cdots \circ \tau'_r$, then $\sigma = \tau_1 \circ \cdots \circ \tau_r$, where $\tau_i = \varphi^{-1}(\tau'_i)$. So we have

$$\det(M) = a_{11} \sum_{\sigma' \in \mathcal{S}(X)} \varepsilon(\sigma') a_{\sigma'(2)2} \cdots a_{\sigma'(n)n}$$

and this last expression is clearly equal to $a_{11}\det(N)$. □

Corollary 12.30. *Let $M \in M_n(K)$, $M = (a_{ij})$, be a triangular matrix (the elements above or below the diagonal are zero). Then $\det(M) = a_{11} \cdot \cdots \cdot a_{nn}$ (the product of the coefficients on the diagonal).*

Proof. By induction on n using Lemma 12.29. □

Before stating the main theorem we need a definition.

Definition 12.31. Let $M \in M_n(K)$, $M = (a_{ij})$. We denote by $M_{k,t}$ the matrix of type $(n-1, n-1)$ obtained from M by removing the k-th row and the t-th column. The determinant of $M_{k,t}$ is called the *minor* relative (associated) to a_{kt}.

We are ready for the main theorem.

Theorem 12.32 (Laplace expansion)**.** *Let $M = (a_{ij}) \in M_n(K)$.*

(1) *We have $\det(M) = \sum_{r=1}^{n}(-1)^{r+j}a_{rj}\det(M_{rj})$ (Laplace expansion according to the j-th column).*

(2) *We have $\det(M) = \sum_{r=1}^{n}(-1)^{r+i}a_{ir}\det(M_{ir})$ (Laplace expansion according to the i-th line).*

Proof. Since $\det(M) = \det({}^tM)$ it is enough to prove (1). The idea is to observe that the j-th column can be written as the sum of n columns:

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{nj} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ a_{ij} \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ a_{nj} \end{pmatrix}.$$

We can rewrite this equation as follows: $C_j = c_1^j + \cdots + c_n^j$, where C_j indicates the j-th column of M.

Looking at the columns we have: $\det(M) = \det(C_1, \ldots, c_1^j + \cdots + c_n^j, \ldots, C_n)$. By multilinearity (fixing all variables except the j-th) we get

$$\det(M) = \sum_{i=1}^{n} \det(C_1, \ldots, c_i^j, \ldots, C_n).$$

So we are led back to calculating n determinants of the following type:

$$D_i := \det(C_1, \ldots, c_i^j, \ldots, C_n) = \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}.$$

Let $R_1, \ldots, R_n$ be the rows of D_i. Exchanging R_i with $R_{i-1}, R_{i-2}, \ldots, R_1$, we obtain after $i-1$ exchanges, the determinant:

$$D_i' = \begin{vmatrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}.$$

Let $S_1, \ldots, S_n$ be the columns of D_i'. Exchanging S_j with $S_{j-1}, \ldots, S_1$ we get, after $j-1$ exchanges, the following determinant:

$$D_i'' = \begin{vmatrix} a_{ij} & a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{in} \\ 0 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

Since we made $i-1+j-1 = i+j-2$ exchanges we have $(-1)^{i+j} D_i'' = D_i$ (the parity of $i+j-2$ is equal to that of $i+j$). Now, by Lemma 12.29, we have $D_i'' = a_{ij}.\Delta$, where Δ is the determinant obtained by deleting the first row and the first column of D_i''. But removing the first row and first column of D_i'' is like removing the i-th row and j-th column of D_i. The i-th row of D_i is the i-th row of M, the j-th column of M and D_i are the only differences between these two determinants, but they are removed. Therefore, $\Delta = \det(M_{ij})$, where, as already said, M_{ij} is the matrix obtained from M by removing the i-th row and the j-th column. In conclusion, we have $\det(M) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij})$. □

This result reduces the computation of an $n \times n$ determinant to the computation of n determinants of order $n-1$. In particular, the calculation of a determinant of any order n can be traced back to the calculation of many determinants of order two. Note that the determinants of order two are also calculated with the Laplace expansion.

Example 12.33. Developing

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

according to the first column we get

$$D = 1\begin{vmatrix} 3 & -1 \\ 3 & -2 \end{vmatrix} - 2\begin{vmatrix} -2 & 1 \\ 3 & -2 \end{vmatrix} - 1\begin{vmatrix} -2 & 1 \\ 3 & -1 \end{vmatrix} = (-6+3) - 2.(4-3) - (2-3) = -4.$$

There is a faster way to proceed. We have seen that the determinant does not change if a linear combination of the rows (resp. columns) is added to a row (resp. column). The idea is to fit as many zeros as possible into a row (resp. columns). Here for example we can replace R_1 with $R_1 + R_3$, an operation which we indicate as follows: $R_1 \to R_1 + R_3$. We obtain

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

Now, we do $C_2 \to C_2 + C_3$ and we get

$$D = \begin{vmatrix} 0 & 0 & -1 \\ 2 & 2 & -1 \\ -1 & 1 & -2 \end{vmatrix}.$$

Developing according to the first line we have

$$D = -\begin{vmatrix} 2 & 2 \\ -1 & 1 \end{vmatrix} = -4.$$

Remark 12.34. Pay attention to the signs. The rule is very simple: the + sign goes in place $(1, 1)$ (first row, first column), then every time you move by one, either according to a row or a column, the sign changes. So in the case 3×3 the signs are:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}.$$

As a first consequence of Laplace's expansion we have a generalization of Lemma 12.29.

Proposition 12.35. *Let $M \in M_{n+p}(K)$ be a square matrix of the form*

$$M = \begin{pmatrix} A & P \\ 0 & B \end{pmatrix}$$

where $A \in M_n(K)$, $B \in M_p(K)$ $P \in M_{n,p}(K)$ and 0 indicates a null matrix of order (p, n). Then $\det(M) = \det(A)\det(B)$.

Proof. We proceed by induction on n. The case $n = 1$ follows from Lemma 12.29. Suppose the claim holds for $n-1$. We develop according to the first column. We have: $\det(M) = \sum_{i=1}^{n}(-1)^{i+1}a_{i1}\det M_{i,1}$. Here, as usual, $M_{i,1}$ is the matrix obtained from M by removing the row i and the first column. So

$$M_{i,1} = \begin{pmatrix} A_{i,1} & P(i) \\ 0 & B \end{pmatrix}$$

where $P(i)$ is the matrix obtained from P by removing the i-th row: it is a $(n-1, p)$ matrix; 0 indicates a null matrix of type $(p, n-1)$ (we removed the first column of the initial null matrix). The matrix $A_{i,1}$ is square of order $n-1$. By induction hypothesis $\det(M_{i,1}) = \det(A_{i,1})\det(B)$. So

$$\det M = \sum_{i=1}^{n}(-1)^{i+1}a_{i1}\det(A_{i,1})\det(B) = \det(B)\sum_{i=1}^{n}(-1)^{i+1}a_{i1}\det(A_{i,1})$$

which indeed is equal to $\det(B)\det(A)$. □

Corollary 12.36. *Let M be a block matrix of the following form:*

$$M = \begin{pmatrix} A_1 & \cdots & \cdots & \cdots \\ 0 & A_2 & \cdots & \cdots \\ 0 & 0 & \ddots & \cdots \\ 0 & 0 & 0 & A_k \end{pmatrix}$$

where $A_1, \ldots, A_k$ are square matrices and the zeros indicate null matrices. Then $\det(M) = \det(A_1)\ldots\det(A_k)$.

Proof. Proceed by induction on k. If $k = 2$ is the previous proposition. Again by the previous proposition we have

$$\det(M) = \det(A_1)\begin{vmatrix} A_2 & \cdots & \cdots \\ 0 & \ddots & \vdots \\ 0 & 0 & A_k \end{vmatrix}.$$

By induction hypothesis the last determinant is $\det(A_2)\ldots\det(A_k)$ and the result follows. □

12.6. Inverse matrix, complementary matrix

Let $A = (a_{ij}) \in M_n(K)$ be a square matrix.

Definition 12.37. The *cofactor* of a_{ij} is $(-1)^{i+j}\det A_{ij}$ where, as usual, A_{ij} is the matrix obtained from A by removing the i-th row and the j-th column.

The *cofactor matrix* of A is $Co(A) = (c_{ij})$ where $c_{ij} = (-1)^{i+j}\det A_{ij}$ is the cofactor of a_{ij}.

The *complementary matrix* A^c of A is the transpose of the cofactor matrix:

$$A^c = {}^tCo(A).$$

The interest of these definitions is revealed by the following result.

Theorem 12.38. *Let $A \in M_n(K)$, then*

(1) $AA^c = A^cA = \det(A)I_n$;

(2) *if A is invertible then* $A^{-1} = \dfrac{1}{\det(A)} \cdot A^c$.

Proof. (1) Let $AA^c = (d_{ij})$. By definition d_{ij} is the product of the i-th row of A with j-th column of A^c. Since $A^c = {}^tCo(A)$, d_{ij} is equal to the scalar product of the i-th row of A with the j-th row of $Co(A)$. That is

$$d_{ij} = (R_i(A) \mid R_j(Co(A)) = \sum_{k=1}^{n} a_{ik}(-1)^{j+k} \det A_{jk}.$$

If $i = j$ we get

$$d_{ii} = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik}.$$

From Theorem 12.32 we have $d_{ii} = \det(A)$. Now, consider the case $i \neq j$. Let $R_1, \ldots, R_n$ be the rows of A and $A_{i(j)}$ the matrix whose rows, R'_i, are given by $R'_t = R_t$ if $t \neq j$, $R'_j = R_i$. Since $A_{i(j)}$ has two equal rows ($R'_i = R'_j$), we have $\det(A_{i(j)}) = 0$. If we develop $\det A_{i(j)}$ according to the j-th line, we get $0 = \det(A_{i(j)}) = \sum_k a_{ik}(-1)^{j+k} \det(A_{jk}) = d_{ij}$. So if $i \neq j$, $d_{ij} = 0$ and $AA^c = \det(A)I_n$.

In the same way we prove $A^cA = \det(A)I_n$.

(2) If A is invertible then $\det(A) \neq 0$ (Corollary 12.25) and therefore

$$\left(\frac{1}{\det(A)} A^c\right) A = I_n$$

concluding the proof. □

Remark 12.39. Theorem 12.38 provides an effective way to compute A^{-1} when A is invertible (i.e. when $\det(A) \neq 0$).

Remark 12.40. In the proof of Theorem 12.38 we never used that K is a field. That is, we have not divided by any element. The proof of Theorem 12.38 is valid even if $K = R$ is only a commutative ring. The determinant of a square matrix is defined by (12.23). Under these conditions Theorem 12.38 says that A is invertible if and only if $\det(A)$ is invertible in R. For example if $A \in M_n(\mathbb{Z})$, $\det(A)$ is invertible in $\mathbb{Z}$ if and only if $\det(A) = \pm 1$.

Remark 12.41. Consider the matrix

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

We have $\det(A) = -2$. We can consider A as a matrix with rational coefficients. We have $-2 \neq 0$, so A is invertible in $M_2(\mathbb{Q})$ ($\mathbb{Q}$ is a field). We can also consider A as a matrix with integer coefficients ($A \in M_3(\mathbb{Z})$). But -2 is not invertible in the ring $\mathbb{Z}$ and therefore A is not invertible in $M_2(\mathbb{Z})$. That is, it is not possible to find a matrix B with integer coefficients such that $AB = I_3$. It is instead possible to find a matrix $A^{-1} \in M_3(\mathbb{Q})$ such that $A.A^{-1} = I_3$. We conclude that the matrix $A^{-1} \in M_3(\mathbb{Q})$ has at least one rational, non-integer coefficient.

In this case it is not even too difficult to compute A^{-1}. We associate to A an endomorphism of K^3 via $E = K^3$, $\mathcal{B} = \mathcal{C}$ (the canonical basis). So let $\mathcal{B} = (e_i)$ and $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = A$. We have $f(e_1) = e_1$, $f(e_2) = 2e_1 + 2e_2$, $f(e_3) = -e_2 - e_3$. Applying f^{-1} to the two sides of the previous equations we obtain $e_1 = f^{-1}(e_1)$, $e_2 = 2f^{-1}(e_1) + 2f^{-1}(e_2) = 2e_1 + 2f^{-1}(e_2)$, $e_3 = -f^{-1}(e_2) - f^{-1}(e_3)$. In conclusion $f^{-1}(e_1) = e_1$, $f^{-1}(e_2) = (e_2 - 2e_1)/2 = -e_1 + e_2/2$, $f^{-1}(e_3) = e_1 - e_2/2 - e_3$. Finally

$$A^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix A^{-1} has two rational, non-integer coefficients. Here we got lucky because the system to solve to calculate $f^{-1}(e_i)$ is particularly easy, this approach is not recommended in general.

12.7. Rank and determinants

So far we have seen that the determinant allows us to say whether an $n \times n$ square matrix has rank n or not. If the determinant is zero, the rank is less than n but we do not know what it is exactly. In this section we show that using determinants it is possible to calculate the rank of any matrix (even a non-square one).

Definition 12.42. Let $M = (a_{ij}) \in M_{n,p}(K)$ and

$$I \subset \{1, 2, \ldots, n\},\ J \subset \{1, 2, \ldots, p\}.$$

With $M_{I,J}$ we indicate the matrix obtained from M by deleting the rows whose index belongs to I and the columns whose index belongs to J. We say that $M_{I,J}$ is a matrix extracted from M (or that $M_{I,J}$ is a submatrix of M).

We have already seen a special case of this notion in Laplace's expansion: the matrices $M_{ij} = M_{I,J}$ with $I = \{i\}, J = \{j\}$.

Definition 12.43. A *minor* of order r of M is the determinant of an $r \times r$ matrix extracted from M. We observe that necessarily $r \leq \min\{n, p\}$.

Example 12.44. The matrix

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 3 & 1 & 5 \end{pmatrix}$$

has three minors of order two.

Lemma 12.45. *Let $f : E \to F$ be a linear map between two K-vector spaces, and $v_1, \ldots, v_r \in E$ vectors such that $f(v_1), \ldots, f(v_r)$ are independent in F. Then $v_1, \ldots, v_r$ are independent.*

Proof. If $\sum \lambda_i v_i = 0$, applying f we have $\sum \lambda_i f(v_i) = 0$ but the vectors $f(v_i)$ are independent, so $\lambda_i = 0$, $\forall i$. □

Proposition 12.46. *Let $M \in M_{n,p}(K)$. We have* $\mathrm{rk}(M) \geq r$ *if and only if there exists a non-zero minor of M of order r.*

Proof. Suppose that there exists a minor of M, of order r, Δ_r, not zero. The rank of M is the rank of its column vectors (or its row vectors), so a permutation of the columns or rows of M does not change its rank. We can therefore assume that the smallest Δ_r is constructed on the first r columns and the first r rows. Let $C_1, \ldots, C_p$ be the column vectors of M; they are vectors of K^n. Consider

$$\pi : K^n \to K^r : (x_1, \ldots, x_n) \to (x_1, \ldots, x_r)$$

the projection onto the first r coordinates. We set $c_i := \pi(C_i)$. Then $\Delta_r = \det_{\mathcal{C}}(c_1, \ldots, c_r)$, where $\mathcal{C}$ denotes the canonical basis of K^r. Since $\Delta_r \neq 0$, this implies that $c_1, \ldots, c_r$ are independent. By Lemma 12.45 this implies that $C_1, \ldots, C_r$ are independent as well, so $\mathrm{rk}(M) \geq r$.

Now, suppose $\mathrm{rk}(M) \geq r$. There are therefore r independent column vectors of M. By rearranging the columns we can assume that $C_1, \ldots, C_r$ are independent. The $n \times r$ matrix, N, whose columns are $C_1, \ldots, C_r$ has therefore exactly rank r. But the rank of N is also the rank of its rows. There are therefore r independent rows of N. By rearranging the rows we can assume that they are the first r. Note that the rows of N are vectors of K^r (while the columns are vectors of K^n). If $r_1, \ldots, r_r$ indicate the first rows of N, we have $\det_{\mathcal{C}}(r_1, \ldots, r_r) \neq 0$ (r independent vectors of K^r). This shows that the minor of M constructed over the first r rows and r columns is non-zero. □

Corollary 12.47. *Let $M \in M_{n,p}(K)$ then M has rank r if and only if there exists a non-zero minor of M of order r and all minors of order $r+1$ are zero.*

Proof. If M has rank r by Proposition 12.46 there exists a non-zero minor of order r, and again for the same proposition all minors of order $r+1$ are zero. The converse is clear. □

So for example if we have a matrix $M \in M_{3,4}(K)$ with a non-zero minor of order two, to see if M has rank 2 or 3 we have to look at the four minors of order 3. If they are all zero the rank is 2 if one of them is non-zero the rank is 3. We show that actually it is not necessary to check all the minors of order 3 but only some of them.

Definition 12.48. Let $M \in M_{n,p}(K)$ and let Δ_r be a minor of M of order r. An *edge* of Δ_r is a minor of order $r+1$ of which Δ_r is a minor.

Example 12.49. Let

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

and

$$\Delta_2 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}.$$

An edge of Δ_2 is a minor of order 3 that contains Δ_2, there are two of them:

$$\begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ -1 & 0 & 2 \end{vmatrix} \text{ and } \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & 0 & 3 \end{vmatrix}.$$

The main result of this section:

Theorem 12.50 (Edge method)**.** *Let $M \in M_{n,p}(K)$. Then $\mathrm{rk}(M) = r$ if and only if there exists a non-zero minor of order r, Δ_r, whose edges are all zero.*

Proof. If $\mathrm{rk}(M) = r$ by Corollary 12.47 there exists a non-zero minor of order r and all minors of order $r+1$ are zero and we are fine.

For the other implication we show the opposite: if $\Delta_r \neq 0$ and $\mathrm{rk}(M) > r$ then there exists a non-zero edge of Δ_r.

We can assume that Δ_r is the minor constructed over the first r columns and the first r rows. So $C_1, \ldots, C_r$, the first r columns of M, are linearly independent. If $\mathrm{rk}(M) \geq r+1$, the space generated by the columns $F = \langle C_1, \ldots, C_p \rangle$ has dimension greater than r. Since $C_1, \ldots, C_r$ are independent, by the incomplete basis theorem we can complete these vectors to a basis of F. In particular, there exists C_j, with $j > r$, such that $C_1, \ldots, C_r, C_j$ are independent. Let N be the matrix of type $(n, r+1)$ whose columns are $C_1, \ldots, C_r, C_j$. Clearly, N has rank $r+1$. So the row space of N has dimension $r+1$. The rows of N are elements of K^{r+1}. Using

Lemma 12.45 (with the projection $\pi : K^{r+1} \to K^r : (x_1, \ldots, x_{r+1}) \to (x_1, \ldots, x_r)$), we see that the first r rows, $r_1, \ldots, r_r$, of N are independent. So there exists r_t, with $t > r$, such that $r_1, \ldots, r_r, r_t$ are independent. The determinant of the matrix $(r_1, \ldots, r_r, r_t)$ is a non-zero edge of Δ_r. □

Example 12.51. Let

$$A = \begin{pmatrix} 1 & 2 & 4 & 5 \\ 0 & 1 & 1 & 3 \\ 1 & 0 & 2 & -1 \end{pmatrix}.$$

We have

$$\Delta_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

So $\mathrm{rk}(A) \geq 2$. Following Corollary 12.47 we need to compute four minors of order 3. The edge method reduces the work by half. Just compute the two edges of Δ_2. Do this to verify that $\mathrm{rk}(A) = 2$.

Exercises

Exercise 12.52. In the $\mathbb{C}$-vector space $\mathbb{C}^2$, calculating a determinant, say whether the vectors u, v are linearly independent where $u = (2+i, 1-i)$, $v = (4+i, 2+3i)$.

Exercise 12.53. In the K-vector space K^2 for what values of $m \in K$ are the two vectors $u = (m, 1)$, $v = (-1, m)$ independent? $(K = \mathbb{R}, \mathbb{C})$.

Exercise 12.54. Let $M \in M_n(K)$. Show that if $M = AB$ with $A \in M_{n,p}(K)$, $B \in M_{p,n}(K)$ and $n > p$, then $\det(M) = 0$. Is it still true that $\det(M) = 0$ if we assume instead $n < p$?

Exercise 12.55. The chessboard matrix $S \in M_n(K)$ is defined as follows $S = (a_{ij}$, with $a_{ij} = 0$ if $i + j$ is even, $a_{ij} = 1$ if $i + j$ is odd. Compute $\det(S)$.

Exercise 12.56. Let $n \geq 3$ be a natural odd number. Prove that there is no matrix $M \in M_n(\mathbb{R})$ such that $M^2 + I_n = 0$. What happens if n is even or if instead of $K = \mathbb{R}$ we take $K = \mathbb{C}$?

Exercise 12.57. Compute

$$D_n = \begin{vmatrix} 1 & 1 & \cdots & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 1 & \dots & \dots & 1 & 0 \end{vmatrix}$$

where $d_{ij} = 1$ if $i \neq j$, $d_{11} = 1$, $d_{ii} = 0$ if $i > 1$.

Exercise 12.58. Let $a \in \mathbb{R}$. For every $x \in \mathbb{R}$ consider the matrix

$$A_n(x) := \begin{pmatrix} x & a & \cdots & \cdots & a \\ a & x & a & \cdots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a & \cdots & \cdots & a & x \end{pmatrix} \in M_n(\mathbb{R})$$

with $a_{ij} = a$ if $i \neq j$, $a_{ii} = x$. Prove that:

$$\det A_n(x) = (x + (n-1)a)(x-a)^{n-1} \ , \forall x \in \mathbb{R}$$

(Hint: start by adding all columns first).

Exercise 12.59. Factor the polynomial

$$Q(x) = \begin{vmatrix} x & a & b & x \\ a & x & x & b \\ b & x & x & a \\ x & b & a & x \end{vmatrix}.$$

Exercise 12.60. Let $A \in M_3(\mathbb{R})$, $A = (a_{ij})$, with $a_{ij} \in \{\pm 1\}, \forall i, j$.

(1) Show that $\det A \leq 6$.

(2) Note that $\det A$ is an even integer.

(3) Show that $\det A \leq 4$. Give an example of a matrix A (with $a_{ij} \in \{\pm 1\}, \forall i, j$) with $\det A = 4$.

Exercise 12.61. Let $f_1, \dots f_n \in \mathbb{R}[X]$ be polynomials whose degree is at most $n-2$. Let $a_1, \dots a_n$ be any real numbers. Show that the determinant of the matrix:

$$M = \begin{pmatrix} f_1(a_1) & f_1(a_2) & \dots & f_1(a_n) \\ f_2(a_1) & f_2(a_2) & \dots & f_2(a_n) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(a_1) & f_n(a_2) & \dots & f_n(a_n) \end{pmatrix}$$

is zero.

Exercise 12.62. Let $E := M_2(\mathbb{R})$ be the $\mathbb{R}$-vector space of 2×2 matrices with real coefficients

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in E.$$

and $f : E \to E : M \to AM$.

(i) Show that f is linear.

(ii) Compute the determinant of the endomorphism f.

Exercise 12.63. (1) In $\mathbb{R}^3$ consider the vectors $v_1 = (1, 0, 0)$, $v_2 = (0, -1, 0)$, $v_3 = (1, 0, -1)$. By calculating a determinant, show that $\mathcal{B} = (v_1, v_2, v_3)$ is a basis of $\mathbb{R}^3$.

(2) Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be such that

$$\mathrm{Mat}(f; \mathcal{C}, \mathcal{C}) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

where $\mathcal{C}$ indicates the canonical basis. Show that M is invertible and calculate M^{-1}. Conclude that f is an automorphism.

(3) Determine $\mathrm{Mat}(f^{-1}; \mathcal{B}, \mathcal{B})$.

Exercise 12.64. Let $M \in M_n(K)$, $M = (m_{ij})$ where

$$m_{ij} = \begin{cases} 1 \text{ if } j = i-1, j < n; \\ 1 \text{ if } j = n, i = 1; \\ 0 \text{ otherwise.} \end{cases}$$

Prove that M is invertible and compute M^{-1} (Hint: see M as the matrix of an endomorphism).

Exercise 12.65. Let $M \in M_n(\mathbb{R})$ be of the following form

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

where A is an $r \times r$ square matrix.
(i) Show, using the edge method, that if A is invertible then

$$\mathrm{rk}(M) = \mathrm{rk}(A) + \mathrm{rk}(C).$$

(ii) If A is not invertible, is it still true that $\mathrm{rk}(M) = \mathrm{rk}(A) + \mathrm{rk}(C)$?
(iii) Show (without using any determinant) that we always have

$$\mathrm{rk}(M) \geq \mathrm{rk}(A) + \mathrm{rk}(C)$$

by considering the columns of M.

Exercise 12.66. Determine the rank of the following matrices:

$$A = \begin{pmatrix} 1 & 7 & 5 & 3 & 2 \\ 0 & 4 & 2 & 2 & 0 \\ 2 & -2 & 4 & 0 & 1 \\ 3 & -1 & 7 & 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -3 & 4 \\ 3 & 1 & 5 \\ -1 & 0 & -1 \\ 0 & 2 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 1 & 0 & -3 & 0 \\ -3 & 0 & 1 & 6 & 1 \\ 2 & 0 & -1 & -4 & 0 \\ -4 & -1 & 0 & 5 & 1 \end{pmatrix}.$$

Exercise 12.67. Let $n \geq 3$ be an odd integer. Show that there is no $M \in M_n(\mathbb{R})$ such that $M^2 + I_n = 0$.

Exercise 12.68. Let A be an $n \times n$ matrix, with integer coefficients ($A \in M_n(\mathbb{Z})$) such that $\det(A) = 2$. Prove that A has at least $n - 1$ odd coefficients.

Exercise 12.69. Let C^n be the $n \times n$ matrix whose diagonal entries are 2, the entries on the superdiagonal and on the subdiagonal are 1, and the remaining entries are 0:

$$C^n = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}.$$

(i) Prove that $\det(C^n) = n + 1$.
(ii) Let $(C^n)^{-1}$ be the inverse of C^n, and $d_{i,i}$ the i-th diagonal entry of $(C^n)^{-1}$. Prove that

$$d_{i,i} = \frac{i(n-i+1)}{n+1}$$

for $i = 1, \ldots, n$. In particular,

$$d_{i,i} = d_{n-i+1,n-i+1}$$

for $i = 1, \ldots, \lfloor \frac{n+1}{2} \rfloor$.
(iii) Show that if $n \geq 8$ then $d_{i,i} \geq 2$ for $i = 3, \ldots, n-2$.

Chapter 13

Diagonalization

We have seen in Proposition 11.25 that, if $f : E \to F$ is a linear map between two K-vector spaces, it is possible to find bases $\mathcal{B}, \mathcal{C}$ of E, F such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{C})$ is particularly simple. Now, we face the same problem for endomorphisms of a space E. More precisely, if $f \in \mathrm{End}(E)$ we ask if there exists a $\mathcal{B}$ of E such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ is simple. Note that we take the same basis in the domain and in the codomain. In other words we are studying similarity of matrices.

What does simple matrix mean? The simplest we can hope for is a diagonal one. In fact, the simplest linear maps are homotheties, i.e. maps of the form $g : E \to E : v \to \alpha v$. If $\mathcal{B}$ is any basis of E, $\mathrm{Mat}(g; \mathcal{B}, \mathcal{B}) = \alpha I_n$. Finding a basis $\mathcal{B}$ such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ is a diagonal matrix means expressing f as a sum of homotheties (relative to certain particular subspaces). As we will see this is not always possible.

The diagonalization problem for square matrices is posed in these terms: Let $M \in M_n(K)$, does there exist a diagonal matrix, D, such that $M \approx D$ (i.e. such that M is similar to D)? Recall that $M \approx D \Leftrightarrow \exists P$ invertible, such that $PDP^{-1} = M$.

This problem relates to the previous one in the following way. Let $\mathcal{B}$ be a basis of E and $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$. If $\mathcal{C}$ is a basis such that $\mathrm{Mat}(f; \mathcal{C}, \mathcal{C}) = D$ is diagonal, then

$$E_{\mathcal{C}} \overset{P}{\to} E_{\mathcal{B}} \overset{M}{\to} E_{\mathcal{B}} \overset{P^{-1}}{\to} E_{\mathcal{C}}$$

and we conclude that $D = P^{-1}MP$. Therefore $M \approx D$.

Conversely, let $D = P^{-1}MP$, where D is diagonal, and E a K-vector space with $\dim(E) = n$. Let $\mathcal{B}$ be a basis of E and $f \in \mathrm{End}(E)$ such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = M$.

The usual diagram

$$E_{\mathcal{C}} \overset{P}{\to} E_{\mathcal{B}} \overset{M}{\to} E_{\mathcal{B}} \overset{P^{-1}}{\to} E_{\mathcal{C}}$$

tells us that $D = \mathrm{Mat}(f; \mathcal{C}, \mathcal{C})$.

We conclude that the problem of diagonalization of endomorphisms is equivalent to the problem of diagonalization of matrices.

Suppose that f is diagonalizable, that is, suppose

$$D = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}.$$

If $\mathcal{B} = (e_1, \ldots, e_n)$ we have $f(e_i) = \lambda_i e_i$. Hence, the vectors e_i and $f(e_i)$ lie on the same line. In other words, the line $\langle e_i \rangle$ is *stable* under f i.e. $f(\langle e_i \rangle) = \langle e_i \rangle$ for $i = 1, \ldots, n$.

We now show that not all endomorphisms are diagonalizable. Let E be a 2-dimensional $\mathbb{R}$-vector space and $\mathcal{B} = (e_1, e_2)$ a basis of E. Let $f \in \mathrm{End}(E)$ be such that $f(e_1) = -e_2$ and $f(e_2) = e_1$. So

$$M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Let $v \in E$, $v \neq 0$, $v = ae_1 + be_2$, such that $f(v) = \lambda v$. We have $f(v) = -ae_2 + be_1$. Since $f(v) = \lambda v$, we get $f(v) = -ae_2 + be_1 = \lambda ae_1 + \lambda be_2$. So $b = \lambda a$ and $-a = \lambda b$. That is, $-a = \lambda^2 a$. We have $a \neq 0$ (otherwise $a = b = 0$ and $v = 0$). Dividing by a we get $\lambda^2 = -1$ but this is impossible in $\mathbb{R}$.

We can also think geometrically. The map f is a rotation of 90 degrees clockwise. This map has no fixed line, so there exists no non-zero vector such that $f(v) = \lambda v$.

13.1. Eigenvectors and eigenvalues

Let us begin with a definition.

Definition 13.1 (Eigenvector)**.** Let E be a K-vector space and $f \in \mathrm{End}(E)$. An *eigenvector* of f is a vector $v \in E$ such that:

(1) $v \neq 0$;

(2) $f(v) = \lambda v$ for some $\lambda \in K$.

Remark 13.2. We have that f is diagonalizable if and only if there exists a basis of E made of eigenvectors.

Definition 13.3 (Eigenvalue)**.** Let E be a K-vector space and $f \in \mathrm{End}(E)$. A scalar $\lambda \in K$ is an *eigenvalue* of f if there exists $v \in E$, $v \neq 0$, such that $f(v) = \lambda v$. In this case v is an eigenvector, we say that v is associated to λ.

Each eigenvector is associated to a single eigenvalue. However, multiple eigenvectors can be associated with an eigenvalue. For example, if $f \in \mathrm{End}(E)$ is not injective, 0 is an eigenvalue and every non-zero vector of $\mathrm{Ker}(f)$ is an eigenvector associated to the eigenvalue 0.

Lemma 13.4. *Let $f \in \mathrm{End}(E)$. Then λ is an eigenvalue of f if and only if $\lambda Id_E - f$ is not injective.*

If λ is an eigenvalue of f, the set of eigenvectors of f associated to the eigenvalue λ is $\mathrm{Ker}(\lambda Id_E - f) \setminus \{0\}$.

Proof. Follows from the definitions. □

Definition 13.5 (Eigenspace)**.** Let f be an endomorphism of the K-vector space E. For every $\lambda \in K$ we set $E_f(\lambda) = \mathrm{Ker}(\lambda Id_E - f)$. If λ is an eigenvalue of f we say that

$$E_f(\lambda) = \mathrm{Ker}(\lambda Id_E - f) = \{v \in E \mid f(v) = \lambda v\}$$

is the *eigenspace* associated to λ.

Remark 13.6. (1) λ is eigenvalue of $f \Leftrightarrow E_f(\lambda) \neq \{0\}$.

(2) Let $E_f(\lambda)$ be an eigenspace of f and $v \in E_f(\lambda)$. We have $f(v) = \lambda v \in E_f(\lambda)$ (since $E_f(\lambda)$ is a subvector space being the kernel of a linear map). So $f(E_f(\lambda)) \subset E_f(\lambda)$, i.e. $E_f(\lambda)$ is stable under f, in particular the restriction of f to $E_f(\lambda)$ is an endomorphism of $E_f(\lambda)$. This endomorphism is nothing but the homothety $E_f(\lambda) \to E_f(\lambda) : v \to \lambda v$.

Lemma 13.7. *Let E be a K-vector space and $f \in \mathrm{End}(E)$. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of f ($i \neq j \Rightarrow \lambda_i \neq \lambda_j$) and $v_1, \ldots, v_m$ the associated eigenvectors ($f(v_i) = \lambda_i v_i, \forall i$). Then the vectors $v_1, \ldots, v_m$ are linearly independent.*

Proof. We proceed by induction on m. The case $m = 1$ is clear ($v_1 \neq 0$). Let us assume the lemma proved for $m - 1$. Let

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0 \ (\star).$$

We have

$$f(\alpha_1 v_1 + \cdots + \alpha_m v_m) = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_m \lambda_m v_m = 0 \ (\star\star).$$

We multiply $(\star)$ by λ_1 and subtract the result from $(\star\star)$. Then

$$(\lambda_2 - \lambda_1)\alpha_2 v_2 + \cdots + (\lambda_m - \lambda_1)\alpha_m v_m = 0.$$

By induction hypothesis the $m-1$ vectors $v_2, \ldots, v_m$ are independent. So $(\lambda_i - \lambda_1)\alpha_i = 0, \forall i \geq 2$. Since $\lambda_i - \lambda_1 \neq 0$, this implies $\alpha_i = 0, \forall i \geq 2$. So $(\star)$ becomes $\alpha_1 v_1 = 0$, and since $v_1 \neq 0$ we get $\alpha_1 = 0$. In conclusion, all α_i are zero and $v_1, \ldots, v_m$ are independent. □

Corollary 13.8. *Let $f \in \mathrm{End}(E)$, where E is a K-vector space of dimension n. If f has n distinct eigenvalues then f is diagonalizable.*

Proof. Follows from Lemma 13.7. □

Corollary 13.9. *Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of $f \in \mathrm{End}(E)$. Then the corresponding eigenspaces $E_f(\lambda_1), \ldots, E_f(\lambda_m)$ are in direct sum.*

Proof. We need to show that $E_f(\lambda_i) \cap \sum_{j \neq i} E_f(\lambda_j) = \{0\}$ for every i. If $v_i = \sum_{j \neq i} v_j$, then $v_1, \ldots, v_m$ are dependent, but this is not possible because the v_i are eigenvectors corresponding to distinct eigenvalues (Lemma 13.7). □

13.2. The characteristic polynomial

As we have already observed, $f \in \mathrm{End}(E)$ is diagonalizable if and only if there exists a basis made of eigenvectors. Each eigenvector belongs to an eigenspace and to find these eigenspaces it is best to first find the eigenvalues.

Proposition 13.10. *Let $f \in \mathrm{End}(E)$, where E is a K-vector space. Let $\mathcal{B}$ be a basis of E and $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$. The following are equivalent:*

(1) *λ is an eigenvalue of f;*

(2) *$\lambda I_n - M$ is not invertible;*

(3) $\det(\lambda I_n - M) = 0$.

Proof. It is an immediate consequence of the fact, already observed before, that λ is an eigenvalue if and only if $\mathrm{Ker}(\lambda Id_E - f) \neq \{0\}$. □

This proposition allows us to determine the eigenvalues of f. The idea is to compute $\det(xI_n - M) = P(x)$, the eigenvalues will be those values λ_i such that $P(\lambda_i) = 0$.

If $M = (a_{ij})$, $xI_n - M = (x\delta_{ij} - a_{ij})$ using the determinant formula we have

$$P(x) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)(x\delta_{\sigma(1)1} - a_{\sigma(1)1}) \cdots (x\delta_{\sigma(i)i} - a_{\sigma(i)i}) \cdots (x\delta_{\sigma(n)n} - a_{\sigma(n)n}). \tag{13.1}$$

In particular, $P(x)$ is a polynomial in x of degree $\leq n$. If $\sigma = Id$ we have $(x - a_{11}) \ldots (x - a_{ii}) \ldots (x - a_{nn}) = x^n - (a_{11} + \cdots + a_{nn})x^{n-1} +$ terms of lower degree.

If $\sigma \neq Id$ there exist at least two indexes i, j such that $\sigma(i) \neq i$ and $\sigma(j) \neq j$, then $\delta_{\sigma(i)i} = \delta_{\sigma(j)j} = 0$ and $\varepsilon(\sigma)(x\delta_{\sigma(1)1} - a_{\sigma(1)1}) \cdots (x\delta_{\sigma(n)n} - a_{\sigma(n)n})$ has degree at most $n-2$.

Finally, the constant term of the polynomial $P(x)$ is $P(0) = \det(-M) = (-1)^n \det(M)$. We proved the following:

Proposition 13.11. *With the previous notations $P(x)$ is a monic polynomial of degree n in x, more precisely:*

$$P(x) = x^n - tr(M)x^{n-1} + \cdots + (-1)^n \det(M)$$

where $tr(M) := a_{11} + \cdots + a_{nn}$ ($tr(M)$ is the trace *of the matrix M). The eigenvalues of f (and hence of M) are the roots of the polynomial $P(x)$.*

This calculation seems to depend on M. What happens if instead of considering M we consider $N = \mathrm{Mat}(f; \mathcal{C}, \mathcal{C})$ where $\mathcal{C} \neq \mathcal{B}$ is another basis of E?

We find the same polynomial. In fact, the two matrices M and N are similar so there exists an invertible P such that $M = P^{-1}NP$. We have $\det(xI_n - M) = \det(P^{-1}xI_nP - P^{-1}NP) = \det(P^{-1}(xI_n - N)P) = \det P^{-1} \det(xI_n - N).\det P = \det(xI_n - N)$. This proves the following:

Proposition 13.12. *With the previous notations:*

(1) *The polynomial $P(x)$ does not depend on the chosen basis; it is therefore an invariant that we can associate to f.*

(2) *The polynomial $P(x)$ is the characteristic polynomial of f (or of M); we will denote it by $P_f(x)$ (or $P_M(x)$).*

(3) *The eigenvalues of f are the roots, in K, of $P_f(x)$.*

(4) *Two similar matrices have the same characteristic polynomial.*

13.3. Roots of a polynomial

Before moving on, let us make a reminder about the roots of a polynomial with coefficients in a field K. Let $P(x) \in K[x]$. An element $a \in K$ is a root of $P(x)$ if $P(a) = 0$. This is equivalent to saying that $(x - a) \mid P(x)$. It is clear that if $(x - a) \mid P(x)$ then $P(a) = 0$. There are various ways of looking at the other implication.

We can compute the Euclidean division of $P(x)$ by $(x - a)$ that is $P(x) = (x-a)Q(x) + r$ where r is a constant. Setting $x = a$, since $P(a) = 0$, we get $r = 0$. So $P(x) = (x - a)Q(x)$.

Or we can think like this: The result is clear if $a = 0$, 0 is the root of $P(x)$ if and only if the constant term of $P(x)$ is zero. In the general case we set $X = x - a$

(therefore $x = X + a$). We have $P(x) = P(X + a) =: R(X)$. Hence 0 is a root of

$$R(X) \Leftrightarrow P(a) = 0 \Leftrightarrow X \mid R(X) \Leftrightarrow (x - a) \mid P(x).$$

The advantage of this proof is that it works for $K = A$, A being any commutative ring.

If a is a root of $P(x)$, the *multiplicity* of a as a root of $P(x)$ is the greatest power of $x - a$ that divides $P(x)$.

Let $P(x) = x^2 + 1$. Clearly $P(x)$ has no root in $\mathbb{R}$. Instead $P(x)$ has two roots in $\mathbb{C}$ and they are $\pm i$. Whether or not the roots of a polynomial exist depends on where we look for the roots. If one looks for solutions in $\mathbb{R}$, then there are none. If one searches for solutions in $\mathbb{C}$, then there are two.

Let $P(x) \in K[x]$, K a field, $P(x)$ of degree n. Then $P(x)$ has at most n roots in K (counted with multiplicity).

If a_1 is a root with multiplicity m_1, then $P(x) = (x - a_1)^{m_1} Q_1(x)$, with $Q_1(a_1) \neq 0$. If a_2 is a root with multiplicity m_2 then $P(x) = (x - a_1)^{m_1}(x - a_2)^{m_2} R(x)$, with $R(a_1)R(a_2) \neq 0$.

In fact, if a_2 is a root, then $(x - a_2) \mid P(x) = (x - a_1)^{m_1} Q(x)$. If $x = a_2$, we have $0 = (a_2 - a_1)^{m_1} Q(a_2)$. Since $a_2 \neq a_1$, $(a_2 - a_1)^{m_1} \neq 0$, because K is a field (this is not necessarily true if $K = A$ is any commutative ring). Therefore, we have $Q(a_2) = 0$ i.e. $(x - a_2) \mid Q(x)$ i.e. $Q(x) = (x - a_2)Q_1(x)$. If a_2 is a root with multiplicity $m_2 > 1$ then $(x-a_2)^{m_2} \mid P(x) = (x-a_1)^{m_1}(x-a_2)Q_1(x)$. Dividing by $x-a_2$ and repeating the previous argument we see that $(x-a_2) \mid Q_1(x)$. Continuing in this way we get $P(x) = (x-a_1)^{m_1}(x-a_2)^{m_2} R(x)$ with $R(a_1) \neq 0$ and $R(a_2) \neq 0$.

Repeating this process for each root we have $P(x) = Q_1(x) \dots Q_t(x)R(x)$, where $Q_i(x) = (x - a_i)^{m_i}$, $R(a) \neq 0 \, \forall a \in K$, and where the a_i are the roots in K of the polynomial $P(x)$. In any case we have $m_1 + \cdots + m_r \leq n$. This proves:

Proposition 13.13. *Let $P(x) \in K[x]$ be a polynomial of degree n. We have*

$$P(x) = (x - a_1)^{m_1} \dots (x - a_r)^{m_r} R(x)$$

where $R(a) \neq 0 \, \forall a \in K$. In particular, $m_1 + \cdots + m_r \leq n$.

In other words, a polynomial of degree n, with coefficients in a field K, has at most n roots (counted with multiplicity) in K.

It is not always true that a polynomial of degree n with real coefficients has n real roots. For example $P(x) = x^2 + 1$ has no real root. The *fundamental theorem of algebra* states that this situation does not arise for polynomials with coefficients in $\mathbb{C}$:

Theorem 13.14 (Fundamental theorem of algebra)**.** *Let $P(x) \in \mathbb{C}[x]$ be a non-constant polynomial. Then $P(x)$ admits a complex root, that is, there exists $a \in \mathbb{C}$ such that $P(a) = 0$.*

This implies that if $P(x) \in \mathbb{C}[x]$ is a polynomial of positive degree then

$$P(x) = c\prod_{i=1}^{r}(x - a_i)^{m_i}$$

where $m_1 + \cdots + m_r = n$ and $c \in \mathbb{C}$.

Proof. We refer to [**DF04**, Theorem 35]. □

Theorem 13.14 tells us that $\mathbb{C}$ is *algebraically closed*.

Definition 13.15. A field K is algebraically closed if one of two following equivalent conditions hold.

(1) For every polynomial $P(x)$, of positive degree, there exists $a \in K$ such that $P(a) = 0$.

(2) For every polynomial $P(x)$ of degree $n \geq 1$ there is a factorization

$$P(x) = c\prod_{i=1}^{r}(x - a_i)^{m_i}$$

where $m_1 + \cdots + m_r = n$ and $c \in K$.

Note that (1) $\Rightarrow$ (2). In fact, if $P(a) = 0$, $(x - a) \mid P(x)$, after dividing the necessary number of times, we obtain $P(x) = (x - a)^{m_1}Q(x)$, with $Q(a) \neq 0$. If Q has positive degree, then there exists a_2 such that $Q(a_2) = 0$. Repeating the previous reasoning we obtain $P(x) = (x - a_1)^{m_1}(x - a_2)^{m_2}Q_1(x)$. Proceeding in this way we get $P(x) = (x-a_1)^{m_1}\cdots(x-a_r)^{m_r}R(x)$, with $R(a) \neq 0, \forall \in K$. Since K is algebraically closed $R(x) = c$ is a constant.

The interesting thing is that every field K is contained in an algebraically closed field, the smallest with respect to inclusion algebraically closed field that contains K is called the *algebraic closure* of K and is denoted by $\overline{K}$. For example, the algebraic closure of $\mathbb{R}$ is $\mathbb{C}$. The algebraic closure of $\mathbb{Q}$, $\overline{\mathbb{Q}}$, is not $\mathbb{C}$ but the field of *algebraic numbers*:

$$\overline{\mathbb{Q}} = \{z \in \mathbb{C} \mid z \text{ is the root of a polynomial in } \mathbb{Q}[x]\}.$$

If $z \in \mathbb{C} \setminus \overline{\mathbb{Q}}$ we say that z is *transcendental*. For instance, π and e (Euler's number) are transcendental. One can show that $\overline{\mathbb{Q}}$ is countable (since $\mathbb{Q}$ is). Since $\mathbb{R}$ (resp. $\mathbb{C}$) is uncountable if one takes a random real (resp. complex) number it will most likely be a transcendental number.

The thing you need to keep in mind is:

Remark 13.16. Each field is contained in its algebraic closure, $\overline{K}$, which is an algebraically closed field.

This means that somewhere every polynomial factors as a product of polynomials of degree one. In particular, a polynomial with real coefficients has all its roots in $\mathbb{C}$.

13.4. The main theorem on diagonalization

We have seen that an eigenvalue λ is a root of the characteristic polynomial, $P_f(x)$, and that each eigenvalue is associated to an eigenspace $E_f(\lambda)$.

Definition 13.17. Let E be a K-vector space, $f \in \mathrm{End}(E)$ and λ an eigenvalue of f.

(1) The *algebraic multiplicity* of λ, $m_a(\lambda)$, is the multiplicity of λ as a root of $P_f(x)$.

(2) The *geometric multiplicity* of λ, $m_g(\lambda)$, is the dimension of the eigenspace $E_f(\lambda)$.

Lemma 13.18. *Let E be a K-vector space, $f \in \mathrm{End}(E)$ and λ an eigenvalue of f. Then the geometric multiplicity of λ is always less than or equal to its algebraic multiplicity:*

$$1 \leq m_g(\lambda) \leq m_a(\lambda).$$

Proof. We have $m_g \geq 1$. Suppose $\dim E_f(\lambda) = t = m_g(\lambda)$. Let $(v_1, \ldots, v_t)$ be a basis of $E_f(\lambda)$. We can complete to a basis of E: $\mathcal{B} = (v_1, \ldots, v_t, w_{t+1}, \ldots, w_n)$. If $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ then

$$M = \begin{pmatrix} \lambda I_t & A \\ 0 & B \end{pmatrix}.$$

Here B is an $(n-t) \times (n-t)$ square matrix. We have

$$xI_n - M = \begin{pmatrix} (x-\lambda)I_t & -A \\ 0 & xI_{n-t} - B \end{pmatrix}$$

and applying Proposition 12.35 we get

$$P_f(x) = \det(xI_n - M) = \det((x-\lambda)I_t)\det(xI_{n-t} - B) = (x-\lambda)^t \det(xI_{n-t} - B).$$

It follows that $(x-\lambda)^t \mid P_f(x)$. Then $t \leq m_a(\lambda)$ and the proposition is proved. □

Now, we can state the main result.

Theorem 13.19. *Let E be a K-vector space and $f \in \mathrm{End}(E)$. The endomorphism f is diagonalizable if and only if the following two conditions hold:*

(1) *The characteristic polynomial $P_f(x)$ has all (counted with multiplicity) its roots in K.*

(2) *For each eigenvalue λ of f we have $m_a(\lambda) = m_g(\lambda)$.*

Proof. If f is diagonalizable there exists a basis made of eigenvectors and therefore $E = E_f(\lambda_1) \oplus \cdots \oplus E_f(\lambda_m)$. So $\sum m_g(\lambda_i) = n := \dim E$. Since $m_a(\lambda_i) \geq m_g(\lambda_i)$ (Lemma 13.18), $\sum m_a(\lambda_i) \geq \sum m_g(\lambda_i) = n$. On the other hand $n \geq \sum m_a(\lambda_i)$, because $P_f(x)$ has degree n. So we have $\sum m_a(\lambda_i) = \sum m_g(\lambda_i) = n$. This shows that $P_f(x)$ has all its roots in K. Since $m_a(\lambda_i) \geq m_g(\lambda_i)$, we have $m_g(\lambda_i) = m_a(\lambda_i)$, $\forall i$.

Vice versa, if the two conditions are satisfied $\sum m_g(\lambda_i) = \sum m_a(\lambda_i) = n$ (the first equality follows from (2), the second from (1)). From $\sum m_g(\lambda_i) = n$ it follows that $E = E_f(\lambda_1) \oplus \cdots \oplus E_f(\lambda_m)$, so there exists a basis of eigenvectors and f is diagonalizable. □

Let $A \in M_3(\mathbb{R})$ how can we establish if A is diagonalizable? We can always think that $A = \mathrm{Mat}(f; \mathcal{C}, \mathcal{C})$ where $f \in \mathrm{End}(E)$, $E = \mathbb{R}^3$, $\mathcal{C}$ the canonical basis. To diagonalize A is equivalent to diagonalize f.

(1) We calculate $P_A(x) = \det(xI_3 - A)$. Be careful since we then want to find the roots, it is convenient to factor.

(2) The roots of $P_A(x)$ are found, let us assume as will often (but not always) be the case that there is a simple root α and a double root β. So $P_A(x)$ has all its roots in $\mathbb{R}$. The first condition of Theorem 13.19 is satisfied.

(3) We need to verify the second condition, that is, we need to see if $E_f(\alpha)$ has dimension one and if $E_f(\beta)$ has dimension two. The verification in the case $E_f(\alpha)$ is immediate: from Lemma 13.18 it immediately follows that $m_g(\alpha) = m_a(\alpha) = 1$.

(4) For $E_f(\beta)$ we proceed like this: by the rank theorem $E_f(\beta) = \mathrm{Ker}(\beta I_3 - f)$ has dimension two if and only if $\beta I_3 - A$ has rank one. The rank of $\beta I_3 - A$ is then calculated: if it is one f is diagonalizable, otherwise f is not diagonalizable (observe that the rank of $\beta I_3 - A$ is at most 2).

Example 13.20. Consider the matrix

$$A = \begin{pmatrix} 4 & 3 & -3 \\ -36 & -20 & 18 \\ -30 & -15 & 13 \end{pmatrix}.$$

Let us see if A is diagonalizable.

(1) Calculation of the characteristic polynomial $P_A(x) = \det(xI_3 - A)$. We have:

$$\begin{vmatrix} x-4 & -3 & 3 \\ 36 & x+20 & -18 \\ 30 & 15 & x-13 \end{vmatrix} =_{C_2 \to C_2 + C_3} = \begin{vmatrix} x-4 & 0 & 3 \\ 36 & x+2 & -18 \\ 30 & x+2 & x-13 \end{vmatrix}$$

$$=_{R_2 \to R_2 - R_3} = \begin{vmatrix} x-4 & 0 & 3 \\ 6 & 0 & -x-5 \\ 30 & x+2 & x-13 \end{vmatrix} = -(x+2)[(x-4)(-x-5)+18]$$

From which we obtain $P_A(x) = (x+2)^2(x-1)$. Having two zeros on the second column we automatically have a factor (therefore a root) of $P_A(x)$.

(2) There are therefore two eigenvalues: $\lambda_1 = -2$ (with multiplicity two) and $\lambda_2 = 1$. The eigenspace relative to $\lambda_2 = 1$ necessarily has dimension one. The matrix is diagonalizable if and only if the eigenspace relative to $\lambda_1 = -2$ has dimension two, that is, if and only if $-2I_3 - A$ has rank one. We have

$$-2I_3 - A = \begin{pmatrix} -6 & -3 & 3 \\ 36 & 18 & -18 \\ 30 & 15 & -15 \end{pmatrix}$$

which clearly has rank one ($C_1 = 2C_2$, $C_3 = -C_2$). So A is diagonalizable.

To calculate $\det(xI_n - A)$ some try to modify A with combinations of columns and rows. Once they obtain a simpler matrix, A', they calculate $\det(xI_n - A')$. This is not good. What is the problem? The problem is that the matrix A' obtained from the matrix A after manipulation on the rows or columns is not necessarily similar to the starting matrix A, in other words the endomorphism considered could be different from f.

Let us continue the previous example. The eigenspace relative to -2 is given by the system:

$$\begin{pmatrix} -6 & -3 & 3 \\ 36 & 18 & -18 \\ 30 & 15 & -15 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since this eigenspace has dimension two it is defined by a single equation, for example: $-6x - 3y + 3z = 0$. For $x = 0$ we get $y = z$ and we have the vector $v_1 = (0, 1, 1)$. For $y = 0$ we get $z = 2x$ and we have the vector $v_2 = (1, 0, 2)$. The vectors v_1, v_2 are independent and form a basis of the eigenspace relative to $\lambda_1 = -2$.

The eigenspace relative to $\lambda_2 = 1$ is given by the system:

$$\begin{pmatrix} -3 & -3 & 3 \\ 36 & 21 & -18 \\ 30 & 15 & -12 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since the eigenspace has dimension one we need two independent equations, for example:

$$\begin{cases} -3x - 3y + 3z = 0; \\ 36x + 21y - 18z = 0. \end{cases}$$

From the first comes $x + y = z$ $(*)$; the second is rewritten as $12x + 7y - 6z = 0$, and plugin in $(*)$ we obtain $y = -6x, z = -5x$. Hence, $v_3 = (1, -6, -5)$ is a basis of the eigenspace relative to $\lambda_2 = 1$.

In conclusion $\mathcal{B} = (v_1, v_2, v_3)$ is an eigenvector basis.

When are two matrices $A, B \in M_n(K)$ similar? In general, the answer to this question is very complicated.

Lemma 13.21. *Let $A, B \in M_n(K)$. Suppose A diagonalizable: $A \approx D$, D diagonal. Then $A \approx B$ if and only if B is diagonalizable with the same eigenvalues (counting the multiplicities) of A, that is if and only if $B \approx D$.*

Proof. The similarity relation is an equivalence relation. □

What happens if neither A nor B is diagonalizable? This is the hard case and we will not develop a general method to decide. Let us consider the two matrices with real coefficients ($K = \mathbb{R}$):

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is clear (Corollary 12.30) that $P_A(x) = P_B(x) = (x-1)^3$. So there is a unique eigenvalue $\lambda = 1$, with algebraic multiplicity three. Neither A nor B is diagonalizable. In fact, if A were diagonalizable it would be similar to I_3, but a matrix similar to I_3 is equal to I_3 and $A \neq I_3$. Same reasoning for B. We can consider the dimension of the eigenspaces. We have $E_A(1) = \mathrm{Ker}(I_3 - A)$. The matrix

$$I_3 - A = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

clearly has rank one, so $\dim E_A(1) = 2$. Instead

$$I_3 - B = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

has rank two, so $\dim E_B(1) = 1$. We conclude that A is not similar to B. In fact, if A and B were similar they would represent the same endomorphism f in different bases (Proposition 11.16) and therefore $\dim E_f(1) = \dim E_A(1) = \dim E_B(1)$, but this is not the case.

We note that if we had found $\dim E_A(1) = \dim E_B(1)$, we would not have been able to conclude $A \approx B$. In this case one can try to find an invertible P such that $AP = PB$ (homogeneous linear system of 9 equations in 9 unknowns).

We observe that the matrices in the example have the same determinant and the same trace but they are not similar.

Exercises

Exercise 13.22. Find two matrices in $M_2(\mathbb{R})$ with the same characteristic polynomial but not similar.

Exercise 13.23. Let E be a K-vector space and u, v two endomorphisms of E. If u is bijective, prove that $u \circ v$ and $v \circ u$ have the same characteristic polynomial (Hint: consider $u^{-1} \circ u \circ v \circ u$).

Exercise 13.24. Consider the matrix

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

1) Calculate the characteristic polynomial of M and conclude that M is not diagonalizable.
2) Conclude that for every $n \geq 3$ there exists a non-diagonalizable $M \in M_n(K)$.

Exercise 13.25. 1) Show that every symmetric matrix, $M \in M_2(\mathbb{R})$, is diagonalizable.
2) Let $A \in M_2(\mathbb{C})$

$$A = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$$

Observe that A is symmetric. Is the matrix A diagonalizable?

Exercise 13.26. 1) Let E be a K-vector space and $f \in \mathrm{End}(E)$ such that f is not injective and there exists a hyperplane $H \subset E$ such that the restriction of f to H is the identity. Is f diagonalizable?
2) For every $n \geq 2$ give an example of an endomorphism f such that there exists a hyperplane $H \subset E$ such that the restriction of f to H is the identity and furthermore such that

(i) f is surjective, $f \neq Id$ and f is diagonalizable.

(ii) f is surjective and f is not diagonalizable.

Exercise 13.27. Let E be a K-vector space where K is a field of characteristic different from two. Let $f \in \mathrm{End}(E)$ be such that $f \circ f = Id_E$. We set $E_+ = \{x \in E \mid f(x) = x\}$ and $E_- = \{x \in E \mid f(x) = -x\}$.
1) Show that $E_+ \oplus E_- = E$.
2) Conclude that f is diagonalizable.

Exercise 13.28. Let $A \in M_3(K)$ be the following matrix:

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

State whether A is diagonalizable where $K = \mathbb{R}, \mathbb{C}, \mathbb{Z}/2\mathbb{Z}$.

Exercise 13.29. Let E be a finite-dimensional $\mathbb{C}$-vector space and f an endomorphism of E. It is assumed that there exists a subvector space $V \subset E$ such that $f(V) \subset V$ (V is stable under f).
1) Prove that V contains an eigenvector of f.
2) Let $g \in \mathrm{End}(E)$ be an endomorphism that commutes with f. Show that every eigenspace of f is stable under g.
3) Deduce that f and g have an eigenvector in common.

Exercise 13.30. (1) Does there exist a linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ such that $f((0,1,-1)) = (2,0,0)$ and such that $V = \{(x,y,z) \in \mathbb{R}^3 \,|\, y = z\}$ is the eigenspace relative to the eigenvalue $\lambda = 2$?

(2) Is such an f, if it exists, uniquely determined?

(3) If the answer to (1) is positive, determine the third eigenvalue.

Exercise 13.31. Consider the matrix

$$A = \begin{pmatrix} 4 & 3 & -3 \\ -36 & -20 & 18 \\ -30 & -15 & 13 \end{pmatrix} \in M_3(\mathbb{R}).$$

Show that A is diagonalizable and give a basis of eigenvectors.

Exercise 13.32. Are there two matrices $M, N \in M_2(\mathbb{R})$, $M \neq N$, that are similar but not diagonalizable? If the answer is yes, give an example. If the answer is no, explain why.

Exercise 13.33. Let $A \in M_n(\mathbb{R})$ be such that $A^2 + I_n = 0$. Show that A is not diagonalizable.

Exercise 13.34. Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be an endomorphism such that

$$\mathrm{Mat}(f; \mathcal{C}, \mathcal{C}) =: A = \begin{pmatrix} 5 & 9 & -3 \\ 0 & -1 & 0 \\ 6 & 9 & -4 \end{pmatrix}$$

where $\mathcal{C} = (e_1, e_2, e_3)$ is the canonical basis.

(1) Show that f is diagonalizable.

(2) Determine a basis of eigenvectors.

(3) Let $g \in \mathrm{End}(\mathbb{R}^3)$ be such that $(g \circ f)(e_1 + e_3) = 0$ and $g(e_2) = e_2$.
   (a) Show that if $g(e_1) = \alpha e_1$ then g is diagonalizable (choose an appropriate basis).
   (b) We assume e_1 is not an eigenvector of g and $g(e_1) \in \langle e_1, e_2 \rangle$. Observe that $g(e_1) = ae_1 + be_2$, with $b \neq 0$.
   (c) With the notations and assumptions of (b) show that g is diagonalizable $\Leftrightarrow a \neq 1$.

Exercise 13.35. State whether the following matrices with real coefficients are diagonalizable or not:

$$A = \begin{pmatrix} 0 & -3 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & -1 & -3 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}, C = \begin{pmatrix} -1 & 0 & -3 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Exercise 13.36. Consider the matrix

$$A = \begin{pmatrix} 3 & 6 & -1 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{pmatrix} \in M_3(\mathbb{R}).$$

Compute the trace of A^8.

Exercise 13.37. Let $A \in M_n(\mathbb{R})$.

(1) Show that if A is diagonalizable then A^2 is also diagonalizable.

(2) Is the converse of (1) true?

Exercise 13.38. (1) Let E be a K-vector space and $f \in \mathrm{End}(E)$. Show that if f is not a homothety (i.e. f is not of the form λId_E, $\lambda \in K$), then there exists $v \in E$ such that v and $f(v)$ are linearly independent.

(2) Show that every matrix $M \in M_2(K)$ is similar to a matrix of the form

$$\begin{pmatrix} 0 & a \\ 1 & b \end{pmatrix}$$

or to λI_2 for some $\lambda \in K$.

(3) Show that if $M \in M_2(K)$ then

$$Tr(M^2) = Tr(M)^2 - 2\det(M)$$

where as usual $Tr(M)$ indicates the trace of matrix M.

Exercise 13.39. Let E be a K-vector space of dimension n and let f an endomorphism of E that is not a homothety.

(1) Show that there exists a basis $(e_1, \ldots, e_n)$ of E such that $f(e_1) = e_2$.

(2) Let $A \in M_n(K)$ be a zero trace matrix (i.e. $Tr(A) = 0$). Show, using (i), that A is similar to a matrix whose diagonal contains only zeros (reason by induction on n).

Exercise 13.40. Without using determinant theory compute

$$|A_n(x)| = \begin{vmatrix} x & a & \cdots & \cdots & a \\ a & x & a & \cdots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a \\ a & \cdots & \cdots & a & x \end{vmatrix}.$$

Exercise 13.41. Consider the following matrices in $M_3(\mathbb{R})$:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(1) Are the matrices A and B diagonalizable?

(2) Are the matrices A and B similar?

Exercise 13.42. Let $t \in \mathbb{R}$ and

$$M_t = \begin{pmatrix} 0 & t & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & t \\ 0 & 0 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{R}).$$

(1) Determine $D := \{t \in \mathbb{R} \mid M_t \text{ is diagonalizable}\}$.

(2) Consider the matrix

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Determine $S := \{t \in \mathbb{R} \mid M_t \text{ is similar to } B\}$.

Exercise 13.43. Let $E = M_2(\mathbb{R})$ be the $\mathbb{R}$-vector space of 2×2 matrices with real coefficients, and

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in E.$$

(1) Show that $f : E \to E : M \to AM$ is linear.

(2) Show that f is diagonalizable if and only if A is diagonalizable.

Exercise 13.44. Let f be an endomorphism of the K-vector space E.
(i) Show that if $\mathrm{Ker}(f) \cap \mathrm{Im}(f) \neq \{0\}$ then f is not diagonalizable.
(ii) Show that f is diagonalizable $\Leftrightarrow \mathrm{Ker}(f) \cap \mathrm{Im}(f) = \{0\}$ and $f_{|\,\mathrm{Im}(f)}$ is diagonalizable. Here $f_{|\,\mathrm{Im}(f)}$ is the restriction of f to $\mathrm{Im}(f)$. Note that $\mathrm{Im}(f)$ is stable for f.

Exercise 13.45. We denote by $\mathcal{I}_r \in M_n(K)$ the matrix

$$\mathcal{I}_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

where I_r is the $r \times r$ identity matrix and the zeros indicate null matrices of the appropriate order. Let $N \in M_n(K)$ and write

$$N = \begin{pmatrix} N_1 & N_2 \\ N_3 & N_4 \end{pmatrix}$$

where N_1 is a square $r \times r$ matrix, $N_2 \in M_{r,n-r}(K)$, $N_3 \in M_{n-r,r}(K)$ and $N_4 \in M_{n-r}(K)$.
(i) Show that

$$\mathcal{I}_r N = \begin{pmatrix} N_1 & N_2 \\ 0 & 0 \end{pmatrix}$$

and

$$N\mathcal{I}_r = \begin{pmatrix} N_1 & 0 \\ N_3 & 0 \end{pmatrix}$$

by considering the linear maps associated to the matrices $\mathcal{I}_r$, N.
(ii) Conclude that $\mathcal{I}_r N$ and $N\mathcal{I}_r$ have the same characteristic polynomial.
(iii) Let $A, B \in M_n(K)$. Let r be the rank of A. We know that there exist invertible matrices P, Q such that $A = P\mathcal{I}_r Q$. Set $\overline{B} = QBP$ so that $B = Q^{-1}\overline{B}P^{-1}$. Show that AB is similar to $\mathcal{I}_r\overline{B}$ while BA is similar to $\overline{B}\mathcal{I}_r$. Conclude that AB and BA have the same characteristic polynomial.

Exercise 13.46. Let K be a field and $A, B \in M_n(K)$. The aim of the exercise is to show that AB and BA have the same characteristic polynomial and therefore also the same trace.
(i) If A (or B) is invertible show that AB and BA are similar. In particular, $P_{AB}(x) = P_{BA}(x)$.
(ii) Assumed K to be infinite. If $\alpha \in K$, observe that $\alpha I + A$ is invertible $\Leftrightarrow \alpha$ is not an eigenvalue of $-A$. We set $q(\alpha, x) = P_{(\alpha I+A)B}(x) - P_{B(\alpha I+A)}(x)$. Show that $q(\alpha, x)$ is a polynomial in α, x.
(iii) An algebraic subset $Z \subset K^2$ is the locus of the zeros of a finite set of polynomials i.e. the set of solutions of a system of polynomial equations: $Z = \{u \in K^2 \mid P_1(u) = \cdots = P_t(u) = 0\}$, where $P_i \in K[x, y]$.

The algebraic identity prolongation theorem says that if K is infinite and if a polynomial $P \in K[x,y]$ vanishes on $U = K^2 \setminus Z$, where Z is an algebraic set, then $P = 0$.

With the notations of (ii) observe that $q(\alpha, x) \in K[\alpha, x]$ vanishes on $U = K^2 \setminus \bigcup_{i=1}^{r} D_i$, where D_i is the line of equation $\alpha = \alpha_i$ and the α_i are the eigenvalues of A. Conclude that $P_{AB}(x) = P_{BA}(x)$.
(iv) If K is finite, let $\overline{K}$ be its algebraic closure. Observe that $\overline{K}$ is infinite. We consider A, B, AB, BA as elements of $M_n(\overline{K})$. The characteristic polynomial of AB, BA considered as elements of $M_n(\overline{K})$ is equal to the characteristic polynomial of AB, BA considered as elements of $M_n(K)$. Conclude that $P_{AB}(x) = P_{BA}(x)$ in $K[x]$ even if K is finite.

Exercise 13.47. (i) Let E be a 3-dimensional $\mathbb{C}$-vector space and f, g two endomorphisms of E such that $g^2 = f$. Let λ be an eigenvalue of f and $W = E_f(\lambda)$ the corresponding eigenspace.

Show that g has an eigenvalue α such that $\alpha^2 = \lambda$. Furthermore, if $\tilde{W} = E_g(\alpha)$ is the corresponding eigenspace, show that $\tilde{W} \subset W$ (use Exercise **??**). Finally, show with an example that we could have $\tilde{W} \neq W$.

(ii) Let $A \in M_3(\mathbb{C})$ be a diagonalizable matrix and

$$SQ(A) = \{M \in M_3(\mathbb{C}) \mid M^2 = A\}.$$

State whether $SQ(A)$ is a finite or infinite set (distinguish according to the multiplicity of the eigenvalues).

Exercise 13.48. Let $K = \mathbb{Z}/2\mathbb{Z}$ and $f : E \to E$ an invertible endomorphism of the K-vector space E of dimension n. Show that if f is diagonalizable then $f = Id$.

Exercise 13.49. Let E be a K-vector space and $f : E \to E$ an endomorphism of rank one.

(1) Show that if $f^2 \neq 0$ then f is diagonalizable.
(2) Show that if $\dim(E) > 1$ there exist non-diagonalizable rank-one endomorphisms of E.

Exercise 13.50. Tell for which values of m the matrix

$$A_m = \begin{pmatrix} 0 & 0 & m \\ 0 & m & 0 \\ m & 0 & 0 \end{pmatrix} \in M_3(\mathbb{R})$$

is diagonalizable, and for which values of m, A_m is similar to

$$B = \begin{pmatrix} 20 & 3 & 2008 \\ 372 & -157 & 34 \\ -285 & 129 & 137 \end{pmatrix}.$$

Exercise 13.51. (1) Let m be a real parameter and

$$A_m = \begin{pmatrix} -2m & 0 & -m \\ 0 & m & 0 \\ m & 0 & 0 \end{pmatrix} \in M_3(\mathbb{R}).$$

Determine

$$D := \{m \in \mathbb{R} \mid A_m \text{ is diagonalizable}\}.$$

(2) Let

$$B = \begin{pmatrix} 3 & 2 & 2 \\ -4 & -3 & -2 \\ 0 & 0 & -1 \end{pmatrix}.$$

For which values of m is A_m similar to B?

Exercise 13.52. Consider the matrix

$$N = \begin{pmatrix}
6 & 3 & 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
-3 & 0 & -2 & -2 & -2 & -2 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\
-3 & -2 & 0 & -2 & -2 & -2 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & -1 & -1 & -1 \\
-3 & -2 & -2 & 0 & -2 & -2 & -1 & 0 & -1 & -1 & 0 & -1 & -1 & 0 & 0 & -1 \\
-3 & -2 & -2 & -2 & 0 & -2 & -1 & -1 & 0 & -1 & -1 & 0 & -1 & 0 & -1 & 0 \\
-3 & -2 & -2 & -2 & -2 & 0 & -1 & -1 & -1 & 0 & -1 & -1 & 0 & -1 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

and the vector

$$v_{K^2} = (25, -9, -9, -9, -9, -9, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1).$$

(1) Show that v_{K^2} is fixed by N.

(2) Show that N^2 is the identity matrix.

(3) Compute the eigenvalues and the eigenvectors of N, and show that N is diagonalizable.

Exercise 13.53 (Jordan form for nilpotent endomorphisms in dimension three)**.** Let $f \neq 0$ be an endomorphism of $\mathbb{R}^3$ such that $f^3 = 0$ (such an endomorphism is called nilpotent).

(1) Show that f is not invertible and that if λ is an eigenvalue of f then $\lambda = 0$. Is f diagonalizable?

(2) Show that if $\mathrm{rk}(f) = 2$ then $f^2 \neq 0$.

(3) Show that if $\mathrm{rk}(f) = 2$ there exists a basis $\mathcal{B}$ such that

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

by considering $v, f(v), f^2(v)$ for a suitable v.

(4) Show that if $\mathrm{rk}(f) = 1$ then $f^2 = 0$ and there exists a basis $\mathcal{D}$ such that

$$\mathrm{Mat}(f; \mathcal{D}, \mathcal{D}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

(5) Conclude that two nilpotent matrices, $M, N \in M_3(\mathbb{R})$, are similar if and only if they have the same rank.

Exercise 13.54 (Triangularization of matrices)**.** Let E be a K-vector space of dimension n, where K is algebraically closed. Let f be an endomorphism of E. The aim of the exercise is to show that there exists a basis $\mathcal{B}$ such that $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ is *triangular*, that is:

$$A = \begin{pmatrix} \lambda_1 & \star & \star & \cdots & \star \\ 0 & \lambda_2 & \star & \star & \vdots \\ 0 & 0 & \ddots & \star & \vdots \\ \vdots & \vdots & \ddots & \ddots & \star \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

where $\lambda_1, \ldots, \lambda_n$ are the n eigenvalues (not necessarily distinct) of f.
(i) Show that f has an eigenvector e_1. Complete e_1 to a basis $\mathcal{B}_1 = (e_1, e_2, \ldots e_n)$ and write the matrix of f in the basis $\mathcal{B}_1$.
(ii) Now, we proceed by induction. The initial case is (i) and now we do the induction step. Suppose we have a basis $\mathcal{B}_k = (e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ with respect to which the matrix of f is of the form

$$\begin{pmatrix} \lambda_1 & \star & \star & \\ \vdots & \ddots & \star & A \\ 0 & \ldots & \lambda_k & \\ & 0_{n-k,k} & & B \end{pmatrix}$$

where $0_{n-k,k}$ stand for the $(n-k) \times k$ zero matrix. Let $V = \langle e_{k+1}, \ldots, e_n \rangle$ and $p : E \to V$ defined by $p(\sum_1^n \alpha_i e_i) = \sum_{k+1}^n \alpha_i e_i$. Let $\overline{f} := p \circ f \circ i : W \hookrightarrow E \to E \to W$, where $i : W \hookrightarrow E$ is the natural inclusion.

Show that $\overline{f}$ admits an eigenvector v_{k+1}. Complete to a basis $(v_{k+1}, \ldots, v_n)$ of W, and show that $\mathcal{B} = (e_1, \ldots, e_k, v_{k+1}, \ldots, v_n)$ is a basis of E. Write the matrix of f with respect to the basis $\mathcal{B}$ and conclude.

Exercise 13.55 (Jordan form for 3×3 matrices)**.** Let E be a 3-dimensional K-vector space, where K is algebraically closed. Let f be a non-diagonalizable endomorphism of E.
(1) Note that there are three cases:
(A) f has two distinct eigenvalues, λ, α; α has multiplicity one, λ has multiplicity two but the corresponding eigenspace, $E_f(\lambda)$, has dimension one.
(B1) f has a unique eigenvalue, λ, of multiplicity three and the eigenspace $E_f(\lambda)$ has dimension two.
(B2) f has a unique eigenvalue, λ, of multiplicity three and the eigenspace $E_f(\lambda)$ has dimension one.
(2) Show that:
(i) In case (A) there exists a basis $\mathcal{B}$ of E such that

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \alpha \end{pmatrix}.$$

(ii) In case (B1) there exists a basis $\mathcal{B}$ of E such that

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

(iii) In case (B2) there exists a basis $\mathcal{B}$ of E such that

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

using the fact that there exists a basis of E in which the matrix of f is triangular and then making suitable base changes).

Exercise 13.56. Let E be a K-vector space of dimension greater than three. Show that there exists a nilpotent endomorphism f of E of rank two with $f^2 = 0$ (cf. Exercise 13.53).

Exercise 13.57 (Nilpotent endomorphisms)**.** Let E be a K-vector space of dimension n, and f a nilpotent endomorphism of E with $f^{m+1} = 0$, $f^m \neq 0$ ($m \geq 1$).
(i) Show that 0 is an eigenvalue of f.
(ii) Show that if λ is an eigenvalue of f then $\lambda = 0$.
(iii) Show that $f^n = 0$ (Hint: consider $v, f(v), \ldots, f^m(v)$ for v appropriate).
(iv) Show that the characteristic polynomial of f is equal to X^n.

Exercise 13.58 (Rank one endomorphisms, take two)**.** Let E be a K-vector space of dimension $n \geq 2$ and $f : E \to E$ an endomorphism of rank one.
(i) Show that there exist $w \in E$, $w \neq 0$ and $\varphi : E \to K$ (i.e. $\varphi \in E^*$) such that

(1) $\mathrm{Ker}(f) = Ker(\varphi)$;

(2) $\forall v \in E$ we have $f(v) = \varphi(v)w$.

(ii) With the notations of (i), show that f is diagonalizable if and only if $w \notin H$ where $H := \mathrm{Ker}(f)$.
(iii) Show that f is diagonalizable if and only if $f^2 \neq 0$.
(iv) Show that a matrix of rank one is similar to

$$\begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}$$

with $\lambda \neq 0$ in the diagonalizable case, $f^2 \neq 0$; or to

$$\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

in the non-diagonalizable case, $f^2 = 0$. In the non-diagonalizable case: let $e_{n-2} = w$, e_n such that $\varphi(e_n) = 1$ and complete the basis of H.
(v) Let $P_f(X), M_f(X)$ be the characteristic and minimal polynomial of f. Show that:
f is diagonalizable $\Leftrightarrow P_f(X) = X^{n-1}(X - \lambda)$ with $\lambda \neq 0 \Leftrightarrow M_f(X) = X(X - \lambda)$.
f is not diagonalizable $\Leftrightarrow P_f(X) = X^n \Leftrightarrow M_f(X) = X^2$.

The aim of the next exercises is to determine the center and generators of the groups $Gl(E), Sl(E)$. For this we will use Exercise 13.58. These results have particular relevance in algebra and arithmetic when K is a finite field.

Exercise 13.59 (Transvections, dilations)**.** Let E be a K-vector space of dimension n and $f : E \to E$ an endomorphism, $f \neq Id$, such that there exists a hyperplane H with $f_{|H} = Id_H$.
(i) The following are equivalent:

(1) $\det(f) = \lambda \neq 1$;

(2) f is diagonalizable;

(3) the line $D := \mathrm{Im}(Id - f)$ is not contained in H.

In this case we say that f is a *dilation* of hyperplane H, line D and ratio λ.
(ii) Show that a dilation f is uniquely determined by H, D, λ.
(iii) The following are equivalent:

(1) $\det(f) = 1$;

(2) f is not diagonalizable;

(3) $D := \mathrm{Im}(Id - f) \subset H$;

(4) $\exists w \in H$, $w \neq 0$ and $\varphi \in E^*$ with $\mathrm{Ker}(\varphi) = H$ such that $f(v) = v + \varphi(v)w$, $\forall v \in E$;

(5) there exists a basis $\mathcal{B}$ such

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}.$$

In this case we say that f is a *transvection* of hyperplane H and line D.
(iv) Let $n = 2$ and $\mathcal{B} = (e_1, e_2)$ be a basis of E. Determine all transvections with $H = D = \langle e_1 \rangle$. Conclude that a transvection is not uniquely determined by H and D.
(v) Let $\varphi \in E^*$, $\varphi \neq 0$ and $w \in \mathrm{Ker}(\varphi)$, $w \neq 0$, observe that φ and w determine a transvection $f = \tau(\varphi, w)$ via $f(v) = v + \varphi(v)w, \forall v \in E$; $\tau(\varphi, w)$ is a transvection of hyperplane $H = \mathrm{Ker}(\varphi)$, line $D = \langle w \rangle$. Conversely, if $f = \tau(\varphi, w)$ is a transvection, f does not determine (φ, w).

Exercise 13.60 (Transvections: inverses, products)**.** Let E be a K-vector space.
(i) Show that if $f \in \mathrm{End}(E)$ is a transvection, then f^{-1} is also a transvection (with the same hyperplane and the same line).
(ii) Show that if f, g are two transvections of hyperplane H, then $f \circ g = Id$ or $f \circ g$ is a transvection of hyperplane H.
(iii) Show that, in general, the product of two transvections is not a transvection.

Exercise 13.61 (Center of $Gl(E)$ and $Sl(E)$)**.** Let E be a K-vector space of dimension n. The aim of this exercise is to determine the center of $Gl(E)$ and of the *special linear group*

$$Sl(E) = \{g \in Gl(E) \mid \det(g) = 1\}.$$

(i) Let $f \in \mathrm{End}(E)$ be such that $\forall v \in E$, v and $f(v)$ are dependent. Show that f is a homothety.

(ii) Let $f = \tau(\varphi, w)$ be a transvection (Exercise 13.59) and $g \in Gl(E)$ an automorphism of E. Show that gfg^{-1} is a transvection of hyperplane $g(H)$ and line $g(D)$ where $H = \mathrm{Ker}(\varphi)$ and $D = \langle w \rangle$.
(iii) Let G be a group. Show that

$$Z(G) = \{x \in G \mid xy = yx, \forall y \in G\}$$

is a normal subgroup of G, remember that a subgroup is normal $H \lhd G$ if $\forall x \in G : xHx^{-1} = H$. The subgroup $Z(G)$ is called the *center* of G.
(iv) Show that $Sl(E)$ is a normal subgroup of $Gl(E)$ by considering the determinant.
(v) Show that if $g \in Gl(E)$ commutes with every transvection, then g is a homothety (use (i) and (ii)). Deduce that

$$Z(Gl(E)) \simeq K^*$$

and

$$Z(Sl(E)) \simeq \mu_n = \{z \in K \mid z^n = 1\}$$

the multiplicative group of the n-th roots of unity.

Exercise 13.62 (Center of $Gl(E)$ and $Sl(E)$, take two)**.** A direct, simpler (but less informative) proof of the determination of $Z(Gl(E)), Z(Sl(E))$. Let E be a K-vector space of dimension n.
(i) Let $f \in \mathrm{End}(E)$ be such that $f \circ g = g \circ f, \forall g \in Sl(E)$. Show that $\forall v \in E, v$ and $f(v)$ are linearly dependent (if $f(w) = v$ with w, v linearly independent, consider g, h, appropriate automorphisms such that $g(w) = w + v$, $g(v) = -w$; $h(w) = w + v$, $h(v) = v$).
(ii) Conclude that every element of $Z(Gl(E))$ (resp. $Z(Sl(E))$) is a homothety (resp. an n-th root of unity) (Exercise 13.61 (i)).

Exercise 13.63 (Transvections)**.** Let E be a K-vector space of dimension n.
(i) Let $u, v \in E \setminus \{0\}$.

(1) If u, v are linearly independent, show that there exists a transvection, f, such that $f(u) = v$.

(2) If u, v are linearly dependent, show that there exist two transvections, f, g, such that $(f \circ g)(u) = v$.

(ii) Let H_1, H_2 be two distinct hyperplanes of E and let $v \notin H_1 \cup H_2$. Show that there exists a transvection, f, such that $f(v) = v$, $f(H_1) = H_2$.

Exercise 13.64 (Generators of $Gl(2, K)$ and $Sl(2, K)$)**.** Let E be a 2-dimensional K-vector space.
(i) Show that every $g \in Sl(E)$ that is not a homothety is written as a product of at most two transvections by using Exercise 13.63.
(ii) Show that $-Id$ can be written as a product of three transvections but not as a

product of two.
(iii) Show that if $f \in Gl(E)$ there exists a dilation (Exercise 13.59) h such that $(h \circ f) \in Sl(E)$. Conclude that dilations and transvections generate $Gl(E)$.

Exercise 13.65 (Generators of $Sl(E), Gl(E)$)**.** Let E be a K-vector space of dimension n.
(i) Show that each element of $Sl(E)$ can be written as a product of transvections (use Exercise 13.63, cf Exercise 13.64).
(ii) Conclude that dilations and transvections generate $Gl(E)$.

Exercise 13.66. Remember that two elements x, y of a group G are *conjugate* if there exists $z \in G$ such that $zxz^{-1} = y$.
Let E be a K-vector space of dimension n and $f, g \in Sl(E)$ two transvections.
(i) Use Exercise 13.59, (iii), and conclude that f and g are conjugate in $Gl(E)$ (i.e. $\exists u \in Gl(E)$ such that $ugu^{-1} = f$).
(ii) Let $\lambda \in K$, $\lambda \neq 0$. Show that if $n \geq 3$, there exists $v \in Gl(E)$ with $\det(v) = \frac{1}{\lambda}$ such that $vgv^{-1} = g$ (consider $\mathrm{Mat}(g; \mathcal{B}, \mathcal{B})$ for a suitable basis $\mathcal{B}$ of E.
(iii) Conclude that if $n \geq 3$, two transvections f and g are conjugate in $Sl(E)$ (i.e $\exists u \in Sl(E)$ such that $ugu^{-1} = f$).

Exercise 13.67 (Simplicity of $PSl(E), \dim(E) \geq 3$)**.** We have seen that the center, Z, of $Sl(E)$ is made up of the homotheties in $Sl(E)$ (Exercises 13.61, 13.62).

Let G be a subgroup of $Sl(E)$ containing Z. The aim of the exercise is to show, under the hypothesis $\dim(E) \geq 3$, that if G is normal (i.e. $fGf^{-1} = G, \forall f \in Sl(E)$) and $G \neq Z$, then $G = Sl(E)$.
(i) Observe that G being normal, if $g \in G$ and $f \in Sl(E)$, then $gfg^{-1}f^{-1} \in G$.
(ii) The idea now is to use (i) to show that G contains a transvection. In fact, show (using Exercise 13.65 and Exercise 13.66: this is where the hypothesis $\dim(E) \geq 3$ comes in) that if G contains a transvection, then $G = Sl(E)$.
(iii) Let $g \in G \setminus Z$. Since g is not a homothety there exist linearly independent u, v such that $g(u) = v$. Let t be a transvection of line $\langle u \rangle$ and set: $f = gtg^{-1}t^{-1}$. Show:

(1) $f \in G$ and $f \neq Id$ (note that gtg^{-1} is a transvection of line $\langle g(u) = v \rangle$).
(2) Let H be a hyperplane containing $\langle u, v \rangle$ (such hyperplanes exist because $\dim(E) \geq 3$). Show that $\forall x \in E$, $f(x) - x \in H$.
(3) Conclude that $f(H) = H$.

(iv) Suppose that there exists a transvection, h, of hyperplane H that does not commute with f ($hf \neq fh$). Let $v := fhf^{-1}h^{-1}$. Show that $v \in G$, $v \neq Id$, v is a transvection of hyperplane H (consider fhf^{-1}, h^{-1} and use Exercise 13.60).
(v) Finally, suppose instead that f commutes with each transvection of hyperplane

H. Show that f is a transvection of hyperplane H.
(vi) We have therefore shown that every normal subgroup of $Sl(E)$ containing Z is equal to Z or $Sl(E)$. Since the center is a normal subgroup, we can consider the quotient group: $PSl(E) := Sl(E)/Z$. Deduce from the above that if $\dim(E) \geq 3$, $PSl(E)$ is a *simple* group (a group is simple if each of its normal subgroups is trivial).

Chapter 14

The Cayley-Hamilton theorem

Let E be a K-vector space of dimension n and $f \in \mathrm{End}(E)$. If $Q(x) \in K[x]$, $Q(x) = a_p x^p + \cdots + a_1 x + a_0$ we set

$$Q(f) := a_p f^p + \cdots + a_1 f + a_0 Id$$

where $f^i = f \circ f \circ \cdots \circ f$, i times. Similarly, if $A \in M_n(K)$ we set

$$Q(A) := a_p A^p + \cdots + a_1 A + a_0 I_n.$$

The purpose of this section is to prove the following:

Theorem 14.1 (Cayley-Hamilton)**.** *Let E be a K-vector space and $f : E \to E$ an endomorphism. Then $P_f(f) = 0$. Equivalently, if $A \in M_n(K)$, then A is a root of its characteristic polynomial that is $P_A(A) = 0$.*

Remark 14.2. (1) If $A \in M_2(K)$, we have $P_A(x) = x^2 - Tr(A)x + \det(A)$. So by the Cayley-Hamilton theorem: $A^2 - Tr(A)A + \det(A)I_2 = 0$.

(2) The theorem might seem obvious, in fact $P_A(x) = \det(xI_n - A)$, so if we take $x = A$ we get $P_A(A) = \det(A - A) = 0$. Where is the mistake?

This result is an important theorem and has various applications. In particular, it can be useful for calculating the powers of a matrix.

We will give two proofs of the theorem. The first uses a very deep result of algebra (each field is a subfield of an algebraically closed field). The second uses Remark 12.40 i.e. the fact that the relation $MM^c = M^c M = \det(M)I_n$ is satisfied for matrices with coefficient in a commutative ring A (we will take $A = K[x]$).

14.1. Triangular matrices and Cayley-Hamilton theorem

We begin with a general result under the hypothesis that the characteristic polynomial has all its roots in the field K.

Proposition 14.3. *Let E be a K-vector space of dimension n and $f : E \to E$ an endomorphism of E. The characteristic polynomial of f is assumed to have all its roots (not necessarily distinct) in K:*

$$P_f(X) = (X - \lambda_1)(X - \lambda_2)\dots(X - \lambda_n).$$

Then there exists a basis $\mathcal{B}$ of E such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ is upper triangular, with $\lambda_1, \dots, \lambda_n$ on the diagonal.

Proof. The proof is by induction on $n = \dim(E)$. If $n = 1$ there is nothing to prove. Suppose the result is true for $n - 1$. Let $e_1 \neq 0$ such that $f(e_1) = \lambda_1 e_1$ and $\mathcal{B}_1 = (e_1, e'_2, \dots, e'n)$ a basis containing e_1. We have

$$\mathrm{Mat}(f; \mathcal{B}_1, \mathcal{B}_1) = \begin{pmatrix} \lambda_1 & a_{12} \dots a_{1n} \\ 0 & \\ \vdots & C \\ 0 & \end{pmatrix}.$$

Let $E' = \langle e'_2, \dots, e'_n \rangle$ and $f' : E' \to E \to E \to E'$ be the composite $f' = p \circ f \circ i$ where $i : E' \to E$ is the inclusion and where $p : E \to E'$ is the projection. Then $\mathcal{B}' = (e'_2, \dots, e'_n)$ is a basis of E' and $\mathrm{Mat}(f'; \mathcal{B}', \mathcal{B}') = C$. Furthermore

$$P_{f'}(X) = (X - \lambda_2)\dots(X - \lambda_n).$$

Then f' satisfies the induction hypothesis and there exists a basis $\mathcal{B}'_2 = (v_2, \dots, v_n)$ of E' such that $\mathrm{Mat}(f'; \mathcal{B}'_2, \mathcal{B}'_2)$ is upper triangular. The vectors $e_1, v_2, \dots, v_n$ are linearly independent, in fact from $\alpha e_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$ it follows $\alpha e_1 \in E' = \langle e'_2, \dots, e'_n \rangle$ and this implies $\alpha = 0$. It then follows that $\alpha_2 = \dots = \alpha_n = 0$. In the basis $\mathcal{B} = (e_1, v_2, \dots, v_n)$ the matrix of f is upper triangular:

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ 0 & \lambda_2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

The elements on the diagonal are necessarily $\lambda_1, \dots, \lambda_n$ as can be seen by calculating the characteristic polynomial. □

Corollary 14.4. *Let $f : E \to E$ be an endomorphism of the K-vector space E. If $P_f(X)$ has all its roots in K (a hypothesis certainly verified if K is algebraically closed) then $P_f(f) = 0$.*

Proof. If $P_f(X)$ has all its roots in K, by Proposition 14.3, there exists a basis $\mathcal{B}$ of E such that $\mathrm{Mat}(f;\mathcal{B},\mathcal{B})$ is upper triangular (with the eigenvalues λ_i on the diagonal):

$$\mathrm{Mat}(f;\mathcal{B},\mathcal{B}) = \begin{pmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ 0 & \lambda_2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

We set $f_i = f - \lambda_i Id$ and $F_i = f_1 \circ f_2 \circ \cdots \circ f_i$, $1 \leq i \leq n$. We observe that $f_i \circ f_j = f_j \circ f_i$. In fact if $P_i(X) = (X - \lambda_i)$, then $P_i(f) = f_i$. Since $P_i(X)P_j(X) = P_j(X)P_i(X)$ we have $f_i \circ f_j = f_j \circ f_i$.

Note that $F_t = F_{t-1} \circ f_t = f_t \circ F_{t-1}$, in fact f_t commutes with f_i. We show, by induction on t, that $F_t(e_i) = 0$ if $i \leq t$. We have $F_1(e_1) = (f - \lambda_1 Id)(e_1) = f(e_1) - \lambda_1 e_1 = 0$ (because $f(e_1) = \lambda_1 e_1$ by construction of the basis $\mathcal{B}$).

Suppose the statement is true for $t-1$ and prove it for t: if $j < t$, $F_t(e_j) = (F_{t-1} \circ f_t)(e_j) = (f_t \circ F_{t-1})(e_j) = f_t(F_{t-1}(e_j)) = 0$ (because $F_{t-1}(e_j) = 0$ by induction hypothesis). Finally, $f_t(e_t) = (f - \lambda_t Id)(e_t) = \sum_{i=1}^{t-1} c_i e_i$, $F_t(e_t) = (F_{t-1} \circ f_t)(e_t) = \sum_{i=1}^{i=t-1} c_i F_{t-1}(e_i) = 0$.

In particular, $F_n(e_i) = 0, \forall i, 1 \leq i \leq n$, that is $F_n = 0$. So

$$F_n = (f - \lambda_1 Id) \circ (f - \lambda_2 Id) \circ \cdots \circ (f - \lambda_n Id) = P_f(f) = 0$$

and the proposition is proved. □

Remark 14.5. The previous corollary proves the Cayley-Hamilton theorem under the hypothesis that $P_f(X)$ has all its roots in K. We observed that this hypothesis is certainly verified if K is algebraically closed. Then the Cayley-Hamilton theorem is proved for K algebraically closed.

To move to the general case (K not necessarily algebraically closed) we will use a very deep algebraic result (already mentioned in Section 13.3).

Theorem 14.6. *Each field K is a subfield of an algebraically closed field. More precisely, there exists a unique (module isomorphism) smallest algebraically closed field containing K, this field, denoted by $\overline{K}$, is the algebraic closure of K.*

Proof. We refer to [**DF04**, Propositions 30, 31]. □

We will now prove the the Cayley-Hamilton theorem.

Proof. (First proof of the Cayley-Hamilton theorem) Let f be an endomorphism of the K-vector space E, and $A = \mathrm{Mat}(f;\mathcal{B},\mathcal{B})$ the matrix of f with respect to a basis $\mathcal{B}$ of E. We have $K \subset \overline{K}$, with $\overline{K}$ algebraically closed (Theorem 14.6). So

$A \in M_n(K) \subset M_n(\overline{K})$ and we can consider A as the matrix of an endomorphism of $\overline{K}^n$ with $n = dim(E)$.

We have $P_A(X) \in K[X] \subset \overline{K}[X]$; now $P_A(X)$, seen as a polynomial in $\overline{K}$, has all its roots in $\overline{K}$. So (Corollary 14.4) $P_A(A) = 0$. Since $P_A(X) = X^n + a_1X^{n-1} + \cdots + a_n$ with $a_i \in K, \forall i$, the relation $P_A(A) = 0$ is also verified in $M_n(K)$, and hence $P_f(f) = 0$. □

14.2. Second proof of the Cayley-Hamilton theorem

Let us make some reminders about the complementary matrix of a matrix $M \in M_n(K)$. The complementary matrix, M^c, is the transpose of the cofactor matrix of M. We have:

$$MM^c = M^cM = \det(M)I_n. \tag{14.1}$$

Actually the definition of M^c and (14.1) hold for matrices with coefficients in a commutative ring R (12.40). In the following we will take $R = K[X]$.

By definition $P_f(X) = \det(XI_n - A)$ where A is the matrix of f with respect to a fixed basis. We set $M := XI_n - A \in M_n(K[X])$. The coefficients of the matrix M are polynomials of degree at most one: $M = (\delta_{ij}X - a_{ij})$.

From 14.1 it follows that

$$MM^c = P_f(X)I_n. \tag{14.2}$$

The coefficients of the matrix M^c are the minors of order $n-1$ of M, so they are polynomials of degree at most $n-1$. Therefore M^c is a polynomial of degree at most $n-1$ with coefficients in $M_n(K)$. Let us take an example to clarify this last statement. Let

$$P = \begin{pmatrix} X^2 - X + 3 & X - 1 \\ 3X^2 + X + 2 & -X^2 + 3 \end{pmatrix}$$

then

$$P = X^2 \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix} + X \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 3 & -1 \\ 2 & 3 \end{pmatrix}.$$

Obviously this writing of P as a polynomial in X with coefficients in $M_n(K)$ is uniquely determined.

We are now ready to give our second proof of the Cayley-Hamilton theorem.

Proof. (Second proof of the Cayley-Hamilton theorem) With the previous notations we have $M^c = B_1X^{n-1} + \cdots + B_{n-1}X + B_n$ where $B_i \in M_n(K)$. The relation 14.2 is written as

$$(XI_n - A)(B_1X^{n-1} + \cdots + B_{n-1}X + B_n) = P_f(X)I_n. \tag{14.3}$$

Let us set $P_f(X) = X^n + a_1X^{n-1} + \cdots + a_n$ $(a_i \in K, \forall i)$, then by 14.3 we get $X^nB_1 + X^{n-1}(B_2 - AB_1) + \cdots + X(B_n - AB_{n-1}) - AB_n = X^nI_n + X^{n-1}a_1I_n + \cdots + a_nI_n$, and hence

$$B_1 = I_n,\ B_2 - AB_1 = a_1I_n,\ \ldots,\ B_n - AB_{n-1} = a_{n-1}I_n,\ -AB_n = a_nI_n.$$

Multiplying (on the left) the first equation by A^n, the second by A^{n-1} and so on, up to the second to last by A, we obtain $A^nB_1 = A^n$, $A^{n-1}B_2 - A^nB_1 = a_1A^{n-1}, \ldots,$ $AB_n - A^2B_{n-1} = a_{n-1}A$, $-AB_n = a_nI_n$. Adding term by term we get

$$A^n + a_1A^{n-1} + \cdots + a_nI_n = P_f(A) = 0.$$

In fact the terms of the terms on the left cancel out. So the matrix of f is a root of the characteristic polynomial, in terms of endomorphisms $P_f(f) = 0$. □

Exercises

Exercise 14.7. Let $A \in M_2(K)$. We assume A to be invertible, show, without doing calculations, that

$$A^{-1} = \frac{1}{\det(A)}(-A + Tr(A)I_2).$$

Exercise 14.8. Consider the matrix

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

and compute A^7.

Exercise 14.9. Consider the matrix

$$B = \begin{pmatrix} 3 & 1 & -3 \\ 20 & 5 & -18 \\ 8 & 2 & -7 \end{pmatrix} \in M_3(\mathbb{R}).$$

(1) Compute the characteristic polynomial $P_B(x)$.

(2) Is B diagonalizable?

The aim of the exercise is to compute B^n. There are various steps. Let us first consider

$$x^n = P_B(x)Q_n(x) + R_n(x)$$

with $R_N(x) = 0$ or $\deg(R_n) < 3$ (Euclidean division of x^n by $P_B(x)$). We can set $R_n(x) = a_n x^2 + b_n x + c_n$. Setting $x = 1$ and $x = -1$ we find two linear relations between a_n, b_n, c_n. To find a third relation use the fact that 1 is a double root of $P_B(x)$.

At this point we have a linear system of three equations with unknowns a_n, b_n, c_n. Solve it and determine $R_n(x)$. Express B^n as a function of B^2, B and I_3.

Exercise 14.10. Let E be a K-vector space of dimension n, and $f \in \mathrm{End}(E)$ a nilpotent endomorphism.

(1) Show that every eigenvalue of f is zero.

(2) Conclude that $f^n = 0$ (associate a matrix $A \in M_n(K)$ to f and use Cayley-Hamilton).

Exercise 14.11. Let f be an endomorphism of $\mathbb{R}^3$, of characteristic polynomial

$$P_f(X) = (X - \alpha)^2(X - \beta)$$

with $\alpha, \beta \in \mathbb{R}$, $\alpha \neq \beta$. We set

$$N_\alpha := \mathrm{Ker}(f - \alpha Id)^2, \quad N_\beta := \mathrm{Ker}(f - \beta Id).$$

(i) Show that $N_\alpha \cap N_\beta = \{0\}$.
(ii) Show that $\dim(N_\alpha) = 2$ (use the Cayley-Hamilton theorem).
(iii) Conclude that $\mathbb{R}^3 = N_\alpha \oplus N_\beta$.

Exercise 14.12. (i) Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be an endomorphism of rank two. It is assumed that $P_f(x)$ (the characteristic polynomial) has all its roots in $\mathbb{R}$. Show that 0 is a simple root of $P_f(x)$ if and only if $\mathrm{Ker}(f) \oplus \mathrm{Im}(f) = \mathbb{R}^3$.
(ii) Give an example of a rank two endomorphism $f \in \mathrm{End}(\mathbb{R}^3)$ such that $P_f(X)$ has all its roots in $\mathbb{R}$ and such that $\mathrm{Ker}(f) \cap \mathrm{Im}(f) \neq \{0\}$ (see Exercise 14.13 for a more general result).

Exercise 14.13 (Characteristic subspaces)**.** Let E be a K-vector space of dimension n and $f \in \mathrm{End}(E)$ an endomorphism of rank $n-1$. The characteristic polynomial $P_f(x)$ is assumed to have all its roots in K. Show that 0 is a simple root of $P_f(x)$ if and only if $\mathrm{Ker}(f) \oplus \mathrm{Im}(f) = E$ (Exercise 14.14).

Exercise 14.14 (Characteristic subspaces)**.** Let E be a K-vector space of dimension n and $f \in \mathrm{End}(E)$. The characteristic polynomial, $P_f(x)$, of f is assumed to have all its roots in K. For every $i \geq 0$ we set

$$K_i = \mathrm{Ker}(f^i), \quad J_i = \mathrm{Im}(f^i)$$

with $f^0 = Id$. Let $s := \min\{i \mid K_i \oplus J_i = E\}$ (Exercise 9.47). Finally, let m be the multiplicity of zero as a root of $P_f(x)$ ($m = 0$ if $P_f(0) \neq 0$).
(i) Show that $s \leq m$. Give an example where the inequality is strict.
(ii) Let $p := \dim \mathrm{Ker}(f)$. Show that $p \leq m$. Give an example where the inequality is strict.
(iii) Show, with an example, that in general $K_p \oplus J_p = E$ (p as in (ii)) does not imply $m = p$ (in Exercise 14.13 instead it is shown that the implication is true if $p = 1$).
(iv) We have $s \leq m$ and $p \leq m$, however there is no order relation between s and p. Show with examples that we can have both $s > p$ and $p > s$.

Exercise 14.15. Let $A \in M_2(\mathbb{R}), A \neq I_2$ such that $A^3 = I_2$. Show that $Tr(A) = -1$.

Exercise 14.16. Let $f : E \to E$ be an endomorphism of the $\mathbb{R}$-vector space E with $\dim(E) = 3$. We assume

$$P_f(X) = (X - \alpha)^2(X - \beta)$$

with $\alpha \neq \beta$. Show that if

$$M_f(X) = (X - \alpha)(X - \beta),$$

then f is diagonalizable (P_f, M_f indicate respectively the characteristic polynomial and the minimal polynomial of f).

Exercise 14.17. Let $f : E \to E$ be an endomorphism of the K-vector space E. Show, without using Cayley-Hamilton, that if λ is a root of the minimal polynomial of f, then λ is an eigenvalue.

Exercise 14.18 (Minimal polynomial, rank one endomorphisms)**.** Let E be a K-vector space of dimension $n > 1$ and $f : E \to E$ an endomorphism, $M_f(X)$ denotes the minimal polynomial of f.
(i) Show that $M_f(X)$ has degree one if and only if $f = \lambda Id$ ($\lambda \neq 0$).
(ii) Assume f of rank one, show that there exists $\alpha \in K$, uniquely determined, such that $f^2 = \alpha f$. Conclude that $M_f(X)$ has degree two.
(iii) Show that for every r, $1 \leq r \leq n$, there exists an endomorphism of E, of rank r, with minimal polynomial of degree two.

Exercise 14.19. Let E be a K-vector space and f an endomorphism of E, f is assumed to be diagonalizable. Show that if $V \subset E$ is a stable subvector space under f ($f(V) \subset V$), then $f_V : V \to V$ is diagonalizable.

Exercise 14.20 (Minimum polynomial, $K = \mathbb{Z}_2$)**.** Let $f : E \to E$ be an endomorphism of the $\mathbb{Z}_2$-vector space, E, of dimension n, f is assumed to be invertible. Show that if the minimal polynomial of f has degree two then f is not diagonalizable.

Exercise 14.21. Let $A \in M_n(K)$ be such that $A^3 = A$.
(i) Show that if $\mathrm{char}(K) \neq 2$ then A is diagonalizable.
(ii) If $\mathrm{char}(K) = 2$ show that there exists $A \in M_3(K)$, with $A^3 = A$, such that A is not diagonalizable.

Exercise 14.22. Let $A \in M_n(\mathbb{R})$ and $M_A(x) \in \mathbb{R}[x]$ its minimal polynomial. We can consider A as a matrix with complex coefficients ($A \in M_n(\mathbb{C})$). The aim of the exercise is to show that the minimal polynomial of $A \in M_n(\mathbb{C})$ is still $M_A(x)$.
(i) Let $K \subset L$ be a field extension, E a K-vector space of dimension n and F an L-vector space. Let $f : E \to F$ be a K-linear map (note that F is also a K-vector space). It is assumed that there exists a basis, (e_i), of the K-vector space E such that $(f(e_i))$ is a basis of the L-vector space F.

Show that if $v_1, \ldots, v_m$ are independent in E (over K), then $f(v_1), \ldots, f(v_m)$ are independent in F (over L).
(ii) Let $E = M_n(K)$, $F = M_n(L)$ and $f : E \to F$ the canonical injection. Observe that the canonical basis of E is transformed, via f, into the canonical basis of the

L-vector space $M_n(L)$. Conclude that if $A_1, \ldots, A_m \in M_n(K)$ are independent, then $A_1, \ldots, A_m$ are independent in $M_n(L)$.
(iii) Let $A \in M_n(K)$ and $M_A(x) \in K[x]$ its minimal polynomial. Show that the minimal polynomial of $A \in M_n(L)$ is still $M_A(x)$ (note that if d is the degree of the minimal polynomial, then $I, A, A^2, \ldots, A^{d-1}$ are independent).
(iv) Let $A \in M_n(\mathbb{R})$. The characteristic polynomial of A is assumed to have all its roots in $\mathbb{R}$. Show that A is diagonalizable on $\mathbb{R} \Leftrightarrow A$ is diagonalizable on $\mathbb{C}$.

Exercise 14.23. Consider two matrices $M, N \in M_3(\mathbb{R})$ having the same characteristic polynomial, namely

$$P_M(X) = P_N(X) = (X - \lambda)^3.$$

Show that M and N are similar if and only if $\mathrm{rk}(\lambda I_3 - M) = \mathrm{rk}(\lambda I_3 - N)$.

Chapter 15

Linear systems

This is the last section of linear algebra, it is a bridge between linear algebra and affine geometry.

In general, a linear system is a system of equations of the following type:

$$\mathcal{S} = \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1; \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = b_p; \end{cases}$$

where the $a_{ij}, b_i \in K$ and the b_i are not necessarily zero.

Such a system may have no solution, for example let us consider the system:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1; \\ a_{11}x_1 + \cdots + a_{1n}x_n = b_2. \end{cases}$$

If $b_1 \neq b_2$ this system has no solution (we can add many other equations to these two and we will always have a system without solutions).

Definition 15.1. A linear system that does not admit a solution is said to be *incompatible*; if the system admits solutions we say that the system is *compatible*.

A linear system can be viewed in several different ways.

(1) Let $\varphi_i : K^n \to K : (x_1, \ldots, x : n) \to \varphi_i(x) = a_{i1}x_1 + \cdots + a_{in}x_n$. Then the linear system $\mathcal{S}$ can be written:

$$\begin{cases} \varphi_1(x) = b_1; \\ \vdots \\ \varphi_p(x) = b_p. \end{cases}$$

Here $x = (x_1, \ldots, x_n) \in K^n$ and $\varphi_i : K^n \to K$ is a linear form, i.e. an element of the dual $(K^n)^*$.

Let $A = (a_{ij})$ be the matrix with p rows and n columns. The system $\mathcal{S}$ can be rewritten as follows:

$$AX = B.$$

Here $X = {}^t(x_1, \ldots, x_n)$ and $B = {}^t(b_1, \ldots, b_p)$. We can see A as the matrix associated to the linear map $f : K^n \to K^p : X \to AX$, with respect to the canonical bases ($A = \mathrm{Mat}(f; \mathcal{C}_n, \mathcal{C}_p)$). A solution of the system is an $X \in K^n$ such that $f(X) = B$. In other words, the solution set of the system is

$$S = \{X \in K^n \mid f(X) = B\} = f^{-1}(B).$$

The system $AX = B$ can be rewritten as follows:

$$x_1 C_1 + \cdots + x_n C_n = B$$

where $C_1, \ldots, C_n$ are the columns of A. The set of solutions are the $x_1, \ldots, x_n$ for which B can be written as a linear combination of $C_1, \ldots, C_n$.

15.1. Structure of the solution set

Contrary to what happens with homogeneous linear systems ($b_i = 0, \forall i$) a general linear system can have no solutions: $S = \{\emptyset\}$. As already said, the simplest example consists of taking $\varphi_2 = \varphi_1$ and $b_1 \neq b_2$. For example:

$$\begin{cases} x - y + 2z = 1; \\ x - y + 2z = 2 : \\ 3x + 2y + z = 0. \end{cases}$$

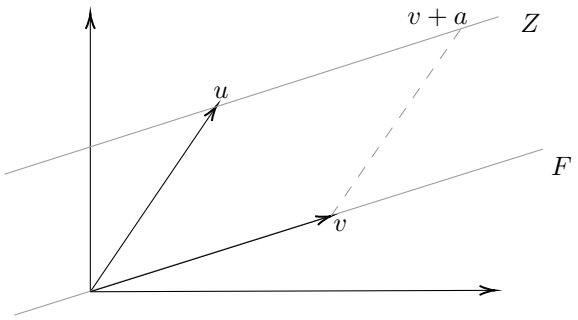
Now, suppose that our system $\mathcal{S}$ has a solution X_0. In the matrix interpretation (or with the linear map) we have $f(X_0) = B$. If $Y \in \mathrm{Ker}(f)$, then $f(X_0 + Y) = f(X_0) + f(Y) = f(X_0) = B$, then $X_0 + Y$ is also a solution of the system. Conversely, if Z is another solution of the system, we have $Z = X_0 + (Z - X_0)$ and $f(Z - X_0) = f(Z) - f(X_0) = 0$, therefore $-X_0 \in \mathrm{Ker}(f)$. This shows that the solution set of the system is of the form

$$S = \{X_0 + Y \mid Y \in \mathrm{Ker}(f)\}.$$

This last writing has the advantage of being suggestive and making us see that S is the translated of the vector subspace $\mathrm{Ker}(f)$. More precisely, let $t : K^n \to K^n : Y \to Y + X_0$, the map t is the *translation* of vector X_0. Then $S = t(\mathrm{Ker}(f))$.

Definition 15.2. Let E be a K-vector space. A subset $Z \subset E$ is an *affine subspace* of E if Z can be written in the form $Z = t_a(F)$, where $F \subset E$ is a subvector space, $a \in E$ and where $t_a : E \to E : v \to a + v$ is the translation of vector a.

The dimension of Z is, by definition, the dimension of F, and F is said to be the *direction* of Z.



We can summarize what was said before:

Proposition 15.3. *Suppose the system $\mathcal{S}$ is compatible and X_0 is a particular solution of $\mathcal{S}$. Then the set of solutions is the affine subspace $X_0 + \mathrm{Ker}(f)$, of dimension $\dim(\mathrm{Ker}(f))$ and direction $\mathrm{Ker}(f)$.*

We observe that $\mathrm{Ker}(f) = \{X \mid AX = 0\}$ is nothing but the set of solutions of the *associated* homogeneous linear system:

$$\mathcal{S}_0 = \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0; \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = 0. \end{cases}$$

In other words, if the system is compatible, the set of solutions is obtained by adding the solutions of the associated homogeneous system (elements of $\mathrm{Ker}(f)$) to a particular solution X_0.

At this point we still need to understand two things:

(1) How do we know if a system is compatible or not?

(2) If the system is compatible, how do you find a particular solution (and also the solutions of the associated homogeneous system)?

15.2. Compatibility condition

Let $\mathcal{S}$ be the linear system $AX = B$ (matrix writing). The matrix A is called the *coefficient matrix*, the rank of the system is the rank of A. If $\mathrm{rk}(A) = r$, then $\dim\langle\varphi_1, \ldots, \varphi_p\rangle = r$ (row interpretation). By rearranging the vectors we can assume that $(\varphi_1, \ldots, \varphi_r)$ is a basis of $\langle\varphi_1, \ldots, \varphi_p\rangle$. This implies that φ_i is a linear combination of $\varphi_1, \ldots, \varphi_r$, if $i > r$:

$$\sum_{k=1}^{r} \lambda_k^{(i)} \varphi_k = \varphi_i, \quad if \ i > r.$$

The compatibility conditions are:

$$\sum_{k=1}^{r} \lambda_k^{(i)} b_k = b_i, i > r. \tag{15.1}$$

In fact, let us consider the system:

$$\mathcal{T} = \begin{cases} \varphi_1(x) = b_1; \\ \vdots \\ \varphi_r(x) = b_r. \end{cases}$$

Then $\mathcal{S}$ and $\mathcal{T}$ are equivalent systems, that is, they have the same set of solutions S. It is clear that every solution of $\mathcal{S}$ is a solution of $\mathcal{T}$. Conversely, if x_0 is a solution of $\mathcal{T}$ then it is also a solution of $\mathcal{S}$ because, if $i > r$, $\varphi_i(x_0) = \sum_{k=1}^{r} \lambda_k^{(i)} \varphi_k(x_0) = \sum_{k=1}^{r} \lambda_k^{(i)} b_k = b_i$.

Lemma 15.4. *A system of r equations, in n unknowns, of rank r is always compatible and the set of solutions is an affine subspace of dimension $n - r$.*

Proof. Let $AX = B$ be a system of r equations in n unknowns, and $f : K^n \to K^r$ such that $A = \mathrm{Mat}(f; \mathcal{C}_n, \mathcal{C}_r)$. Since $\mathrm{rk}(f) = \mathrm{rk}(A) = r$, f is surjective, so there exists $X_0 \in K^n$ such that $f(X_0) = B$. On the other hand, by the Rank Theorem $\dim \mathrm{Ker}(f) = n - r$. □

From this lemma it follows that $\mathcal{T}$ is compatible, so $S \neq \emptyset$ and $\mathcal{S}$ is compatible as well. Finally, it is clear that $\mathcal{S}$ is incompatible if the compatibility conditions are not satisfied.

This analysis will be useful later but fortunately there is a simpler way to check if a system is compatible. Let $\mathcal{S} : AX = B$, $A = (a_{ij})$. We denote by $(A|B)$ the matrix obtained by adding the column of known terms to A. This matrix is called the *complete matrix* of the system:

$$(A|B) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{p1} & \cdots & a_{pn} & b_p \end{pmatrix}.$$

Lemma 15.5. *The system $\mathcal{S}$ is compatible if and only if* $\mathrm{rk}(A) = \mathrm{rk}(A|B)$.

Proof. Let us consider the column interpretation of the system: $x_1 C_1 + \cdots + x_n C_n = B$. The system is compatible if and only if B can be written as a linear combination of the column vectors, that is, if and only if $B \in \langle C_1, \ldots, C_n \rangle$. This can be rewritten as follows: $\langle C_1, \ldots, C_n \rangle = \langle C_1, \ldots, C_n, B \rangle$ and this is true if and only if $\mathrm{rk}(A) = \mathrm{rk}(A|B)$. □

This provides a fairly convenient criterion for seeing if a system is compatible. We still have to figure out how to find solutions. Lemma 15.4 suggests that the easiest case could be that of a system of n equations in n unknowns, of rank n: in this case the set of solutions has dimension zero, that is, there is a single solution. Such a system is called a *Cramer system*.

15.3. Cramer systems

In this section $\mathcal{S} : AX = B$ will indicate a *Cramer system* i.e. a system of n equations in n unknowns, of rank n. So $A \in M_n(K)$ and A is invertible. We know that this system has a unique solution (Lemma 15.4). Let $X_0 = (\alpha_1, \ldots, \alpha_n)$ be the unique solution of the system. Obviously $X_0 = A^{-1}B$. Let us see a faster way to compute X_0. We denote by A_i the matrix obtained from A by replacing the i column with B.

Lemma 15.6. *With the previous notations we have, for every i, $1 \leq i \leq n$:*

$$\alpha_i = \frac{\det(A_i)}{\det(A)}.$$

Proof. Let us consider the column interpretation:

$$\alpha_1 C_1 + \cdots + \alpha_n C_n = B.$$

The solution $(\alpha_1, \ldots, \alpha_n)$ is nothing but the coordinates of the vector B in the basis $(C_1, \ldots, C_n)$ ($\mathrm{rk}(A) = n$). We have $\det(A_i) = \det(C_1, \ldots, B, \ldots, C_n)$ (B in the entry i). We plug-in the above expression for B:

$$\det(A_i) = \det(C_1, \ldots, \alpha_1 C_1 + \cdots + \alpha_n C_n, \ldots, C_n).$$

Since the determinant is a multilinear function of the columns we have:

$$\det(A_i) = \sum_j \alpha_j \det(C_1, \ldots, C_j, \ldots, C_n).$$

If $i \neq j$ then

$$\det(C_1, \ldots, C_j, \ldots, C_n) = 0$$

since two columns are equal. If $i = j$ then

$$\det(C_1, \ldots, C_j, \ldots, C_n) = \det(A)$$

and the result follows. □

The general case reduces to the case of a Cramer system. Let $\mathcal{S} : AX = B$ be our usual system. First, we compute the rank of the matrix A and check the compatibility condition is verified $\mathrm{rk}(A) = \mathrm{rk}(A|B)$. Suppose the condition is satisfied and set $r = \mathrm{rk}(A)$. There is therefore a non-zero minor of order r: $\Delta_r \neq 0$. By reordering, if needed, the rows and columns we can assume that Δ_r is the minor

constructed on the first r rows and columns. The system is then equivalent to the system:

$$\mathcal{S}' = \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1; \\ \vdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n = b_r. \end{cases}$$

Let A' be the matrix of the system $\mathcal{S}'$. Since $\mathrm{rk}(A') = r$ there are r linearly independent columns of A'. After reordering we can assume that they are the first r. We can rewrite the system as follows:

$$\mathcal{S}' = \begin{cases} a_{11}x_1 + \cdots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \cdots a_{1n}x_n; \\ \vdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - \cdots a_{rn}x_n. \end{cases}$$

Assigning values to $x_{r+1}, \ldots, x_n$ we obtain a Cramer system, which we know how to solve. We conclude that the set of solutions, S, is the set of all the solutions of these Cramer systems, varying $x_{r+1}, \ldots, x_n$.

Another way of doing this consists in solving the Cramer system $\mathcal{S}'$ formally with Lemma 15.6:

$$x_i = \frac{\det(A''_i)}{\det(A'')}.$$

Here

$$A'' = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}$$

and A''_i is the matrix obtained from A'' by replacing the i-th column with the column

$$B'(x_{r+1}, \ldots, x_n) = \begin{pmatrix} b_1 - a_{1,r+1}x_{r+1} - \cdots a_{1n}x_n \\ \vdots \\ b_r - a_{r,r+1}x_{r+1} - \cdots a_{rn}x_n \end{pmatrix}.$$

The r equations above thus express $x_1, \ldots, x_r$ as functions of the $(n-r)$ parameters $x_{r+1}, \ldots, x_n$.

Finally, yet another way to describe the set of solutions: set $x_{r+1} = \cdots = x_n = 0$ and solve the corresponding Cramer system $\mathcal{S}'$ with Lemma 15.6. This gives a solution of $\mathcal{S}'$, z_0. We have $S = z_0 + \mathrm{Ker}(A')$.

In practice the use of the Lemma 15.6 is not necessarily the best way to proceed. The same applies to the relation $\mathrm{rk}(A) = \mathrm{rk}(A|B)$ to check the compatibility of the system. In fact, if the system is simple enough, the best thing to do is to solve by substitution. Calculating $\mathrm{rk}(A)$ and then $\mathrm{rk}(A|B)$ can be quite cumbersome and if we then find that the system is compatible we will still have to solve it.

Determinants are useful because they show us that, theoretically, we can calculate (describe) the set of solutions with a well-established procedure that we can, for example, implement on the computer. Furthermore, in certain general situations they can be useful for exploiting any particular properties of the system such as symmetries.

Finally, there is yet another method for solving a linear system: *the Gauss pivot method*. It is a systematization of the resolution method by substitution, which like any systematization, has its advantages and disadvantages.

To conclude we summarize what we have seen so far in the following theorem (often attributed to Rouché-Capelli):

Theorem 15.7. *Let $AX = B$ be a linear system of p equations in n unknowns, of rank r.*

(1) *The system is compatible if and only if $r = \mathrm{rk}(A) = \mathrm{rk}(A|B)$.*

(2) *If the system is compatible, the set of solutions, S, is an affine subspace of dimension $n - r$, more precisely $S = z_0 + \mathrm{Ker}(A)$, where z_0 is a particular solution of the system.*

(3) *Every compatible system can be traced back to a Cramer system.*

Example 15.8. Let us consider the linear system

$$\begin{cases} 3x + 2y + 2z + 3t = 1; \\ 3y + 2z + t = 5; \\ 3x + 11y + 8z + 6t = 16. \end{cases}$$

The minor constructed over the first two rows and columns is

$$\Delta_2 = \begin{vmatrix} 3 & 2 \\ 0 & 3 \end{vmatrix} = 9.$$

So the rank of the system is at least 2. We calculate the rank of

$$A = \begin{pmatrix} 3 & 2 & 2 & 3 \\ 0 & 3 & 2 & 1 \\ 3 & 11 & 8 & 6 \end{pmatrix}$$

with the edged method, considering the minor Δ_2. The two edges are zero, so $\mathrm{rk}(A) = 2$. Then it remains to compute the rank of

$$(A|B) = \begin{pmatrix} 3 & 2 & 2 & 3 & 1 \\ 0 & 3 & 2 & 1 & 5 \\ 3 & 11 & 8 & 6 & 16 \end{pmatrix}.$$

We have already seen that the two edges of Δ_2 made with the third and fourth columns of A are zero, so all that remains is to calculate:

$$\begin{vmatrix} 3 & 2 & 1 \\ 0 & 3 & 5 \\ 3 & 11 & 16 \end{vmatrix}.$$

This determinant is zero, so the system is compatible, of rank two.

It follows that the system is equivalent to the Cramer system:

$$\begin{cases} 3x + 2y = 1 - 2z - 3t; \\ 3y = 5 - 2z - t. \end{cases}$$

Applying Lemma 15.6 we get that the solutions of this system are

$$x = \frac{\begin{vmatrix} 1-2z-3t & 2 \\ 5-2z-t & 3 \end{vmatrix}}{9}, \quad y = \frac{\begin{vmatrix} 3 & 1-2z-3t \\ 0 & 5-2z-t \end{vmatrix}}{9}.$$

So $x = \frac{-2z-7t-7}{9}, y = \frac{15-6z-3t}{9}$ and

$$S = \left\{ \left(\frac{-2z-7t-7}{9}, \frac{5-2z-t}{3}, z, t \right) \mid z, t \in \mathbb{R} \right\}.$$

Looking carefully at the system (or the matrix $(A|B)$) we can see that the third equation is the first plus three times the second (or look at the rows of $(A|B)$). So the system is compatible and of rank two. From the second equation we obtain $t = -3y - 2z + 5$; then doing the twice the first plus twice the second minus the third we obtain $3x - y + 2t = -4$, i.e. $x = \frac{-7y-4z+6}{3}$. we conclude that

$$S = \left\{ \left(\frac{-7y-4z+6}{3}, y, z, -3y-2z+5 \right) \mid y, z \in \mathbb{R} \right\}.$$

Example 15.9. A situation in which determinants are useful. Let us consider, in $\mathbb{C}$, the system with parameter $\lambda \in \mathbb{C}$:

$$\mathcal{S}(\lambda) = \begin{cases} \lambda x + y - 3\lambda z = 1; \\ x - \lambda y + 2\lambda z = 2. \end{cases}$$

The minor constructed with the first two columns and rows is

$$D_1 = \begin{vmatrix} \lambda & 1 \\ 1 & -\lambda \end{vmatrix} = -\lambda^2 - 1.$$

We have $D_1 = 0 \Leftrightarrow \lambda = \pm i$. The minor constructed with the first and third columns is

$$D_2 = \begin{vmatrix} \lambda & -3\lambda \\ 1 & 2\lambda \end{vmatrix} = \lambda(2\lambda + 3)$$

and $D_2 = 0 \Leftrightarrow \lambda = 0$ or $\lambda = -\frac{3}{2}$. So for every $\lambda \in \mathbb{C}$ there exists a non-zero minor of order two, we therefore have a system of two equations, three unknowns, of rank two and we conclude that the set of solutions S has dimension one.

Example 15.10. The pivot method. The method is based on the following fact, let us consider the system

$$\mathcal{S} = \begin{cases} \varphi_1(x) = b_1; \\ \vdots \\ \varphi_i(x) = b_i; \\ \vdots \\ \varphi_p(x) = b_p. \end{cases}$$

We replace the i-th equation $\varphi_i(x) = b_i$ with a linear combination of all equations: $\sum_j \lambda_j \varphi_j(x) = \sum_j \lambda_j b_j$. Then, if $\lambda_i \neq 0$, the system $\mathcal{S}'$ thus obtained is equivalent to $\mathcal{S}$ (they have the same set of solutions). It is clear that a solution of $\mathcal{S}$ is also a solution of $\mathcal{S}'$. Conversely, if x_0 is a solution of $\mathcal{S}'$ then x_0 verifies the equations of $\mathcal{S}$ except perhaps the i-th one. From the relation $\sum_l \lambda_j \varphi_j(x_0) = \sum_j \lambda_j b_j$, given that $\varphi_k(x_0) = b_k$ if $k \neq i$, it follows $\lambda_i \ fG_i(x_0) = \lambda_i b_i$, dividing by $\lambda_i \neq 0$ we obtain $\varphi_i(x_0) = b_i$ (if $\lambda_i = 0$ we can not conclude anything).

Having said this, it is now a question of putting the matrix of the system in triangular form. Let us see how it works on an example. Consider the system

$$\mathcal{S} = \begin{cases} x + 2y + 2z = 2; \\ x + 3y - 2z = -1; \\ 3x + 5y + 8z = 8. \end{cases}$$

We keep the first equation (the pivot) which we will need to eliminate x from the other equations. Then we replace the second equation with the second minus the first and replace the third with the third minus three times the second, we obtain:

$$\mathcal{S} = \begin{cases} x + 2y + 2z = 2; \\ y - 4z = -3; \\ -y + 2z = 2. \end{cases}$$

Now, we use the second equation to eliminate y from the third, replacing the third with the third plus the second:

$$\mathcal{S} = \begin{cases} x + 2y + 2z = 2; \\ y - 4z = -3; \\ -2z = -1. \end{cases}$$

The last equation gives us $z = \frac{1}{2}$. Then by the second equation $y = -1$, and finally by the first equation $x = 3$.

Exercises

Exercise 15.11. Let $\mathcal{S}(\lambda)$ be the following linear system in four real unknowns, with parameter $\lambda \in \mathbb{R}$:

$$\begin{cases} x + \lambda y - z + t = \lambda; \\ 2x + 2y - \lambda z + 2\lambda t = 1; \\ -x + y - z + \lambda t = \lambda. \end{cases}$$

Without solving it, show that $\mathcal{S}(\lambda)$ is compatible for every $\lambda \in \mathbb{R}$.

Exercise 15.12. Let $\mathcal{S}(\lambda)$ be the following linear system in four real unknowns, with parameter $\lambda \in \mathbb{R}$:

$$\begin{cases} x + \lambda y - z - t = 1; \\ -x - y + \lambda z = \lambda; \\ 2x + y - \lambda z + t = -1. \end{cases}$$

(1) Show, without solving it, that the linear system $\mathcal{S}(\lambda)$ is compatible for every value of the parameter $\lambda \in \mathbb{R}$.

(2) Determine the set of solutions when $\lambda = 1$.

Exercise 15.13. Consider for λ a real parameter the system:

$$\mathcal{S}(\lambda) = \begin{cases} \lambda x + y - z + t = 1; \\ x - y + z - t = \lambda; \\ 2x + \lambda z = 2. \end{cases}$$

Let $S(\lambda)$ be the set of solutions of $\mathcal{S}(\lambda)$.

(1) Without solving the system, determine $E = \{\lambda \in \mathbb{R} \mid \dim S(\lambda) = 2\}$.

(2) Solve the system for $\lambda = 1$ and determine $\dim S(1)$.

Exercise 15.14. Consider the system

$$\mathcal{S}(\lambda) = \begin{cases} \lambda x - y + 2\lambda z = 1; \\ x - 2\lambda y + 3z = \lambda; \\ 2\lambda x + y + \lambda z = 2; \end{cases}$$

with $\lambda \in \mathbb{R}$.

(1) Determine $R = \{\lambda \mid \mathcal{S}(\lambda) \text{ is compatible}\}$.

(2) Solve $\mathcal{S}(1)$.

Exercise 15.15. Consider the system

$$\mathcal{S}(\lambda) = \begin{cases} \lambda x + y - 3\lambda z = 1; \\ x - \lambda y + 2\lambda z = 2; \end{cases}$$

with $\lambda \in \mathbb{C}$ and $(x, y, z) \in \mathbb{C}^3$. Without solving the system show that it is compatible for any $\lambda \in \mathbb{C}$.

Now, consider the system

$$\mathcal{S}'(\lambda) = \begin{cases} \lambda x + y - 3\lambda z = 1; \\ x - \lambda y + 2\lambda z = 2; \\ \alpha x + \beta y + \gamma z = \delta; \end{cases}$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. Do there exist $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ such that
(i) $\mathcal{S}'(\lambda)$ is compatible for any $\lambda \in \mathbb{C}$?
(ii) $\mathcal{S}'(\lambda)$ is compatible for any $\lambda \in \mathbb{R}$?

Exercise 15.16. Consider the linear system $AX = B$ where $A \in M_{m,n}(\mathbb{R})$, $A \neq 0$, and where

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m.$$

Let $r = \mathrm{rk}(A)$. Tell whether the following statements are true or false, justifying the answer.

(1) The system admits solutions for every $B \in \mathbb{R}^m$ if and only if $r = m$.

(2) The system admits one and only one solution if and only if it is a Cramer system (i.e. $n = m = r$).

(3) Let $n = m$. If there exists $B_1 \in \mathbb{R}^m$ such that the system admits more than one solution, then there exists $B_2 \in \mathbb{R}^m$ such that the system is incompatible.

(4) If there exists $B_1 \in \mathbb{R}^m$ such that the system admits infinite solutions, then $\forall B \in \mathbb{R}^m$ the system admits infinite solutions.

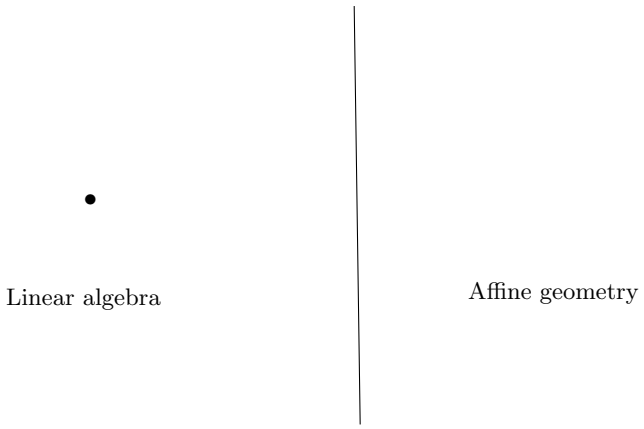
(5) If the system admits at most one solution for every $B \in \mathbb{R}^m$ then $r = n$.

(6) If $n = m$ and if there exists $B_1 \in \mathbb{R}^m$ such that the system is incompatible, then there exists $B \in \mathbb{R}^m$ such that the system admits infinite solutions.

Part 3

Affine Geometry

The main difference between linear algebra and affine geometry can be represented with the following picture:



In linear algebra there is a special element: the zero vector (the origin). Affine geometry is more democratic: all points are equal. This normalization is achieved through translations. If E is a vector space and $a \in E$, the translation with respect to the vector a is the map $t_a : E \to E : v \to a + v$. In particular, $t_a(0) = a$: the origin is moved to the point a. Every point can be the origin, so being the origin is no longer a privilege.

There are various ways of presenting affine geometry. For example, we can say that an affine space is a triple (X, E, φ), where X is a set (the set of points), E is a K-vector space (the set of vectors) and φ is a *simply transitive action* of the group $(E, +)$ on X. This is how they taught it to us and we understood very little! Therefore we will follow another path and start studying the affine subspaces of a vector space. It is much simpler, more intuitive and that's what we need in practice. It is a particular case of the general situation ($X = E$ and the action of $(E, +)$ on $X = E$ is given by the translations). We will then see that this particular case is equivalent to the general case. But everything has pros and cons, there is a price to pay. In fact we will have to consider an element of E both as a point and as a vector ($X = E$). This can be confusing, especially at first, but the context will help us make the necessary differences.

Chapter 16

Affine subspaces of a vector space

Let us formalize what we have already said.

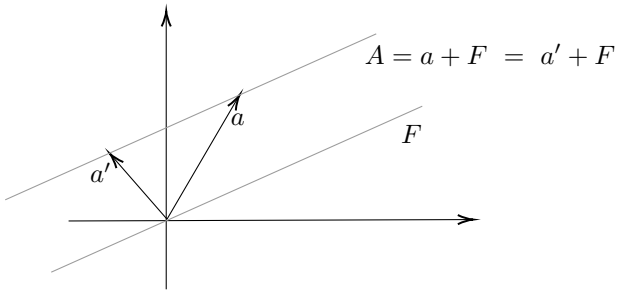
Definition 16.1. Let E be a K-vector space and $a \in E$, the translation with respect to the vector a is the map $t_a : E \to E : v \to a + v$.

Remark 16.2. The translation is not (if $a \neq 0$) a linear map: $t_a(0) = a \neq 0$. A translation is bijective ($(t_a)^{-1} = t_{-a}$). Obviously $t_0 = Id_E$.

Definition 16.3. Let E be a K-vector space. Any subset of E constructed by translation from a subvector pace of E is called an affine subspace of E. In other words $A \subset E$ is an affine subspace of E if there exist $a \in E$ and a subvector space $F \subset E$ such that $A = t_a(F) = \{x \in E \mid \exists v \in F, x = a + v\}$.

Remark 16.4. We will write $A = a + F = \{a + v \mid v \in F\}$.

The following picture shows that A is uniquely determined by the translation vector a.



Moreover we see that A uniquely determines the sub-vector space F.

Lemma 16.5. *Let E be a K-vector space, F, F' two subvector spaces of E, and a, a' two elements of E. If $a + F = a' + F'$ then $F = F'$.*

Proof. We observe that $a = a + 0 \in a + F$. So $a \in a' + F'$, that is, $a = a' + f'$, with $f' \in F'$. It follows that $a - a' \in F'$. Let $f \in F$, then $a + f \in a' + F'$, that is, $a + f = a' + v'$, with $v' \in F'$. So $f = (a' - a) + v'$. Since $a' - a = -(a - a') \in F'$, we have $f \in F'$. This shows that $F \subset F'$. Similarly, we get that $F' \subset F$. □

This shows that in the writing $A = a + F$ the subvector space F is uniquely determined.

Definition 16.6. Let $A = a + F$ be an affine subspace of the vector space E. The subspace F is uniquely determined by A and is called the *direction* of A. We will write $F = \mathrm{dir}(A)$. The dimension of A is the dimension of F.

If F is a line, a plane, a hyperplane we say that A is a line, a plane, a hyperplane (affine). A 0-dimensional affine subspace is a *point* (it is also a vector, an element of E). Obviously every subvector space of E is also an affine subspace ($F = t_0(F)$).

Coming back to the equation $A = a + F$, we have already said that the point a is not uniquely determined, more precisely we have:

Lemma 16.7. *Let $A = a + F$ be an affine subspace of the vector space E. For every point $b \in A$ we have*

(1) $b - a \in F$;

(2) $A = b + F$.

Proof. 1) Since $b \in A$, $b = a + f$, with $f \in F$. Therefore $b - a = f \in F$.

2) Let $a + f$ be an element of A. Then $a + f = b + (a - b) + f \in b + F$. This shows $a + F \subset b + F$. Vice versa $b + f = a + (b - a) + f \in a + F$. So $A = a + F = b + F$. □

Remark 16.8. The first point of Lemma 16.7 says that given two points of an affine subspace A, their difference belongs to the direction of A. This is a useful and important fact.

The intersection of two affine subspaces is not always an affine subspace. In fact the intersection can be empty: let $A = a + F$ and $b \notin A$. If $B = b + F$, then $A \cap B = \emptyset$. If $w \in A \cap B$, then $w = a + f = b + g$, with $f, g \in F$. It follows that $b = a + (f - g) \in A$, against our hypothesis $b \notin A$. More generally if $F' \subset F$ and if $B = b + F'$, $A \cap B = \emptyset$. As we will see later, A and B are *parallel*.

Lemma 16.9. *Let $(A_i)_{i\in I}$ be a family of affine subspaces of E and $A = \bigcap_{i\in I} A_i$. If $A \neq \emptyset$, then A is an affine subspace of E of direction $\bigcap_{i\in I} F_i$, where $F_i = \mathrm{dir}(A_i)$, $\forall i \in I$.*

Proof. Suppose $A \neq \emptyset$ and take $a \in A$, then $a \in A_i$, $\forall i \in I$. So we can write (Lemma 16.7) $A_i = a + F_i = t_a(F_i)$. We have $A = \bigcap_{i\in I} t_a(F_i) = t_a(\bigcap_{i\in I} F_i)$ (because t_a is one-to-one). So $A = a + \bigcap_{i\in I} F_i$ is an affine subspace. □

Definition 16.10. Let E be a K-vector space and $v_1, \ldots, v_k \in E$. The affine subspace generated by $v_1, \ldots, v_k$ is the smallest affine subspace of E containing $v_1, \ldots, v_k$. In other words it is the intersection of all the affine subspaces containing $v_1, \ldots, v_k$. We will denote it by $[v_1, \ldots, v_k]$.

Remark 16.11. Do not confuse $\langle v_1, \ldots, v_k\rangle$ (the subvector space generated by $v_1, \ldots, v_k$) with $[v_1, \ldots, v_k]$. For example $\langle v\rangle = \{\lambda v \mid \lambda \in K\}$, while $[v] = \{v\}$.

Lemma 16.12. *With the previous notations we have*

$$[v_1, \ldots, v_k] = v_i + \langle v_1 - v_i, \ldots v_{i-1} - v_i, v_{i+1} - v_i, \ldots, v_k - v_i\rangle$$

for every $i, 1 \leq i \leq k$.

Proof. We set $A_i = v_i + \langle v_1 - v_i, \ldots v_{i-1} - v_i, v_{i+1} - v_i, \ldots, v_k - v_i\rangle$. We have $v_i \in A_i$ and if $j \neq i$, $v_j = v_i + (v_j - v_i) \in A_i$. So A_i contains $v_1, \ldots, v_k$. Let A be an affine subspace containing $v_1, \ldots, v_k$. We can write $A = v_i + F$ (Lemma 16.7). Since $v_j \in A$, we have $v_j = v_i + f$, $f \in F$. So $v_j - v_i = f \in F$. This shows $\langle v_1 - v_i, \ldots v_{i-1} - v_i, v_{i+1} - v_i, \ldots, v_k - v_i\rangle \subset F$, then $A_i \subset A$. We conclude that A_i is the smallest affine subspace containing $v_1, \ldots, v_k$. □

Definition 16.13. The points $v_1, \ldots, v_k$ are said to be *affinely independent* if $\dim[v_1, \ldots, v_k] = k - 1$ (the maximum possible dimension). The points $v_1, \ldots, v_k$ are said to be *aligned* (respectively *coplanar*) if $\dim[v_1, \ldots, v_k] = 1$ (respectively $\dim[v_1, \ldots, v_k] = 2$).

Two points v_0, v_1 are affinely independent if and only if they are distinct. In this case $[v_0, v_1]$ is a line (the only line passing through v_0 and v_1). Three points v_0, v_1, v_2 are affinely independent if and only if they generate a plane (i.e. if they are not aligned). Four points are affinely independent if and only if they are not coplanar and so on.

16.1. Affine maps, affinity group

Let us now move on to the definition of affine map.

Definition 16.14. Let E, F be two K-vector spaces. A map $\psi : E \to F$ is an affine map if there exist a linear map $v : E \to F$ and a translation t_a ($a \in F$) such that

$$\psi = t_a \circ v.$$

Pay attention to the order, the linear map is applied first and then the translation.

Remark 16.15. An affine map $\psi : K \to K$ is of the form $\psi(x) = ax + b$. Its graph is a line that does not pass through the origin if $b \neq 0$.

An affine map is linear if and only if $a = 0$. In particular, every linear map is an affine map.

A translation $t_a : E \to E$ is an affine map (just take $F = E$ and $v = Id_E$ in the previous definition).

Lemma 16.16. *With the above notations, the decomposition of ψ as a composition of a linear map and a translation is unique i.e. v and t_a are uniquely determined by ψ.*

Proof. Suppose $\psi(x) = v(x)+a = v'(x)+a'$. We have $\psi(0) = a = a'$. Furthermore $v'(x) = \psi(x) - a' = \psi(x) - a = v(x)$. So $v = v'$. □

Definition 16.17. With the previous notations the linear map v is called *the linear part* of the affine map ψ and we denote it by $v = L(\psi)$.

Proposition 16.18. *Let E, F, G be K-vector spaces and $\psi : E \to F$, $\varphi : F \to G$ affine maps. The map $\varphi \circ \psi : E \to G$ is an affine map. Furthermore $L(\varphi \circ \psi) = L(\varphi) \circ L(\psi)$.*

Proof. Let $\psi = t_a \circ v$ and $\varphi = t_b \circ u$. Then $\varphi(\psi(x)) = u(v(x) + a) + b = u(v(x)) + (u(a) + b)$. Therefore $\varphi \circ \psi = t_{u(a)+b} \circ (u \circ v)$. □

Proposition 16.19. *Let E, F be two K-vector spaces and $\psi : E \to F$ an affine map with $L(\psi) = v$. If $A \subset E$ is an affine subspace then $\psi(A)$ is an affine subspace of F with direction $v(F)$ where $F = \dir(A)$.*

Proof. Let $A = a+F$. If $x \in A$, $x = a+f$, $f \in F$ and $\psi(x) = v(a+f)+b$ ($\psi = t_b \circ v$). So $\psi(x) = v(a) + b + v(f)$, and $\psi(A)$ is the affine subspace $v(a) + b + v(F)$. □

An useful observation:

Lemma 16.20. *An affine map ψ is injective (surjective, bijective) if and only if its linear part $L(\psi)$ is.*

Proof. We have $\psi = t_a \circ L(\psi)$ and the claim follows from the fact that the translation t_a is bijective. □

We now introduce an important class of affine maps: affine, bijective maps from E to itself. Such a map is called an *affinity* of E.

Theorem 16.21. *Let E be a K-vector space. The set of affine, bijective maps from E to E is a group for the composition of maps. This group is known as $G_A(E)$ and is called the affinity group of E.*

Proof. If ψ, φ are two affinities of E, then $\psi \circ \varphi$ is also an affinity of E. The identity Id_E is an affinity. The composition of maps is associative. It only remains to show that if ψ is an affinity then ψ^{-1} is also an affinity. Let $\psi = t_a \circ v$. Clearly (Lemma 16.20) v is a linear isomorphism. Let $\psi(x) = v(x) + a =: z$. We want $\psi^{-1}(z) = x$ If we apply v^{-1} to z (forced choice by Proposition 16.18) we get $x + v^{-1}(a)$. We conclude that $\psi^{-1} = t_{-v^{-1}(a)} \circ v^{-1}$. □

Remark 16.22. From a certain point of view, affine geometry is the study of geometric properties invariant under $G_A(E)$, more precisely:

- two subsets, X, Y, of E are *affinely equivalent* if there is an affinity of E that sends X to Y;
- an affine property of a subset $X \subset E$ is a property common to all subsets affinely equivalent to X.

Proposition 16.23. *Two affine subspaces A, A' of E are affinely equivalent if and only if* $\dim(A) = \dim(A')$.

Proof. Let $A = a + F$, $A' = a' + F'$ and suppose $\dim(A) = \dim(A')$, i.e. $\dim(F) = \dim(F')$. There exists a linear isomorphism $v : E \to E$ such that $v(F) = F'$. In fact, let $(e_1, \ldots, e_r)$, $(e'_1, \ldots, e'_r)$ be bases of F, F'. Let us complete them to bases of E: $\mathcal{B} = (e_1, \ldots, e_n)$, $\mathcal{B}' = (e'_1, \ldots, e'_n)$. We define v by $v(e_i) = e'_i$. Then v is an

isomorphism (transforms the basis $\mathcal{B}$ into the basis $\mathcal{B}'$) and $v(F) = F'$. We have $v(A) = v(a) + F'$. So $\psi = t_b \circ v$, with $b = a' - v(a)$ and $\psi(A) = A'$. Hence A and A' are affinely equivalent.

Now, let ψ be an affinity such that $\psi(A) = A'$. We know (Proposition 16.19) that $\psi(A)$ has as its direction $L(\psi)(\dir A)$. So $L(\psi)(\dir A) = \dir A'$. Since $L(\psi)$ is an isomorphism, the two directions have the same dimension, i.e. $\dim(A) = \dim(A')$. □

Remark 16.24. 1) The dimension is therefore an affine property of affine subspaces.
2) We can identify two notable subgroups of $G_A(E)$:

- the linear group $Gl(E)$ of the linear automorphisms of E ($\psi = t_0 \circ v$ with $v \in Gl(E)$ is an affinity);
- the group $\mathcal{T}(E)$ of the translations of E ($\psi = t_a \circ Id_E$ is an affinity).

Furthermore, the map $L : G_A(E) \to Gl(E) : \psi \to L(\psi)$ is a morphism of groups.

Exercises

Exercise 16.25. Let E, F be two K-vector spaces, $A \subset F$ an affine subspace and $u : E \to F$ an affine map. Prove that, if not empty, $u^{-1}(A)$ is an affine subspace of E.

Exercise 16.26. Let E be a K-vector space and v, w two vectors of E, with $v \neq 0$.
1) Show that there exists $\varphi \in \mathrm{End}(E)$ such that $\varphi(v) = w$.
2) Let $A = \{f \in \mathrm{End}(E) \mid f(v) = w\}$. Show that A is an affine subspace of $\mathrm{End}(E)$, and compute $\dim(A)$.

Exercise 16.27. Let E be a K-vector space and $g : E \to E$ an affine map. A point $p \in E$ is said to be fixed for g if $g(p) = p$.
1) Let $f : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \to (x + y - z + 2, -x + y + z - 2, x - y + z)$. Prove that f is an affinity and determine the set of its fixed points.
2) Prove that the set of the fixed points of g is an affine subspace of E.
3) Prove that g has a unique fixed point if and only if 1 is not an eigenvalue of the linear part of g.

Exercise 16.28. Let E be a K-vector space and $G_A(E)$ the affinity group of E. Let $T \subset G_A(E)$ be the subgroup of translations; furthermore, if $x \in E$ we set $F_x = \{f \in G_A(E) \mid f(x) = x\}$.
1) Prove that F_x is a subgroup of $G_A(E)$.
2) Show that $F_x \cap T = \{Id\}$ and: $\forall\, t \in T$, $\forall\, g \in G_A(E)$, $g^{-1}\ circt \circ g \in T$.
3) Show that $\forall\, f \in F_x$, $\forall\, g \in G_A(E)$, $g^{-1} \circ f \circ g$ has a fixed point.

Chapter 17

Equations of affine subspaces

17.1. Cartesian equations

The first result of this section highlights the correspondence between the geometric theory of affine subspaces and the algebraic theory of linear systems. This correspondence will become more evident in the following.

Proposition 17.1. *Let E be a K-vector space of dimension n and $A \subset E$ an affine subspace of dimension p. Then there exists a linear system of $n-p$ equations in n unknowns and of rank $n-p$ whose solution set is A.*

Proof. Let $A = a + F$ with $\dim F = p$. We know that F is defined by $n-p$ independent linear equations. We quickly remember how it works: let $F^\circ \subset E^*$, $F^\circ = \{\varphi \in E^* \mid F \subset \mathrm{Ker}(\varphi)\}$. Then $\dim F^\circ = n-p$ and if $(\varphi_1, \ldots, \varphi_{n-p})$ is a basis of F°, F is the set of solutions of the homogeneous linear system

$$\begin{cases} \varphi_1(v) = 0; \\ \vdots \\ \varphi_{n-p}(v) = 0. \end{cases}$$

Let us set $\varphi_i(a) = b_i$, then a is a solution of the linear system

$$\mathcal{S} = \begin{cases} \varphi_1(v) = b_1; \\ \vdots \\ \varphi_{n-p}(v) = b_{n-p}. \end{cases}$$

Since the set of solutions of the linear system $\mathcal{S}$ is obtained by adding to a particular solution the set of solutions of the associated homogeneous system, we conclude that the set of solutions of the system $\mathcal{S}$ is $a + F = A$. □

Remark 17.2. The equations of the system $\mathcal{S}$ are *Cartesian equations* of the affine subspace A. As in the vector case these equations are not uniquely determined, they depend on the choice of a basis of F° and on the choice of the point a. Furthermore, linear combinations of the equations $\varphi_i = b_i$ can be added to the system, as we know, this will not change the set of solutions.

Example 17.3. Let in $\mathbb{R}^3$ be the affine line $R = a + \langle v \rangle$, where $a = (1, 1, 0)$ and where $v = (1, -1, 2)$. We are looking for Cartesian equations of R. Since $\dim(R) = 1$ we are looking for $3 - 1 = 2$ equations. We follow the procedure described above. The orthogonal of $\langle v \rangle$ is generated by two linear forms that vanish on v, for example $\varphi_1(x, y, z) = x + y, \varphi_2(x, y, z) = 2x - z$. These two forms are independent so $\langle v \rangle^\circ = \langle \varphi_1, \varphi_2 \rangle$. We have $\varphi_1(a) = 2, \varphi_2(a) = 2$. The system that defines R is

$$\begin{cases} x + y = 2; \\ 2x - z = 2. \end{cases}$$

We will often use the following result:

Lemma 17.4. *Let $\mathcal{S} := \{\varphi_1(v) = b_1, \ldots, \varphi_r(v) = b_r\}$ be the Cartesian equations of the affine subspace A, and $\mathcal{S}' := \{\psi_1(v) = c_1, \ldots, \psi_p(v) = c_p\}$ the Cartesian equations of the affine subspace A'. Then $A \cap A'$ is the (possibly empty) set of solutions of the linear system $\mathcal{S} + \mathcal{S}' := \{\varphi_1(v) = b_1, \ldots, \varphi_r(v) = b_r, \psi_1(v) = c_1, \ldots, \psi_p(v) = c_p\}$.*

Proof. Clear. □

So in the previous example we can say that the line R is the intersection of the affine plane, H_1, of equation $x + y - 2 = 0$ with the affine plane, H_2, of equation $2x - z - 2 = 0$.

17.2. Parametric representation of affine subspaces

Let E be a K-vector space of dimension n and $A = a + F$ an affine subspace of E, of dimension p. An element $v \in E$ belongs to A if $v = a + f$, with $f \in F$. Let $(f_1, \ldots, f_p)$ be a basis of F, then $v \in A \Leftrightarrow \exists\, (\lambda_1, \ldots, \lambda_p) \in K^p$ such that $v = a + \lambda_1 f_1 + \cdots + \lambda_p f_p$. So we have

$$A = \{a + \lambda_1 f_1 + \cdots + \lambda_p f_p \mid (\lambda_1, \ldots, \lambda_p) \in K^p\}. \tag{17.1}$$

This is a *parametric representation* of A, the λ_i are the *parameters*, they are equal in number to $\dim(A) = p$. Obviously this representation depends on the choice of a basis of F.

Let $(e_1, \dots, e_n)$ be a basis of E. We have $a = \sum_i \alpha_i e_i$, $f_j = \sum_i \beta_{ij} e_i$, then A is the set of $v \in E$ whose coordinates $(x_1, \dots, x_n)$ are of the form

$$\text{(17.2)} \qquad \begin{cases} x_1 = \alpha_1 + \lambda_1\beta_{11} + \cdots + \lambda_p\beta_{1p}; \\ x_2 = \alpha_2 + \lambda_1\beta_{21} + \cdots + \lambda_p\beta_{2p}; \\ \vdots \\ x_n = \alpha_n + \lambda_1\beta_{n1} + \cdots + \lambda_p\beta_{np}. \end{cases}$$

We just rewrote (17.1) using the coordinates of $a, f_1, \dots, f_p$. This parametric representation also depends on the choice of a basis of E.

Example 17.5. Let us take the line from Example 17.3. We have $w \in R \Leftrightarrow w = \lambda v$, so (taking the canonical basis of $\mathbb{R}^3$) a parametric representation of R is

$$\begin{cases} x = 1 + \lambda; \\ y = 1 - \lambda; \\ z = 2\lambda. \end{cases}$$

Let us now look for a parametric representation of the plane H_1 with Cartesian equation $x + y - 2 = 0$. Consider the point $a = (1, 1, 0) \in H_1$. The direction of H_1 is given by the associated homogeneous equation

$$x + y = 0.$$

The solution set is a subvector space of dimension two and basis $((0, 0, 1), (1, -1, 0))$. A parametric representation is given by

$$\begin{cases} x = 1 + \mu; \\ y = 1 - \mu; \\ z = \lambda. \end{cases}$$

Remark 17.6. Let $A \subset E$ be an affine subspace of dimension p of a vector space of dimension n. A parametric representation of A is given by n equations with p parameters. A Cartesian representation is given by $n-p$ equations (in n unknowns).

Be careful not to confuse the number of parameters with the number of Cartesian equations.

17.3. Hyperplanes and lines: cartesian, parametric, round trip

Let H be a hyperplane of the K-vector space E. In Cartesian form H is defined by an equation $\varphi(v) = a$. To find a parametric representation of H we need to find a basis of $\mathrm{Ker}(\varphi)$ and a particular solution of the system $\varphi(v) = a$.

After choosing a basis E is identified as K^n and the equation $\varphi(v) = a$ will be written as $a_1x_1 + \cdots + a_nx_n = a$. To find a basis of

$$\mathrm{Ker}(\varphi) = \{(x_1, \dots, x_n) \mid a_1x_1 + \cdots + a_nx_n = 0\}$$

we give values to the parameters. For example we can proceed like this: $\exists\, i$ such that $a_i \neq 0$. For each j, $j \neq i$, we set

$$w_j = \left(0, \ldots, 0, 1, 0, \ldots, -\frac{a_j}{a_i}, 0, \ldots, 0\right)$$

with 1 in place j and $-\frac{a_j}{a_i}$ in place i. It is easily verified that these $n-1$ vectors form a basis of $\mathrm{Ker}(\varphi)$.

A particular solution of $\varphi(v) = a$ is given, for example, by

$$p = \left(0, \ldots, 0, 1, 0, \ldots, \frac{a - a_j}{a_i}, 0, \ldots, 0\right).$$

Example 17.7. In $\mathbb{R}^3$ consider the plane of equation $x - 2y + 3z = -1$. We apply the previous procedure with $i = 1$. The vectors $(2, 1, 0), (-3, 0, 1)$ form a basis of the direction and $(1, 1, 0)$ is a point on the plane. A parametric representation is given by

$$\begin{cases} x = 1 + 2\lambda - 3\mu; \\ y = 1 + \lambda; \\ z = \mu. \end{cases}$$

Now suppose we have a parametric representation of the hyperplane: $H = a + \langle h_1, \ldots, h_{n-1}\rangle$. We want to derive a Cartesian equation of H.

Note that $x \in H \Leftrightarrow x - a \in \langle h_1, \ldots, h_{n-1}\rangle$ and this is equivalent to saying that the $n = \dim E$ vectors $x - a, h_1, \ldots, h_{n-1}$ are dependent which translates into $\det(x - a, h_1, \ldots, h_{n-1}) = 0$ and this is our Cartesian equation.

In practice, however, we can do it differently by trying to eliminate the parameters. Let us consider an example.

Example 17.8. Let $H \subset \mathbb{R}^3$ be the hyperplane given by

$$\begin{cases} x = 1 + \lambda - \mu; \\ y = 2 + 2\lambda + \mu; \\ z = -1 - \lambda + 2\mu. \end{cases}$$

Adding the first and third equations we get $\mu = x + z$. Plugging this value of μ into the first equation yields $\lambda = 2x + z - 1$. Substituting these expressions of λ, μ into the second equation gives $5x - y + 3z = 0$, a Cartesian equation of H.

The other approach consists in calculating the following determinant:

$$\begin{vmatrix} x - 1 & 1 & -1 \\ y - 2 & 2 & 1 \\ z + 1 & -1 & 2 \end{vmatrix}$$

which is indeed equal to $5x - y + 3z$.

Now, let $R \subset E$ be a parametric representation of the line $R = a + \langle u \rangle$. We have $w \in R \Leftrightarrow w - a$ and u are dependent. After choosing a basis and setting $w = (x_1, \ldots, x_n), a = (a_1, \ldots, a_n)$ this translates into saying that the matrix

$$M = \begin{pmatrix} x_1 - a_1 & x_2 - a_2 & \cdots & x_n - a_n \\ u_1 & u_2 & \cdots & u_n \end{pmatrix}$$

has rank one, where after choosing a basis the u_i are coordinates on E. Since there exists i such that $u_i \neq 0$, to say that M has rank one is equivalent to say that the $n-1$ edges of u_i are zero. This gives $n-1$ Cartesian equations of the line R.

In practice we will try to eliminate the parameter. For example, if $R \subset \mathbb{R}^3$ is the line of equations $x = 1 + 2\lambda$, $y = 3 - \lambda$, $z = 2 + 3\lambda$ then $\lambda = 3 - y$ from which we derive the Cartesian equations $x = 7 - 2y, z = 11 - 3y$.

Now, suppose we have our line R given by $n-1$ Cartesian equations $\varphi_1(v) = \beta_1$, $\ldots$,$\varphi_{n-1}(v) = \beta_{n-1}$. To find the direction of R we must find a non-trivial solution of the homogeneous system $\varphi_1(v) = 0, \ldots, \varphi_{n-1}(v) = 0$.

Lemma 17.9. *Let $\mathcal{S}$ be the homogeneous system of $n-1$ equations in n unknowns*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0; \\ \vdots \\ a_{n-1,1}x_1 + \cdots + a_{n-1,n}x_n = 0; \end{cases}$$

and

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n} \end{pmatrix}$$

the $(n-1, n)$ matrix of the coefficients. Let A_i be the matrix obtained from A by removing the i-th column, and $\Delta_i = (-1)^{i+1} \det(A_i)$. If the system has rank $n-1$ then $(\Delta_1, \ldots, \Delta_n)$ is a non-trivial solution of $\mathcal{S}$.

Proof. For $1 \leq k \leq n$ we consider the determinant of the $n \times n$ matrix obtained by adding its k-th row to A:

$$\begin{vmatrix} a_{k1} & \cdots & a_{kn} \\ a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n} \end{vmatrix}.$$

This determinant is zero because it has two equal rows. Developing according to the first line we get $a_{k1}\Delta_1 + \cdots + a_{kn}\Delta_n = 0$. So $(\Delta_1, \ldots, \Delta_n)$ is a solution of $\mathcal{S}$. Since $\mathcal{S}$ has rank $n-1$, A has a non-zero minor of order $n-1$, i.e. there exists i such that $\Delta_i \neq 0$. $\square$

This lemma allows us to find the direction of R. A point remains to be found and this will be done by giving values to the unknowns. We can also try to eliminate the unknowns by setting $x_1 = \lambda$. Let us consider an example.

Example 17.10. Let $R \subset \mathbb{R}^3$ be given by

$$\begin{cases} x + y - z = 2; \\ 2x + y + z = -4. \end{cases}$$

Setting $x = \lambda$ we obtain $y - z = 2 - \lambda, y + z = -4 - 2\lambda$. Adding up we get $2y = -2 - 3\lambda$, so $y = -1 - \frac{3}{2}\lambda$. Then $z = \lambda - 1 - \frac{3}{2}\lambda - 2 = -3 - \frac{1}{2}\lambda$. Be careful though. Setting $x = \lambda$ assumes that the first coordinate of a vector in $\operatorname{dir}(R)$ is non-zero. For example consider

$$\begin{cases} x + y - z = 2; \\ 2x - y + z = -4. \end{cases}$$

Only one sign was changed. After setting $x = \lambda$ we obtain $y - z = 2 - \lambda$, $-y + z = -4 - 2\lambda$, adding $2 = -3\lambda$ and the parameter is constant. This means that the first coordinate of a vector in $\operatorname{dir}(R)$ is zero. In fact, let $y = \mu$, then $x - z = 2 - \mu$ and $2x + z = -4 + \mu$, adding $3x = -2$, i.e. $x = -\frac{2}{3}$ and then $z = -\frac{16}{3} + \mu$. A parametric representation is

$$\begin{cases} x = -\frac{2}{3}; \\ y = \mu \\ z = -\frac{16}{3} + \mu. \end{cases}$$

The direction of R is $\langle (0, 1, 1) \rangle$. This can also be seen by solving the homogeneous system

$$\begin{cases} x + y - z = 0; \\ 2x - y + z = 0; \end{cases}$$

adding $x = 0$ and then $y = z$. To find a point we cut with a coordinate plane. In this case the line is parallel to the plane $x = 0$ so it is preferable to set $y = 0$ or $z = 0$. For $y = 0$ we get $x - z = 2$ and $2x + z = -4$, then $x = -\frac{2}{3}$ and $z = -\frac{8}{3}$. So

$$R = \left(-\frac{2}{3}, 0, -\frac{8}{3} \right) + \langle (0, 1, 1) \rangle.$$

Exercises

Exercise 17.11. In $\mathbb{R}^3$ give equations (parametric and Cartesian) of the affine subspace generated by P, Q, R where
1) $P = (0, 0, 2)$, $Q = (0, 1, 2)$, $R = (0, 2, 2)$;
2) $P = (0, 0, 2)$, $Q = (-1, 0, 2)$, $R = (0, 2, 2)$.

Exercise 17.12. In $\mathbb{R}^3$ let D be the plane of equation $x + y + 1 = 0$ and let R be the line of equations $x - 5y + 6z = 0$, $5x - y + 6z + 1 = 0$.
1) Show that $R \cap D = \emptyset$.
2) Give parametric equations of D and R.

Exercise 17.13. Let $A \subset \mathbb{R}^4$ be defined by the equations $x + y - z + t = 1$, $2x + y + 2z - t = -1$.
1) Determine the dimension of the affine subspace A.
2) Give parametric equations of A.

Exercise 17.14 (Magic matrices: the end)**.** We set

$$S = \{M \in M_3(\mathbb{R}) \mid M \text{ is symmetric}\}$$

and

$$A = \{M \in M_3(\mathbb{R}) \mid M \text{ is antisymmetric}\}.$$

Show that S, A are subvector spaces, calculate their dimension and show that $M_3(\mathbb{R}) = S \oplus A$.

Therefore each matrix is written in a unique way as the sum of a symmetric matrix and an antisymmetric matrix. Make this decomposition explicit i.e. write $M = (a_{ij}) \in M_3(\mathbb{R})$ as the sum of a symmetric matrix and an antisymmetric matrix.

Let $\mathcal{M} \subset M_3(\mathbb{R})$ be the subvector space of magic matrices (Exercise 11.30). We set $\mathcal{M}_+ = \mathcal{M} \cap S$ and $\mathcal{M}_- = \mathcal{M} \cap A$. Prove that $\mathcal{M}_+$ and $\mathcal{M}_-$ are subvector spaces of $\mathcal{M}$ and that $\mathcal{M} = \mathcal{M}_+ \oplus \mathcal{M}_-$.

It follows from Exercise 11.30 that $\mathcal{M}_- = \{\alpha A_0 \mid \alpha \in \mathbb{R}\}$ where

$$A_0 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

If $\alpha \in \mathbb{R}$ we set $\mathcal{M}_+(\alpha) = \{M \in \mathcal{M}_+ \mid s(M) = \alpha\}$. Determine $\mathcal{M}_+(0)$ by considering a suitable linear system, computing its rank and observing that

$$S_0 = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \in \mathcal{M}_+(0).$$

Show that every $\mathcal{M}_+(\alpha)$ is a line parallel to $\mathcal{M}_+(0)$.

Let U be the matrix whose coefficients are all equal to 1. Clearly $U \in \mathcal{M}_+(3)$. Show that the line $\langle U \rangle$ intersects every $\mathcal{M}_+(\alpha)$. Conclude that $\dim(\mathcal{M}_+) = 2$.

In conclusion $\mathcal{M}$ is a subvector space of dimension 3. A basis is (A_0, S_0, U). So every magic matrix is of the form

$$\begin{pmatrix} b+c & -a-b+c & a+c \\ a-b+c & c & -a+b+c \\ -a+c & a+b+c & -b+c \end{pmatrix}$$

with $a, b, c \in \mathbb{R}$. Here the sum is $3c$. Let us observe a curious fact: if $M = (a_{ij})$ is a magic matrix with sum s the central coefficient a_{22} is always equal to $\frac{s}{3}$.

Let $M \in M_3(\mathbb{R})$ be a magic matrix. We denote by R_i (resp. C_i, D_i), the sum of the elements of the i-th row (resp. of the i-th column, diagonal) of M. Since M is magic we have $R_i = R_1, 2 \leq i \leq 3$, $D_j = R_1, 1 \leq j \leq 2$, $C_k = R_1, 1 \leq k \leq 3$. A homogeneous linear system of 7 equations in 9 unknowns is thus obtained. Compute the rank of the system and conclude that $\dim(\mathcal{M}) = 3$.

Chapter 18

Parallelism and incidence

As we have seen when studying duality the main obstacle to a good notion of duality in the affine plane lies in the fact that two lines may not intersect, in this case they are parallel. Parallelism is certainly an emblematic phenomenon of affine geometry.

Definition 18.1. Two affine subspaces A, A' of a K-vector space, E, are *parallel* if $\dim(A) \subset \dim(A')$ or $\dim(A') \subset \dim(A)$; in this case we write $A \parallel A'$.

Remark 18.2. (i) A point is parallel to any affine subspace, in particular two points are parallel. In the following we will limit ourselves (tacitly) to considering spaces of positive dimension.

(ii) The parallelism relation is not an equivalence relation on the set of affine subspaces of positive dimension. In fact $A \parallel A'$ and $A' \parallel A''$ do not necessarily imply $A \parallel A''$. Take for example A' a plane and A, A'' two intersecting lines contained in A'. It is true, however, that parallelism is an equivalence relation on the set of affine subspaces of fixed positive dimension.

Lemma 18.3. *Let A, A' be two affine subspaces of the K-vector space E. If $A \subset A'$ then $A \parallel A'$.*

Proof. Let $a \in A$, then $a \in A'$ and we can write (Lemma 16.7) $A = a + F, A' = a + F'$. If $f \in F$, $a + f \in a + F'$. Hence $f \in F'$, therefore $F \subset F'$ and $A \parallel A'$. $\square$

Lemma 18.4. *Let A, A' be two parallel affine subspaces of the K-vector space E, with $\dim(A) \leq \dim(A')$.*
1) If A and A' have a point in common then $A \subset A'$.
2) If $\dim(A) = \dim(A')$ and if A and A' have a point in common then $A = A'$.

Proof. 1) Let $a \in A \cap A'$. We know (Lemma 16.9) that $A \cap A' = a + (F \cap F')$, where F, F' are the directions of A, A'. Since $A \parallel A'$ and $\dim(A) \leq \dim(A')$ we have $F \subset F'$. This implies $F \cap F' = F$, so $A \cap A' = A$ i.e. $A \subset A'$.
2) From 1) we have $A \subset A'$ and $A' \subset A$. □

Corollary 18.5. *Let A be an affine subspace of fixed positive dimension of the K-vector space E. Let $p \in E$. There exists one and only one affine subspace $A' \subset E$ such that: $p \in A'$, $\dim(A) = \dim(A')$ and $A \parallel A'$.*

Proof. Let $A = a + F$. For the existence it is enough to consider $A' = p + F$. The uniqueness follows from point (2) of Lemma 18.4. □

Remark 18.6. We identify $E = \mathbb{R}^2$ with the *plane of elementary geometry.* The previous corollary states that given a line and a point in the plane, there exists one and only one line passing through the point and parallel to the given line. In short, we have proved Euclid's famous axiom of parallels. How is this possible?

We all know that using the axioms of Euclidean geometry it is not possible to prove the axiom of parallels (there are geometries that are equally valid, from a logical point of view, as Euclidean geometry in which there is no parallel or there are infinitely many parallels). The fact is that we are using a much stronger axiomatic system than that of classical Euclidean geometry.

In linear algebra the Grassmann relation allows us to estimate the dimension of the intersection of two subvector spaces, in affine geometry we have:

Lemma 18.7. *Let E be a K-vector space of dimension n and A, A' two affine subspaces of E of dimension s, t. If $A \cap A' \neq \emptyset$ then $\dim(A \cap A') \geq s + t - n$.*

Proof. Let $a \in A \cap A'$, then $A \cap A' = a + (F \cap F')$, where F, F' are the directions of A, A'. The Grassmann relation gives

$$\dim(F \cap F') = \dim(F) + \dim(F') - \dim(F + F')$$

and the result follows since $\dim(F + F') \leq n$. □

Considering two parallel affine hyperplanes we see that the hypothesis $A \cap A' \neq \emptyset$ is necessary. By adding an assumption about the relative position of the directions we can guarantee that the intersection of two affine subspaces is non-empty.

Proposition 18.8. *Let E be a K-vector space of dimension n. Let $A = a + F$, $A' = a' + F'$ be two affine subspaces of dimensions s, t. The following are equivalent:*
1) $F + F' = E$;
2) $A \cap A' \neq \emptyset$ and $\dim(A \cap A') = s + t - n$.

Proof. 1) $\Rightarrow$ 2). Since $E = F + F'$, $\exists\, f \in F, f' \in F'$ such that $a - a' = f + f'$. Then $a - f = a' + f' \in A \cap A'$. If $b \in A \cap A'$ we can write $A = b + F, A' = b + F'$ and $A \cap A' = b + (F \cap F')$. Since $\dim(F + F') = n$, the result follows from the Grassmann relation.
2) $\Rightarrow$ 1). We have $A \cap A' = b + (F \cap F')$ and $\dim(F \cap F') = s + t - n$. From the Grassmann relation it follows that $\dim(F + F') = n$, so $F + F' = E$. □

Corollary 18.9. *Let H be a hyperplane of the K-vector space E. If $A \subset E$ is an affine subspace (of positive dimension) not parallel to H, then $H \cap A \neq \emptyset$ and* $\dim(H \cap A) = \dim(A) - 1$.

Proof. Let $A = a + F$ and $H = h + G$. Since A and H are not parallel F is not contained in G. Let $v \in F \setminus G$, then $\langle v \rangle \oplus G = E$, i.e. $F + G = E$. We conclude by Proposition 18.8. □

This corollary is very important. In fact it allows us to do proofs by induction on the dimension.

18.1. Intersection of two hyperplanes

Let H, H' be two hyperplanes of E given by the Cartesian equations $\varphi(v) = \beta$, $\varphi'(v) = \beta'$.

Lemma 18.10. *With the previous notations we have:*
1) $H = H'$ if and only if there exists $\lambda \in K$ such that $\varphi = \lambda\varphi'$ and $\beta = \lambda\beta'$.
2) H and H' are parallel if and only if there exists $\lambda \in K$ such that $\varphi = \lambda\varphi'$.
3) H and H' are not parallel if and only if φ and φ' are linearly independent. In this case $H \cap H'$ is an affine subspace of dimension $n-2$, defined by the Cartesian equations $\{\varphi(v) = \beta, \varphi'(v) = \beta'\}$.

Proof. Follows from Corollary 18.9. □

If instead the hyperplanes are given in parametric representation $H = a + \langle h_1, ..., h_{n-1} \rangle$, $H' = a' + \langle h'_1, ..., h'_{n-1} \rangle$, we have:

Lemma 18.11. *With the previous notations the following are equivalent:*
1) $H \parallel H'$.
2) for every $i, 1 \leq i \leq n-1$, $\det(h_1, \ldots, h_{n-1}, h'_i) = 0$.
3) for every $i, 1 \leq i \leq n-1$, $\det(h'_1, \ldots, h'_{n-1}, h_i) = 0$.

In practice it is perhaps more convenient to derive Cartesian equations.

18.2. Intersection of a hyperplane and a line

Let H, R be a hyperplane and a line of the K-vector space E of dimension n. If $H \parallel R$ there are two possible cases: (i) $R \subset H$, (ii) $R \cap H = \emptyset$. To distinguish the two cases it is enough to check whether any point of R belongs to H or not. If the point belongs to H we are in the first case, otherwise we are in the second case.

If $H \not\parallel R$ (H is not parallel to R) then (Corollary 18.9) $H \cap R = \{p\}$.

Lemma 18.12. *If H is given by the Cartesian equation $\varphi(v) = \beta$ and if R is given by the equations $\varphi_1(v) = \beta_1, ..., \varphi_{n-1}(v) = \beta_{n-1}$, then the following are equivalent:*
1) $R \parallel H$.
2) $\det(\varphi_1, \ldots, \varphi_{n-1}, \varphi) = 0$.

Proof. By definition $H \parallel R \Leftrightarrow \dim(R) \subset \dim(H)$. We have $\dim(R) = \langle \varphi_1, ..., \varphi_{n-1} \rangle^\circ$ and $\dim(H) = \langle \varphi \rangle^\circ$. Now $\langle \varphi_1, \ldots, \varphi_{n-1} \rangle^\circ \subset \langle \varphi \rangle^\circ \Leftrightarrow \langle \varphi \rangle \subset \langle \varphi_1, \ldots, \varphi_{n-1} \rangle$ (duality reverses inclusions). This last inclusion implies that $\varphi, \varphi_1, \ldots, \varphi_{n-1}$ are linearly dependent and this occurs if and only if their determinant in any basis is zero. □

Once a basis of E is chosen, E is isomorphic to K^n and the intersection $H \cap R$ is given by the linear system

$$\begin{cases} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = \beta_1; \\ \vdots \\ \alpha_{n-1,1}x_1 + \cdots + \alpha_{n-1,n}x_n = \beta_{n-1}; \\ \alpha_1 x_1 + \cdots + \alpha_n x_n = \beta. \end{cases}$$

It is a system of n equations in n unknowns. Let A be the coefficient matrix. If A has rank n (the equations $\varphi_1, \ldots, \varphi_{n-1}, \varphi$ are independent) then the system is a Cramer system and admits a unique solution (we are in the case H and R are parallel). If $\det(A) = 0$ then the equations are dependent and $\mathrm{rk}(A) = n - 1$ (since $\varphi_1, \ldots, \varphi_{n-1}$ are independent). The system is compatible if and only if $\mathrm{rk}(A \mid B) = n - 1$ ($(A \mid B)$ is the complete matrix of the system). If compatible, the solution set is an affine subspace of dimension one (it is the line R, we are in the case $R \subset H$). If the system is incompatible $H \cap R = \emptyset$, we are in the case $R \parallel H$, $R \not\subset H$.

In practice, rather than calculating the rank of A and that of $(A \mid B)$, it may be more advantageous to solve the system directly by substitutions.

Let us now move on to the parametric point of view. We have $H = a + F$, $R = b + \langle u \rangle$. Clearly $R \parallel H \Leftrightarrow u \in F$. If $(f_1, \ldots, f_{n-1})$ is a basis of F we have: $u \in F \Leftrightarrow \det(u, f_1, \ldots, f_{n-1}) = 0$. If $R \parallel H$ we have $R \subset H \Leftrightarrow b \in H$.

If H and R are not parallel we know that $H \cap R = \{p\}$. So

$$p = b + \lambda u = a + \lambda_1 f_1 + \cdots + \lambda_{n-1} f_{n-1}$$

and $b - a = \lambda_1 f_1 + \cdots + \lambda_{n-1} f_{n-1} - \lambda u$. The parameters $\lambda_1, ..., \lambda_{n-1}, -\lambda$ are the coordinates of the vector $b - a$ in the basis $(f_1, ..., f_{n-1}, u)$. We got a Cramer system.

18.3. Incidences in the plane and in the 3-dimensional space

Let E be a 2-dimensional K-vector space. The only interesting case from the point of view of incidence of affine subspaces is that of two lines. But this is the case of the intersection of two hyperplanes that we have already discussed. Let us remember the conclusion: the two lines R, D are parallel (with the two cases $R = D$ or $R \cap D = \emptyset$) or they are not and intersect in a point.

Now, let E be a 3-dimensional K-vector space. The proper affine subspaces of positive dimension are lines and planes. In this case the planes are hyperplanes. We still have to study the mutual position of two lines.

Definition 18.13. Two affine lines $R, D \subset E$ are said to be *coplanar* if there exists an affine plane that contains both of them. The lines R, D are said to be *skew* if they are not parallel and if $R \cap D = \emptyset$.

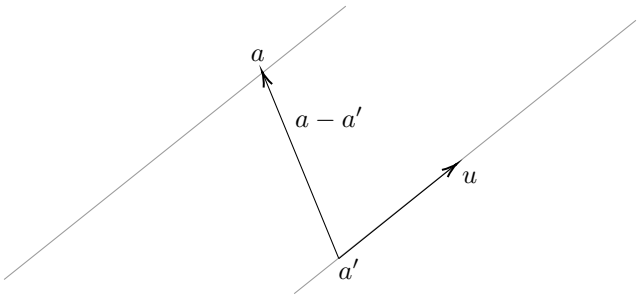
We have often seen this tragic error: *two lines are skew if they are contained in distinct planes.* Be careful, any two distinct lines are always contained in distinct planes!

Lemma 18.14. *Let E be a K-vector space of dimension n, R and R' two affine lines of E.*
1) If R and R' are coplanar, not parallel, then their intersection consists of one and only one point.
2) If R and R' meet at a point then they are coplanar.

Proof. 1) This follows from the study in the affine plane. Suppose $R, R' \subset \Pi$, Π affine subspace of dimension two. Let $R = a + \langle u \rangle$, $R' = a' + \langle u' \rangle$. We have $u, u' \in \mathrm{dir}(\Pi)$ (Lemma 18.3). Since R and R' are not parallel, u, u' are independent and therefore form a basis of $\mathrm{dir}(\Pi)$. Hence $a - a' = \lambda u + \mu u'$ (note that $a - a' \in \mathrm{dir}(\Pi)$, Lemma 16.7) i.e. $p := a - \lambda u = a' + \mu u' \in R \cap R'$. Since R and R' are distinct we conclude that $R \cap R' = \{p\}$.
2) Set $R \cap R' = \{p\}$. We can write $R = p + \langle u \rangle$, $R' = p + \langle u' \rangle$. Then R and R' are contained in the plane $p + \langle u, u' \rangle$ (u, u' are independent because the lines are not parallel). □

Lemma 18.15. *Let E be a K-vector space of dimension n, and R, R' two affine lines of E. If R and R' are parallel, then they are coplanar.*

Proof. Clearly we can assume $R \neq R'$. We can write $R = a + \langle u \rangle$ and $R' = a' + \langle u \rangle$. Let $w = a - a'$. We show that w and u are linearly independent. Let $\alpha w + \mu u = 0$. If $\alpha = 0$, then $\mu = 0$ and we are done. If $\alpha \neq 0$, then $a - a' = w = \lambda u$, i.e. $a = a' + \lambda u \in R'$, but this is a contradiction since $R \cap R' = \emptyset$. Let $H = a + \langle a - a', u \rangle$. It is clear that $R \subset H$. If $p = a' + \lambda u \in R'$, then $p = a - (a - a') + \lambda u \in H$, so $R' \subset H$. □



Corollary 18.16. *Two lines of an affine space are skew if and only if they are not coplanar.*

Proof. It follows from Lemma 18.14 and Lemma 18.15. □

Let us now consider two lines R, D, of a 3-dimensional space. After choosing a basis we can assume R given by the Cartesian equations: $\alpha_{11}x + \alpha_{12}y + \alpha_{13}z = \beta_1$, $\alpha_{21}x + \alpha_{22}y + \alpha_{23}z = \beta_2$ and D given by the equations: $\alpha'_{11}x + \alpha'_{12}y + \alpha'_{13}z = \beta'_1$, $\alpha'_{21}x + \alpha'_{22}y + \alpha'_{23}z = \beta'_2$. Note that the matrix

$$M = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \beta_2 \\ \alpha'_{11} & \alpha'_{12} & \alpha'_{13} & \beta'_1 \\ \alpha'_{21} & \alpha'_{22} & \alpha'_{23} & \beta'_2 \end{pmatrix}$$

is the complete matrix of the system that determines $R \cap D$.

Proposition 18.17. *With the previous notations, the following are equivalent:*
1) $\det(M) = 0$*;*
2) $R \parallel D$ or R and D intersect at a point;
3) R and D are coplanar.

Proof. 1) ⇒ 2). Let A be the 4×3 matrix obtained from M by removing the last column; A is the coefficient matrix of the system that determines $R \cap D$. We have $\mathrm{rk}(A) \geq 2$ (the two equations of R (or D) are independent). Furthermore

$\mathrm{rk}(A) = 2 \Leftrightarrow \langle f_1, f_2\rangle = \langle f_1', f_2'\rangle$ (f_i (f_i') are the equations of R (D)), i.e. if and only if $\langle f_1, f_2\rangle^\circ = \langle f_1', f_2'\rangle^\circ$, i.e. if and only if $R \parallel D$. If $\mathrm{rk}(A) = 3$ then $\mathrm{rk}(A) = \mathrm{rk}(M)$ (because $\det(M) = 0$) and the system is compatible, the solution set has zero dimension (the system is equivalent to a Cramer system), then the two lines intersect at a point.

2) $\Rightarrow$ 1) If $R \parallel D$ then $\mathrm{rk}(A) = 2$ and this implies $\mathrm{rk}(M) \leq 3$, so $\det(M) = 0$. If $R \cap D = \{p\}$, the system is compatible so $\mathrm{rk}(A) = \mathrm{rk}(M)$. Since $\mathrm{rk}(A) \leq 3$ (A has three columns), we have $\det(M) = 0$.

2) $\Leftrightarrow$ 3). Follows from Lemma 18.14 and Lemma 18.15. □

Using this result is not recommended in practice. It is not even recommended to calculate the ranks of A and M.

Instead, it is better to troubleshoot the system directly (for replacements). If the system is compatible, we will see that the lines meet at a point (and the coordinates of the point of intersection are found). If the system is incompatible, the two lines are parallel or skew. By solving the homogeneous systems $f_1(v) = 0, f_2(v) = 0$ and $f_1'(v) = 0, f_2'(v) = 0$ the two directions are found and it is easy to determine the mutual position of the two lines.

Exercises

Exercise 18.18. In $\mathbb{R}^3$ consider the plane H of equation $2x + y - 2z = -1$, the line L of equations $\{z = x - 1, y + 1 = 2z\}$ and the line $D = (1, 1, 1) + \langle(1, 2, 2)\rangle$.
1) Determine the mutual positions of H, L, D.
2) Without doing calculations, show that there exists a unique line R such that $R \parallel H$, $b := (3, 5, 5) \in R$, $R \cap L \neq \emptyset$.
3) Give Cartesian equations of R.

Exercise 18.19. In $\mathbb{R}^3$ let $u = (1, -1, 2)$, $v = (0, -1, 1)$, $w = (1, 2, 0)$, $p = (1, 2, -2)$, $q = (1, 1, 1)$, $r = (3, -1, 1)$, $L = p + \langle u\rangle$, $R = q + \langle v\rangle$ and $D = r + \langle w\rangle$.
(1) Determine the mutual positions of L, R, D.
(2) Show that for every $a \in D$, $a - p$ and u are linearly independent.
(3) For each $a \in D$ let $H(a)$ be the affine subspace generated by a and L. Determine

$$X := \{a \in D \mid H(a) \parallel R\}.$$

Exercise 18.20. In $\mathbb{R}^4$, with coordinates x, y, z, t, consider the planes $\Pi_1 = (1, 0, 0, 0) + \langle(-1, 1, 0, 0), (0, 1, 1, 1)\rangle$ and Π_2 given by the equations $x - 1 + y = t, z = t$. Finally, let R be the line of equations $x + y = 3$, $z = 0$, $t = 1$.
1) Determine the mutual positions of Π_1, Π_2, R.
2) Let $L_a = (1, 1, a, 1) + \langle(-1, 1, 0, 0)\rangle$ where a is a real parameter. Determine

$$T = \{a \mid \text{ there exists a hyperplane containing } \Pi_1, \Pi_2 \text{ and } L_a\}.$$

Exercise 18.21. In $\mathbb{R}^3$ let D, L be two non-parallel and coplanar lines. Denote by H the plane containing D and L. Let R be a line not contained in H and which meets D at a point p, with $p \neq D \cap L$. Show that R and L are skew.

Exercise 18.22. In $\mathbb{R}^3$ let R, S be two skew lines and let p be a point not belonging to $R \cup S$. Prove that there is a unique line, L, passing through p and coplanar with both R and S. Also show that L meets at least one of the two lines R, S.

Exercise 18.23. In $\mathbb{R}^4$ consider the affine subspace E defined by

$$E = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + 2y + z = -1 \text{ and } x - y - t = 1\}$$

and the line $L = (1, 0, 1, 0) + \langle(0, 1, 0, 1)\rangle$.

(1) Determine the mutual positions of E and L.

(2) Let $p = (1, -1, 0, 1)$ and Π the affine subspace generated by p and L. Determine $E \cap \Pi$.

Exercise 18.24. In $\mathbb{R}^3$ consider the two lines $L = q + \langle u\rangle$, $R = p + \langle v\rangle$. Show that L and R are skew if and only if the three vectors $q - p, u, v$ are linearly independent.

Exercise 18.25. Consider three lines $D_i = a_i + \langle u_i \rangle$, $1 \leq i \leq 3$ in $\mathbb{R}^3$, such that two of them are skew.
(i) Show with an example that u_1, u_2, u_3 are not necessarily independent.
(ii) Show that u_1, u_2, u_3 are independent if and only if for each plane $H \subset \mathbb{R}^3$, $H \cap (D_1 \cup D_2 \cup D_3) \neq \emptyset$.

Exercise 18.26. In $\mathbb{R}^3$ consider the plane E of equation $x - 2y + z = 1$, the line D of equations: $z = 1 - x$, $y = -1$ and the line $R = (1, 0, 2) + \langle (1, 0, -1) \rangle$.
(i) Determine a Cartesian representation of the affine subspace, H, generated by R and D.
(ii) Determine the mutual positions of R, E, D.
(iii) Let $p = (1, 0, 0)$. Show that there exists one and only one plane, Π, such that $p \in \Pi$, $\Pi \parallel D$, $\Pi \cap R \neq \emptyset$. Give a Cartesian equation of Π.

Exercise 18.27. In $\mathbb{R}^3$ consider the plane E of equation $x + y - z = 1$, the line D of equations: $z = x$, $1 + y = x$ and the line $R = (1, 2, 0) + \langle (1, 1, 1) \rangle$.
(i) Determine a Cartesian representation of the affine subspace, H, generated by R and D.
(ii) Determine $L = E \cap H$ and deduce the mutual positions of R, E, D.
(iii) Let $w = (2, 2, 2)$. Show that there exists an affinity, g, such that $g(w) = w$, $g(L) = L$ and $g(E) = H$.

Exercise 18.28. In $\mathbb{R}^4$ consider the affine subspace E defined by $E = \{(x, y, z, t) \in \mathbb{R}^4 \,|\, x + 2y + z = -1 \text{ and } x - y - t = 1\}$, and the line $L = (1, 0, 1, 0) + \langle (0, 1, 0, 1) \rangle$.

(1) Determine the mutual positions of E and L.

(2) Let $p = (1, -1, 0, 1)$ and Π be the affine subspace generated by p and L, determine $E \cap \Pi$.

Exercise 18.29. In $\mathbb{R}^3$ consider two lines R, D such that $R \cap D = \{p\}$. Let L be a line incident to R: $L \cap R = \{q\}$, with $p \neq q$. It is assumed that L is not contained in the plane $H = \langle R, D \rangle$. Show that L and D are skew.

Exercise 18.30. In $\mathbb{R}^3$ consider the plane H of equation $2x - 4y - z = 3$, the line R of equations $x - 3y = -2$, $z - 2y = -3$, the line $D = (-1, 0, -1) + \langle (3, 1, 2) \rangle$ and the line $L = (1, 0, -1) + \langle (1, 1, 1) \rangle$. Let E be the affine subspace generated by R and D.
(i) Determine the reciprocal positions of H, E, L.
(ii) Show that there exists no line, l, passing through $P = (1, 0, -1)$ such that $l \cap R \neq \emptyset$ and $l \cap D \neq \emptyset$.
(iii) Say whether there exists an affinity, φ, of $\mathbb{R}^3$ such that $\varphi(L) = D$ and $\varphi(D) = R$.

Exercise 18.31. Consider two parallel planes H, H' in $\mathbb{R}^3$. Let $R, D \subset H'$ be two parallel lines. Show that there exists an affinity, φ, of $\mathbb{R}^3$ such that $\varphi(H) = H$, $\varphi(R) = D$, and $\varphi(D) = R$ (consider $p \in D, q \in R, m \in H$ and use the fact that $\varphi(p) - \varphi(q) = f(p - q)$ where f is the linear part of φ).

Exercise 18.32. Let R_1, R_2, R_3 be three distinct lines of $\mathbb{R}^3$.
(i) R_1, R_2, R_3 are assumed two by two skew. Show that there exist infinitely many lines, L, such that $L \cap R_i \neq \emptyset$, $1 \leq i \leq 3$.
(ii) The lines R_i are again assumed to be two by two skew. Let $p \in R_1$. Show with an example that you can not always find a line L with $p \in L$ and $L \cap R_i \neq \emptyset$, $i = 2, 3$.
(iii) Is (i) still true if the lines are not two by two skew?

Exercise 18.33. Consider the following affine subspaces in $\mathbb{R}^3$: $H = \{(x, y, z) \mid x + y - 3z + 2 = 0\}$, $L = \{(x, y, z) \mid x - y = 2 \text{ and } 2z - x - y = -4\}$, $D = (0, 1, 2) + \langle (1, 0, 1) \rangle$.

(1) Determine dimensions and mutual positions of H, L, D.

(2) In an affine space of dimension three consider two skew lines, L', D' and a plane H' such that $L' \cap H' = \{p'\}$, $D' \cap H' = \{q'\}$. Let $l = [p', q']$ and $r \in l \setminus \{p', q'\}$. Show that there is no line R such that
   (a) $r \in R$;
   (b) $R \not\subseteq H'$;
   (c) $R \cap L' \neq \emptyset$, $R \cap D' \neq \emptyset$.

(3) Conclude that there is no line, R, $R \not\subseteq H$, passing through $r = (-12, -5, -5)$ and meeting both L and D.

Exercise 18.34. Let

$$\mathcal{S}(\lambda) = \begin{cases} x + y - \lambda z - t = 2; \\ -x + \lambda y - t = 1; \\ 2x + y - \lambda z + t = 3. \end{cases}$$

(1) Show, without solving it, that the linear system $\mathcal{S}(\lambda)$ is compatible for every value of the parameter $\lambda \in \mathbb{R}$.

(2) Let $S(\lambda) \subset \mathbb{R}^4$ be the set of solutions of $\mathcal{S}(\lambda)$. Show that if $\lambda \neq 0$ then $S(\lambda)$ is never parallel to $S(0)$.

Exercise 18.35. Consider the following system in $\mathbb{R}^4$

$$\mathcal{S}(\lambda) = \begin{cases} \lambda x + y + z - t = 1; \\ 2x + \lambda y + z = \lambda; \\ -x + \lambda y + t = 2; \end{cases}$$

where $\lambda \in \mathbb{R}$ is a real parameter.

(1) Show, without solving the system, that $\mathcal{S}(\lambda)$ is compatible $\forall\lambda \in \mathbb{R}$ and that the solution set, $S(\lambda)$, is a line.

(2) Is it possible to find $\lambda, \mu \in \mathbb{R}$, $\lambda \neq \mu$, such that $S(\lambda)$ and $S(\mu)$ are parallel?

Exercise 18.36. Remember that two affine subspaces are skew if they do not intersect and if they are not parallel.
(i) Show that two general affine planes in $\mathbb{R}^4$ intersect at a point (in other words: two planes, in general, intersect at a point; or again: given a plane E, in the set of all the planes of $\mathbb{R}^4$, those which intersect E at a point are the most numerous ones, they form a dense open space.)
(ii) Are there two skew affine planes in $\mathbb{R}^4$?

Exercise 18.37. Let E, F be two skew planes in $\mathbb{R}^4$, show that $\dim(\mathrm{dir}(E) \cap \mathrm{dir}(F)) = 1$.

Exercise 18.38. (1) Let E, F be two parallel planes in $\mathbb{R}^4$, with $E \cap F = \emptyset$. Show that there is one and only one hyperplane containing $E \cup F$.

(2) Let E, F be two planes in $\mathbb{R}^4$ whose intersection is a line. Show that there is one and only one hyperplane containing $E \cup F$.

Exercise 18.39. Let E, F be two affine planes in $\mathbb{R}^4$. In the following, E and F are assumed to be skew.

(1) Show that it is possible to find linear forms f, g, h and constants $\alpha, \beta, \gamma, \delta$ such that E is given by the equations $f(v) = \alpha$, $g(v) = \beta$ and F by the equations $f(v) = \gamma$, $h(v) = \delta$. Show that under these conditions $\alpha \neq \gamma$.

(2) Tell whether there is a plane G parallel to both E and F.

(3) Show that there exists a plane L such that $\dim(L \cap E) = \dim(L \cap F) = 1$. Show that for such a plane L, necessarily $L \cap E$ and $L \cap F$ are parallel.

(4) Show that there are infinitely many planes L with $\dim(L \cap E) = \dim(L \cap F) = 1$.

Exercise 18.40 (Continuation of the Exercise 18.39)**.** Let E, F be two skew planes in $\mathbb{R}^4$, and A, B two planes such that $\dim(A \cap E) = \dim(A \cap F) = 1$, $\dim(B \cap E) = \dim(B \cap F) = 1$.

(1) Show that if $A \cap B \neq \emptyset$, then $A \cap B$ is a straight line, more precisely $A \cap B = A \cap E = B \cap E$ or $A \cap B = A \cap F = B \cap F$.

(2) Show that if $A \cap B = \emptyset$, then A and B are skew. In particular, A and B can not be parallel.

Exercise 18.41. In $\mathbb{R}^4$ consider the affine subspaces E, F where $E = (0,1,0,0) + \langle(-2,2,1,0),(-2,2,1,1)\rangle$ and

$$F = \begin{cases} x + y = 2; \\ -z + t = -1. \end{cases}$$

(1) Determine the dimensions of E, F and their mutual positions.

(2) Let $a = (1,1,0,-1)$. Show that there exists a unique line, R, such that $a \in R \subset F$ and $R \parallel E$.

(3) Let $b \notin E \cup F$ and $G = \langle b, R\rangle$ be the plane generated by b and R. Show that if $a - b \notin \mathrm{dir}(E)$ then the linear system corresponding to $G \cap E$ has rank three.

Exercise 18.42. Consider the following affine subspaces in $\mathbb{R}^4$: $E = \{(x,y,z,t) \in \mathbb{R}^4 \mid x - y + z + t = 1\,,\ x + y + 2t = -1\}$, $F = \{(x,y,z,t) \in \mathbb{R}^4 \mid 2x - z + t = 2\,,\ y + z - t = 1\}$, $G = p + \langle e_1, e_2\rangle$, where $p = (\frac{5}{3}, \frac{-1}{3}, \frac{1}{6}, \frac{-7}{6})$ and where $(e_i), 1 \leq i \leq 4$ is the canonical basis.

(1) Determine the mutual positions of E, F, G.

(2) Determine the equations of a plane Π verifying the following conditions:
- $q = (0,0,0,1) \in \Pi$;
- $e_1 \in \mathrm{dir}(\Pi)$;
- Π is skew with E, F and G.

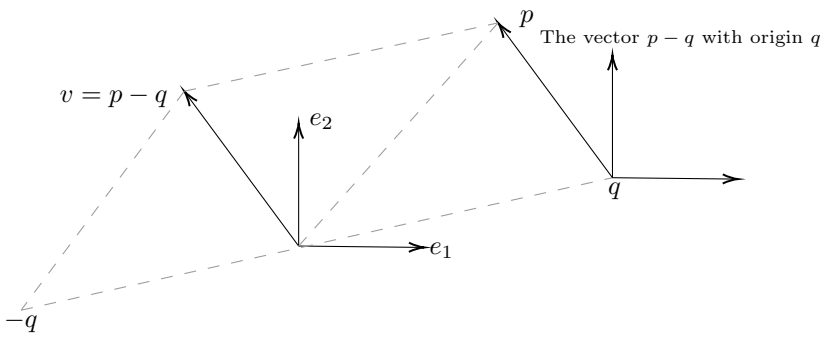
(3) Show that the plane Π is uniquely determined.

Exercise 18.43. Let $n \geq 3$ and s, t two integers such that $1 \leq s \leq t < n - 1$. Show that there exist in $\mathbb{R}^n$ two skew affine subspaces, E, F, with $\dim(E) = s$, $\dim(F) = t$. Show that the condition $t < n - 1$ is necessary.

Chapter 19

Affine coordinate systems and affinities

Let E be a K-vector space and $q \in E$ a point. If $p \in E$ we can consider the vector $p-q=v$, we can think of it as the vector with origin q and pointing in the direction of p.



It is clear that $p = q + v$ i.e. $p = q + (p - q)$; in this equality p and q should be thought of as points and v as a vector. Once q has been fixed (as the origin of the affine space E) the point p is completely determined by the vector v. If $(e_1, \ldots, e_n)$ is a basis of E we can consider the coordinates of v in this basis: $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$, and the point p is completely determined by the scalars $(\lambda_1, \ldots, \lambda_n)$.

Definition 19.1. An *affine coordinate system* on the affine space E is a pair $(q, \mathcal{B})$ where q is a point of E and $\mathcal{B}$ is a basis of vector space E. The point q is called the origin of the coordinate system.

If $p \in E$, the coordinates of the vector $v = p - q$ in the basis $\mathcal{B}$ are the affine coordinates of the point p in the coordinate system $(q, \mathcal{B})$.

Remark 19.2. As can be seen from the drawing above, the coordinate system $(q, \mathcal{B})$, $\mathcal{B} = (e_1, e_2)$ is nothing but the basis $\mathcal{B}$ translated to q. The coordinates of the origin q in the coordinate system $(q, \mathcal{B})$ are $(0, \ldots, 0)$.

19.1. Change of coordinates

Let $(q, \mathcal{B})$, $(q', \mathcal{B}')$ be two affine coordinate systems on E. The problem consists in determining the coordinates of a point p in the coordinate system $(q', \mathcal{B}')$ using those in the coordinate system $(q, \mathcal{B})$. For this it is enough to know the coordinates of q' in the coordinate system $(q, \mathcal{B})$.

In fact, let $p = q + v$, then $p = q' + ((q - q') + v) = q' + (v - (q' - q))$. The coordinates of p with respect to $(q, \mathcal{B})$ are the coordinates $(\lambda_1, \ldots, \lambda_n)$ of the vector v in the basis $\mathcal{B}$; the coordinates of p with respect to $(q', \mathcal{B}')$ are the coordinates $(\lambda'_1, \ldots, \lambda'_n)$ of the vector $v - (q' - q)$ in the basis $\mathcal{B}'$.

Let $(\beta_1, \ldots, \beta_n)$ be the coordinates of the vector $q' - q$ in the basis $\mathcal{B}$ i.e. the coordinates of q' in the coordinate system $(q, \mathcal{B})$. If $M = \mathrm{Mat}(Id_E; \mathcal{B}, \mathcal{B}')$ then

$$M \begin{pmatrix} \lambda_1 - \beta_1 \\ \vdots \\ \lambda_n - \beta_n \end{pmatrix} = \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix}.$$

In conclusion:

Proposition 19.3. *If $(\lambda_1, \ldots, \lambda_n)$ are the coordinates of p in the coordinate system $(q, \mathcal{B})$, then the coordinates $(\lambda'_1, \ldots, \lambda'_n)$ of p in the coordinate system $(q', \mathcal{B}')$ are given by*

$$M \begin{pmatrix} \lambda_1 - \beta_1 \\ \vdots \\ \lambda_n - \beta_n \end{pmatrix} = \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix}$$

where $(\beta_1, \ldots, \beta_n)$ are the coordinates of q' in the coordinate system $(q, \mathcal{B})$ and where $M = \mathrm{Mat}(Id_E; \mathcal{B}, \mathcal{B}')$.

Remark 19.4. The previous formula can also be rewritten (with obvious notations) as $\underline{\lambda}' = M\underline{\lambda} + \underline{\delta}$, where $\underline{\delta} = -M\underline{\beta}$. This shows that the application $\underline{\lambda} \to \underline{\lambda}'$ is an affinity.

19.2. Affinities

A characterization of affinities:

Lemma 19.5. *Let E be a K-vector space. A map $\psi : E \to E$ is an affinity if and only if there exists a linear isomorphism $v : E \to E$ such that: $\forall (p, q) \in E^2$: $\psi(p) - \psi(q) = v(p - q)$. In this case $L(\psi) = v$.*

Proof. Suppose the condition is verified. Let us fix a point q. From $\psi(x) - \psi(q) = v(x-q)$ we obtain $\psi(x) = (t_a \circ v)(x)$, where $a = \psi(q) - v(q)$. Since v is invertible, ψ is an affinity. Conversely, if $\psi = t_a \circ v$ is an affinity it is clear that $\psi(p) - \psi(q) = v(p-q)$ for every pair of points (p, q). □

Lemma 19.6. *Let E be a K-vector space and $q \in E$ a point. Then $\forall q' \in E$, $\forall v \in Gl(E)$, there exists a unique affinity $\psi : E \to E$ such that $\psi(q) = q'$ and $L(\psi) = v$.*

Proof. Set $\psi(x) = v(x - q) + q'$. We see that $\psi = t_b \circ v$, with $b = q' - v(q)$, so ψ is an affinity, also $\psi(q) = q'$. If $\varphi = t_c \circ w$ is an affinity that satisfies the conditions of the lemma then $w = v$ and $\varphi(q) = v(q) + c = q'$. Hence $c = q' - v(q) = b$ and $\varphi = \psi$. □

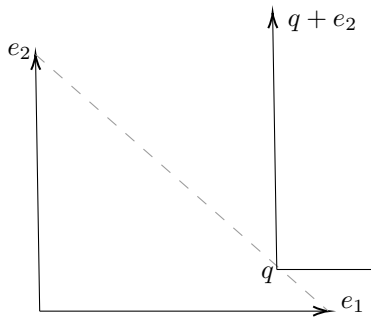
Remark 19.7. Lemma 19.6 shows that an affinity is completely determined by its linear part and the image of any point. It is the analogue of the linear algebra result stating that a linear map is determined by the images of the basis vectors.

Let us now move on to the so-called *fundamental theorem of affine geometry.*

Theorem 19.8. *Let E be a K-vector space of dimension n, $q_0, q_1, \ldots, q_n \in E$ and $p_0, p_1, \ldots, p_n \in E$ two collections or $n + 1$ affinely independent points. There is a unique affinity $\psi : E \to E$ such $\psi(q_i) = p_i$ for $0 \leq i \leq n$.*

Proof. Since the points q_i (resp. p_i) are affinely independent, the vectors $q_i - q_0$ (resp. $p_i - p_0$) are linearly independent and form a basis $\mathcal{B}$ (resp. $\mathcal{C}$) of E. Let $v : E \to E$ be the linear map such that $v(p_i - p_0) = q_i - q_0, 1 \leq i \leq n$. The map v is a linear isomorphism because it transforms the basis $\mathcal{B}$ into the basis $\mathcal{C}$. Let ψ be the unique affinity such that $L(\psi) = v$ and $\psi(q_0) = p_0$ (Lemma 19.6). Then $\psi(q_i) - \psi(q_0) = v(q_i - q_0) = p_i - p_0$. So $\psi(q_i) = p_i, 0 \leq i \leq n$. □

Remark 19.9. 1) Another way of formulating the theorem: *An affinity of E is completely determined by its values on $n + 1$ affinely independent points, where $n = \dim(E)$.*
2) If $(q, \mathcal{B})$, $\mathcal{B} = (e_1, \ldots, e_n)$, is an affine coordinate system, the $q, e_1, \ldots, e_n$ are not necessarily affinely independent; it is instead true that the points $q, q+e_1, \ldots, q+e_n$ are affinely independent.



The three points $q,\ e_1,\ e_2$ are collinear but the three points $q,\ q + e_1,\ q + e_2$ are not.

If $p_1, \ldots, p_{n+1}$ are $n + 1$ affinely independent points then for every i, $1 \leq i \leq n + 1$, $(p_i, \mathcal{B}_i)$ with $\mathcal{B}_i = (p_1 - p_i, \ldots, p_{n+1} - p_i)$ is a coordinate system.

In conclusion $(p_0, (p_1 - p_0, \ldots, p_n - p_0))$ is a coordinate system if and only if the points $p_0, \ldots, p_n$ are affinely independent.

Exercises

Exercise 19.10. In $\mathbb{R}^3$ consider the points $q = (1,1,1)$, $v_1 = (-1,0,0)$, $v_2 = (1,2,0)$, $v_3 = (2,0,-3)$ and $b = (4,-1,0)$.
1) Show that $\mathcal{B} = (v_1, v_2, v_3)$ is a basis of $\mathbb{R}^3$.
2) Determine the coordinates of the point b in the coordinate system $(q, \mathcal{B})$.

Exercise 19.11. In $\mathbb{R}^2$ consider the line D of equation $x+y = 1$. Give the equation of D in the coordinate system $(q, \mathcal{B})$ where $q = (1,-1)$, $\mathcal{B} = (u, v)$ with $u = (0,-1)$, $v = (2,1)$ (verify that this is indeed a coordinate system).

Exercise 19.12. Show that two distinct planes of $\mathbb{R}^4$ can
a) intersect at a point (this is the general case);
b) be skew (they do not intersect and are not parallel);
c) intersect in a line.

Exercise 19.13. Let $E, F \subset \mathbb{R}^4$ be two skew planes of $\mathbb{R}^4$. Show that there exists an affinity φ of $\mathbb{R}^4$ such that $\varphi(E) = F, \varphi(F) = E$ (show that we can assume $E = a + \langle u, v\rangle$, $F = b + \langle u, w\rangle$ and that $(a-b, u, v, w)$ is a basis of $\mathbb{R}^4$. Finally, use the fact that an affinity is determined by its linear part and the image of a point).

Exercise 19.14. Consider two parallel and disjoint planes H, H' in $\mathbb{R}^3$. Let $R, D \subset H'$ be two disjoint parallel lines. Show that there exists an affinity, φ, of $\mathbb{R}^3$ such that $\varphi(H) = H$, $\varphi(R) = D$, and $\varphi(D) = R$ (consider $p \in D, q \in R, m \in H$ and use the fact that $\varphi(p) - \varphi(q) = f(p-q)$ where f is the linear part of φ).

Chapter 20

Affine spaces: general theory

One can define on a set X an affine space structure (on a K-vector space) in an axiomatic way.

Definition 20.1. Let E be a K-vector space of dimension n. An affine space on E is a triple (X, E, φ) where

- X is a set;
- $\varphi : X \times E \to X$ is a map satisfying the following conditions:
  A1 $\forall P \in X, \forall (v, w) \in E^2$ we have $\varphi(\varphi(P, v), w) = \varphi(P, v + w)$;
  A2 $\forall (P, Q) \in X^2$, $\exists! v \in E$ such that $\varphi(P, v) = Q$.

If the vector space is specified by the context, we say that X is an affine space. To better understand this definition let us set some writing conventions and terminology:

- The elements of X (the points) will be denoted by capital letters.
- The elements of E (the vectors) will be denoted by lowercase letters.
- If $P \in X, v \in E$ we set $\varphi(P, v) = P + v$. Pay attention: this $+$ sign has nothing to do with addition in E, it is just a convention for writing φ. With these conventions we have:
  A1 $\forall P \in X$, $\forall (v, w) \in E^2$: $(P + v) + w = P + (v + w)$ (on the left side the two $+$ have different meanings).
  A2 $\forall (P, Q) \in X^2$ there exists a unique $v \in E$ such that $P + v = Q$ (this vector v is $Q - P$).

Example 20.2. Let us take $X = E$ and $\varphi : E \times E \to E : (x, v) \to x + v$ (addition in E). Conditions (A1), (A2) are satisfied and (E, E, φ) is an affine space; this is the structure we have studied so far.

20.1. First properties

Lemma 20.3. *For every $(P, v) \in X \times E$ we have $P + v = P$ if and only if $v = 0$.*

Proof. We show that $P + 0 = P$, the equivalence will result from uniqueness in condition (A2). By (A2) there exists v such that $P+v = P$. So $P+0 = (P+v)+0$; by (A1) we get $(P + v) + 0 = P + (v + 0) = P + v$. In conclusion $P + 0 = P$. □

Definition 20.4. For every $v \in E$ the map $\varphi_v : X \to X : P \to P + v$ is called the *translation* of vector v.

Lemma 20.5. *For every $v \in E$ the translation t_v is a bijection.*

Proof. Using Lemma 20.3 we get that $(t_v)^{-1} = t_{-v}$. □

Definition 20.6. A subset Y of X is an affine subspace of X if there exists a point $Q \in X$ and a subspace $V \subset E$ such that $Y = \{P \in X \mid \exists v \in V, P = Q + v\}$. We say that V is the direction of Y. We will write $Y = Q + V$.

Remark 20.7. As in the vector case we prove $Q \in Y$ and $Q+V = Q'+V' \Rightarrow V = V'$, so the direction is uniquely determined and $\forall Q \in Y, Y = Q+V$. The dimension of Y is the dimension of V. It follows from (A2) that $X = P + E, \forall P \in X$.

Definition 20.8. Let X, X' be two affine spaces on E, E'. A map $f : X \to X'$ is an affine map if there exists a linear map $v : E \to E'$ such that $f(P)-f(Q) = v(P-Q)$ for any $(P, Q) \in X^2$.

As in the vector case v is uniquely determined by f. Note that it makes no sense to require that f is the composition of a linear morphism $E \to E'$ with a translation $X' \to X'$.

20.2. Coordinate system and back to vector spaces

We have seen in the vector case how the choice of an origin causes the zero vector to lose its privileged status (translating the vector space structure to the new origin). In the general case, the choice of an origin in X allows us to give X a vector space structure (isomorphic to E) and therefore leads us back to the vector case.

Let X be an affine space on E and $Q \in X$ a point, we denote by φ_Q the partial map $\varphi_Q : E \to X : v \to Q + v$.

Lemma 20.9. *For every point $Q \in X$ the map φ_Q is a bijection.*

Proof. It is a consequence of (A2). For every $P \in X$ there exists one and only one v such that $\varphi_Q(v) = P$. □

The bijection φ_Q allows us to transport the vector space structure of E to X. An addition on X is defined as follows: $P + P' := \varphi_Q(\varphi_Q^{-1}(P) + \varphi_Q^{-1}(P'))$ and an external multiplication via: $\lambda P = \varphi_Q(\lambda\varphi_Q^{-1}(P))$.

For this vector space structure on X, the map φ_Q is linear, so it is a linear isomorphism i.e. we have identified X to E. Note that the vector $\varphi_Q^{-1}(Q)) =: w$ is such that $\varphi_Q(w) = Q + w = Q$, and (Lemma 20.3) $w = 0$. For this reason we say that Q was chosen as the *origin* of X.

Definition 20.10. Let X be an affine space on E. An affine coordinate system $(Q, \mathcal{B})$ on X consists in choosing an origin Q and a basis $\mathcal{B}$ of E.

Take $P \in X$, by (A2) there exists a unique vector v such $\varphi_Q(v) = Q + v = P$. The coordinates $(\lambda_1, \ldots, \lambda_n)$ of v in the basis $\mathcal{B}$ are the coordinates of P in the affine coordinate system $(Q, \mathcal{B})$.

By writing $\mathbb{A}^n(K)$ we indicate K^n seen as an affine space on itself.

Proposition 20.11. *Let E be a K-vector space of dimension n and X an affine space on E. There exists an affine isomorphism $f : X \to \mathbb{A}^n(K)$ i.e. f is a bijective affine map such that f^{-1} is an affine map.*

Proof. Let $(Q, \mathcal{B})$, $\mathcal{B} = (e_1, \ldots, e_n)$ be an affine coordinate system on X. If $P \in X$ has coordinates $(\lambda_1, \ldots, \lambda_n)$ we set $f(P) = (\lambda_1, \ldots, \lambda_n)$. The map f is bijective ($f^{-1}(\alpha_1, \ldots, \alpha_n)$ is the unique point P' such that $P' = Q + w$, with $w = \alpha_1 e_1 + \cdots + \alpha_n e_n$). The associated linear map is $g : E \to K^n : w \to (\alpha_1, \ldots, \alpha_n)$, g is clearly bijective (it transforms the basis $\mathcal{B}$ into the canonical basis of K^n).

We verify that f is affine: if $(\lambda_i), (\lambda_i')$ are the coordinates of P, P', we have $P - P' = \sum_i (\lambda_i - \lambda_i')e_i$ and the condition $f(P) - f(P') = g(P - P')$ is verified. In the same way we verify that f^{-1} is affine (take the standard coordinate system in K^n i.e. the origin in 0 and the canonical basis). □

20.3. Group actions

Let G be a group and X a set. An *action* of G on X is a morphism of groups

$$\psi : G \to \mathcal{S}(X)$$

where $\mathcal{S}(X)$ is the group of permutations of X.

The action is *transitive* if $\forall (P, Q) \in X^2$ there exists a $g \in G$ such that $\psi(g)(P) = Q$ (X is said to be a *homogeneous* space under G); furthermore, if the element g is unique the action is said to be *simply transitive*.

Let X be an affine space on the K-vector space E. Consider

$$\psi : (E, +) \to \mathcal{S}(X) : v \to t_v.$$

The map is well-defined since translation is bijective (Lemma 20.5). The map ψ is a group morphism due to (A1) and the action is simply transitive due to (A2).

Exercises

Exercise 20.12. Let X be a set and G a group acting on X via the group morphism $\varphi : G \to \mathcal{S}(X)$. If $x \in X$ we will write $\varphi(g)(x) = g(x)$. The *orbit* of the point x is

$$O_x = \{g(x) \mid g \in G\}.$$

1) Show that the set of orbits is a partition of X. In other words, the relation $xRy \Leftrightarrow \exists g \in G$ such that $g(x) = y$ is an equivalence relation and the equivalence classes are the orbits. The quotient set is denoted by X/G.
2) If (X, E, φ) is an affine space, describe the quotient set X/G, where $G = (E, +)$.

Exercise 20.13. Let G be a group, X a set and $\varphi : G \to \mathcal{S}(X)$ an action of G on X. Suppose that the action is transitive i.e. $\forall (x, y) \in X^2$, $\exists g \in G$ such that $g(x) = y$.
1) Describe the quotient set X/G (Exercise 20.12).
2) Show that if φ is injective and G is abelian, the action is simply transitive i.e. $\forall (x, y) \in X^2$, $\exists! g$ such that $g(x) = y$.

Exercise 20.14. Let E be a K-vector space and $f : E \to E$ an affine map. A point $p \in E$ is fixed for f if $f(p) = p$.

(1) Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \to (x + y + 1, y - z + 2, x - y - z - 1)$. Determine the set of fixed points of φ.

(2) If $f : E \to E$ is an affine map, show that the set of fixed points of f, if not empty, is an affine subspace of E.

(3) Show that f has a unique fixed point if and only if 1 is not an eigenvalue of the linear part of f.

Exercise 20.15. Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ be an affinity.

(1) Let $F(\varphi) = \{x \mid \varphi(x) = x\}$ be the set of fixed points of φ. Show that $F(\varphi)$, if non-empty, it is an affine subspace of $\mathbb{R}^3$.

(2) Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \to (x + y + z - 1, 3y + 2z - 2, -y + 1)$. Show that φ is an affinity and determine $F(\varphi)$.

(3) Let φ be an affinity of $\mathbb{R}^3$. Show that if $F(\varphi)$ is a plane, then the characteristic polynomial of $L(\varphi)$, the linear part of φ, has all its roots in $\mathbb{R}$. Conclude that if $\det(L(\varphi)) \neq 1$ then $L(\varphi)$ is diagonalizable.

(4) Let φ be an affinity of $\mathbb{R}^3$ such that $F(\varphi)$ is a plane, H. Show that if $\det(L(\varphi)) \neq 1$, then there exist infinitely many lines, D, such that $D \not\subset H$ and $\varphi(D) = D$ (these lines are globally invariant or stable but not fixed point by point).

Exercise 20.16. Consider two skew planes in $\mathbb{R}^4$, E, F. Show that there exists an affinity, f, of $\mathbb{R}^4$ such that $f(E) = F$ and $f(F) = E$.

Part 4

Euclidean geometry

Euclidean geometry is a mathematical system attributed to ancient Greek mathematician *Euclid*, which he described in the *Elements*. In this geometry theorems are derived from a small number of simple axioms. In the beginning of the first book of the Elements, Euclid gives the following five postulates (axioms) for plane geometry:

(i) Through every pair of points in the plane there passes a straight line.

(ii) It is possible to extend a line segment continuously in a straight line.

(iii) It is possible to describe a circle with any center and radius.

(iv) All right angles are equal to one another.

(v) If a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than two right angles.

Although Euclid explicitly only asserts the existence of the objects, in his reasoning he also implicitly assumes them to be unique.

In 1795, *John Playfair* $(1748 - 1819)$ offered an alternative version of the fifth postulate. This alternative version gives rise to the identical geometry as Euclid's. Playfair's version of the fifth postulate reads as follows:

(v') Given a line L and a point p not lying on L there is a unique line L' passing through p which does not meet L.

The modern version of Euclidean geometry is the theory of Euclidean spaces, where distance and angles are measured by a suitable generalizations of Pythagoras theorem.

Chapter 21

Bilinear forms

Bilinear forms are a special case of multilinear maps. Symmetric bilinear forms are particularly important in geometry because they allow us to define distances, orthogonality and angles.

Definition 21.1. Let E be a K-vector space. A map $f : E \times E \to K$ is a bilinear form if for every $x \in E$ the partial maps $f_{x,} : E \to K : y \to f(x, y)$, $f_{,x} : E \to K : y \to f(y, x)$ are linear.

Definition 21.2. A bilinear form $f : E \times E \to K$ is

(1) Symmetric if $\forall\, (x, y) \in E^2$: $f(x, y) = f(y, x)$.

(2) Antisymmetric if $\forall\, (x, y) \in E^2$: $f(x, y) = -f(y, x)$.

(3) Alternating if $\forall\, x \in E$: $f(x, x) = 0$.

Remark 21.3. We have already seen that an alternating form is antisymmetric. If $\mathrm{char}(K) \neq 2$, an antisymmetric form is alternating. In fact, from $f(x, x) = -f(x, x)$ we obtain $2f(x, x) = 0$ and since 2 is invertible in K it follows that $f(x, x) = 0$. Instead, if the $\mathrm{char}(K) = 2$ then $-1 = 1$ and an antisymmetric form is symmetric.

The set $\mathrm{Bil}(E)$ of bilinear forms on E is a K-vector space. The set $S^2(E)$ of symmetric bilinear forms is a subvector space of $\mathrm{Bil}(E)$.

Let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E and $f : E \times E \to K$ a bilinear form. Let $v = x_1e_1 + \cdots + x_ne_n$, $u = y_1e_1 + \cdots + y_ne_n$ be two vectors of E. Using the bilinearity of f we have

$$f(v, u) = f(x_1e_1 + \cdots + x_ne_n, u) = x_1f(e_1, u) + \cdots + x_nf(e_n, u).$$

Furthermore, $f(e_k, u) = f(e_k, y_1e_1 + \cdots + y_ne_n) = \sum_j y_j f(e_k, e_j)$. In conclusion

$$f(v,u) = \sum_{1\leq i\leq n,\, 1\leq j\leq n} x_i y_j f(e_i, e_j). \tag{21.1}$$

It turns out that f is completely determined by the n^2 scalars $f(e_i, e_j)$.

Definition 21.4. The matrix associated to the bilinear form f with respect to the basis $\mathcal{B}$ of E is the $n \times n$ matrix $mat_{\mathcal{B}}(f) = (f(e_i, e_j)), 1 \leq i \leq n, 1 \leq j \leq n$.

We can now express the relation (21.1) in matrix form. Set

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

then

Lemma 21.5. *With the previous notations:*
1) $f(v,u) = {}^tXAY$, *where* $A = mat_{\mathcal{B}}(f)$.
2) If $M \in M_n(K)$ *the map* $g : E \times E \to K : (v,u) \to {}^tXMY$ *is a bilinear form on* E *such that* $mat_{\mathcal{B}}(g) = M$.

Proof. 1) The product of matrices is associative therefore $({}^tXA)Y = {}^tX(AY)$. By definition AY is the matrix with coefficients $\alpha_i = y_1 f(e_i, e_1) + \cdots + y_n f(e_i, e_n)$. Now, ${}^tX(AY)$ is the scalar product of the row ${}^tX$ with the column AY and we find the expression (21.1).
2) It follows from the properties of the matrix product. □

Proposition 21.6. *Let* $\mathcal{B}$ *be a basis of the* K*-vector space* E. *The map* $mat_{\mathcal{B}} : \mathrm{Bil}(E) \to M_n(K) : f \to mat_{\mathcal{B}}(f)$ *is an isomorphism of* K*-vector spaces.*

The image of $S^2(E)$ *is the subspace of symmetric matrices and the image of* $\mathcal{A}^2(E)$ *is the subspace of antisymmetric matrices.*

Proof. The map $mat_{\mathcal{B}}$ is clearly linear, it is injective because with the previous notations f is completely determined by the scalars $f(e_i, e_j)$ and, for the same reason it is surjective ((2) of Lemma 21.5).

The form f is symmetric if and only if $f(e_i, e_j) = f(e_j, e_i), \forall\, i, j$ (same thing for antisymmetric forms but with a minus), i.e. if and only if $mat_{\mathcal{B}}(f)$ is symmetric (resp. antisymmetric). □

Obviously, the isomorphism $mat_{\mathcal{B}}$ depends on the basis $\mathcal{B}$ and the matrices representing the same bilinear form in different bases are generally different. Let us see what the relation between two such matrices is.

Let $\mathcal{B} = (e_i)$, $\mathcal{B}' = (e'_i)$ be two bases of E, and (x_i) (resp. (x'_i)) the coordinates of the vector $u \in E$ in the basis $\mathcal{B}$ (resp. $\mathcal{B}'$). Similarly, we indicate with $(y_i), (y'_i)$ the coordinates of the vector v in the bases $\mathcal{B}, \mathcal{B}'$.

Let $f : E \times E \to K$ be a bilinear form with $A = mat_{\mathcal{B}}(f)$, $A' = mat_{\mathcal{B}'}(f)$. Finally, let $P = \mathrm{Mat}(Id_E; \mathcal{B}, \mathcal{B}')$. We know that $PX = X', PY = Y'$.

We have $f(u, v) = {}^tX'A'Y' = {}^t(PX)A'(PY) = {}^tX({}^tPA'P)Y = {}^tXAY$. Since this is true for every X, Y we conclude that ${}^tPA'P = A$ (in fact f is determined by the scalars $f(e_i, e_j)$). We used the fact that ${}^t(AB) = {}^tB\,{}^tA$ (Exercise 21.33). We observe that since P is invertible we also have: $A' = {}^t(P^{-1})AP^{-1}$ (because $({}^tP)^{-1} = {}^t(P^{-1})$, Exercise 21.33). We proved:

Proposition 21.7. *Two matrices $A, A' \in M_n(K)$ represent the same bilinear form if and only if there exists an $n \times n$ invertible matrix, P, such that ${}^tPA'P = A$.*

Definition 21.8. Two matrices $A, A' \in M_n(K)$ are *congruent* if there exists an invertible matrix $P \in M_n(K)$ such that ${}^tPA'P = A$. In this case we write $A \equiv A'$.

Lemma 21.9. *The congruence relation $\equiv$ is an equivalence relation on the set $M_n(K)$.*

Proof. Exercise 21.34. □

Lemma 21.10. *Two congruent matrices have the same rank.*

Proof. In fact two congruent matrices are equivalent. If ${}^tPA'P = A$, since ${}^tP$ and P are invertible A and A' are equivalent, therefore they have the same rank. □

It follows from this lemma that the rank of the matrix representing a given bilinear form f does not depend on the basis but only on f. This justifies the following:

Definition 21.11. Let $f : E \times E \to K$ be a bilinear form. The *rank* of f is the rank of the matrix $mat_{\mathcal{B}}(f)$ where $\mathcal{B}$ is any basis of E.
The form is said to be *non-degenerate* if it has maximum rank that is rank equal to $\dim(E)$, otherwise f is said to be *degenerate*.

21.1. Symmetric bilinear forms and quadratic forms

Let $f : E \times E \to K$ be a symmetric bilinear form. By definition $f(u, v) = f(v, u), \forall\, (u, v) \in E^2$ and the matrix $mat_{\mathcal{B}}(f)$ which represents f with respect to any basis $\mathcal{B}$ is symmetric. We observe that if A is symmetric and P is invertible then ${}^tPAP$ is also symmetric (it could not be otherwise since the fact of being symmetric does not depend on a basis).

We now define a map $q : E \to K$ via $q(v) := f(v, v)$.

Definition 21.12. The map q is the *quadratic form* associated to f.

Remark 21.13. The quadratic form q satisfies the following properties:

Q1 $\forall\, v \in E, \forall\, \lambda \in K$: $q(\lambda v) = \lambda^2 q(v)$;

Q2 $\forall\, (u, v) \in E^2$: $q(u + v) - q(v) - q(u) = 2f(u, v)$;

Q'2 In particular, the map $E \times E \to K : (u, v) \to q(u + v) - q(u) - q(v)$ is a symmetric bilinear form.

Definition 21.14. Let E be a K-vector space. A quadratic form on E is a map $q : E \to K$ that satisfies (Q1) and (Q'2).

Lemma 21.15. *Let E be a K-vector space. If* $\mathrm{char}(K) \neq 2$ *there is a one-to-one correspondence between $S^2(E)$ and $Q(E)$, the set of quadratic forms on E.*

Proof. The correspondence $\varphi : S^2(E) \to Q(E)$ is given by $\varphi(f) = q$ where $q(v) := f(v, v)$. We have $\varphi^{-1}(q) = f$ with $f(u, v) = \frac{q(u+v)-q(u)-q(v)}{2}$. □

Remark 21.16. Lemma 21.15 is false in characteristic two.

The theory of symmetric bilinear forms and quadratic forms is very different depending on whether $\mathrm{char}(K) = 2$ or $\mathrm{char}(K) \neq 2$. From now on we will always assume $\mathrm{char}(K) \neq 2$ (and every now and then, in certain sentences where this assumption is essential, we will remember it). Under the hypothesis $\mathrm{char}(K) \neq 2$ we have (Lemma 21.15) a perfect correspondence between symmetric bilinear forms and quadratic forms.

Remark 21.17. We remember what the characteristic of a field K is. We have $f : \mathbb{Z} \to K : n \to 1_k + \cdots + 1_k$ (n terms, $n > 0$; $f(0) = 0$ and $f(-n) = -f(n)$). If f is injective, we say that K has characteristic 0 (in this case K contains a field isomorphic to $\mathbb{Q}$, consider the inverses in K of the $f(n)$). If f is not injective the kernel of the ring morphism f is an ideal, I, of $\mathbb{Z}$. Every ideal of $\mathbb{Z}$ is of the form $a\mathbb{Z}$. By the factorization theorem for morphisms of groups, rings and so on, the image $\mathrm{Im}(f) \subset K$ is isomorphic to $\mathbb{Z}/a\mathbb{Z}$. So $\mathrm{Im}(f)$ is a finite, integral ring (since $\mathrm{Im}(f) \subset K$). It follows that $\mathrm{Im}(f)$ is a field, in particular a is a prime number.

In conclusion, the characteristic of a field, if not zero, is always a prime number p. We have $px = x + \cdots + x = 0, \forall x \in K$.

21.2. Quadratic forms and homogeneous polynomials of degree two

Let us remember the definition of homogeneous polynomial.

Definition 21.18. A polynomial $P \in K[x_1, \ldots, x_n]$ in n variables, with coefficients in K is *homogeneous of degree* d if

$$P(x_1, \ldots, x_n) = \sum a_{i_1 \ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$$

with $i_1 + \cdots + i_n = d$, for every $a_{i_1 \ldots i_n} \neq 0$. In other words P is a sum of monomials of degree d.

Remark 21.19. 1) In particular, if P is homogeneous of degree d then

$$P(\lambda x_1, \ldots, \lambda x_n) = \lambda^d . P(x_1, \cdots, x_n).$$

2) A homogeneous polynomial of degree 2 in the variables $x_1, \ldots, x_n$ is written as

$$P(x_1, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j.$$

If $i = 1$ there are n possible values for j. If $i = 2$, there are $n - 1$. In general there are $n - (i - 1)$ possible values of j. So by varying i from 1 to n we have

$$n + (n-1) + \cdots + 1 = \frac{n(n+1)}{2}$$

possible values for j, i.e. P is determined by $\frac{n(n+1)}{2}$ coefficients a_{ij}. We thus see that the K-vector space $K[x_1, \ldots, x_n]_2$ of homogeneous polynomials of degree two has dimension

$$\dim(K[x_1, \ldots, x_n]_2) = \frac{n(n+1)}{2}.$$

Compare it with the dimension of $S_n(K)$, the subspace of $n \times n$ symmetric matrices.

Let us go back to quadratic forms. Let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E and $v = x_1 e_1 + \cdots + x_n e_n$. By definition

$$q(v) = f(v, v) = {}^t X A X = \sum_{1 \leq i \leq n,\ 1 \leq j \leq n} a_{ij} x_i x_j$$

so $q(v)$ can be seen as a homogeneous polynomial of degree two in the coordinates of v. Let us try to be more precise. Let p, t be such that $1 \leq p < t \leq n$. In the sum we have a term $a_{pt} x_p x_t$ and a term $a_{tp} x_t x_p$. But the matrix A is symmetric so $a_{pt} = a_{tp}$ and we can rewrite

$$q(v) = \sum_{1 \leq i \leq n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j.$$

Now let $Q(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ be the homogeneous polynomial of degree two

$$Q(x_1, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} q_{ij} x_i x_j$$

with $q_{ii} = a_{ii}$, $q_{ij} = 2a_{ij}$ if $i \neq j$. We have that $q(v)$ is the scalar $Q(x_1, \ldots, x_n)$, where the x_i are the coordinates of v in the basis $\mathcal{B}$. Under these conditions we say

that the polynomial Q represents the quadratic form q in the basis $\mathcal{B}$. Conversely, let

$$Q(x_1, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} b_{ij} x_i x_j.$$

We set $a_{ii} = b_{ii}$ and $a_{ij} = a_{ji} = \frac{b_{ij}}{2}$ if $i < j$ ($\mathrm{char}(K) \neq 2$). The matrix $A = (a_i j)$ is symmetric and $Q(X) = {}^tXAX$. Therefore, once a basis of E is chosen, if $\mathrm{char}(K) \neq 2$, homogeneous polynomials of degree two, quadratic forms, symmetric bilinear forms and symmetric matrices are avatars of the same thing.

21.3. Orthogonality with respect to a symmetric bilinear form

We introduce the notion of orthogonality with respect to a symmetric bilinear form and make the connection with the orthogonality seen in linear algebra (duality).

Definition 21.20. Let E be a K-vector space and $f : E \times E \to K$ a symmetric bilinear form. A vector v is *f-orthogonal* to a vector u if $f(u, v) = 0$. By symmetry $f(v, u) = 0$ and we say that u and v are f-orthogonal.

Remark 21.21. For simplicity of notation we will say more simply that u and v are orthogonal omitting the form f. We must remember, however, that this notion depends on f. Two vectors that are orthogonal for one form may not be for another.

Definition 21.22. Let $X \subset E$ be a subset of the K-vector space E. We denote with $X^\perp$ the set of vectors orthogonal to each element of X:

$$X^\perp := \{v \in E \mid f(u, v) = 0, \forall\ , u \in X\}.$$

Remark 21.23. 1) It is easily verified that $X^\perp$ is a subspace of E (even if X is not).
2) The subspace $E^\perp = \{v \in E \mid f(u, v) = 0, \forall u \in E\}$ is called the *radical* of f.
3) If $v \in E$ we will write $v^\perp$ instead of $\{v\}^\perp$.

Definition 21.24. A vector $v \in E$ is *isotropic* if $f(v, v) = q(v) = 0$. In other words v is isotropic if v is orthogonal to itself ($v \in v^\perp$).

Definition 21.25. Let F, G be two subspaces of E. We say that F and G are orthogonal if $F \subset G^\perp$.

Remark 21.26. If $F \subset G^\perp$ then $\forall u \in F, v \in G$, $f(u, v) = 0$. So $F \subset G^\perp \Leftrightarrow G \subset F^\perp$.

Let us now move on to the connection with the notion of orthogonality seen in linear algebra.

Let $F \subset E$ be a subspace and f a symmetric bilinear form. We have two orthogonals of F:

- $F^\circ \subset E^*$, $F^\circ = \{\varphi \in E^* \mid F \subset \mathrm{Ker}(\varphi)\}$;
- $F^\perp \subset E$, $F^\perp = \{v \in E \mid f(u,v) = 0, \forall u \in F\}$.

Let $f : E \times E \to K$ be a bilinear form (not necessarily symmetric). For each $u \in E$ the partial maps $f_{u,} : E \to K : v \to f(u,v)$, $f_{,u} : E \to K : v \to f(v,u)$ are linear forms, i.e. $f_{u,}, f_{,u} \in E^*$. Let $\mathcal{B} = (e_1, \ldots, e_n)$ be a basis of E, $\mathcal{B}^*$ the dual basis and $A = mat_{\mathcal{B}}(f)$.

Lemma 21.27. *With the previous notations we have that* $\mathrm{Mat}(f_{\bullet,}; \mathcal{B}, \mathcal{B}^*) = {}^tA$ *and* $\mathrm{Mat}(f_{,\bullet}; \mathcal{B}, \mathcal{B}^*) = A$.

Proof. Let $M = (m_{ij})$ be the matrix of $f_{\bullet,}$ in the bases $\mathcal{B}, \mathcal{B}^*$. By definition m_{ij} is the i-th coordinate in the basis $\mathcal{B}^*$ of $f_{\bullet,}(e_j) = f_{e_j,}$. If $f_{e_j,} = \sum_k m_{kj} e_k^*$, then $m_{ij} = \sum_k m_{kj} e_k^*(e_i) = f_{e_j,}(e_i)$. But $f_{e_j,}(e_i) = f(e_j, e_i) = a_{ji}$. So $m_{ij} = a_{ji}, \forall\, (i,j)$, therefore $M = {}^tA$. In the same way we prove $\mathrm{Mat}(f_{,\bullet}; \mathcal{B}, \mathcal{B}^*) = A$. □

If f is symmetric $f_{\bullet,} = f_{,\bullet} =: f^*$ (and ${}^tA = A$). So for a symmetric bilinear form we have the linear map $f^* : E \to E^* : u \to f_u$, where $f_u : E \to K : v \to f(u,v) = f(v,u)$. Furthermore, if $\mathcal{B}$ is a basis of E and $A = mat_{\mathcal{B}}(f)$, then $\mathrm{Mat}(f^*; \mathcal{B}, \mathcal{B}^*) = A$.

Proposition 21.28. *Let $f : E \times E \to K$ be a symmetric bilinear form and $U \subset E$ a subvector space. Then $f^*(U)^\circ = U^\perp$.*

Proof. By definition $f^*(U) = \{f_u \mid u \in U\}$. So $f^*(U)^\circ = \{v \in E \mid f_u(v) = 0, \forall\, u \in U\} = \{v \in E \mid f(u,v) = 0, \forall\, u \in U\} = U^\perp$. □

The following result will be very useful.

Lemma 21.29. *Let $v \in E$, then $v^\perp$ is a hyperplane if and only if $v \notin \mathrm{Ker}(f^*)$. In particular, if v is not isotropic then $v^\perp$ is a hyperplane and $E = v^\perp \oplus \langle v \rangle$.*

Proof. From Proposition 21.28 we have: $v^\perp = f^*(v)^\circ = \mathrm{Ker}(f_v)$. So $v^\perp = E \Leftrightarrow f^*(v) = 0$, otherwise $v^\perp$ is a hyperplane.

If v is not isotropic $f(v,v) = f_v(v) \neq 0$, then $f_v = f^*(v) \neq 0$ and $v^\perp$ is a hyperplane. We have $v^\perp \cap \langle v \rangle = \{0\}$ since v is not isotropic. □

Another important result:

Proposition 21.30. *Let $f : E \times E \to K$ be a symmetric bilinear form. The rank of f is equal to the rank of f^*. In particular f is non-degenerate if and only if f^* is an isomorphism.*

Proof. It follows immediately from Lemma 21.27 and the definition of rank of f ($\mathrm{rk}(f) := \mathrm{rk}(A) = \mathrm{rk}(f^*)$). □

Remark 21.31. 1) If f is a non-degenerate symmetric bilinear form, the isomorphism $f^* : E \to E^*$ identifies E with its dual (in general there is no *canonical* isomorphism between E and E^*).
2) If f is a non-degenerate bilinear form (not necessarily symmetric), since $\mathrm{rk}(A) = \mathrm{rk}({}^tA)$, we have two isomorphism between E and E^* induced by f ($f_\bullet$, and $f_{,\bullet}$).

In this identification the orthogonality with respect to f corresponds, via f^*, to the orthogonality of linear algebra and we can use the duality results, for example:

Proposition 21.32. *Let $f : E \times E \to K$ be a non-degenerate symmetric bilinear form and let $U \subset E$ be a subspace. Then 1)* $\dim(U^\perp) = \dim(E) - \dim(U)$*;*
2) $(U^\perp)^\perp = U$.

Proof. 1) By Proposition 21.28 $\dim(U^\perp) = \dim(f^*(U)^\circ)$. We know from linear algebra that $\dim(f^*(U)^\circ) = \dim(E^*) - \dim(f^*(U))$. We have $\dim(E) = \dim(E^*)$ and $\dim(f^*(U)) = \dim(U)$ since f^* is an isomorphism.
2) We have $U \subset (U^\perp)^\perp = \{v \in E \mid f(v,u) = 0, \forall\, u \in U^\perp\}$. But by 1) these two subspaces have the same dimension, so they are equal. □

Exercises

Exercise 21.33. 1) Let $A, B \in M_n(K)$, show that ${}^t(AB) = {}^tB\,{}^tA$.
2) If P is invertible show that $({}^tP)^{-1} = {}^t(P^{-1})$.

Exercise 21.34. Prove Lemma 21.9.

Exercise 21.35. Let E be a K-vector space and $f : E \times E \to K$ a symmetric, non-degenerate bilinear form. Let V be a subspace of E and let f' the restriction of f to $V \times V$. Is f' non-degenerate?

Exercise 21.36. Let E be a K-vector space and $f : E \times E \to K$ a symmetric bilinear form. The form f is said to be *anisotropic* if it does not possess non-zero isotropic vectors.

(1) Show that if f is anisotropic then f is non-degenerate. Is the vice versa true?

(2) If f is anisotropic and $U \subset E$ is a subvector space, show that $E = U \oplus U^{\perp}$.

Exercise 21.37. The aim of the exercise is to show that if $A, B \in M_n(K)$, then AB and BA have the same characteristic polynomial.

Let $\mathcal{I}_r \in M_n(K)$ denote the matrix

$$\mathcal{I}_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

where I_r is the $r \times r$ identity matrix and the zeros indicate null matrices of the appropriate size. Take $N \in M_n(K)$ and write

$$N = \begin{pmatrix} N_1 & N_2 \\ N_3 & N_4 \end{pmatrix}$$

where N_1 is an $r \times r$ square matrix, $N_2 \in M_{r,n-r}(K)$, $N_3 \in M_{n-r,r}(K)$ and $N_4 \in M_{n-r}(K)$.

(i) Show that

$$\mathcal{I}_r N = \begin{pmatrix} N_1 & N_2 \\ 0 & 0 \end{pmatrix}$$

and

$$N\mathcal{I}_r = \begin{pmatrix} N_1 & 0 \\ N_3 & 0 \end{pmatrix}$$

(consider linear maps associated to the matrices $\mathcal{I}_r$, N).

(ii) Conclude that $\mathcal{I}_r N$ and $N\mathcal{I}_r$ have the same characteristic polynomial.

(iii) Let $A, B \in M_n(K)$, and r the rank of A. We know that there exist invertible matrices P, Q such that $A = P\mathcal{I}_r Q$. Set $\overline{B} = QBP$ so that $B = Q^{-1}\overline{B}P^{-1}$.

Show that AB is similar to $\mathcal{I}_r\overline{B}$ while BA is similar to $\overline{B}\mathcal{I}_r$. Conclude that AB and BA have the same characteristic polynomial.

Exercise 21.38. 1) Show that

$$Tr : M_n(K) \times M_n(K) \to K : (A, B) \to Tr(AB)$$

is a non-degenerate symmetric bilinear form.
2) Determine $S^\perp$ where S indicates the subspace of symmetric matrices and the orthogonal is taken with respect to Tr.

Exercise 21.39. Let $f : E \times E \to K$ be a symmetric bilinear form and $V \subset E$ a subvector space that does not contain any non-zero isotropic vectors. Show that $E = V \oplus V^\perp$.

Exercise 21.40. Let E be a K-vector space, $\text{char}(K) \neq 2$ and $f : E \times E \to K$ a non-degenerate symmetric bilinear form. A subspace U of E is said to be *totally isotropic* if for every $x \in U$, $f(x, x) = 0$.

(1) Prove that if U is totally isotropic then $\dim(U) \leq \frac{\dim(E)}{2}$.
(2) For each $n \geq 1$, give an example of a non-degenerate symmetric bilinear form on a space of dimension $2n$, with a totally isotropic subspace of dimension n.

Exercise 21.41. Let E be a K-vector space and $f : E \times E \to K$ a *non-degenerate* symmetric bilinear form.

(1) Show that for every $w \in E, w \neq 0$, $w^\perp$ is a hyperplane (remember that $U^\perp = f^*(U)^\circ$, where $f^* : E \to E^* : u \to f_u$).
(2) Show that if $v^\perp = w^\perp$, then v and w are dependent.
(3) Assume $\dim(E) = 3$ and $u \in E$ isotropic ($u \neq 0$). Let $u^\perp = \langle u, v\rangle$. Determine $u^\perp \cap v^\perp$.

Exercise 21.42. Let φ, ψ be two linear forms on the K-vector space E.

(1) Prove that $f : E \times E \to K : (u, v) \to \varphi(u)\psi(v) + \varphi(v)\psi(u)$ is a symmetric bilinear form.
(2) Let $E = \mathbb{R}^2$, and φ, ψ be the linear forms determined by the following conditions: $\varphi((1, 1)) = 1$, $\text{Ker}(\varphi) = \{(x, y) \mid x + 2y = 0\}$, $\psi((1, 1)) = 2$, $\text{Ker}(\psi) = \{(x, y) \mid x + y = 0\}$ (coordinates in the canonical basis $\mathcal{C}$). Determine $mat_{cc}(f)$ (f defined as in (1)). State whether f is degenerate.

Exercise 21.43. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to z^2 - 4xy + 2xz - 2yz.$$

(1) Determine the rank and signature of q using Gauss method.

(2) Show that there are two linearly independent vectors that are isotropic with respect to q.

(3) Let v, w be two linearly independent isotropic vectors. Show that if $f(v, w) = 0$ (f the symmetric bilinear form associated to q), then the orthogonal of $\langle v, w\rangle$ is contained in $\langle v, w\rangle$.

Exercise 21.44. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to -x^2 - 3y^2 - 2z^2 - 2xy + 2xz + 6yz.$$

(1) Determine $A = mat_{\mathcal{C}}(f)$, where $\mathcal{C} = (e_i)$ is the canonical basis and f is the associated symmetric bilinear form. Compute the rank of A.

(2) Determine $e_1^{\perp}$.

(3) Let $v = (0, 1, 1)$. Determine $v^{\perp}$.

(4) Determine the signature of q.

(5) Find the signature of q with Gauss method.

Exercise 21.45. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to x^2 - 2z^2 + 2xy - 2xz.$$

(1) Determine, with Gauss method, the rank and signature of q.

(2) Show that there are three linearly independent isotropic vectors. Are these three vectors two by two orthogonal?

Exercise 21.46. For each $M \in M_n(\mathbb{R})$ we denote by $\Sigma(M)$ the sum of the elements of M. Let

$$S = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

and $\mathcal{A} := \{X \in M_3(\mathbb{R}) \mid X = -\,{}^tX\}$ the subvector space of antisymmetric matrices.

(1) Show that $\varphi : \mathcal{A} \times \mathcal{A} \to \mathbb{R} : (X, Y) \to \Sigma(XSY)$ is a symmetric bilinear form.

(2) Determine the signature of φ.

Chapter 22

Diagonalization of quadratic forms

22.1. Orthogonal bases

Definition 22.1. Let $f : E \times E \to K$ be a symmetric bilinear form on the K-vector space E. A basis $\mathcal{B} = (e_1, \ldots, e_n)$ of E is *f-orthogonal* (or simply orthogonal) if $f(e_i, e_j) = 0$ if $i \neq j$.

Remark 22.2. The basis $\mathcal{B}$ is orthogonal if and only if $\mathrm{Mat}_{\mathcal{B}}(f)$ is diagonal. In fact $\mathrm{Mat}_{\mathcal{B}}(f) = (f(e_i, e_j))$. In this case if q is the associated quadratic form and if $v = \sum x_i e_i$, then $q(v) = \alpha_1 x_1^2 + \cdots + \alpha_n x_n^2$ where $\alpha_i = f(e_i, e_i)$, so the homogeneous polynomial representing f does not have mixed terms $x_i x_j, i \neq j$.

A first important result:

Theorem 22.3. *Let E be a finite-dimensional K-vector space, with* $\mathrm{char}(K) \neq 2$*, and $f : E \times E \to K$ a symmetric bilinear form. Then there exists a basis of E which is f-orthogonal.*

Proof. If f is identically zero the result is clear. We can therefore assume $f \neq 0$. We proceed by induction on $\dim(E)$. The initial case $\dim(E) = 1$ is fine. The essential point is to show the existence of a non-isotropic vector u. In fact, once u has been found, we can write $E = \langle u \rangle \oplus H$, $H = u^{\perp}$ (Lemma 21.29). By induction hypothesis there exists an orthogonal basis, $\mathcal{B}' = (e_1, \ldots, e_{n-1})$, for $f_{|H} : H \times H \to K$. Then $\mathcal{B} = (e_1, \ldots, e_{n-1}, e_n = u)$ is an orthogonal basis for f.

We show the existence of a non-isotropic vector. Since $f \neq 0$ there exist $v, w \in E$ such that $f(v, w) \neq 0$. We have $f(v + w, v + w) = f(v, v) + f(w, w) + 2f(v, w)$.

Since $2f(v,w) \neq 0$ (this is where the hypothesis $\mathrm{char}(K) \neq 2$ is used), one of the three terms $f(v+w, v+w), f(v,v), f(w,w)$ is non-zero. □

Remark 22.4. The assumption $\mathrm{char}(K) \neq 2$ is necessary, the theorem is false if $\mathrm{char}(K) = 2$. Let E be a 2-dimensional K-vector space with $K = \mathbb{Z}/2\mathbb{Z}$ and $\mathcal{B} = (e_1, e_2)$ a basis of E. Let $f : E \times E \to K$ be defined by the matrix

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and $v = \alpha e_1 + \beta e_2$ be a vector of E. We have $f(v,v) = 2\alpha\beta = 0$. If $\mathcal{C} = (w_1, w_2)$ were an orthogonal basis we would have $f(w_1, w_2) = 0$ and as seen above $f(w_i, w_i) = 0$, therefore $\mathrm{Mat}_{\mathcal{C}}(f)$ would be the zero matrix, in contradiction with $f(e_1, e_2) = 1$.

Another way to formulate the theorem:

Corollary 22.5. *Let K be a field with* $\mathrm{char}(K) \neq 2$. *Every symmetric matrix $A \in M_n(K)$ is congruent to a diagonal matrix.*

Proof. Indeed in an f-orthogonal basis, $\mathcal{B}$, $\mathrm{Mat}_{\mathcal{B}}(f)$ is diagonal. □

Remark 22.6. This does not imply that every symmetric matrix is diagonalizable. Actually, such statement is false. Consider the matrix

$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \in M_2(\mathbb{C}).$$

We have $P_A(x) = x^2$, if A were diagonalizable it would be similar to the null matrix. So A is symmetric and not diagonalizable.

We will see (spectral theorem) that every symmetric matrix with real entries is diagonalizable.

Finally, yet another way of expressing the theorem: In an f-orthogonal basis the associated quadratic form does not contain terms of the type $x_i x_j$ with $i \neq j$. That is, if $v = x_1 e_1 + \cdots + x_n e_n$ ($\mathcal{B} = (e_i)$ an f-orthogonal basis), then $q(v) = \alpha_1 x_1^2 + \cdots + \alpha_n x_n^2$ (where α_i are the terms on the diagonal of $\mathrm{Mat}_{\mathcal{B}}(f)$).

22.2. Orthonormal families and Sylvester's theorem

Let us now move on to the second result. From Theorem 22.3 it follows that there always exists an f-orthogonal basis, the matrix of f in this basis is diagonal with $a_i = a_{ii}$ on the diagonal. We have (after possibly reordering) $a_i \neq 0$ if $1 \leq i \leq r$, where $r = \mathrm{rk}(f)$, $a_i = 0$ if $i > r$. Now, we want to normalize, i.e. replace the non-zero a_i with 1. In other terms we want to write the associated quadratic form as a sum of squares: $q(v) = x_1^2 + \cdots + x_r^2$. It is possible to normalize if the a_i have a square root in K. In particular:

Theorem 22.7. *Let E be a K-vector space and $f : E \times E \to K$ a symmetric bilinear form. We assume K algebraically closed, with* $\mathrm{char}(K) \neq 2$. *Then there exists a basis $\mathcal{B}$ of E such that*

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

where r is the rank of f and where the zeros indicate null matrices of appropriate orders.

Proof. Let $\mathcal{B}' = (e'_i)$ be an orthogonal basis (Theorem 22.3). We set $e_i = \frac{e'_i}{\sqrt{a_i}}$ for $1 \leq i \leq r$, $e_i = e'_i$ if $i > r$. Clearly, $\mathcal{B}$ is an orthogonal basis. We have $f(e_i, e_i) = \frac{f(e'_i, e'_i)}{(\sqrt{a_i})^2} = 1$, $1 \leq i \leq r$. □

Remark 22.8. There are non-algebraically closed fields in which every element has a square root. These are called *quadratically closed field*. For instance, the *field of constructible numbers* is quadratically closed but not algebraically closed. A real number λ is *constructible* if and only if, given a line segment of unit length, a line segment of length $|\lambda|$ can be constructed with compass and straightedge in a finite number of steps. Equivalently, λ is constructible if and only if there is a closed-form expression for λ using only integers and the operations for addition, subtraction, multiplication, division, and square roots. Theorem 22.7 holds more generally when K is quadratically closed.

This normalization process has to do with the basic notion orthonormal family.

Definition 22.9. A family $v_1, \ldots, v_k$ of vectors of E is *orthonormal* if the v_i are two by two orthogonal and if $f(v_i, v_i) = 1, \forall i$ (this, as we will see, means requiring that the v_i have length 1).

Lemma 22.10. *Let $v_1, \ldots, v_k \in E$ be vectors forming an orthonormal family, then $v_1, \ldots, v_k$ are independent.*

Proof. Let $\sum_i \lambda_i v_i = 0$. We have $0 = f(v_k, \sum_i \lambda_i v_i) = \lambda_k f(v_k, v_k)$ (because $f(v_k, v_i) = 0$ if $k \neq i$). Since $f(v_k, v_k) = 1 \neq 0$, we have $\lambda_k = 0$. □

Theorem 22.7 says that if $\mathrm{rk}(f) = r$, then there exists an orthonormal family formed by r vectors (and it is not possible to have more than r).

The ideal situation is when we have an orthonormal basis $\mathcal{B} = (e_i)$. In this case $\mathrm{Mat}_{\mathcal{B}}(f) = I_n$; this is possible if and only if $\mathrm{rk}(f) = n$ (K algebraically closed). In this case if $v = \sum x_i e_i, w = \sum y_i e_i$, $f(v, w) = (x_1, \ldots, x_n)^t (y_1, \ldots, y_n) = x_1 y_1 + \cdots + x_n y_n$. We find the usual scalar product that we used in defining the product of matrices and which we will indicate with $(v \mid w)$.

Remark 22.11. Homogeneous polynomials of degree two (quadratic forms) define *quadrics* (in projective space, in the plane they are called *conics*). Theorem 22.7 says that if K is algebraically closed, the classification of quadrics is given by the rank.

More precisely, if K is algebraically closed, a form of rank r can be transformed to $q(v) = x_1^2 + \cdots + x_r^2$, where the x_i are the coordinates of v.

22.3. Sylvester's theorem

Sylvester's theorem concerns the case $K = \mathbb{R}$. By Theorem 22.3 there is always an orthogonal basis but since there exist real numbers that do not have a real square root (the negative real numbers) we can not always normalize. If $a_i < 0$ then $-a_i > 0$ and $-a_i$ has a square root. We can write $a_i = -(-a_i) = -(\alpha_i^2)$. Resuming the proof of the Theorem 22.7 we have

$$f(\frac{e_i'}{\alpha_i}, \frac{e_i'}{\alpha_i}) = \frac{a_i}{\alpha_i^2} = -1.$$

We can therefore normalize but with $+1$ and -1. However, there is a problem. There are infinite orthogonal bases (review the proof of the Theorem 22.3) and the number of positive terms on the diagonal might vary. Sylvester's theorem says that the number of positive terms does not depend on the orthogonal basis chosen, it only depends on f.

Theorem 22.12 (Sylvester)**.** *Let E be an $\mathbb{R}$-vector space and $f : E \times E \to \mathbb{R}$ a symmetric bilinear form. There exists an integer p ($0 \leq p \leq r := \mathrm{rk}(f)$) such that for each orthogonal basis $\mathcal{B}$ the number of positive terms in the matrix $\mathrm{Mat}_{\mathcal{B}}(f)$ is equal to p. In particular, there exists a basis $\mathcal{B}$ such that*

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where, as usual, the zeros indicate null matrices of appropriate orders. The pair $(p, r-p)$ is called the signature of f.

Proof. It remains to be seen that the number of positive terms is always the same in all orthogonal bases. Let $\mathcal{B} = (e_i)$, $\mathcal{C} = (f_i)$ be two orthogonal bases. Suppose $q(e_i) > 0, 1 \leq i \leq p$, $q(e_i) \leq 0, i > p$ and $q(f_j) > 0, 1 \leq j \leq t$, $q(f_j) \leq 0, j > t$. If $p > t$ we consider $F := \langle e_1, \cdots, e_p \rangle$ and $G = \langle f_{t+1}, \cdots, f_n \rangle$. We have $\dim(F) + \dim(G) = p + (n - t) > n$, because $p - t > 0$. For Grassmann there exists $w \in F \cap G, w \neq 0$. Since $w \in F$, $w = a_1 e_1 + \cdots + a_p e_p$ and $q(w) = a_1^2 q(e_1) + \cdots + a_p^2 q(e_p) > 0$ (note that $a_i \neq 0$ for some i). But w also belongs to G:

$w = b_{t+1}f_{q+1} + \cdots + b_n f_n$ and $q(w) \leq 0$ (since $q(f_j) \leq 0$ if $j > t$). A contradiction. So $p = t$.

As explained before we can normalize in the following way: we take the basis $\mathcal{B}$, for $1 \leq i \leq p$ set $v_i = \frac{e_i}{\sqrt{a_i}}$ (where $a_i = q(e_i)$) and for $p+1 \leq i \leq r$, set $v_i = \frac{e_i}{\sqrt{-q(e_i)}}$. Finally, for $i > r$ we set $v_i = e_i$. Then in the basis (v_i) the matrix of f has the desired form. □

Definition 22.13. Let E be an $\mathbb{R}$-vector space let $f : E \times E \to \mathbb{R}$ a symmetric bilinear form. The form f is:

- positive if $q(v) = f(v,v) \geq 0, \forall\, v \in E$;
- positive definite if $q(v) = f(v,v) \geq 0, \forall\, v \in E$, with $q(v) = 0 \Leftrightarrow v = 0$;
- negative if $q(v) = f(v,v) \leq 0, \forall\, v \in E$;
- negative definite if $q(v) = f(v,v) \leq 0, \forall\, v \in E$, with $q(v) = 0 \Leftrightarrow v = 0$;

otherwise the form is said to be indefinite.

The positive definite forms are the *scalar products* and will play an important role. These definitions are peculiar to the $\mathbb{R}$ field (it makes no sense to talk about negative/positive numbers over $\mathbb{C}$).

22.4. Diagonalization of quadratic forms in practice

In this section we will see how to diagonalize a quadratic forms in practice (in the real case it will also be a matter of determining the signature of the form).

The idea is to resume the proof of the Theorem 22.3. Then we look for a non-isotropic vector v_1 and write $E = \langle v_1 \rangle \oplus H_1$ where $H_1 = v_1^\perp$. Now, we try to apply the previous procedure with H_1 instead of E. Be careful though, it could be that $f_{|H_1} = 0$. But in this case we are done: if $\mathcal{C} = (u_i)$ is any basis of H_1, $\mathcal{C}$ is an orthogonal basis for $f_{|H_1}$ and $\mathcal{B} = (v_1, u_i)$ is an orthogonal basis for f. We can therefore assume that $f_{|H_1} \neq 0$ and in this case there exists a non-isotropic vector $v_2 \in H_1$. We determine $v_2^\perp$ (which is a hyperplane of E) and $H_2 := H_1 \cap v_2^\perp$. We observe that $\dim(H_2) = \dim(H_1) - 1$ since $v_2 \in H_1$ but $v_2 \notin v_2^\perp$ because v_2 is not isotropic. We can write $H_1 = \langle v_2 \rangle \oplus H_2$ and therefore $E = \langle v_1 \rangle \oplus \langle v_2 \rangle \oplus H_2$ (we have $\dim(H_2) = \dim(E) - 2$ because $H_1 \neq v_2^\perp$ since $v_2 \notin v_2^\perp$). We observe that v_1 and v_2 are orthogonal and that each vector of H_2 is orthogonal to both v_1 and v_2. In the case $\dim(E) = 3$ we are finished.

We show by induction that in the general case the process can be completed. Suppose we have $E = \langle v_1 \rangle \oplus \cdots \oplus \langle v_t \rangle \oplus H_t$ with the v_i two-by-two orthogonal and non-isotropic, and with $H_t = v_1^\perp \cap \cdots \cap v_t^\perp$. If $f_{|H_t} = 0$, we take any basis of H_t: $(w_{t+1}, \ldots, w_n)$. Then $\mathcal{B} = (v_1, \ldots, v_t, w_{t+1}, \ldots, w_n)$ is an orthogonal basis of E. In

fact $f(v_i, w_j) = 0$ because $w_j \in H_t = v_1^\perp \cap \cdots \cap v_t^\perp$, $f(w_k, w_j) = 0$ since $f_{|H_t} = 0$ and $f(v_i, v_l) = 0$ being the v_i two-by-two orthogonal. The vectors $(v_1, \ldots, v_t)$ are independent since they are two-by-two orthogonal and not isotropic (Lemma 22.10 and Lemma 22.15 below). Hence the vectors (v_i, w_j) form a basis.

If $f_{|H_t} \neq 0$, we take $w = v_{t+1}$ non-isotropic in H_t. We compute $w^\perp$ and set $H_{t+1} = w^\perp \cap H_t$. We have $H_t = \langle v_{t+1} \rangle \oplus H_{t+1}$ (because w is non isotropic, so $w \notin w^\perp$), and $E = \langle v_1 \rangle \oplus \cdots \oplus \langle v_{t+1} \rangle \oplus H_{t+1}$.

We will inevitably end up with $E = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$, where the v_i two-by-two orthogonal, and not isotropic (therefore Lemma 22.10 and Lemma 22.15 below) or there will be some i such that $f_{|H_i} = 0$. But in this case we conclude by taking any basis of H_i. In any case we found an orthogonal basis.

Example 22.14. Let us see how it works in a concrete case. Let E be a $\mathbb{R}$-vector space of dimension three, $\mathcal{B} = (e_1, e_2, e_3)$ a basis of E and f the symmetric bilinear form defined by

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}.$$

(i) We have $q(e_1) = f(e_1, e_1) = a_{11} = 1$, so e_1 is not isotropic.

(ii) Furthermore

$$e_1^\perp = \left\{ (x, y, z) \mid (x, y, z) A \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0 \right\}$$

where $A = \mathrm{Mat}_{\mathcal{B}}(f)$. From the study of linear maps we know that

$$A \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

is given by the first column of A, so we can directly write

$$e_1^\perp = \{(x, y, z) \mid (x, y, z)\, {}^t(1, 1, -1) = x + y - z = 0\}.$$

To find a basis we give values to x, y, z: if $x = 0, y = z$ and we can take $(0, 1, 1)$. If $y = 0$, we can take $(1, 0, 1) = e_1 + e_3$.

(iii) It follows from (ii) that $a \in e_1^\perp$. The associated quadratic form is

$$q(x, y, z) = x^2 + 2y^2 + 3z^2 + 2xy - 2xz.$$

So $q(a) = 1+3-2 = 2 \neq 0$ and a is not isotropic. We have $A(e_1+e_3) = Ae_1+Ae_3 = {}^t(0, 1, 2)$ (sum of the first column of A with the third). So $a^\perp = \{(x, y, z) \mid y = -2z\}$. A basis is given by e_1 and $(0, -2, 1)$.

(iv) We have $e_1^\perp \cap a^\perp = \{(x,y,z) \mid x+y-z=0, y=-2z\}$, i.e. $x=3z, y=-2z$. In conclusion, $e_1^\perp \cap a^\perp = \langle b \rangle$ where $b=(3,-2,1)$.

By construction the vectors e_1, a, b are two-by-two orthogonal. Be careful, two-by-two orthogonal vectors are not necessarily independent.

Lemma 22.15. *Let $v_1, \ldots, v_k \in E$ be two-by-two orthogonal vectors. We also assume v_i is non-isotropic, $1 \leq i \leq k$. Then the v_i are independent.*

Proof. It comes from the proof of the Lemma 22.10. □

Going back to our example we see that $q(b)=2$. We conclude that (e_1, a, b) is an orthogonal basis. The matrix of f in this basis is diagonal with $1, 2, 2$ on the diagonal. The signature is $(3,0)$ (3 positive terms, 0 negative terms). In particular, f is non-degenerate (positive definite).

22.5. Gauss method for diagonalizing quadratic forms

The Gauss method can be stated as follows:

Theorem 22.16. *Given a quadratic form*

$$q(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

on a K-vector space of dimension n ($\mathrm{char}(K) \neq 2$), it is possible to write q in the form

$$q(x_1, \ldots, x_n) = \sum_{i=1}^{n} \lambda_i L_i(x_1, \ldots, x_n)^2$$

where the L_i are linear, linearly independent forms.

The important point is that the L_i are independent. In fact, let A be the matrix whose rows are the coefficients of the L_i. If the L_i are independent, the matrix A is invertible and if we set

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

then $X_i := L_i(x_1, \ldots, x_n)$ is a change of variables and after this change of variables our form is written as

$$q(X_1, \ldots, X_n) = \sum_{i=1}^{n} \lambda_i X_i^2.$$

The rank is given by the number of non-zero λ_i, the signature (in the real case) by the signs of the λ_i.

The interest of Gauss's method lies more in its proof than in its statement. The method is based on two algebraic identities:

- $a^2 + 2ab = (a+b)^2 - b^2$ (completion of the square);
- $ab = (\frac{a+b}{2})^2 - (\frac{a-b}{2})^2$.

Proof. (of Theorem 22.16) Let us see how it works. We proceed by induction on n. The case $n = 1$ is clear. Let us do the induction step $n-1 \to n$. We distinguish two cases:

(i) one of the a_{ii} is non-zero;

(ii) $a_{ii} = 0, \forall i$.

In case (i) we can assume $a := a_{11} \neq 0$ and write $q = ax_1^2 + 2x_1R + S$, where R is a linear form in $x_2, \ldots, x_n$ and S is a quadratic form in $x_2, \ldots, x_n$. Now, we complete the square:

$$q = a\left(x_1 + \frac{R}{a}\right)^2 - \left(\frac{R}{a}\right)^2 + S.$$

Finally, $q = a(x_1 + \frac{R}{a}))^2 + T$, where $T = S - (\frac{R}{a})^2$. We observe that T is a quadratic form in $x_2, \ldots, x_n$. By induction hypothesis $T = \sum_{i=2}^{n} \lambda_i L_i(x_2, \ldots, x_n)$, with the L_i linearly independent. We conclude that

$$q = \sum_{i=1}^{n} \lambda_i (L_i(x_1, \ldots, x_n))^2$$

with $L_1 = x_1 + \frac{R}{a}, \lambda_1 = a$ and $L_i(x_1, \ldots, x_n) = L_i(x_2, \ldots, x_n)$ if $i > 1$. Since $L_1 \notin \langle L_2, \ldots, L_n \rangle$ (x_1 does not appear in $L_i, i > 1$), the forms $L_1, \ldots, L_n$ are independent.

In case (ii) there are only rectangular terms. We can assume $q = ax_1x_2 + x_1R + x_2S + T$, with $a \neq 0$, R, S linear forms in $x_3, \ldots, x_n$ and with T a quadratic form in $x_3, \ldots, x_n$. We have

$$q = a\left(\left(x_1 + \frac{S}{a}\right)\left(x_2 + \frac{R}{a}\right)\right) - \frac{RS}{a} + T = aD_1D_2 + M$$

where $D_1 = x_1 + \frac{S}{a}$, $D_2 = x_2 + \frac{R}{a}$ and $M = T - \frac{RS}{a}$. We observe that M is a quadratic form in $x_3, \ldots, x_n$. By induction hypothesis $M = \sum \lambda_i L_i^2$, with $L_3, \ldots, L_n$ linearly independent.

Now, we use the second algebraic identity: $D_1D_2 = (\frac{D_1+D_2}{2})^2 - (\frac{D_1-D_2}{2})^2$. Set $L_1 = \frac{D_1+D_2}{2}$, $L_2 = \frac{D_1-D_2}{2}$, so $q = \sum \lambda_i L_i^2$, with $\lambda_1 = \lambda_2 = a$. Furthermore, D_1 and D_2 are linearly independent, so $L_1 = \frac{D_1+D_2}{2}$ and $L_2 = \frac{D_1-D_2}{2}$ are also linearly independent.

Since $\langle L_1, L_2 \rangle \cap \langle L_3, \ldots, L_n \rangle = \{0\}$, the L_i are linearly independent. This concludes the proof. □

Example 22.17. Let us see how it works in practice with the quadratic form of Example 22.14: $q(x, y, z) = x^2 + 2y^2 + 3z^2 + 2xy - 2xz$.

We collect the terms in x: $q = x^2 + 2xy - 2xz + (2y^2 + 3z^2)$, and we complete the square: $x^2 + 2x(y - z) = (x + y - z)^2 - y^2 - z^2 + 2yz$. So

$$q = (x + y - z)^2 - y^2 - z^2 + 2yz + (2y^2 + 3z^2) = (x + y - z)^2 + y^2 + 2z^2 + 2yz.$$

Now, we have to deal with the form $y^2 + 2z^2 + 2yz$ in the two variables y, z. We have $y^2 + 2z^2 + 2yz = (y + z)^2 + z^2$. In conclusion

$$q(x, y, z) = (x + y - z)^2 + (y + z)^2 + z^2.$$

The rank is three and the signature is $(3, 0)$.

This approach is much faster. Be careful though, if not rigorously applied the method can lead to inextricable calculations.

Exercises

Exercise 22.18. Let E be an $\mathbb{R}$-vector space of dimension three, $\mathcal{B} = (e_i)$ a basis of E and $f : E \times E \to \mathbb{R}$ a symmetric bilinear form such that

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}.$$

1) Show that e_1 is not isotropic for f.
2) Determine $e_1^{\perp}$.
3) Note that $a = e_1 + e_3 \in e_1^{\perp}$. Show that a is not isotropic and determine $a^{\perp}$.
4) If b is a non-zero vector of $e_1^{\perp} \cap a^{\perp}$, prove that $\mathcal{B}' = (e-1, a, b)$ is an orthogonal basis of E and write $\mathrm{Mat}_{\mathcal{B}'}(f)$.

Exercise 22.19. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to -x^2 - 3y^2 - 2z^2 - 2xy + 2xz + 6yz.$$

(1) Determine $A = \mathrm{Mat}_{\mathcal{C}}(f)$ where $\mathcal{C} = (e_i)$ is the canonical basis and where f is the associated symmetric bilinear form. Calculate the rank of A.
(2) Determine $e_1^{\perp}$.
(3) Let $v = (0, 1, 1)$. Determine $v^{\perp}$.
(4) Determine an orthogonal basis for f.

Exercise 22.20. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to z^2 - 4xy + 2xz - 2yz.$$

1) Write the matrix, A, of the associated bilinear form in the canonical basis and calculate its rank.
2) Compute $q(e_1 + e_2)$ ((e_i) indicates the canonical basis). Deduce the signature of q (look at A).
3) Find the signature of q using the Gauss method.

Exercise 22.21. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to x^2 - 2z^2 + 2xy - 2xz.$$

(1) Determine, with the Gauss method, the rank and the signature of q.
(2) Show that there are three linearly independent isotropic vectors. Are these three vectors two-by-two orthogonal?

Exercise 22.22. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to -2x^2 - 2y^2 - 4xy + 4xz + 4yz.$$

i) Write the matrix, A, of the associated symmetric bilinear form and calculate its rank.
ii) Determine $E^{\perp} = \mathrm{Ker}(A)$. Observe that $v = e_2 + e_3$ is orthogonal to e_1 and deduce the signature of q.
iii) Find the signature of q with the Gauss method.
iv) Show that there exist three linearly independent vectors u_1, u_2, u_3, with $q(u_i) < 0, \forall i$.

Exercise 22.23. Consider the quadratic form

$$q : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \to 4x^2 + 5y^2 + z^2 - 4xy + 4zy.$$

1) Write the associated matrix and calculate its rank.
2) Calculate $e_1^{\perp}$ ((e_i) indicates the canonical basis).
3) Let $v = (1, 2, 0)$. Compute $v^{\perp}$ and determine an orthogonal basis. Deduce the signature of q.
4) Find the previous result with the Gauss method.

Exercise 22.24. Consider the quadratic form

$$q : \mathbb{R}^4 \to \mathbb{R} : (x, y, z, t) \to -x^2 + 3y^2 + t^2 + 2xy - 2xz - 2yt.$$

1) Write the matrix of the associated symmetric bilinear form.
2) Calculate $e_4^{\perp}$ and $e_1^{\perp}$ ((e_i) is the canonical basis).
3) Let $v = e_1 + e_2 + e_4$, calculate $v^{\perp}$ and then determine the signature of q.
4) Find the signature of q with the Gauss method.

Chapter 23

Euclidean vector spaces

23.1. Metric spaces

In this section we introduce in general terms the notions of distance and norm which will be fundamental in the study of Euclidean spaces.

Definition 23.1. Let X be a set, a *distance* on X is a map

$$d : X \times X \to \mathbb{R}$$

such that

(1) $\forall (x, y) \in X^2, d(x, y) \geq 0$ with equality if and only if $x = y$.

(2) $\forall (x, y) \in X^2, d(x, y) = d(y, x)$.

(3) $\forall (x, y, z) \in X^3, d(x, z) \leq d(x, y) + d(y, z)$ (*triangular inequality*).

Definition 23.2. A *metric space* is a pair (X, d) where X is a set and d is a distance on X.

Example 23.3. On $\mathbb{R}$ we have the distance $d(x, y) = |x-y|$. On $\mathbb{R}^2$ using Pythagoras theorem we have the Euclidean distance

$$d((x, y), (x', y')) = \sqrt{(x - x')^2 + (y - y')^2}.$$

On any set X we have the trivial distance defined by

$$d(x, y) = \begin{cases} 1 \text{ if } x \neq y; \\ 0 \text{ if } x = y. \end{cases}$$

Definition 23.4. Let (X, d) be a metric space and $x_0 \in X$. If $r \in \mathbb{R}$, $r > 0$, the *open disk* with center x_0 and radius r is

$$D(x_0, r) := \{x \in x) < r\}.$$

A subset $U \subset X$ is *open* if $\forall x \in U$, there exists $r > 0$ such that $D(x, r) \subset U$. A subset $F \subset X$ is *closed* if its complement $X \setminus F$ is open.

Remark 23.5. (1) An open disk $D(x_0, r)$ is an open subset.

(2) The empty set is open: by convention or since the implication $\mathcal{P} \Rightarrow \mathcal{Q}$, with $\mathcal{P} = x \in \emptyset$ and $\mathcal{Q} = \exists\, r > 0$ such that $D(x, r) \subset \emptyset$ is always true, being $\mathcal{P}$ always false.

(3) The open subsets of X satisfy the following properties:
- T1 X and $\emptyset$ are open.
- T2 Any intersection of open sets is an open set: if $(U_i)_{i \in I}$ are open, then $\bigcup_{i \in I} U_i$ is open.
- T3 The intersection of a finite number of open sets is an open set.

Definition 23.6. Let Z be a set and $\mathcal{T} = (V_j)_{j \in J}$ a family of subsets of Z. Suppose that $\mathcal{T}$ satisfies the following conditions:

T1 Z and $\emptyset$ are in $\mathcal{T}$.

T2 Any union of elements of $\mathcal{T}$ is still an element of $\mathcal{T}$.

T3 An intersection of a finite number of elements of $\mathcal{T}$ is an element of $\mathcal{T}$.

We then say that $\mathcal{T}$ defines a *topology* on Z and that $(Z, \mathcal{T})$ is a *topological space*; the elements of $\mathcal{T}$ are the open sets of this topology.

So a metric space is a special case of topological space. In a metric space the notion of *limit* can be defined: the sequence $(x_n)_{n \in \mathbb{N}}$ of points of X tends to the point a if every open disk with center a contains all the points of the sequence except at most a finite number, that is if $\forall r > 0$, $\exists n_0$ such that $n \geq n_0 \Rightarrow x_n \in D(a, r)$.

We can also define the notion of *continuous map* between metric spaces.

Definition 23.7. Let $(X, d), (Y, \delta)$ be two metric spaces. A map $f : X \to Y$ is *continuous* if and only if for any open subset $V \subset Y$ the set $f^{-1}(V) \subset X$ is open.

An important notion for what follows:

Definition 23.8. Let $(X, d), (Y, \delta)$ be two metric spaces. A map $f : X \to Y$ is an isometry if f preserves distances, that is $d(x_1, x_2) = \delta(f(x_1), f(x_2))$ for any $(x_1, x_2) \in X^2$.

It is easily seen that an isometry is always injective. If E is a $\mathbb{R}$-vector space one way to define a distance on E is via the notion of *norm*:

Definition 23.9. A norm $||.||$ on E is a map

$$||.|| : E \to \mathbb{R} : v \to ||v||$$

such that

(1) $\forall v \in E$ we have $||v|| \geq 0$, with equality if and only if $v = 0$;

(2) $\forall \lambda \in \mathbb{R}, v \in E$ we have $||\lambda v|| = |\lambda|||v||$;

(3) $||u + v|| \leq ||u|| + ||v||$ (triangular inequality).

In order to define a norm on a K-vector space we need an absolute value notion on K. If $K = \mathbb{R}$ (or $\mathbb{C}$) there is no problem, but in general we do not have an absolute value on an arbitrary field.

We can think of the norm $||v||$ as the length of the vector v. A vector space with a norm is also called a normed vector space. The important fact is:

Lemma 23.10. *Let E be a real vector space with a norm $||.||$. Then*

$$d : E \times E \to \mathbb{R} : (u, v) \to ||u - v||$$

is a distance on E.

Proof. Follows from the properties of norms and the definition of distance. □

Remark 23.11. If E is am $\mathbb{R}$-vector space with a norm $||.||$, the associated distance satisfies $d(x + z, y + z) = d(x, y)$ and $d(\lambda x, \lambda y) = |\lambda| d(x, y)$. Not all distances on E come from a norm. In fact, the trivial distance, for example, does not come from a norm: we have $d(4u, 2v) = 1$ if $4u \neq 2v$, if d came from a norm we would have $1 = ||4u - 2v|| = ||2(2u - v)|| = 2||2u - v|| = 2$, a contradiction.

23.2. Euclidean distance

Throughout this paragraph E will be a vector space over $\mathbb{R}$. Recall that a symmetric bilinear form f on E is positive definite if $\forall v \in E$, $f(v, v) \geq 0$ with equality if and only if $v = 0$.

Definition 23.12. An *Euclidean vector space* is a pair (E, f) where E is a real vector space and f is a positive definite symmetric bilinear form on E. We also say that f is a *scalar product*.

Our goal is to use f to define a norm on E, this will give us a distance. Since f is positive definite there are orthonormal bases for f. We will see that in an orthonormal basis the distance obtained is nothing other than the *Euclidean distance* on $\mathbb{R}^n$.

Lemma 23.13. *Let (E, f) be an Euclidean space. Two vectors $v, w \in E$ are linearly dependent if and only if $f(w, w)v = f(v, w)w$.*

Proof. Suppose v and w are dependent. If $v = 0$ or $w = 0$ the lemma is clear. Let $v = \lambda w$, then $f(v, w) = \lambda f(w, w)$. Since $w \neq 0$, $f(w, w) \neq 0$ (f is positive definite). It follows that $\lambda = \frac{f(v,w)}{f(w,w)}$ and the result follows.

Conversely, if $f(w,w)v - f(v,w)w = 0$ then v, w are dependent except perhaps if $f(w,w) = f(v,w) = 0$. But $f(w,w) = 0$ implies $w = 0$ and, in any case, the two vectors are dependent. □

Proposition 23.14 (Schwarz inequality)**.** *Let (E, f) be an Euclidean space. For every $v, w \in E$ we have*

$$f(v,w)^2 \leq f(v,v)f(w,w),$$

Furthermore, equality holds if and only if v, w are linearly dependent and in this case $f(w,w)v = f(v,w)w$.

Proof. We can assume that v, w are not zero. For each $t \in \mathbb{R}$ let $\varphi(t) = f(v + tw, v + tw)$. We have $\varphi(t) \geq 0, \forall t$. Now $\varphi(t) = f(v,v) + 2tf(v,w) + t^2 f(w,w)$. Since the equation $\varphi(t) = 0$ can not have two distinct roots (one of these would certainly correspond to a non-zero vector), the discriminant Δ is greater than or equal to zero. So $\Delta = 4f(v,w)^2 - 4f(v,v)f(w,w) \leq 0$ and we have our inequality. If $\Delta = 0$, there exists t_0 such that $\varphi(t_0) = 0$, this implies $v + t_0 w = 0$, so v and w are dependent. □

Corollary 23.15. *Let (E, f) be an Euclidean space. The map*

$$||.|| : E \to \mathbb{R} : v \to \sqrt{f(v,v)}$$

is a norm. Also $||u+v|| = ||u|| + ||v||$ if and only if u and v are linearly dependent. In particular

$$d : E \times E \to \mathbb{R} : (u,v) \to ||u - v||$$

is a distance on E.

Proof. We show that $||v + w|| \leq ||v|| + ||w||$ (the other checks are immediate). We have $||v + w||^2 = f(v + w, v + w) = ||v||^2 + ||w||^2 + 2f(v,w)$. By the Schwarz inequality $f(v,w)^2 \leq ||v||^2||w||^2$ (with equality if and only if v, w are dependent). So $f(v,w) \leq ||v||||w||$. Therefore $||v+w||^2 \leq ||v||^2 + ||w||^2 + 2||v||||w|| = (||v|| + ||w||)^2$ or $||v + w|| \leq ||v|| + ||w||$ (with equality if and only if v, w are dependent). □

If f is a scalar product there exist orthonormal bases for f (Sylvester's theorem). If $\mathcal{B} = (e_i)$ is an orthonormal basis, we have $mat_{\mathcal{B}}(f) = I_n$. If $u = \sum x_i e_i$ and $v = \sum y_i e_i$, then $f(u,v) = x_1 y_1 + \cdots + x_n y_n =: (u \mid v)$. So every scalar product f, in an orthonormal basis (or rather f-orthonormal) is expressed as the *usual scalar product.*

In an orthonormal basis we have:

(1) $||u|| = \sqrt{(u \mid u)} = \sqrt{x_1^2 + \cdots + x_n^2}$;

(2) $d(u,v) = ||u - v|| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$;

(3) $x_i = (u \mid e_i)$;

(4) $(x_1y_1 + \cdots + x_ny_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2)$.

Item (4) is the Schwarz inequality which can also be formulated as follows:

$$|(u \mid v)| \leq ||u||||v||.$$

23.3. First properties of Euclidean spaces

In an Euclidean space Pythagoras theorem holds:

Theorem 23.16 (Pythagoras)**.** *Let (E, f) be an Euclidean space. If v, w are two orthogonal vectors of E then $||v + w||^2 = ||v||^2 + ||w||^2$.*

Proof. We have $||v + w||^2 = f(v + w, v + w) = f(v, v) + f(w, w) = ||v||^2 + ||w||^2$ (in fact $f(v, w) = 0$ because the vectors are orthogonal). □

Lemma 23.17. *Let (E, f) be an Euclidean space. If $F \subset E$ is a subvector space of E then $E = F \oplus F^\perp$.*

Proof. This follows from the fact that the form is non-degenerate (so $\dim F^\perp = \dim E - \dim F$) and has no non-zero isotropic vectors so $F \cap F^\perp = \{0\}$. □

With the notations of the previous lemma, each $v \in E$ is written uniquely as $v = f + f'$, with $f \in F, f' \in F^\perp$. The map $\pi_F : E \to F : v \to f$ is linear and it is called the *orthogonal projection* on F.

Proposition 23.18 (Gram-Schmidt)**.** *Let (E, f) be an Euclidean space and let $(v_1, \ldots, v_n)$ be a basis of E. There exists an orthonormal basis of E, $(e_1, \ldots, e_n)$, such that for every p: $f(e_p, v_p) > 0$ and $\langle v_1, \ldots, v_p\rangle = \langle e_1, \ldots, e_p\rangle$.*

Proof. We show by induction on p the existence of an orthonormal family $e_1, \ldots, e_p$ such that $f(e_i, v_i) > 0$ and $\langle e_1, \ldots, e_p\rangle = \langle v_1, \ldots, v_p\rangle$. If $p = 1$, just take $e_1 = \frac{v_1}{||v_1||}$. Suppose we have constructed $e_1, \ldots, e_{p-1}$ with the required properties. We set $F = \langle e_1, \ldots, e_{p-1}\rangle = \langle v_1, \ldots, v_{p-1}\rangle$. We have $E = F \oplus F^\perp$ and $v_p = v + w$ with $v \in F, w \in F^\perp$. Let $e_p = \frac{w}{||w||}$.

We see that e_p satisfies the requests. We have

$$f(e_p, v_p) = \frac{1}{||w||} f(w, v + w) = \frac{1}{||w||} f(w, w) > 0.$$

Clearly, $||e_p|| = 1$ and $f(e_i, e_p) = 0$ if $i < p$ (because $w \in F^\perp$). Now $||w||e_p = w = v_p - v$ with $v \in F$, then $w \in \langle v_1, \ldots, v_p\rangle = F \oplus \langle v_p\rangle$ and $v_p \in \langle e_1, \ldots, e_p\rangle = F \oplus \langle w\rangle$. Therefore $\langle v_1, \ldots, v_p\rangle = \langle e_1, \ldots, e_p\rangle$. □

Corollary 23.19. *In an Euclidean space every family of orthonormal vectors can be completed to an orthonormal basis.*

Proof. Let $u_1, \ldots, u_p$ be our orthonormal family. The vectors u_i are independent and we can complete them to a basis $(v_1, \ldots, v_n)$ ($v_i = u_i$ for $1 \leq i \leq p$). Now, we apply Proposition 23.18 observing that we can take $e_i = u_i$ if $i \leq p$. □

Exercises

Exercise 23.20. 1) Let (X, d) be a metric space. Show that each open subset of X is a union of open disks.
2) Let $(X, d), (Y, \delta)$ be two metric spaces. A map $f : X \to Y$ is continuous at the point x if for every open set $V \subset Y$ containing $f(x)$, there exists an open set $U \subset X$, containing x, such $f(U) \subset V$. Note that V, U can be assumed to be open disks. Observe that in the case of a maps $f : \mathbb{R} \to \mathbb{R}$ we find the usual definition with epsilons and deltas.
3) A map $f : X \to Y$ is continuous if it is continuous at every point $x \in X$. Show that this is equivalent to the following formulation: f is continuous if and only if for every open set $V \subset Y$, $f^{-1}(V)$ is an open set of X.

Exercise 23.21. Let $f : X \to Y$ be an isometry between two metric spaces.
1) Show that f is one-to-one.
2) Show that if f is bijective then f^{-1} is also an isometry.
3) Show that f is continuous.

Exercise 23.22. Prove Lemma 23.10.

Exercise 23.23. If $u = (x, y) \in \mathbb{R}^2$ we set $||u|| = \max\{|x|, |y|\}$. Show that $||.||$ is a norm not coming from a scalar product.

Exercise 23.24. In the Euclidean space $\mathbb{R}^3$ with the usual scalar product (the canonical basis is an orthonormal basis), apply the Gram-Schmidt procedure to the basis $v = (1, 1, 1)$, $v' = (1, 1, 0)$, $v'' = (1, 0, 0)$.

Exercise 23.25. Let v, w be two independent vectors of an Euclidean plane. Determine a real number α such that $v + \alpha w$ has the minimum length among all vectors of the form $v + \beta w, \beta \in \mathbb{R}$. Is the number α uniquely determined?

Exercise 23.26. Let E be an Euclidean space and v_0, w_0 two non-zero vectors of E. Consider the map

$$f : E \to E : v \to (v_0 \mid v)w_0.$$

1) State whether f is linear.
2) Determine the matrix of f with respect to an orthonormal basis. Compute the dimension of $\mathrm{Ker}(f)$.

Exercise 23.27. Let E be an Euclidean vector space. An endomorphism $f : E \to E$ is said to be *positive* if for every $v \in E$ we have $(f(v) \mid v) \geq 0$.
1) Let $W \subset E$ be a subvector space and $p : E \to E$ the orthogonal projection on W (each $v \in E$ is written in a unique way as $v = w + w'$, with $w \in W, w' \in W^{\perp}$,

and $p(v) = w$). Show that p is positive.
2) Prove that for every $u, v \in E$ we have $(p(v) \mid u) = (v \mid p(u))$.
3) Show, with an example, that there exist non-zero endomorphisms $g : E \to E$ such that $(g(v) \mid v) = 0, \forall v \in E$ (begin with the case $\dim(E) = 2$).

Chapter 24

Vector isometries

From now on we will work in orthonormal bases, in such a basis the scalar product f reduces to the standard scalar product $(\cdot \mid \cdot)$. For this reason we will no longer mention f. For example we will simply say that E is an Euclidean (vector) space.

Definition 24.1. Let E be an Euclidean space. An *isometry* $\Phi : E \to E$ is a distance-preserving map i.e. $\forall v, w \in E$, $||\Phi(v) - \Phi(w)|| = ||v - w||$.

We observe that there are isometries that are not linear. For example a translation t_a with $a \neq 0$.

Lemma 24.2. *Let Φ be an isometry of the Euclidean vector space E. If $\Phi(0) = 0$ then $\forall v, w \in E$, $||\Phi(v)|| = ||v||$ and $(\Phi(v) \mid \Phi(w)) = (v \mid w)$.*

Proof. We have $||\Phi(v) - \Phi(0)|| = ||v - 0||$ and the hypothesis $\Phi(0) = 0$ implies $||\Phi(v)|| = ||v||$. Since $||\Phi(v) - \Phi(w)||^2 = ||v - w||^2$, we have $(\Phi(v) - \Phi(w) \mid \Phi(v) - \Phi(w)) = (v - w \mid v - w)$. Expanding

$$||\Phi(v)||^2 + ||\Phi(w)||^2 - 2(\Phi(v) \mid \Phi(w)) = ||v||^2 + ||w||^2 - 2(v \mid w).$$

Since Φ preserves the norm we obtain the statement. □

Lemma 24.3. *Let Φ be an isometry of the Euclidean space E such that $\Phi(0) = 0$, and $\mathcal{B} = (e_i)$ an orthonormal basis. Then $(\Phi(e_i))$ is an orthonormal basis.*

Proof. It follows from the previous lemma, in fact we have $(\Phi(e_i) \mid \Phi(e_j)) = (e_i \mid e_j) = \delta_{ij}$. □

Lemma 24.4. *Let Φ be an isometry of the Euclidean space E. If $\Phi(0) = 0$ then Φ is linear.*

Proof. Let $\mathcal{B} = (e_i)$ be an orthonormal basis and let $v = \sum x_i e_i$. By Lemma 24.3 $(\Phi(e_i))$ is an orthonormal basis, in this basis we have $\Phi(v) = \sum(\Phi(v) \mid \Phi(e_i).\Phi(e_i)$. By Lemma 24.2 $(\Phi(v) \mid \Phi(e_i)) = (v \mid e_i)$. On the other hand $(v \mid e_i) = x_i$. So $\Phi(v) = \sum x_i \Phi(e_i)$ and Φ is linear. □

Hence, an isometry that fixes the origin is a linear map. Such an isometry is called a *vector isometry.*

We observe that a vector isometry is a bijective endomorphism (since an isometry is always injective, and being linear it is also bijective).

The previous result has a very important consequence: Let $f : E \to E$ be an isometry and $a = f(0)$. Set $\Phi = t_{-a} \circ f$. Since translations are isometries and the composition of two isometries is clearly an isometry, Φ is an isometry. We have $\Phi(0) = 0$, so Φ is a linear isometry. In conclusion $f = t_a \circ \Phi$, that is:

Theorem 24.5. *Every isometry of the Euclidean space E is an affinity $t_a \circ \Phi$ whose linear part is a vector isometry.*

In particular an isometry of E is always bijective. The previous theorem shows the way to classify the isometries of E. We must first classify vector isometries and then understand how they can be combined with translations.

Definition 24.6. A matrix $A \in M_n(\mathbb{R})$ is *orthogonal* if ${}^tAA = I_n$ (so A is invertible and $A^{-1} = {}^tA$).

Proposition 24.7. *Let E be an Euclidean vector space and $g : E \to E$ an endomorphism of E. The following are equivalent:*

(1) *g is a vector isometry;*

(2) *g preserves the norm and the scalar product;*

(3) *g transforms orthonormal bases into orthonormal bases;*

(4) *g transforms an orthonormal basis into an orthonormal basis;*

(5) *if $\mathcal{B}$ is an orthonormal basis then* $\mathrm{Mat}(g; \mathcal{B}, \mathcal{B})$ *is an orthogonal matrix.*

Proof. (1) $\Leftrightarrow$ (2): the implication $\Rightarrow$ is Lemma 24.2, for the other implication since g preserves the norm and is linear, we have $||v - w|| = ||g(v - w)|| = ||g(v) - g(w)||$, so g preserves distances.
(2) $\Leftrightarrow$ (3): the implication $\Rightarrow$ is Lemma 24.3. For the other implication let $\mathcal{B} = (e_i)$ be an orthonormal basis. Then if $u = \sum x_i e_i$, $||u||^2 = \sum x_i^2$. Now, $g(u) = \sum x_i g(e_i)$ and since $(g(e_i))$ is an orthonormal basis, $||g(u)||^2 = \sum x_i^2$. So g preserves the norm. We conclude as in the proof of the Lemma 24.2.
(3) $\Leftrightarrow$ (4): One implication is clear. For the other we show (4) $\Rightarrow$ (2) exactly as we showed (3) $\Rightarrow$ (2).

(5) $\Rightarrow$ (3) Let $\mathcal{B}$ be an orthonormal basis and let $A = \mathrm{Mat}(g; \mathcal{B}, \mathcal{B})$. By hypothesis ${}^tAA = I_n$, then the scalar product of the i-th row of ${}^tA$ with the j-th column of A is equal to δ_{ij}. Since the i-th row of ${}^tA$ is the i-th column of A we get $(g(e_i) \mid g(e_j)) = \delta_{ij}$. Therefore $(g(e_i))$ is an orthonormal basis.
(3) $\Rightarrow$ (5) We know that $(g(e_i))$ is an orthonormal basis and conclude as above. □

Lemma 24.8. *Let g be a vector isometry of the Euclidean space E. If λ is an eigenvalue of g then $\lambda = \pm 1$.*

Proof. In fact there exists $v \neq 0$ such that $g(v) = \lambda v$, taking norms $||g(v)|| = |\lambda|||v||$, but g preserves norms so $||g(v)|| = ||v||$. It follows that $|\lambda| = 1$ (note that $||v|| \neq 0$ because $v \neq 0$). □

24.0.1. The group of vector isometries. Let E be an Euclidean vector space. We denote by $O(E)$ the set of vector isometries of E.

Lemma 24.9. *The set $O(E)$ is a group for composition of maps.*

Proof. It is clear that if $f, g \in O(E)$ then $f \circ g \in O(E)$. We have seen that a vector isometry is bijective, we show that f^{-1} is also an isometry. Let $u, v \in E$, we can write $u = f(u'), v = f(v')$. We have $||f^{-1}(u) - f^{-1}(v)|| = ||u' - v'|| = ||f(u') - f(v')||$ (f is an isometry). We conclude that $||f^{-1}(u) - f^{-1}(v)|| = ||u - v||$. Of course the neutral element is Id_E. □

The group $O(E)$ is called the *orthogonal group* of E. We have seen (Proposition 24.7) that if $g \in O(E)$ and if $\mathcal{B}$ is an orthonormal basis, then $A = \mathrm{Mat}(g; \mathcal{B}, \mathcal{B})$ is an orthogonal matrix, i.e. ${}^tA.A = I_n$. From this relation it follows that $\det(A)^2 = 1$. So $\det(g) = \det(A) = \pm 1$.

We will say that a vector isometry g is *positive* (resp. *negative*) if $\det(g) = 1$ (resp. $\det(g) = -1$). We will denote by $O^+(E)$ (resp. $O^-(E)$) the set of positive (resp. negative) vector isometries of E.

The set $O^+(E)$ is a group for the composition of applications while $O^-(E)$ is not. Why?

If $\mathcal{B}$ is an orthogonal basis the isomorphism $\mathrm{Mat}(-; \mathcal{B}, \mathcal{B}) : End(E) \to M_n(\mathbb{R})$ sends $O(E)$ to $O_n(\mathbb{R})$, the group of orthogonal matrices (furthermore $O^{\pm}(E)$ has as its image $O_n^{\pm}(\mathbb{R})$).

24.0.2. Vector isometries of the line and the plane. If $\dim E = 1$ the classification of vector isometries is easily done. In fact, every endomorphism of a space of dimension one is diagonalizable with a single eigenvalue. Therefore the only vector isometries of a space of dimension one are $\pm Id_E$.

This trivial result is important for the sequel. Now, suppose that E is a 2-dimensional Euclidean vector space. Let us start with a lemma:

Lemma 24.10. *Let E be a 2-dimensional Euclidean vector space, and $\mathcal{B} = (e_1, e_2)$ an orthonormal basis of E. Let g be a vector isometry of E and $M = \mathrm{Mat}(g; \mathcal{B}, \mathcal{B})$ with*

$$M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Then:
(a) $ac + bd = 0$ and (b) $a^2 + b^2 = c^2 + d^2 = 1$. Furthermore:
(i) $M \in O_2^+(\mathbb{R}) \Leftrightarrow M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $a^2 + b^2 = 1$;
(ii) $M \in O_2^-(\mathbb{R}) \Leftrightarrow M = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ with $a^2 + b^2 = 1$.

Proof. (a) The first relation follows from the fact that $(g(e_1) \mid g(e_2)) = 0$. The last two follow from $||g(e_1)|| = ||g(e_2)|| = 1$. Another way to proceed is to take the product ${}^tMM$ and equate it to I_2 (the matrix M is orthogonal).

We have the relations

(1) $ac + bd = 0$;

(2) $a^2 + b^2 = 1$;

(3) $c^2 + d^2 = 1$.

If $a = 0$, from 2) we get $b^2 = 1$, therefore $b = \varepsilon$, ($\varepsilon = \pm 1$). From 1) it follows $d = 0$ and from 3) $c = \varepsilon'$ ($\varepsilon' = \pm 1$). In conclusion, if $a = 0$,

$$M = \begin{pmatrix} 0 & \varepsilon \\ \varepsilon' & 0 \end{pmatrix}$$

and $\det(M) = -\varepsilon\varepsilon'$.

We can therefore assume $a \neq 0$. From 1) we have $c = -\frac{bd}{a}$. From 3) $\frac{d^2(a^2+b^2)}{a^2} = 1$. Using 1) gives $d^2 = a^2$, so $d = \varepsilon a$ ($\varepsilon = \pm 1$). We have $c = -\frac{bd}{a}$ hence $c = -\varepsilon b$. In conclusion we have

$$M = \begin{pmatrix} a & -\varepsilon b \\ b & \varepsilon a \end{pmatrix}.$$

Furthermore, $\det(M) = \varepsilon(a^2 + b^2) = \varepsilon$. Now $M \in O_2^+(\mathbb{R}) \Leftrightarrow \varepsilon = 1$ and $M \in O_2^-(\mathbb{R}) \Leftrightarrow \varepsilon = -1$. □

Remark 24.11. From the previous lemma it follows that $O_2^+(\mathbb{R})$ is an abelian group. This is notable because in general $O_n^+(\mathbb{R})$ is not abelian. The geometric description of positive isometries will make this fact clearer.

Let E be an Euclidean space of dimension n and $H \subset E$ a vector hyperplane. We have $E = H \oplus H^\perp$. Each $x \in E$ is written as $x = y + z$ with $y \in H, z \in H^\perp$. Let $r_H : E \to E : x \to y - z$. It is easily seen that r_H is a linear map. We have

$$||x||^2 = (y + z \mid y + z) = (y \mid y) + (z \mid z) = (y - z \mid y - z) = ||r_H(x)||^2$$

since y and z are orthogonal. Therefore, r_H is an isometry. We observe that $r_{H|H} = Id_H$ while $r_{H|H^\perp} = -Id_{H^\perp}$. It follows that $\det(r_H) = -1$, so $r_H \in O^-(E)$. Clearly $r_H \circ r_H = Id_E$.

Definition 24.12. With the previous notations, the isometry r_H is the *orthogonal reflection* with respect to H.

Proposition 24.13. *Let E be an Euclidean plane and $v \in O^-(E)$, then v is an orthogonal reflection with respect to a line.*

Proof. Let $\mathcal{B}$ be an orthonormal basis, from Lemma 24.10

$$\mathrm{Mat}(v; \mathcal{B}, \mathcal{B}) = A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

with $a^2 + b^2 = 1$. Since A is symmetric with ${}^tAA = I_2$ (orthogonality), we get $A^2 = I_2$. Hence, A is diagonalizable with eigenvalues ± 1. Let $\mathcal{B}' = (e'_1, e'_2)$ be an eigenvector basis: $v(e'_1) = e'_1, v(e'_2) = -e'_2$. We see that v is a reflection with respect to the line $\langle e'_1 \rangle$ parallel to $\langle e'_2 \rangle$ ($v(\alpha e'_1 + \beta e'_2) = \alpha e'_1 - \beta e'_2$). It remains to be seen that e'_1 and e'_2 are orthogonal. We have $(v(e'_1) \mid v(e'_2)) = (e'_1 \mid e'_2)$ (the isometry v preserves the scalar product) but we also have $(v(e'_1) \mid v(e'_2)) = (e'_1 \mid -e'_2) = -(e'_1 \mid e'_2)$ (e'_i are eigenvectors). Therefore $(e'_1 \mid e'_2) = 0$. □

Let us now move on to positive isometries.

Definition 24.14. A *rotation* is an isometry belonging to $O^+(E)$.

This unusual definition is a ploy to avoid talking about angles. But a rotation is just what you think. However, with this definition we have:

Proposition 24.15. *Each rotation is the product of two orthogonal reflections with respect to two lines, one of which can be chosen at will.*

Proof. Let $\rho \in O^+(E)$ and let r_L be the orthogonal reflection with respect to the line L, then $\rho \circ r_L \in O^-(E)$ and $\rho \circ r_L = r_D$ i.e. $\rho = r_D \circ r_L$. □

Remark 24.16. Now it is clear that $O^+(E)$ is commutative: two rotations commute between each other.

The Schwarz inequality $|(v \mid w)| \leq ||v|| \cdot ||w||$ implies $-1 \leq \frac{|(v|w)|}{||v|| \cdot ||w||} \leq 1$. From the properties of the cos function it follows that there exists a unique real number θ, such that $0 \leq \theta \leq \pi$ and $\cos\theta = \frac{|(v|w)|}{||v|| \cdot ||w||}$.

Lemma 24.17. *Let E be an Euclidean plane, v and w two non-zero vectors of E with $||v|| = ||w||$. There is a unique rotation, r, such that $r(v) = w$.*

Proof. We can assume $||v|| = ||w|| = 1$ (otherwise consider $\frac{v}{||v||}$, $\frac{w}{||w||}$). Let us consider the line $v^\perp$. There are two vectors $v', v'' \in v^\perp$. In particular $\mathcal{B} = (v, v')$ is an orthonormal basis of E. We have $w = av + bv'$ for suitable scalars a, b. Also $||w|| = a^2 + b^2 = 1$. The rotation r with

$$M := \mathrm{Mat}(r; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

satisfies $r(v) = w$. If r' is a rotation such that $r'(v) = w$ then $N := \mathrm{Mat}(r'; \mathcal{B}, \mathcal{B})$ has the first column equal to that of M. By Lemma 24.10 $M = N$, that is $r = r'$. □

Having said this, we see that $\frac{(v|w)}{||v|| \cdot ||w||} = a = \cos\theta$ where θ is the angle formed by v and w. We would like to consider the matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and say that $r = r_\theta$ is the rotation of angle θ.

However, there is an ambiguity in our construction: there exist two vectors of norm 1 in $v^\perp$ (v' and $-v'$). In the basis $(v, -v')$ the matrix would be

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

The fact is that only a and $|b|$ are completely determined by r ($Tr(r) = 2a$, $\det r = a^2 + b^2 = 1$). To avoid this ambiguity about the sign of b we need to orient the angles.

24.1. Vector isometries of the Euclidean space

Definition 24.18. Let E be an Euclidean space and v an isometry of E. A subspace $F \subset E$ is *fixed* by v if $v_{|F} = Id_F$. The subspace F is *stable*, under v, if $v(F) = F$.

Stated differently, F is fixed if F is contained in the eigenspace relative to the eigenvalue 1. If F is stable we have $v_{|F} : F \to F$, that is, $v_{|F}$ is an isometry of F. If $\dim F = 1$ and if F is stable, then $v_{|F} = \pm Id_F$ and F is an eigenspace.

Lemma 24.19. *Let E be an Euclidean space, v an isometry of E and $F \subset E$ a subspace. If F is stable then $F^\perp$ is also stable.*

Proof. Take $x \in F^\perp$ and $y \in F$. We need to show that $v(x) \in F^\perp$ and $(v(x) \mid y) = 0$. Since F is stable $v_{|F}$ is an isometry of F, in particular it is bijective and we can write $y = v(z), z \in F$. We have $(v(x) \mid y) = (v(x) \mid v(z)) = (x \mid z) = 0$. □

Let us now see a general, very useful result. Let E be an Euclidean space and let f be a vector isometry of E. We set

$$E_+ = \{v \in E \mid f(v) = v\} \text{ and } E_- = \{v \in E \mid f(v) = -v\}.$$

Be careful, one can have $E_+ = \{0\}$ (same thing for E_-). For sure E_+ and E_- are in direct sum and $E_+ \oplus E_-$ is stable under f. So if $W = (E_+ \oplus E_-)^\perp$, we have

$$E = E_+ \oplus E_- \oplus W$$

Furthermore W is stable under f (Lemma 24.19).

Proposition 24.20. *With the previous notations we have that:*

(1) *The dimension of W is even. Furthermore $f_{|W}$ is a positive isometry of W.*

(2) *If f is positive and if $\dim(E)$ is odd then $\dim(E_+) \geq 1$ (i.e. 1 is an eigenvalue and there is an eigenvector for 1).*

(3) *If f is negative, $\dim(E_-) \geq 1$ (i.e. -1 is always an eigenvalue and there is a corresponding eigenvector).*

(4) *If f is negative and if $\dim(E)$ is even then $\dim(E_+) \geq 1$ (i.e. both -1 and 1 are eigenvalues, with corresponding eigenvectors).*

For the proof we will use the following lemma:

Lemma 24.21. *1) Let $P(x) \in \mathbb{R}[x]$ be a polynomial of odd degree, then $P(x)$ has a real root.*
2) If a monic polynomial $P(x) \in \mathbb{R}[x]$ has no real root (therefore $P(x)$ has even degree), then $P(0) > 0$ (the constant term is strictly positive).

Proof. Exercise 24.31. □

Proof. (of Proposition 24.20) 1) Since W is stable $f_{|W} : W \to W$ is an isometry of W. This isometry has no eigenvector. In fact, an eigenvector would be $w \in W$, $w \neq 0$, such that $f(w) = \pm w$, but then we would have $w \in E_+ \oplus E_-$, a contradiction. So $P_{f_{|W}}(x)$ has no real root. Therefore (Lemma 24.21) the degree of $P_{f_{|W}}(x)$ is even and $\dim(W)$ is even: $\dim(W) = 2m$. We have $P_{f_{|W}}(0) = (-1)^{2m} \det(f_{|W}) = \det(f_{|W})$. From point 2) of Lemma 24.21 it follows that $\det(f_{|W}) > 0$, i.e. $\det(f_{|W}) = 1$ and $f_{|W}$ is positive.

So taking bases of E_+, E_-, W, the matrix of f is a block matrix:

$$\begin{pmatrix} I_r & 0 & 0 \\ 0 & -I_p & 0 \\ 0 & 0 & A \end{pmatrix}$$

with the convention that if E_+, E_- or W is trivial, the corresponding matrix does not appear. Furthermore, if $\dim(W) > 0$, $\det(A) = 1$.
2) If f is positive, necessarily $\dim(E_-) = p$ is even (possibly 0). So if $\dim(E)$ is odd, necessarily $\dim(E_+) = r$ is odd, and $\dim(E_+) \geq 1$.
3) If f is negative, since $f_{|E_+}$ and $f_{|W}$ are positive, $\dim(E_-)$ is necessarily odd, so $\dim(E_-) \geq 1$.
4) If f is negative, necessarily $\dim(E_-)$ is odd. So if $\dim(E)$ is even, since $\dim(W)$ is even, we must have $\dim(E_+)$ odd. □

See Exercise 24.32 for a simplified version of this result. We can finally classify the vector isometries of the 3-dimensional space.

Proposition 24.22. *Let E be an Euclidean space with $\dim E = 3$.*
(i) Let v be a positive vector isometry of E. Then $v = Id_E$ or v is a rotation with axis a line.
(ii) Let v be a negative vector isometry of E. Then there exists a line, D, such that $v_{|D} = -Id_D$ and such that $v_{|D^\perp}$ is a rotation.

Proof. (i) From Proposition 24.20, the eigenspace $E_v(1) =: F$ has dimension ≥ 1. Recall that F is fixed. If $\dim(F) = 3$, then v is the identity. Suppose from now on $v \neq Id$. From Lemma 24.19 $v_{|F^\perp}$ is an isometry of $F^\perp$. If $\dim(F) = 2$ then $\dim(F^\perp) = 1$ and $v_{|F^\perp} = \pm Id$. It can not be $v_{|F^\perp} = Id$ because $F^\perp \cap F = \{0\}$ and $F = E_v(1)$. So $v_{|F^\perp} = -Id$. Taking a basis of eigenvectors gives $\det(v) = -1$ against the hypothesis that v is positive. So the only possibility is $\dim(F) = 1$, $\dim(F^\perp) = 2$. Since $\det(v) = \det(v_{|F})\det(v_{|F^\perp})$, we see that $\det(v_{|F^\perp}) = 1$, i.e. $v_{|F^\perp}$ is a rotation. Therefore v is a rotation with axes F.
(ii) From Proposition 24.20, -1 is an eigenvalue of v, let F be the relative eigenspace. If $F = E$, $v = -Id_E$ (and we can take any D). If $\dim F = 2$ we have $E = F \oplus F^\perp$ and the associated matrix in an orthonormal basis $\mathcal{B}$ made with vectors of F, $F^\perp$ is of the form

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & a \end{pmatrix}.$$

Since $F, F^\perp$ are stable, the matrix necessarily has this form. The determinant is a so $a = -1$ but this is a contradiction since we would have $v = -Id$ with $\dim F = 2$.

Suppose $\dim F = 1$. Taking an orthonormal basis with vectors in $F, F^\perp$ the matrix of v is of the form

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$

The determinant is $-(ad - bc)$ so $ad - bc = 1$ and $v_{|F^\perp}$ is a rotation. □

In case (ii) it is customary to distinguish three cases depending on $v_{|F^\perp}$: (a) identity (b) minus identity (c) $\neq \pm Id$. In case (a) v is the orthogonal symmetry with respect to the plane $F^\perp$; in case (b) it is the symmetry with respect to the origin.

24.2. A structure theorem for vector isometries

We now introduce an useful construction.

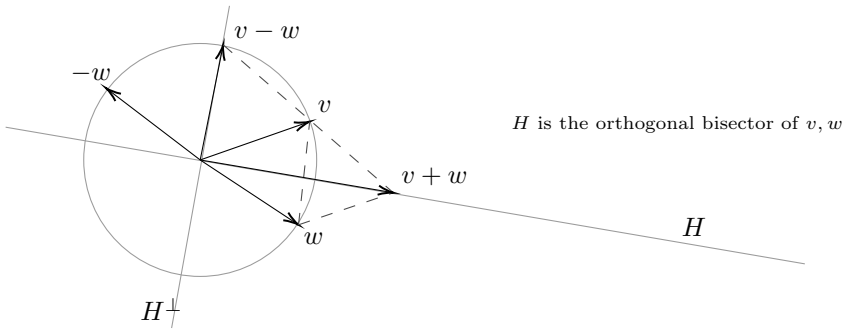
Lemma 24.23. *Let E be an Euclidean vector space of dimension n. Let $v, w \in E$ with $||v|| = ||w||$, and $H = \{x \in E \mid d(x, v) = d(x, w)\}$. Then $H = \{x \mid (x \mid v - w) = 0\} = (v - w)^\perp$, and H is a vector hyperplane.*

Proof. We have $x \in (v - w)^\perp \Leftrightarrow (x \mid v) = (x \mid w)$. Since $d(x, v) = ||x - v|| = (||x||^2 + ||v||^2 - 2(x \mid v))^{1/2}$ and taking into account that $||v|| = ||w||$, we see that $d(x, v) = d(x, w) \Leftrightarrow (x \mid v) = (x \mid w)$. □

Definition 24.24. With the previous notations the hyperplane $H = (v - w)^\perp$ is called the *orthogonal bisector* of v and w.

Be careful, the bisector of v, w is defined only if the two vectors have the same norm.

Remark 24.25. Let $H = (v-w)^\perp$ be the orthogonal bisector of v, w, then $v-w \in H^\perp$. Therefore, if r_H indicates the orthogonal reflection with respect to H, we have $r_H(v - w) = w - v = r_H(v) - r_H(w)$ and $(v + w \mid v - w) = ||v||^2 - ||w||^2 = 0$. So $v + w \in H$ and $r_H(v + w) = v + w = r_H(v) + r_H(w)$. Combining this with the previous result we get $r_H(v) = w$ and $r_H(w) = v$.



Theorem 24.26. *Let E be an Euclidean vector space of dimension $n \geq 2$. Each vector isometry of E can be written as a product of at most n orthogonal reflections.*

We have already seen that the theorem is true for $n = 2$. The theorem will be a consequence of a more precise statement:

Theorem 24.27. *Let E be an Euclidean vector space of dimension $n \geq 2$, and $v \neq Id_E$ a vector isometry of E. If there exists a subvector space $V \subset E$ with $\dim(V) = n - r$ such that $v_{|V} = Id_V$, then v is written as the product of at most r orthogonal reflections with respect to hyperplanes containing V.*

Note that Theorem 24.27 implies Theorem 24.26: for every hyperplane H we have $Id_E = r_H \circ r_H$. If $v \neq Id_E$ just take $V = \{0\}$ in Theorem 24.27.

Proof. (of the Theorem 24.27) We proceed by induction on r. If $r = 1$ there exists a hyperplane H such that $v_{|H} = Id_H$. We have $E = H \oplus H^\perp$. Taking an orthonormal basis of H and completing it to an orthonormal basis $\mathcal{B}$ of E with a vector of norm 1 in $H^\perp$, we see that $\mathrm{Mat}(v; \mathcal{B}, \mathcal{B}) = (1, \ldots, 1, a)$. So $\det g = a = \pm 1$ (an orthogonal matrix has determinant ± 1 since ${}^tAA = I_n$). Since $v \neq Id_E$ we have $a = -1$ and $v = r_H$.

Let us now consider the induction step. We have $E = V \oplus V^\perp$. The subspace V is stable, so $V^\perp$ is also stable (Lemma 24.19). By taking orthonormal bases in V and $V^\perp$ and combining them we obtain an orthonormal basis, $\mathcal{B}$, of E: $\mathcal{B} = (e_i)$ such that $(e_1, \ldots, e_{n-r})$ is a basis of V and $(e_{n-r+1} =: f, \ldots, e_n)$ is a basis of $V^\perp$. Let H be the orthogonal bisector of f and $v(f)$ ($||f|| = ||v(f)||$ since v is an isometry).

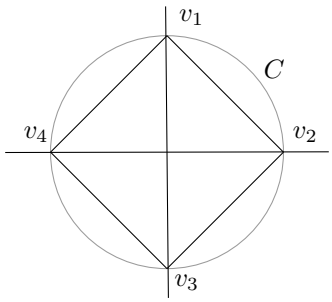
Let $v' = r_H \circ v$. We show that the isometry v' restricted to $V' = V \oplus \langle f \rangle$ is the identity. We observe that $V \subset H$: if $w \in V$, $(w \mid f - v(f)) = 0$ because $f \in V^\perp$ and also $v(f) \in V^\perp$ (because $V^\perp$ is stable under v). This implies $v'_{|V} = Id_V$ (since both v and r_H restricted to V are the identity). Now, $r_H(v(f)) = f$ since H is the bisector of $f, v(f)$. In conclusion $v'_{|V'}$ is the identity.

Since $\dim(V') = (n-r)+1 = n-(r-1)$, by induction hypothesis $v' = r_H \circ v = r_{H_1} \circ \cdots \circ r_{H_k}$, with $k \leq r-1$. Also $V' \subset H_i, \forall i$. It follows that $v = r_H \circ r_{H_1} \circ \cdots \circ r_{H_k}$. Therefore, v is written as a product of at most r reflections with respect to hyperplanes containing V. □

Remark 24.28. For instance, in $\mathbb{R}^3$ we have that a positive isometry is written as the product of two orthogonal reflections. Instead, if f is a negative isometry then either f is an orthogonal reflection, or it is written as the product of three orthogonal reflections.

Exercises

Exercise 24.29. Let $Q \subset \mathbb{R}^2$ be the square of vertices $v_1, \ldots, v_4$, located on the circumference C with center the origin, of radius 1:



Let G be the set of vector isometries of $\mathbb{R}^2$ that send Q into itself:

$$G = \{g \mid g \text{ is a vector isometry and } g(Q) = Q\}.$$

1) Show that G is a group (group of symmetries of the square).
2) If $g \in G$, for every i, $g(v_i)$ is a vertex of Q, i.e. g sends vertices to vertices (reason using the drawing). Two vertices, v_i, v_j are *opposite* if $d(v_i, v_j) = 2$, otherwise they are *consecutive*. If $g \in G$, show that g sends opposite (resp. consecutive) vertices to opposite (resp. consecutive) vertices.
3) If g fixes three vertices, then $g = Id$. If g fixes exactly two vertices, show that these vertices are opposite and conclude that g is the orthogonal symmetry with respect to one of the two axes ($g = s_x$ or $g = s_y$).
4) Let $g \in G$ be a vector isometry which does not fix any vertex. If g is positive it is a rotation. Let r be the rotation that sends v_1 to v_2 (and therefore v_2 to v_3 and so on). Show that the only rotations in G are $r, r^2, r^3, r^4 = Id$.

If g is negative it is an orthogonal symmetry. Considering $g(v_1)$ and using 3) conclude that either g is symmetry with respect to the line $x = y$ ($g = s$) or g is symmetry with respect to the line $x + y = 0$ ($g = \sigma$).

Remark 24.30. It can be shown that every isometry of the square is necessarily a vector isometry, so G is the group of all symmetries of the square. We have

$$G = \{Id, r, r^2, r^3, s, \sigma, s_x, s_y\}.$$

So $\#(G) = 8$ and G is not abelian ($rs(v_1) \neq sr(v_1)$). It is easily seen that $G = D_4$, the dihedral group.

Exercise 24.31. Let $P(x) \in \mathbb{R}[x]$ be a polynomial of odd degree. Considering the limits when x tends to $\pm\infty$ conclude that $P(x)$ has always a real root.

2) Let $P(x) \in \mathbb{R}[x]$. We can consider $P(x)$ as a polynomial with complex coefficients. Show that if $z \in \mathbb{C}$ is a root of $P(x)$ then $\overline{z}$ is also a root of $P(x)$. Indeed, show that if z is a root with multiplicity i, then $\overline{z}$ is also a root with multiplicity i. Conclude that the irreducible factorization in $\mathbb{R}[x]$ of $P(x)$ is of the form:

$$P(x) = \prod Q_i(x)^{a_i} \prod (x - \alpha_j)^{b_j}$$

where the $Q_i(x)$ are irreducible polynomials of degree two and where $\alpha_j \in \mathbb{R}$.
3) Show, in two different ways, that if $P(x) \in \mathbb{R}[x]$ has no real root then the sign of $P(0)$ is the sign of the coefficient of the term with the highest power (so if $P(x)$ is monic, $P(0) > 0$).

Exercise 24.32. Let $n \geq 1$ be an odd integer, and $A \in M_n(\mathbb{R})$ an orthogonal matrix. Show that
(i) if $\det(A) = 1$ then $\det(A - I_n) = 0$ i.e. if A is positive 1 is an eigenvalue (Hint: write $\det(A - I_n) = \det(A - {}^tAA)$).
(ii) if $\det(A) = -1$ then $\det(A + I_n) = 0$ i.e. if A is negative, -1 is an eigenvalue.

Exercise 24.33. We work in $\mathbb{R}^3$ considered as an Euclidean vector space. If H is a vector plane, r_H indicates the orthogonal reflection with respect to H.

(1) Let H be a plane and $L \subset H$ a line. Set $R = H \cap L^\perp$. Show that $\dim(R) = 1$ and that $H^\perp = R^\perp \cap L^\perp$ (note that if $V \subset W$, then $W^\perp \subset V^\perp$). Conclude that $r_{H|L^\perp} = r_R$ (orthogonal reflection in the plane $L^\perp$ with respect to the line R).

Let $v = r_{H_1} \circ r_{H_2} \circ r_{H_3}$ be a vector isometry. Set $L = H_1 \cap H_2 \cap H_3$. The aim of the exercise is to show that if v is not an orthogonal reflection, then the planes H_i are independent i.e. $\dim(L) = 0$.

(2) Show that if $\dim(L) = 2$ i.e. the three planes are equal, then v is an orthogonal reflection.

(3) Suppose $\dim(L) = 1$.
   (a) Show that $v_{|_{L^\perp}} = r_D$ is the orthogonal reflection in the $L^\perp$ plane with respect to a line $D \subset L^\perp$ (Hint: use (1)).
   (b) Conclude that v is the orthogonal reflection with respect to the plane $H = \langle D, L \rangle$.

(4) From (2) and (3) it follows that if v is not an orthogonal reflection then $\dim(L) = 0$. Find this result using a general structure theorem.

Exercise 24.34. Let E be an Euclidean space of dimension n and $\mathcal{B}$ an orthonormal basis of E. If $f \in \mathrm{End}(E)$ with $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$; ${}^tf$ denotes the unique endomorphism of E such that $\mathrm{Mat}({}^tf; \mathcal{B}, \mathcal{B}) = {}^tA$ ($ {}^tf$ is the adjoint of f).

(1) Show that $\forall v, w \in E$ we have $(f(v) \mid w) = (v \mid {}^tf(w))$.

(2) Show that $(f(v) \mid v) = 0, \forall v \in E$ if and only if $f + {}^t f = 0$.

(3) Conclude that an antisymmetric matrix $A \in M_n(\mathbb{R})$ has no real non-zero eigenvalues.

Chapter 25

Euclidean affine spaces

Like any vector space, an Euclidean space E has a natural structure of affine space. Thanks to the scalar product we know how to measure distances and angles. This is a natural framework also because, as we have seen, isometries are affinities.

From now on the term Euclidean space will mean an Euclidean vector space with an affine space structure.

Definition 25.1. The reference system $(Q; e_1, \dots, e_n)$ of the Euclidean space E is said to be orthonormal if $(e_1, \dots, e_n)$ is an orthonormal basis.

Every Euclidean space of dimension n is isomorphic to $\mathbb{R}^n$ and in $\mathbb{R}^n$ we have the standard coordinate system $(O; e_1, \dots, e_n)$ where O is the origin and (e_i) is the canonical basis.

Definition 25.2. Let E be an Euclidean space. Two affine subspaces, F, G, are *orthogonal* if $\mathrm{dir}(F)$ and $\mathrm{dir}(G)$ are orthogonal i.e. $\mathrm{dir}(F) \subset \mathrm{dir}(G)^\perp$ or vice versa. The two subspaces are *perpendicular* if $\mathrm{dir}(F)^\perp \subset \mathrm{dir}(G)$ or vice versa.

In general the two notions are distinct, they coincide if $\dim F + \dim G = \dim E$. In fact in this case $\dim F^\perp = \dim G$ and if one is contained in the other then they are equal.

For example in $\mathbb{R}^3$ the line D of equations $x = y = 0$ (z axis) and the line L of equations $y = z = 0$ (x axis) are orthogonal but not perpendicular.

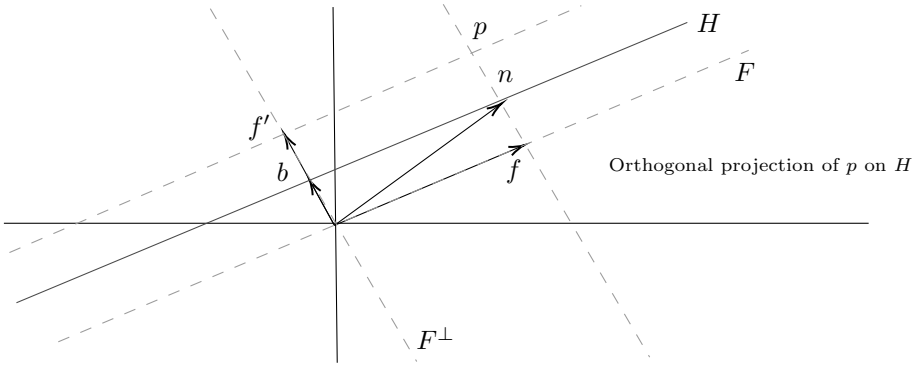
Proposition 25.3. *Let $H = a + F$ be a hyperplane of the Euclidean space E. For each point $p \in E$ there is one and only one line, R_p, passing through p and perpendicular to H. Let $\pi_H : E \to H : p \to R_p \cap H$.*
1) The map π_H is an affine map (orthogonal projection on H).

2) Let $d(p,H) := \inf\{d(p,x) \mid x \in H\}$ *be the distance from* p *to* H*. We have* $d(p,H) = ||p - \pi_H(p)||$.

Proof. Clearly, $R_p = p + F^\perp$, R_p is not parallel to H so the intersection $R_p \cap H$ is a point. This shows that π_H is well defined.

1) Let $b = H \cap F^\perp$, so that $H = b + F$. We have $E = F \oplus F^\perp$ so we can write $p = f + f'$, $f \in F, f' \in F^\perp$. We have $\pi_H(p) = f + b$ (picture below). In fact, the line through p and $f + b$ has direction $\langle p - (f + b) = f' - b\rangle = F^\perp$ (they are two points of the line $F^\perp$; if $f' = b$, $p \in H$ and $\pi_H(p) = p$). We see that $\pi_H = t_b \circ f$ where f is the (orthogonal) projection onto F. So π_H is an affine map.

2) Follows from Pythagoras theorem. □



We now define the orthogonal reflection with respect to an affine hyperplane of an Euclidean space E.

Definition 25.4. Let $H = a + F$ be an affine hyperplane of the Euclidean space E. If $p \in E$ we set $r_H(p) = -p + 2n$ where n is the orthogonal projection of p onto H ($n = \pi_H(p)$). The map r_H is the *orthogonal reflection* with respect to the hyperplane H.

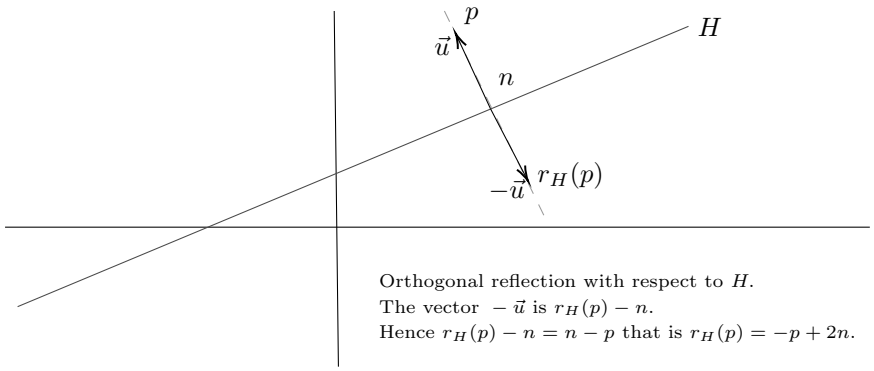
Let us go back to the notations of the previous proof: $b = F^\perp \cap H$ and $p = f + f'$. We have $r_H(p) = -p + 2\pi_H(p) = -f - f' + 2(f + b) = f - f' + 2b$, i.e. $r_H(p) = (t_{2b} \circ r_F)(p)$, where r_F is the orthogonal reflection with respect to the hyperplane F. So r_H is an affinity whose vector part is a vector isometry, therefore r_H is an isometry.

Remark 25.5. If $H = a+F$ we have $r_H = t_e \circ r_F$ where $e = a - r_F(a)$. This follows from the fact that $a - r_F(a) = 2b$. In fact we have $a = f + b$ and $r_F(a) = f - b$, therefore $a - r_F(a) = b + f - (f - b) = 2b$.

For a direct proof we can reason like this: Let us try to write r_H in the form $t_e \circ v$ where v is a vector isometry, since $r_H(H) = H$, we have $v(F) = F$ and

we can set $v = r_F$. We consider the decomposition $E = F \oplus F^\perp$ and write $a = a' + a'', a' \in F, a'' \in F^\perp$. We have $t_e \circ r_F = a' - a'' + e$. Since $r_H(a) = a$, we get $e = 2a'' = a' + a'' - (a' - a'') = a - r_F(a)$.

Let $p = f + f'$ be any point. We observe that $a + f - a' = n$ and $b = a''$, then $r_H(p) = -p + 2n = f - f' + a - (a' - a'') = (t_e \circ r_F)(p)$, with $e = a - r_F(a)$.



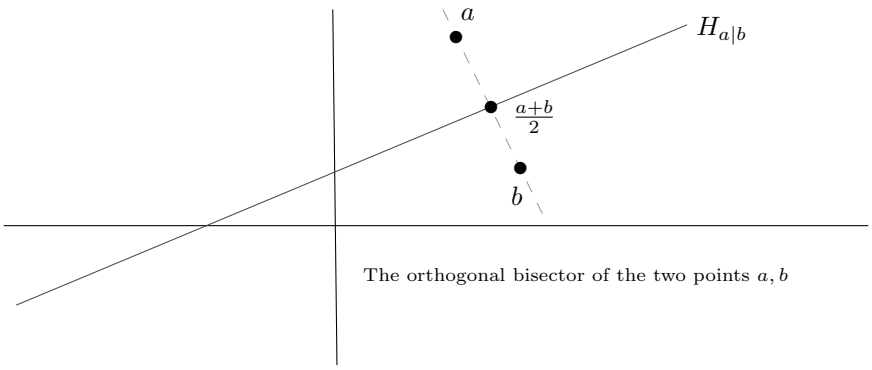
Definition 25.6 (Orthogonal bisector of two points)**.** Let $a \neq b$ be two distinct points of E. We define

$$H_{a|b} := \{p \in E \mid d(a,p) = d(b,p)\}.$$

Then $H_{a|b}$ is an affine hyperplane, more precisely $H_{a|b} = \frac{a+b}{2} + \langle (a-b)^\perp \rangle$. The hyperplane $H_{a|b}$ is the *orthogonal bisector* of a and b.

Lemma 25.7. *Let $a \neq b$ be two distinct points of E, and $H_{a|b}$ their orthogonal bisector. If σ is the orthogonal reflection with respect to $H_{a|b}$ we have $\sigma(a) = b$ and $\sigma(b) = a$.*

Proof. In fact, a point p belongs to $H := H_{a|b}$ if and only if $||p - a|| = ||p - b||$, that is, if and only if $(p - a \mid p - a) = (p - b \mid p - b)$. Hence $p \in H$ if and only if $(a \mid a) - (b \mid b) = 2(p \mid a - b)$. So H is the hyperplane of equation $\varphi(p) = \alpha$ where φ is the linear form $\varphi : E \to k : p \to 2(p \mid a - b)$ and where $\alpha = (a \mid a) - (b \mid b)$. From this description it immediately follows that $\mathrm{dir}(H) = \mathrm{Ker}(\varphi) = (a - b)^\perp$. On the other hand, $\frac{a+b}{2} \in H$ $((\frac{a+b}{2} - a \mid \frac{a+b}{2} - a) = (\frac{a+b}{2} - b \mid \frac{a+b}{2} - b))$. So $H = \frac{a+b}{2} + \langle (a-b)^\perp \rangle$.



The orthogonal projections, $\pi_H(a), \pi_H(b)$ of a, b onto H are equal to $\frac{a+b}{2}$. This can be seen from the picture. However, the perpendicular to H passing through a is $R_a = a + \langle a - b \rangle$. We have $b = a - (a - b) \in R_a$, so $R_b = R_a$ and $\frac{a+b}{2} = a - \frac{a-b}{2} \in R_a$. Hence, $\pi_H(a) = \pi_H(b) = \frac{a+b}{2}$. Finally, $\sigma(a) = -a + 2(\frac{a+b}{2}) = b$ and $\sigma(b) = -b + 2(\frac{a+b}{2}) = a$. □

25.1. A structure theorem for isometries

Using Theorem 24.26 on the structure of vector isometries we can easily prove that any isometry can be written as a product of at most $n + 2$ orthogonal reflections. In fact, it is enough to prove that:

Lemma 25.8. *Any translation can be written as the product of two orthogonal reflections.*

Proof. Let t_u be a translation, $F = u^\perp$, a, b two points such that $a - b = \frac{u}{2}$, and $H = a + F, H' = b + F$. By Remark 25.5 we have $r_{H'}(x) = r_F(x) + b - r_F(b)$. So $r_H(r_{H'}(x)) = r_F(r_F(x) + b - r_F(b)) + a - r_F(a) = x + r_F(b) - b + a - r_F(a)$. Now $r_F(b) = r_F(a) - r_F(\frac{u}{2}) = r_F(a) + \frac{u}{2}$ and we conclude that $r_H(r_{H'}(x)) = x + u = t_u(x)$. □

Corollary 25.9. *Any isometry of an Euclidean space of dimension n can be written as a product of at most $n + 2$ orthogonal reflections of which at most two are not vector reflections.*

Proof. Let g be an isometry, we have $g = t_a \circ v$ where v is a vector isometry. By Theorem 24.26 v is written as a product of at most n orthogonal vector reflections. We conclude by the previous lemma. □

Lemma 25.10. *Let E be an Euclidean space of dimension n and $(a_1, \ldots, a_p)$, $(b_1, \ldots, b_p)$, $p \leq n + 1$ two p-tuples of points such that $d(a_i, a_j) = d(b_i, b_j)$, $\forall i, j$.*

Then there exists an isometry s, product of at most p orthogonal reflections, such that $s(a_i) = b_i$.

Proof. We proceed by induction on p. Let $p = 1$. If $a_1 = b_1$, $s = Id = r_H^0$, any H. If $a_1 \neq b_1$, $s = r_H$ where $H = H_{a_1|b_1}$, the orthogonal bisector of the two points.

Suppose the result is true for $p-1$. So there exists g, product of at most $p-1$ orthogonal reflections, such that $g(a_i) = b_i, 1 \leq i \leq p-1$. If $g(a_p) = b_p$ we are fine, so let us assume $g(a_p) \neq b_p$. For $i < p$ we have $d(g(a_p), b_i) = d(g(a_p), g(a_i)) = d(a_p, a_i)$ (g is an isometry), so $d(g(a_p), b_i) = d(b_p, b_i)$. Therefore $b_1, \ldots, b_{p-1} \in H$, where H is the orthogonal bisector of $g(a_p), b_p$. Consider the isometry $f = r_H \circ g$. It is a product of at most p orthogonal reflections. For $i < p$ we have $f(a_i) = r_H(b_i) = b_i$ (since $b_i \in H$). Then $f(a_p) = r_H(g(a_p)) = b_p$ being H is the orthogonal bisector of $g(a_p), b_p$. □

Theorem 25.11. *Any isometry of an Euclidean space E of dimension n can be written as a product of at most $n+1$ orthogonal reflections.*

Proof. Let f be an isometry and $a_1, \ldots, a_{n+1} \in E$ affinely independent points. Set $b_i = f(a_i)$. We have $d(a_i, a_j) = d(f(a_i), f(a_j)) = d(b_i, b_j)$. By the previous lemma there exists an isometry g, product of at most $n+1$ orthogonal reflections, such that $g(a_i) = b_i$. The two isometries f, g coincide on the a_i, therefore they are equal. This can be seen in the following way: $f(a_i) - f(a_1) = L(f)(a_i - a_1)$, so the linear parts $L(f), L(g)$ coincide on the basis $a_i - a_1, 2 \leq i \leq n+1$ and therefore $f(a_1) = b_1 = g(a_1)$. Hence, $f = g$ (an affinity is determined by its linear part and the image of a point). □

25.2. Classification of plane isometries

Let E be an Euclidean space and $g \in \mathrm{Isom}(E)$ an isometry of E. We have seen that $g = t_a \circ v$ where v is a vector isometry. We will say that g is *direct* or *positive* if v is positive i.e. $\det v = 1$, otherwise we will say that g is *inverse* or *negative*, in this case $\det v = -1$. From now on we will assume $\dim E = 2$.

Let r_θ be the vector rotation of angle θ, $0 \leq \theta \leq \pi$. If $\theta = 0$, r_θ is the identity. Now, let $p \in E, p \neq 0$. The map $r_{p,\theta} := t_p \circ r_\theta \circ t_{-p}$ is the *rotation* of center p, angle θ. As a composition of positive isometries, $r_{p,\theta}$ is a positive isometry. Observe that $r_{p,\theta}(p) = p$, furthermore if $\theta \neq 0$, p is the only fixed point of $r_{p,\theta}$ (if $\theta = 0$, $r_{p,\theta} = Id_E$). In fact $r_{p,\theta}(x) = x$ if and only if $r_\theta(x-p) = x-p$ but a vector rotation other than Id_E has no eigenvector. In fact the associated matrix is of the form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

with $a^2 + b^2 = 1$ and $b \neq 0$. Then the characteristic polynomial is $(x-a)^2 + b^2 > 0, \forall x$. So the only possibility is $x = p$.

Proposition 25.12. *Let $g \neq Id_E$ be a direct isometry of the Euclidean plane E. If g fixes a point, then g is a rotation with the fixed point as center; otherwise g is a non-zero vector translation and has no fixed points.*

Proof. Suppose g fixes a point: $g(p) = p$. Consider $f := t_{-p} \circ g \circ t_p$. The map f is a positive isometry (composition of positive isometries), furthermore $f(0) = 0$. So f is a vector rotation, $f = r_\theta$, where θ is the angle of the rotation. So $g = t_p \circ r_\theta \circ t_{-p}$ is a rotation with center p and angle θ.

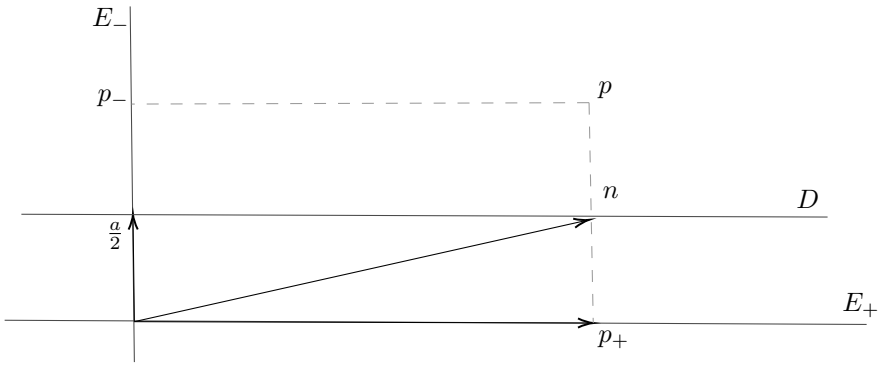
Now, let $g = t_a \circ v$ be a direct isometry of E. If g is not a translation, $v \neq Id_E$ and $f = Id_E - v$ is invertible (1 is not an eigenvalue of a positive vector isometry different from Id). Therefore there exists $x \in E$ such that $f(x) = a$, that is, $x - a = v(x)$. It follows that $g(x) = x$ and g is a rotation with center x. □

Let us now move on to inverse (negative) isometries. Recall that for each affine line L of E we have defined the orthogonal reflection with respect to L.

Definition 25.13. A *glissreflection* of the plane E is the composition, $t_w \circ r_L$, of an orthogonal reflection r_L with a non-zero vector translation w parallel to L.

Let $g = t_a \circ v$ be a negative isometry of E. The vector isometry v is an orthogonal reflection r_L where $L = E_+ = E_v(1)$ is the eigenspace relative to 1. Furthermore, $L^\perp = E_- = E_v(-1)$ and we have $E = E_+ \oplus E_-$. So each $x \in E$ is written uniquely as $x = x_+ + x_-$ and $v(x) = x_+ - x_-$.

Let us see whether g can have fixed points: $g(x) = x \Leftrightarrow x_+ - x_- + a = x_+ + x_- \Leftrightarrow a = 2x_-$. So g has a fixed point if and only if $a \in E_-$. Furthermore, in this case the line $D = \frac{a}{2} + E_+$ is fixed (point by point): $g(\frac{a}{2} + x_+) = -\frac{a}{2} + x_+ + a = \frac{a}{2} + x_+$. Under these conditions $g = r_D$. In fact, let $p = p_+ + p_-$ be any point. The orthogonal projection of p onto D is $n = \frac{a}{2} + p_+$.



So $r_D(p) = -p + 2n = -p_+ - p_- + a + 2p_+ = p_+ - p_- + a = g(p)$. Assume that g has no fixed points $a \notin E_-$ that is $a_+ \neq 0$. The line D is globally invariant: $g(\frac{a}{2} + x_+) = \frac{a_+}{2} - \frac{a_-}{2} + x_+ + a_+ + a_- = \frac{a}{2} + x_+ + a_+$. So $g_{|D} = t_{a_+}$. It follows that the negative isometry $t_{-a_+} \circ g$ fixes the line D point by point. From what we saw before, this implies $t_{-a_+} \circ g = r_D$, i.e. $g = t_{a_+} \circ r_D$ and g is a glissreflection. To sum up:

Proposition 25.14. *Let g be a negative isometry of the plane. If g has a fixed point then g is an orthogonal reflection (and has a line of fixed points). If g has no fixed points, then g is a glissreflection.*

This completes the classification of the isometries of the plane.

Exercises

Exercise 25.15. Let E be an Euclidean vector space. An endomorphism, f, of E is said to be positive if for every v in E we have $(f(v) \mid v) \geq 0$.

(1) Let $W \subset E$ be a vector subspace and $p : E \to E$ the orthogonal projection on W (i.e. $E = W \oplus W^{\perp}$ and if $v = w + w'$, $w \in W$, $w' \in W^{\perp}$, then $p(v) = w$). Show that p is positive.

(2) Prove that $\forall (u, v) \in E^2$ we have $(p(v) \mid u) = (v \mid p(u))$.

(3) Show with an example that there exist non-zero endomorphisms, g, such that $(g(v) \mid v) = 0, \forall v \in E$.

Exercise 25.16. Let E be an Euclidean vector space and v_0, w_0 two non-zero vectors of E. Consider the map $f : E \to E : v \to (v_0 \mid v)w_0$.

(1) Show that f is linear.

(2) Determine the matrix of f with respect to an orthonormal basis. Determine the dimension of the kernel of f.

Exercise 25.17. Let v, w be two non-zero vectors of an Euclidean plane. Determine a real number α such that $v + \alpha w$ has the minimum length (i.e. norm) among all vectors of the form $v + \beta w$, $\beta \in \mathbb{R}$. Is the number α uniquely determined?

Exercise 25.18 (Vector isometries: eigenspaces)**.** Let E be an Euclidean space of dimension n and $f : E \to E$ a vector isometry.

(1) Show that the only real eigenvalues of f are ± 1.

(2) Let E_+, E_- denote the eigenspaces relative to 1 and -1 and set $W = (E_+ \oplus E_-)^{\perp}$. Show that $E = E_+ \oplus E_- \oplus W$.

(3) Show that f is positive (i.e. $\det(f) = 1$) if and only if $\dim(E_-)$ is even. (Observe that the restriction of f to W is an endomorphism of the real vector space W whose characteristic polynomial has no real root. Conclude that $\dim W$ is even and that $f_{|W}$ is a positive isometry of W).

(4) Assume n is odd. Show that if f is positive (resp. negative), then $E_+ \neq \{0\}$ (resp. $E_- \neq \{0\}$). Actually, f negative implies $E_- \neq \{0\}$, whatever the parity of n is.

Exercise 25.19 (Affine orthogonal reflections)**.** Let E be an Euclidean space. If $H_0 \subset E$ is a vector hyperplane of E, the orthogonal reflection with respect to H_0, r_{H_0}, is defined as follows: we have $E = H_0 \oplus H_0^{\perp}$, if $u \in E$, $u = v + v'$, $v \in H_0$, $v' \in H_0^{\perp}$; we set $r_{H_0}(u) = v - v'$. The map r_{H_0} is a negative vector isometry, whose fixed point locus is H_0.

Our aim is to extend this notion to an affine hyperplane $H = a + V$. For each affine hyperplane H we define a map $\pi_H : E \to H$ as follows: if $p \in E$, $\pi_H(p)$ is the intersection point of the line $R_p = p + \langle(\mathrm{dir}(H))^\perp\rangle$ with the hyperplane H.

(1) Show that the map π_H is well defined.

(2) Now we define a map $\sigma_H : E \to E$ as follows: if $p \in E$, then $\sigma_H(p) = -p + 2\pi_H(p)$, i.e. $\sigma_H(p) - \pi_H(p) = -(p - \pi_H(p))$. The map σ_H is called the orthogonal reflection with respect to the hyperplane H.

Let n be a unit vector normal to H (i.e. $n \in \mathrm{dir}(H)^\perp$ and $||n|| = 1$, and let $a \in H$ be a point of H. Show that $\sigma_H(p) = p - 2(p - a \mid n)n$.

(3) Show that the set of fixed points of σ_H is H.

(4) Show that $\sigma_H^2 = Id$ and conclude that σ_H is bijective.

(5) If $H = a + V$, show that $\sigma_H = t_a \circ r_V \circ t_{-a}$, where t_e indicates the translation of vector e and where r_V indicates the orthogonal reflection with respect to the vector hyperplane V. Conclude that σ_H is a negative isometry.

(6) Write σ_H in the form $t_e \circ f$ where f is a vector isometry and where $e \in E$.

Exercise 25.20. Let E be an Euclidean space of dimension n, and $a_0, \ldots, a_n$ affinely independent points. Furthermore, let a, b be two points such that $d(a, a_i) = d(b, a_i), 0 \leq i \leq n$. Show that $a = b$.

Exercise 25.21. Let E be an Euclidean space, $V \subset E$ a vector hyperplane of E, and r_V the orthogonal reflection with respect to V.

(1) Show that the image of $Id - r_V$ is the line $V^\perp$.

(2) Let u be a non-zero vector of E, show that there exists a unique hyperplane V of E such that $w - r_V(W) = u$. More precisely, show that $V = u^\perp$, while $w = w_V + \frac{u}{2}$ where w_V is any vector of V.

(3) If $H = a + V$ is an affine hyperplane, remember that the orthogonal reflection with respect to H, σ_H, is written as $\sigma_H = t_e \circ r_V$ where $e = a - r_V(a)$.

Using this fact and the previous point, show that each translation t_u is written as a product of two orthogonal reflections.

(4) Show that if $t_u = \sigma_H \circ \sigma_{H'}$, then $\mathrm{dir}(H) = \mathrm{dir}(H') = u^\perp$.

Chapter 26

The spectral theorem

Using the theory of Euclidean spaces, we prove that every real symmetric matrix is diagonalizable.

26.1. Self-adjoint operators and spectral theorem

Definition 26.1. Let E be an Euclidean space and $v \in \mathrm{End}(E)$. We say that v is a *symmetric* or *self-adjoint* operator if for every orthonormal basis $\mathcal{B}$ the matrix $\mathrm{Mat}(v; \mathcal{B}, \mathcal{B})$ is symmetric.

Lemma 26.2. *The endomorphism $v \in \mathrm{End}(E)$ is self-adjoint if and only if $\forall x, y \in E$ we have $(x \mid v(y)) = (v(x) \mid y)$.*

Proof. We have $(x \mid v(y)) = {}^tX(AY)$, while $(v(x) \mid y) = {}^t(AX)Y = {}^tX\,{}^tAY$ (here X indicates the column matrix built on the coordinates of x, ditto for Y, y). Since X, Y are arbitrary we have equality if and only if $A = {}^tA$, that is, if and only if A is symmetric. □

Lemma 26.3. *Let E be an Euclidean space, $v \in \mathrm{End}(E)$ a symmetric operator and $W \subset E$ a subspace such that $v(W) \subset W$. Then $v_{|W}$ is a symmetric operator on W.*

Proof. Let $\mathcal{B}' = (e_1, \ldots, e_r)$ be an orthonormal basis of W. Complete it to an orthonormal basis $\mathcal{B}$ of E (Gram-Schmidt). Set $A = \mathrm{Mat}(v; \mathcal{B}, \mathcal{B})$. The matrix A is symmetric. Then the submatrix $A' = (a_{ij})_{1 \leq i,j \leq r}$ is also symmetric. □

Let $A \in M_n(\mathbb{R})$ be a real symmetric matrix. The characteristic polynomial $P_A(x)$ of A is a polynomial with real coefficients but we can also consider it as an element of $\mathbb{C}[x]$ and as such it has all its roots in $\mathbb{C}$ being $\mathbb{C}$ algebraically closed.

Remember that the conjugate of a complex number $z = a + ib$ is $\overline{z} = a - ib$. A complex number is real if and only if $z = \overline{z}$. The modulus of z is $|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$.

Lemma 26.4. *Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Every root in $\mathbb{C}$ of $P_A(x)$ is real i.e. $P_A(x)$ has all its roots in $\mathbb{R}$.*

Proof. We can associate to A an endomorphism f of $\mathbb{C}^n$ via the isomorphism $\mathrm{Mat}(-; \mathcal{B}, \mathcal{B}) : End(\mathbb{C}^n) \to M_n(\mathbb{C})$, where $\mathcal{B}$ is a basis of $\mathbb{C}^n$. Let λ be a root in $\mathbb{C}$ of $P_A(x)$, λ is an eigenvalue of f. So there exists $Y \in \mathbb{C}^n$, $Y \neq 0$ such that $AY = \lambda Y$. Now let us take the conjugate, coordinate by coordinate $\overline{AY} = \overline{\lambda Y}$ (note that $\overline{AB} = \overline{A}\,\overline{B}$). Since the coefficients of A are all real $\overline{A} = A$.

Therefore $A\overline{Y} = \overline{\lambda Y}$. Now, we compute ${}^t\overline{Y}AY$ in two different ways:

- ${}^t\overline{Y}AY = ({}^t\overline{Y}\,{}^tA)Y = {}^t(A\overline{Y})Y = {}^t(\overline{\lambda Y})Y = \overline{\lambda}\,{}^t\overline{Y}Y$;
- ${}^t\overline{Y}AY = {}^t\overline{Y}(AY) = {}^t\overline{Y}(\lambda Y) = \lambda\,{}^t\overline{Y}Y$.

If ${}^tY = (z_1, \ldots, z_n)$ then ${}^t\overline{Y}Y = |z_1|^2 + \cdots + |z_n|^2 > 0$ (since $Y \neq 0$ so one of the z_i is non-zero). We conclude that $\overline{\lambda} = \lambda$, i.e. λ is real. □

Theorem 26.5. *Let E be an Euclidean space and $v \in \mathrm{End}(E)$ a symmetric operator. There is an orthonormal basis, $\mathcal{B}$, of E such that $\mathrm{Mat}(v; \mathcal{B}, \mathcal{B})$ is diagonal.*

Proof. We proceed by induction on $\dim(E)$. If $\dim(E) = 1$ there is nothing to prove. Let $\mathcal{B}'$ be an orthonormal basis of E. The matrix $\mathrm{Mat}(v; \mathcal{B}', \mathcal{B}')$ is symmetric. From the previous lemma there exists a real eigenvalue λ and therefore an associated eigenvector, x. Set $H := x^\perp$. We show that $v(H) \subset H$ (H stable under v). For $y \in H$ we have to show $(x \mid v(y)) = 0$. Since v is self-adjoint: $(x \mid v(y)) = (v(x) \mid y)$. But $(v(x) \mid y) = (\lambda x \mid y) = \lambda(x \mid y) = 0$ since $y \in H = x^\perp$.

It follows that the restriction of v to H is a symmetric operator of H (Lemma 26.3). By induction hypothesis there exists an orthonormal basis of H, $\mathcal{B}_H$, such that the matrix of $v_{|H}$ is diagonal. We complete $\mathcal{B}_H$ to an orthonormal basis, $\mathcal{B}$, of E by adding the vector $\frac{x}{||x||}$. In the basis $\mathcal{B}$ the matrix of v is diagonal. □

An equivalent formulation:

Theorem 26.6. *Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Then there exists an orthogonal matrix P such that $P^{-1}AP$ is diagonal.*

Definition 26.7. A matrix $A \in M_n(K)$ is called *orthogonally diagonalizable* if there exist an orthogonal matrix P and a diagonal matrix D such that $A = QD\,{}^tQ$.

Most matrices, even most diagonalizable matrices, are not orthogonally diagonalizable. Indeed, a necessary condition to be orthogonally diagonalizable is to be symmetric.

Proposition 26.8. *Let $A \in M_n(K)$ be an orthogonally diagonalizable. Then A is symmetric.*

Proof. By hypothesis we have $A = QD\,{}^tQ$ with P orthogonal and D diagonal. Hence

$$ {}^tA = {}^t({}^tQ)\,{}^tD\,{}^tQ = QD\,{}^tQ = A$$

since ${}^tD = D$ being D diagonal. □

Finally, summing up we get the following:

Theorem 26.9 (Spectral theorem)**.** *Let E be an Euclidean space and $v \in \mathrm{End}(E)$ an operator. There is an orthonormal basis, $\mathcal{B}$, of E such that $\mathrm{Mat}(v; \mathcal{B}, \mathcal{B})$ is diagonal if and only if v is self-adjoint.*

Equivalently, a matrix $A \in M_n(\mathbb{R})$ is orthogonally diagonalizable if and only if it is symmetric.

Proof. Follows from Theorems 26.5, 26.6 and Proposition 26.8. □

Remark 26.10. We have already seen in Remark 22.6 that there exist symmetric matrices with complex entries which are not diagonalizable.

A matrix $M \in M_n(\mathbb{C})$ is said to be *Hermitian* if $M = {}^t\overline{M}$. The spectral theorem in all generality states that every Hermitian matrix is diagonalizable (a real symmetric matrix is Hermitian).

26.2. Spectral theorem and real quadratic forms

The spectral theorem offers us a third method for determining the signature of a real quadratic form.

Let q be a real quadratic form and let A be the associated symmetric matrix. By the spectral theorem there exists an orthogonal matrix P such that $P^{-1}AP = D$ is diagonal (A is similar to D). The elements on the diagonal of D are the eigenvalues, i.e. the roots of $P_A(x)$.

Since the matrix P is orthogonal $P^{-1} = {}^tP$ we also have ${}^tPAP = D$ (A is congruent to D). This means that there exists an orthogonal basis $\mathcal{B}$ such that $mat_{\mathcal{B}}(f) = D$ (f is the symmetric bilinear form associated to q). So the signature of q is given by the signs of the elements on the diagonal of D i.e. the signature is given by the signs of the eigenvalues.

In conclusion, to determine the signature of q, the characteristic polynomial $P_A(x)$ is calculated and the sign of its roots are determined.

Example 26.11. Consider the quadratic form

$$q(x, y, z) = x^2 + y^2 + z^2 - 4(xy + xz + yz).$$

The associated matrix is

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

and we have

$$P_A(x) = \begin{vmatrix} x-1 & 2 & 2 \\ 2 & x-1 & 2 \\ 2 & 2 & x-1 \end{vmatrix} = (x-3)^2(x+3).$$

The signature of q is $(2, 1)$ (two positive terms, one negative).

Be careful though, this method is not necessarily faster or safer than the others. Furthermore, regarding the diagonalization of a quadratic form, it can only be used when $K = \mathbb{R}$.

Exercises

Exercise 26.12. Consider the matrix

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

and find an orthogonal matrix P and a diagonal matrix D such that $A = PD\,{}^tP$.

Exercise 26.13. Let $A, B \in M_n(\mathbb{R})$ be two orthogonally diagonalizable matrices such that $AB = BA$. Prove that AB is orthogonally diagonalizable.

Exercise 26.14. Consider the matrices

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}.$$

Show that A and B are orthogonally diagonalizable but AB is not.

Exercise 26.15. Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Prove that there exists a matrix $B \in M_n(\mathbb{C})$ such that $A = {}^tBB$ (Hint: write $A = {}^tPDP$ and consider the diagonal matrix $\sqrt{D}$, whose entries are the square roots of the entries of D, and the matrix $\sqrt{D}P$).

Exercise 26.16. Define the exponential of a symmetric matrix $A \in M_n(\mathbb{R})$ by the convergent series

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Show that $e^A = {}^tPe^DP$ where D is diagonal and P is orthogonal.

Part 5

Jordan canonical form

A *Jordan canonical form* is an upper triangular matrix of the following form

$$\begin{pmatrix} J_1 & 0 & \dots & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & J_r \end{pmatrix}$$

called a Jordan matrix representing a linear operator on a finite-dimensional vector space with respect to some basis. Each block J_i, called a *Jordan block*, is a square matrix of the form

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_i & 1 \\ 0 & \dots & \dots & 0 & \lambda_i \end{pmatrix}.$$

In other terms the only non zero entries a matrix in Jordan canonical form are either on the diagonal or on the superdiagonal.

Let $f : E \to E$ be an endomorphism of the K-vector space E. We will see that there is a basis $\mathcal{B}$ of E such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ is in Jordan form if and only if the endomorphism f has all its eigenvalues in K. In this case, the diagonal entries of the Jordan form are the eigenvalues of f, and the number of times each eigenvalue occurs is the algebraic multiplicity of the eigenvalue.

The diagonal form for diagonalizable matrices is a special case of the Jordan canonical form. As we have seen there are non diagonalizable matrices but, if K is algebraically closed, any square matrix has a Jordan canonical form. For instance, the matrix

$$\begin{pmatrix} 0 & -6 & -4 \\ 5 & -11 & -6 \\ -6 & 9 & 4 \end{pmatrix}$$

is not diagonalizable but admits the following Jordan canonical form

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

Chapter 27

The minimal polynomial

In the following we will denote by E a K-vector space of dimension n and by $f : E \to E$ an endomorphism of E. If

$$P(X) = a_n X^n + \cdots + a_1 X + a_0 \in K[X]$$

then

$$P(f) = a_n f^n + \cdots + a_1 f + a_0 Id \in \mathrm{End}(E)$$

and we get a map

$$\varphi_f : K[X] \to \mathrm{End}(E) : P \to P(f).$$

It turns out that φ_f is a ring morphism (φ_f is also K-linear). Therefore $(PQ)(f) = P(f) \circ Q(f)$ (to verify this take $P(X) = X^m, Q(X) = X^t$ and conclude by linearity).

An important fact: since $K[X]$ is commutative, we have

$$P(f) \circ Q(f) = Q(f) \circ P(f), \forall P, Q \in K[X]$$

that is $\mathrm{Im}(\varphi_f)$ is a commutative subring of $\mathrm{End}(E)$.

The morphism φ_f is not injective. Indeed, By the Cayley-Hamilton theorem $P_f(x) \in \mathrm{Ker}(\varphi_f)$.

The non-injectivity of φ_f can be seen directly. Since $\dim(\mathrm{End}(E)) = n^2$, $Id, f, f^2, \ldots, f^{n^2}$ are dependent: $a_0 Id + \cdots + a_{n^2} f^{n^2} = 0$, i.e. $Q(f) = 0$ with $a_0 + \cdots + a_{n^2} X^{n^2} = Q(X) \neq 0$. This shows that the map φ_f is not one-to-one. On the other hand, $K[X]$ is a K-vector space of infinite dimension, while $\mathrm{End}(E)$ has finite dimension.

Set $J = \mathrm{Ker}(\varphi_f)$. The subset J of $K[X]$ is an ideal, that is if $P, Q \in J$ then $P + Q \in J$ and $\forall P \in J, T \in K[X], TP \in J$. Another crucial fact:

Lemma 27.1. *Any ideal, J, of $K[X]$ is generated by a single polynomial that is $\exists T \in K[X]$ such that $J = (T) = \{PT \mid P \in K[X]\}$.*

Proof. The simplest way to see this is to use Euclidean division. Recall that if P, M are two polynomials with $\deg(P) \geq \deg(M)$, then there exist Q, R such that $P = QM + R$ and $\deg(R) < \deg(M)$. Let $M \in J$ be of minimum degree. If $P \in J$, we divide by M: $P = QM + R$. We have $R = P - QM \in J$. Since $\deg(R) < \deg(M)$, due to the minimality of the degree of M, the only possibility is $R = 0$. So $P = QM$. □

Remark 27.2. An integral ring in which each ideal can be generated by a single element is called a PID (*principal ideal domain*).

Every Euclidean ring, i.e. a ring with an Euclidean division, is a PID. For example $\mathbb{Z}$ is a PID. There are principal ideal domains that are not Euclidean, for instance $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ is a non Euclidean PID.

So every ideal $J \subset K[X]$ can be written as $J = (M)$ where M is a polynomial of minimal degree in J. The polynomial is not uniquely determined, if $a \in K$, $a \neq 0$, then $J = (aM)$. In particular, we can always find a such that aM is monic i.e. the coefficient of the highest power of X in aM is equal to 1:

$$aM = X^t + a_{t-1}X^{t-1} + \cdots + a_0.$$

The monic generator of J is uniquely determined.

In what follows, when we write $J = (P)$ we will assume that P is the monic generator.

Definition 27.3. The *minimal polynomial*, M_f, of the endomorphism $f : E \to E$ is the monic generator of the ideal $J = \mathrm{Ker}(\varphi_f)$.

Remark 27.4. 1) The minimal polynomial is the monic polynomial, P, of smallest degree such that $P(f) = 0$.
2) If P verifies $P(f) = 0$, then $M_f | P$ (M_f divides P). In particular, $M_f \mid P_f$ due to the Cayley-Hamilton theorem.

In the following we will study the relation between the minimal polynomial M_f and the characteristic polynomial P_f.

27.1. Minimal and characteristic polynomial

As already observed $M_f \mid P_f$ that is $M_f Q = P_f$. A priori this does not imply that if λ is a root of P_f then λ is also a root of M_f (λ could be a root of Q). Actually, as we will now see, M_f and P_f have the same roots.

Proposition 27.5. *Let $f : E \to E$ be an endomorphism of the K-vector space E. An element $\lambda \in K$ is a root of P_f (i.e. λ is an eigenvalue of f) if and only if λ is a root of M_f.*

Proof. Exercise 27.9. □

Corollary 27.6. *Let $f : E \to E$ be an endomorphism of the K-vector space E. It is assumed that $P_f(X)$ has all its roots in K (hypothesis verified if K is algebraically closed):*

$$P_f(X) = (X - \lambda_1)^{\alpha_1}(X - \lambda_2)^{\alpha_2} \dots (X - \lambda_r)^{\alpha_r}$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$, and $\alpha_1 + \dots + \alpha_r = \dim(E)$. Then

$$M_f(X) = (X - \lambda_1)^{\beta_1}(X - \lambda_2)^{\beta_2} \dots (X - \lambda_r)^{\beta_r} \text{ with } 1 \leq \beta_i \leq \alpha_i, \forall i.$$

Proof. If $Q(X)$ is an irreducible factor of $M_f(X)$, then Q is an irreducible factor of $P_f(X)$ (since $M_f | P_f$), so Q has degree one. From Proposition 27.5 it follows that

$$M_f(X) = (X - \lambda_1)^{\beta_1}(X - \lambda_2)^{\beta_2} \dots (X - \lambda_r)^{\beta_r}$$

with $\beta_i \geq 1, \forall i$. Finally, since $M_f | P_f$, $\beta_i \leq \alpha_i$, $\forall i$. □

Corollary 27.7. *Let $f : E \to E$ be an endomorphism of the K-vector space E. It is assumed that $P_f(X)$ has all its roots in K (hypothesis verified if K is algebraically closed):*

$$P_f(X) = (X - \lambda_1)^{\alpha_1}(X - \lambda_2)^{\alpha_2} \dots (X - \lambda_r)^{\alpha_r}$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$, and $\alpha_1 + \dots + \alpha_r = dim(E)$. If f is diagonalizable then

$$M_f(X) = (X - \lambda_1)(X - \lambda_2) \dots (X - \lambda_r).$$

In particular, M_f does not have multiple roots.

Proof. Let $\mathcal{B} = (e_{1_1}, \dots, e_{1_{\alpha_1}}; e_{2_1}, \dots, e_{2_{\alpha_2}}; \dots; e_{r_1}, \dots, e_{r_{\alpha_r}})$ be a basis of eigenvectors ($f(e_{i_j}) = \lambda_i e_{i_j}$), and set $F = (f - \lambda_1 Id) \circ (f - \lambda_2 Id) \circ \dots (f - \lambda_r Id)$. We have $F(e_{i_j}) = (f - \lambda_1 Id) \circ \dots \circ (f - \lambda_i Id) \circ \dots (f - \lambda_r Id)(e_{i_j}) = (f - \lambda_1 Id) \circ \dots \circ (f - \lambda_r Id) \circ (f - \lambda_i Id)(e_{i_j}) = 0$. So $F = 0$. It follows that $M_f | (X - \lambda_1) \dots (X - \lambda_r)$. We conclude by Corollary 27.6. □

Remark 27.8. It can be shown that if $P_f(x)$ has all its roots in K and if $M_f(x)$ has no multiple roots, then f is diagonalizable (i.e. these conditions are necessary and sufficient).

Exercises

Exercise 27.9. Prove Proposition 27.5.

Exercise 27.10. 1) Let $P(x) \in \mathbb{R}[x]$. Show that if $P(x)$ is irreducible then $\deg(P) \leq 2$ (Work on $\mathbb{C}$).
2) Let E be an $\mathbb{R}$-vector space and $f \in \mathrm{End}(E)$. Show that M_f and P_f have the same irreducible factors.

Exercise 27.11. A matrix $A \in \mathcal{M}_n(K)$ is scalar if $A = \alpha I_n$ for some $\alpha \in K$.
1) Show that $A \in \mathcal{M}_n(K)$ verifies $\deg(M_A(x)) = 1$ if and only if A is scalar.
2) Give an example of a 3×3 non-scalar matrix, A, whose minimal polynomial has degree two (therefore $M_A(x) \neq P_A(x)$).

Exercise 27.12. 1) Let E be a K-vector space. Throughout the exercise we assume $\dim(E) \geq 2$. Show that there exist $f, g \in \mathrm{End}(E)$ such that $f \circ g \neq g \circ f$ (take a basis (e_i) of E and define f, g by their values on the vectors e_i, just consider e_1 and e_2). In particular, $\mathrm{End}(E)$ is a non-commutative ring.
2) Conclude that if $f \in \mathrm{End}(E)$, the morphism $\varphi_f : K[x] \to \mathrm{End}(E) : P \to P(f)$ is never surjective.
3) The center, Z of $\mathrm{End}(E)$ is

$$Z = \{f \in \mathrm{End}(E) \mid f \circ g = g \circ f, \forall g \in \mathrm{End}(E)\}$$

i.e. Z is the set of endomorphisms that commute with all other endomorphisms. A homothety is an endomorphism of the form αId, clearly a homothety belongs to Z. Show that if $f \in \mathrm{End}(E)$ is not a homothety, then there exists $v \in E$ such that v and $f(v)$ are linearly independent (taking a basis of E and reasoning by contradiction). Taking a suitable basis, show that there exists $g \in \mathrm{End}(E)$ which does not commute with f. Conclude that $Z = \{\alpha Id \mid \alpha \in K\} \simeq K$.

Exercise 27.13. 1) Let $M_f(x)$ be the minimal polynomial of $f \in \mathrm{End}(E)$. Set $m := \deg(M_f)$. Observe that $Id = f^0, f, f^2, \ldots, f^m$ are linearly dependent and conclude that $(Id, f, f^2, \ldots, f^{m-1})$ is a basis of $\mathrm{Im}(\varphi_f)$. So $\dim(\mathrm{Im}(\varphi_f)) \leq n, \forall f$ $(n = \dim(E))$.
2) An endomorphism f is *nilpotent* if there exists t such that $f^t = 0$. What is the shape of the minimal polynomial of a nilpotent endomorphism? Determine the characteristic polynomial $P_f(x)$ and conclude that $f^n = 0$ $(n = \dim(E))$.
3) If f is nilpotent its *index* (of nilpower) is $\iota(f) = \min\{t \mid f^t = 0\}$. For example, if $(e_1, \ldots, e_n)$ is a basis of E and f is defined by $f(e_i) = e_{i+1}$ (with the convention $e_k = 0$ if $k > n$), then f is nilpotent with $\iota(f) = n$. Show that for $2 \leq k \leq n$ there exists f nilpotent with $\iota(f) = k$.

4) Conclude that for every k with $1 \leq k \leq n$ there exists $f \in \mathrm{End}(E)$ such that $\dim(\mathrm{Im}(\varphi_f)) = k$.

Chapter 28

Cyclic spaces and subspaces

28.1. The $K[X]$-module E_f

The map $\varphi_f : K[X] \to \text{End}(E) : P \to P(f)$ allows us to define an external multiplication on E:

$$K[X] \times E \to E : (P, v) \to Pv := P(f)(v).$$

With this external multiplication E becomes a $K[X]$-module, this means that the external multiplication verifies the axioms of the multiplication for vector spaces, the only difference (but it is notable) is that $K[X]$ is not a field but just a ring.

In the following we will write E_f to indicate that we are considering E as a $K[X]$-module via f and we will write Pv instead of $P(f)(v)$.

Definition 28.1. Let A be a commutative ring, an A-module M is an abelian group with an external multiplication i.e. $A \times M \to M : (\alpha, m) \to \alpha m$ such that

- $(\alpha + \beta)m = \alpha m + \beta m$;
- $\alpha(m + n) = \alpha m + \alpha n$;
- $\alpha(\beta m) = (\alpha\beta)m$;
- $1_A m = m$;

for any $\alpha, \beta \in A, m, n \in M$.

Example 28.2. 1) If $A = K$ is a field a K-module is a K-vector space.
2) An ideal $I \subset A$ is an A-module (so A is an A-module, more generally A^n is an A-module).
3) The $\mathbb{Z}$-modules are exactly the abelian groups.

4) If $A = K[x]$ an A-module is a K-vector space with an endomorphism (the situation we are considering).

Definition 28.3. Let M, N be two A-modules, a map $\varphi : M \to N$ is an A-linear morphism if

- $\varphi(m + m') = \varphi(m) + \varphi(m')$ (group morphism);
- $\varphi(\alpha m) = \alpha\varphi(m)$ (A-linearity).

We highlight some differences between the theory of modules and that of vector spaces. For example, in a vector space we have $\lambda v = 0 \Rightarrow \lambda = 0$ or $v = 0$ ($\lambda \in K$) which is not true for A-modules.

In fact, if M_f is the minimal polynomial of f we have $M_f v = 0, \forall v \in E$, this means that E_f is a $K[X]$-module with *torsion*. In general an A-module F has torsion if $\forall m \in F, \exists a \in A, a \neq 0$ such that $am = 0$. A priori a depends on m and it is not certain that there exists a single a that annihilates F. We therefore have two good news: the first is that in our case there is an element of the ring that annihilates the module, the second is that our ring, $A = K[X]$, is a PID. In fact, there is a classification of modules (in particular with torsion) on a PID and what follows is a particular case of this theory.

Definition 28.4. A submodule of E_f is a subvector space $V \subset E$ such $P(V) \subset V$, $\forall P \in K[X]$.

The following fact is immediate:

Lemma 28.5. *A subvector space $V \subset E$ is a submodule of E_f if and only if V is stable under f (i.e. $f(V) \subset V$).*

The notion of stable subspace is important by virtue of the following:

Lemma 28.6. *Let $F \subset E$ be a stable subspace of E_f. Then $P_{f|F} \mid P_f$ and $M_{f|F} \mid M_f$.*

Proof. Let us take a basis $\mathcal{B} = (e_1, \ldots, e_t)$ of F and complete it to a basis $\mathcal{B} = (e_1, \ldots, e_t, \ldots, e_n)$ of E. Since $f(e_i) \in F$ for $1 \leq i \leq t$, $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ is of the form

$$M = \begin{pmatrix} R & D \\ 0 & T \end{pmatrix}.$$

Here R is a $t \times t$ matrix and T is an $(n-t) \times (n-t)$ matrix. Furthermore, $R = \mathrm{Mat}(f_{|F}; \mathcal{B}, \mathcal{B})$. We have $\det(\lambda I_n - M) = \det(\lambda I_t - R)\det(\lambda I_{n-t} - T)$. So $P_{f|F} \mid P_f$. If $v \in F$ then $M_f v = 0$, and hence $M_f \in (M_{f|F})$ i.e. $M_{f|F} \mid M_f$. $\square$

Definition 28.7. For $v \in E_f$ we denote by $\langle v\rangle_f$ is the submodule of E_f generated by v:

$$\langle v\rangle_f := \{Pv \mid P \in K[x]\}.$$

Lemma 28.8. *Let $v \in E_f$, $v \neq 0$. The subspace $\langle v\rangle_f$ is stable. If $t := \dim\langle v\rangle_f$, then $\mathcal{B} = (v, f(v), \dots, f^{t-1}(v))$ is a basis of $\langle v\rangle_f$.*

Proof. Being a submodule, $\langle v\rangle_f$ is clearly a stable subvector space of E_f (note that $f(Pv) = x(Pv) = (xP)(v)$). Set

$$t := \max\{k \mid v, f(v), \dots, f^{k-1}(v) \text{ are independent}\}.$$

We have $1 \leq t$ because $v \neq 0$ and $t \leq n = \dim E$ (so t is well defined). If $m \geq t$, $f^m(v)$ is a linear combination of $v, f(v), \dots, f^{t-1}(v)$. In fact $f^t(v)$ is a linear combination of $v, \dots, f^{t-1}(v)$, by definition of t; now

$$f^{t+1}(v) = f(f^t(v)) = f\left(\sum_{i=0}^{t-1} a_i f^i(v)\right) = \sum_{i=1}^{t-1} a_{i-1} f^i(v) + a_{t-1} f^t(v)$$

and we see that $f^{t+1}(v)$ is a linear combination of $v, \dots, f^{t-1}(v)$. Proceeding in this way we get the statement by induction. Since $Pv = a_l f^l(v) + \cdots + a_1 f(v) + a_0 v$, it follows that the vectors $v, f(v), \dots, f^{t-1}(v)$ generate $\langle v\rangle_f$. So $(v, f(v), \dots, f^{t-1}(v))$ is a basis of $\langle v\rangle_f$. □

Definition 28.9. A subspace $F \subset E_f$ is *cyclic* if there exists $v \in E$ such that $F = \langle v\rangle_f$. The space E_f is cyclic if there exists $v \in E$ such that $E = \langle v\rangle_f$.

28.2. Companion matrices and cyclic spaces

Let $Q(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in K[x]$ be a *monic* polynomial of degree $n = \dim(E)$. To the polynomial Q we associate the matrix

$$\mathcal{C}(Q) := \begin{pmatrix} 0 & 0 & \dots & \dots & 0 & -a_0 \\ 1 & 0 & \dots & \dots & \dots & -a_1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

Definition 28.10. The matrix $\mathcal{C}(Q)$ is called the *companion matrix* of the polynomial Q.

Lemma 28.11. *Let $\mathcal{C}(Q)$ be the companion matrix of the polynomial Q. Then $M_{\mathcal{C}(Q)} = P_{\mathcal{C}(Q)} = Q$ (the characteristic polynomial and the minimal polynomial of $\mathcal{C}(Q)$ are both equal to Q).*

Proof. Exercise 28.16. □

We have a clear relation between companion matrices and cyclic spaces:

Proposition 28.12. *Let E be a K-vector space and $f \in \mathrm{End}(E)$.*
1) E_f is cyclic if and only there exists a basis $\mathcal{B}$ of E such that $\mathrm{Mat}(f;\mathcal{B},\mathcal{B}) = \mathcal{C}(Q)$ and in this case $Q = P_f$.
2) If E_f is cyclic then $P_f = M_f$.
3) Two $n \times n$ companion matrices are similar if and only if they are equal.

Proof. 1) If there exists a basis $\mathcal{B} = (e_i)$ such that $\mathrm{Mat}(f;\mathcal{B},\mathcal{B})$ is a companion matrix $\mathcal{C}(Q)$, then E_f is clearly cyclic ($E = \langle e_1 \rangle_f$) and $Q = P_f$ (Lemma 28.11). Conversely, if E_f is cyclic, $E_f = \langle v \rangle_f$ and if $\mathcal{B} = (v, f(v), \ldots, f^{n-1}(v))$, then $\mathrm{Mat}(f;\mathcal{B},\mathcal{B})$ is a companion matrix $\mathcal{C}(Q)$, with $Q = P_f$ (Lemma 28.11).
2) Follows from 1) and Lemma 28.11.
3) If $\mathcal{C}(Q)$ and $\mathcal{C}(P)$ are similar they represent the same endomorphism f in different bases. By Lemma 28.11 we have $P = P_f = Q$. □

Remark 28.13. 1) Proposition 28.12 provides a necessary condition for an endomorphism f to be cyclic: we must have $P_f = M_f$. We will see later that this condition is also sufficient.
2) A similarity class is associated to each monic polynomial P: the class of $\mathcal{C}(P)$. The map associating to P the class of $\mathcal{C}(P)$ is injective and its image describes the classes of cyclic endomorphisms.

Exercises

Exercise 28.14. An A-module M is of *finite type* (or *finitely generated*) if there exists a surjective morphism $A^n \to M$ for some n. We say that M is *free*, of rank n, if there exists an isomorphism $A^n \to M$. In these situations the image of the canonical basis of A^n is a set of generators (resp. a basis) of M.

Give an example of an A-module M of finite, non-free type (i.e. M has a finite set of generators but does not have a basis).

Exercise 28.15. Let A be an integral (commutative) ring and M an A-module. An element $m \in M$, $m \neq 0$, is torsion if there exists $a \in A, a \neq 0$ such that $am = 0$. Set

$$T(M) = \{m \in M \mid m \text{ is torsion}\} \cup \{0\}.$$

1) Show that $T(M)$ is a submodule of M. The A-module M is torsion-free if $T(M) = \{0\}$.
2) Let G be a $\mathbb{Z}$-module. Show that if G is a finite set (i.e. G is a finite abelian group) every element of G is torsion ($T(G) = G$). Is G finitely generated?
3) If A is integral and M is free of rank n then M is torsion-free. What happens if A is not integral?

Exercise 28.16. Prove Lemma 28.11.

Exercise 28.17. Let E be a K-vector space and $f \in \mathrm{End}(E)$. The endomorphism f is *simple* if the only stable subspaces of E are the trivial ones ($\{0\}, E$). Note that this goes back to requiring that E_f has no non-trivial $K[x]$-submodules (it is also said that E_f is simple). Show that:
1) A simple space is cyclic.
2) The space E_f is simple if and only if P_f is irreducible.

Exercise 28.18. (A third proof of the Cayley-Hamilton theorem) Let E be a K-vector space and $v \in E$, $v \neq 0$.
1) Let $F = \langle v \rangle_f$. Show that $P_{f|F}(v) = 0$.
2) Conclude that $P_f(f) = 0$.

Chapter 29

Characteristic subspaces and the killer lemma

We introduce the decomposition of E into characteristic subspaces. As a first consequence we obtain a new diagonalization criterion (Theorem 29.6). We then use this decomposition to prove the killer lemma which will be crucial in the following.

Remark 29.1. If $F \subset E$ is a subvector space, we denote by

$$\mathrm{Ann}(F) = \{P \in K[x] \mid Pv = 0, \forall v \in F\}$$

its annihilator. It is easily verified that $\mathrm{Ann}(F)$ is an ideal of $K[x]$, so $\mathrm{Ann}(F) = (Q)$ (monic generator as always). Clearly $Q = M_{f|F}$. If $F = \langle v \rangle = \{\lambda v \mid \lambda \in K\}$, we write $\mathrm{Ann}(v)$ instead of $\mathrm{Ann}(\langle v \rangle)$. We observe that $\mathrm{Ann}(v) = \mathrm{Ann}(\langle v \rangle_f)$.

Indeed, let $\mathrm{Ann}(v) = (Q)$, then $Q(f^i(v)) = (Qx^i)(v) = (x^i Q)(v) = 0$. So $Q \in \mathrm{Ann}(\langle v \rangle_f)$. On the other hand, it is clear that $\mathrm{Ann}(\langle v \rangle_f) \subset \mathrm{Ann}(v)$. Therefore, $\mathrm{Ann}(v) = \mathrm{Ann}(\langle v \rangle_f) = (M_{f|\langle v \rangle_f})$. Since $P_{f|\langle v \rangle_f} = M_{f|\langle v \rangle_f}$, if $\mathrm{Ann}(v) = (Q)$, $\dim \langle v \rangle_f = \deg(Q)$.

Lemma 29.2 (Kernel lemma)**.** *Let $P_1, \ldots, P_t \in K[x]$ be polynomials prime to each other i.e. $(P_i, P_j) = (1)$ if $i \neq j$, E a K-vector space, $f \in \mathrm{End}(E)$ and $P := P_1 \cdots P_t$. Then*

$$\mathrm{Ker}(P(f)) = \bigoplus_{i=1}^{t} \mathrm{Ker}(P_i(f)).$$

Proof. The proof is by induction on t. If $t = 1$, there is nothing to prove. Set $t = 2$. Since $(P_1, P_2) = (1)$, there exist Q_1, Q_2 such that $Q_1 P_1 + Q_2 P_2 = 1$. So

$(Q_1P_1)(f) + (Q_2P_2)(f) = Id_E$. Therefore,

$$\forall v \in E:\ (Q_1P_1)v + (Q_2P_2)v = v\ \ (*).$$

Let $w \in \mathrm{Ker}(P_1P_2(f))$. We have $P_2(Q_1P_1)w) = Q_1(Pw) = 0$ ($P = P_1P_2$). Similarly, $P_1((Q_2P_2)w) = Q_2(Pw) = 0$. So, due to $(*)$ we have $w \in \mathrm{Ker}(P_1) + \mathrm{Ker}(P_2)$.

Using $(*)$ it is easy to see that the sum $\mathrm{Ker}(P_1) + \mathrm{Ker}(P_2)$ is direct. At this point we have $\mathrm{Ker}(P) \subset \mathrm{Ker}(P_1) \oplus \mathrm{Ker}(P_2)$.

Since $\mathrm{Ker}(P_1) \oplus \mathrm{Ker}(P_2) \subset \mathrm{Ker}(P)$ (if $w = w_1 + w_2 \in \mathrm{Ker}(P_1) \oplus \mathrm{Ker}(P_2)$, then $Pw = Pw_1 + Pw_2 = P_2(P_1w_1) + P_1(P_2w_2) = 0$), we conclude that $\mathrm{Ker}(P(f)) = \mathrm{Ker}(P_1(f)) \oplus \mathrm{Ker}(P_2(f))$.

Let us assume the result is true for $t-1$ and show it for t. Set $R_1 = P_1 \dots P_{t-1}$, $R_2 = P_t$. We have $(R_1, R_2) = (1)$ and hence $\mathrm{Ker}(R_1R_2) = \mathrm{Ker}(P) = \mathrm{Ker}(R_1) \oplus \mathrm{Ker}(R_2)$. By induction hypothesis (since $P_1, \dots, P_{t-1}$ are prime two to each other), we have $\mathrm{Ker}(R_1) = \bigoplus_1^{t-1} \mathrm{Ker}(P_i)$. So $\mathrm{Ker}(P(f)) = \bigoplus_{i=1}^{t} \mathrm{Ker}(P_i(f))$. □

Theorem 29.3 (Characteristic subspaces)**.** *Let E be a K-vector space and $f \in \mathrm{End}(E)$. Let us consider the irreducible factorization of $M_f(x)$ in $K[x]$:*

$$M_f(x) = M_1^{a_1}(x)M_2^{a_2} \cdots M_r^{a_r}(x).$$

We can assume M_i monic, $\forall i$. Set $E_i := \mathrm{Ker}(M_i^{a_i}(f))$. Then every E_i is stable and $E = E_1 \oplus \cdots \oplus E_r$. The E_i are the characteristic subspaces or generalized eigenspaces of f.

Proof. It is clear that the subspaces E_i are stable. More generally every subspace of the form $\mathrm{Ker}(P)$ is stable (if $v \in \mathrm{Ker}(P)$, $Pf(v) = (Px)v = x(Pv) = 0$).

If $P_i = M_i^{a_i}$, the polynomials $P_1, \dots, P_r$ are prime to each other. Set $P = P_1 \cdots P_r$. By Lemma 29.2 $\mathrm{Ker}(P) = \bigoplus_{i=1}^r \mathrm{Ker}(P_i)$. But $P = M_f$, so $\mathrm{Ker}(P) = E$. □

If $\mathcal{B}_i$ is a basis of E_i, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_r)$ is a basis of E and since E_i is stable $f(\mathcal{B}_i) \in E_i$, then the matrix $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ of f in the basis $\mathcal{B}$ is a block diagonal matrix:

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & A_r \end{pmatrix} \tag{29.1}$$

where $A_i = \mathrm{Mat}(f_i; \mathcal{B}_i, \mathcal{B}_i)$ and $f_i = f_{|E_i}$.

Lemma 29.4. *With the above notations, if $a \geq a_l$, then $\mathrm{Ker}(M_l^a(f)) = E_l$.*

Proof. Set $\delta_i = a_i$ if $i \neq l$, $\delta_l = a$ and set $P = \prod_i M_i^{\delta_i}$. Since $M_f \mid P$, we have $P(f)(v) = 0, \forall v \in E$. By Lemma 29.2 we have $E = \bigoplus_i \mathrm{Ker}(M_i^{\delta_i}(f)) = (\bigoplus_{i \neq l} E_i) \oplus \mathrm{Ker}(M_l^a)$. So $\dim(\mathrm{Ker}(M_l^a(f))) = \dim(E_l)$. Since, clearly $E_l \subset \mathrm{Ker}(M_l^a(f))$, it follows that $\mathrm{Ker}(M_l^a(f)) = E_l$. □

We also have:

Lemma 29.5. *With the previous notations let $f_i = f_{|E_i}$. Then the minimal polynomial of the endomorphism $f_i : E_i \to E_i$ is $M_i^{a_i}(x)$.*

Proof. Let $m_i(x)$ be the minimal polynomial of f_i. Since $E_i = \mathrm{Ker}(M_i^{a_i}(f))$, $m_i \mid M_i^{a_i}$, then $m_i = M_i^{\delta_i}$, $\delta_i \leq a_i$.

On the other hand, if $m = m_1 \dots m_r$, since $E = E_1 \oplus \cdots \oplus E_r$, we have $m(f)(v) = 0, \forall v \in E$, so $M_f \mid m$ and this implies $\delta_i = a_i, \forall i$. □

Suppose that the characteristic polynomial of f has all its roots in K (hypothesis verified if K is algebraically closed). In this case we have (Corollary 27.6):

$$\begin{array}{rcl} P_f(X) & = & \prod_{i=1}^r (X - \lambda_i)^{\alpha_i}; \\ M_f(X) & = & \prod_{i=1}^r (X - \lambda_i)^{\beta_i}; \end{array}$$

with $1 \leq \beta_i \leq \alpha_i$ ($\lambda_i \neq \lambda_j$ if $i \neq j$).

Theorem 29.6. *With the previous hypotheses and notations (in particular P_f has all its roots in K), if V_i is the eigenspace relative to λ_i then:*

(1) *$V_i \subset E_i$;*

(2) *$\dim(E_i) = \alpha_i$, $1 \leq i \leq r$;*

(3) *f is diagonalizable if and only if $M_f(X)$ has no multiple roots.*

Proof. (1) Follows from $\mathrm{Ker}(f - \lambda_i Id) \subset \mathrm{Ker}((f - \lambda_i Id)^{\beta_i})$.

(2) The minimal polynomial of f_i is $M_i^{\beta_i}(X) = (X - \lambda_i)^{\beta_i}$ (Lemma 29.5). The block form (29.1) shows that the characteristic polynomial of f is the product of the characteristic polynomials of the f_i. In particular P_{f_i} has all its roots in K, this implies $P_{f_i}(X) = (X - \lambda_i)^{\gamma_i}$ (the roots of P_{f_i} are roots of the minimal polynomial of f_i). It follows that $\dim(E_i) = \gamma_i = \alpha_i$.

(3) The endomorphism f is diagonalizable if and only if $E = \bigoplus_i V_i$ if and only if $V_i = E_i, \forall i$ if and only if $E_i = \mathrm{Ker}(f - \lambda_i Id)$ if and only if $\beta_i = 1, \forall i$ (Lemma 29.5)

concluding the proof. □

We see that in these hypotheses the characteristic polynomial provides the dimension of the characteristic subspaces.

29.1. The killer lemma

We have seen so far how simple life is with cyclic endomorphisms. Two cyclic endomorphisms are similar if and only if they have the same characteristic polynomial. Furthermore, we can associate a very particular matrix to each cyclic endomorphism: the companion matrix of its characteristic polynomial. To extend this type of result to all endomorphisms the idea is to write E as a direct sum of cyclic subspaces.

The first step consists in showing that there exists $v \in E$ such that $M_{f_{|F}} = M_f$, where $F = \langle v \rangle_f$ (killer lemma). We have already seen that $M_{f_{|F}} \mid M_f$, it is therefore legitimate to ask whether there is a v for which equality holds.

Lemma 29.7. *Let E be a K-vector space, $f \in \mathrm{End}(E)$, $v_1, \ldots, v_r \in E$ vectors such that the subspaces $\langle v_1 \rangle_f, \ldots, \langle v_r \rangle_f$ are in direct sum, and $\mathrm{Ann}(v_i) = (P_i)$. Then if $v = v_1 + \cdots + v_r$, we have $\mathrm{Ann}(v) = (Q)$, with $Q = \mathrm{lcm}(P_i)$.*

Proof. Set $P = \mathrm{lcm}(P_i)$ and $\mathrm{Ann}(v) = (Q)$ (we can assume P and Q are monic). We have $Pv_i = 0, \forall i$ since $P_i \mid P, \forall i$. It follows that $Pv = 0$, so $Q \mid P$.

We have $0 = Qv = Qv_1 + \cdots + Qv_r$. Since $Qv_i \in \langle v_i \rangle_f$ and since the sum of the $\langle v_i \rangle_f$ is direct, this implies $Qv_i = 0, \forall i$. So $P_i \mid Q, \forall i$. Therefore $P \mid Q$ (by definition $(P) = \cap_i(P_i)$). We conclude that $P = Q$. □

Lemma 29.8 (Killer Lemma)**.** *Let E be a K-vector space (K any field) and $f \in \mathrm{End}(E)$. Then there exists $v \in E$ such that $\mathrm{Ann}(v) = (M_f)$.*

Proof. Let $M_f = M_1^{a_1} \cdots M_r^{a_r}$ be the factorization of M_f into irreducible factors. By Theorem 29.3 we have $E = E_1 \oplus \cdots \oplus E_r$, where the $E_i = \mathrm{Ker}(M_i^{a_i})$ are the characteristic subspaces. Since the minimal polynomial of $f_i = f_{|E_i}$ is $m_i = M_i^{a_i}$ (Lemma 29.5), there exists $v_i \in E_i$ such that $M_i^{a_i - 1} v_i \neq 0$ (otherwise $m_i \mid M_i^{a_i - 1}$). If $\mathrm{Ann}(v_i) = (P_i)$, then $P_i \mid m_i = M_i^{a_i}$, and $P_i = M_i^{b_i}$ with $b_i \leq a_i$ (because M_i is irreducible). But since $M_i^{a_i - 1} v_i \neq 0$ we have $a_i = b_i$.

In conclusion, there exists $v_i \in E_i$ such that $\mathrm{Ann}(v_i) = (M_i^{a_i})$, $1 \leq i \leq r$. Since $\langle v_i \rangle_f \subset E_i$ (because E_i is stable (Theorem 29.3), the subspaces $\langle v_1 \rangle_f, \ldots, \langle v_r \rangle_f$ are in direct sum. By Lemma 29.7, if $v = v_1 + \cdots + v_r$, $\mathrm{Ann}(v) = (\mathrm{lcm}(M_i^{a_i})) = (M_1^{a_1} \ldots M_r^{a_r}) = (M_f)$. □

Exercises

Exercise 29.9. Let $A \in M_n(K)$ be such that $A^t = I_n$, K is assumed to be algebraically closed and of characteristic zero. Show that A is diagonalizable.

Exercise 29.10. (The killer lemma with K infinite). Let E be a K-vector space and $f \in \mathrm{End}(E)$, K is assumed to be infinite.
1) If $v \in E$, $\mathrm{Ann}(v) = (M_{f_{|F}})$ where $F = \langle v \rangle_f$. Conclude that if $\mathrm{Ann}(v) = (Q)$ then Q is a monic divisor of M_f.
2) Let $P_1, \ldots, P_t$ be the monic divisors of M_f (they are finite in number). Show that

$$E = \bigcup_{i=1}^{t} \mathrm{Ker}(P_i).$$

3) Show that, if K is infinite, E can not be equal to the union of a finite number of proper subvector spaces.
4) Conclude the proof of the killer lemma.

Chapter 30

Frobenius canonical form

30.1. Similarity invariants

The purpose of this section is to prove the following theorem:

Theorem 30.1. (Rational or Frobenius canonical form) *Let E be a K-vector space and $f \in \mathrm{End}(E)$.*

(1) *There exist cyclic subspaces $F_1 = \langle v_1 \rangle_f$, ..., $F_r = \langle v_r \rangle_f$ such that:*
   (i) *$E = F_1 \oplus \cdots \oplus F_r$;*
   (ii) *if $P_i = M_{f|F_i}$, then $P_1 = M_f$ and $P_r \mid P_{r-1} \mid \cdots \mid P_2 \mid P_1$;*
   (iii) *there exists a basis $\mathcal{B}$ of E such that:*

$$\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \mathcal{C}(P_1) & & \\ & \ddots & \\ & & \mathcal{C}(P_r) \end{pmatrix}.$$

(2) *The polynomials $P_1, \ldots, P_r$ are uniquely determined by f and are the similarity invariants of f.*

(3) *$P_f = P_1 \cdots P_r$.*

Remark 30.2. Theorem 30.1 therefore states the existence of a basis in which the matrix of f is block diagonal, where the blocks are the companion matrices of the similarity invariants. Since the similarity invariants depend only on f we get a canonical representation of f. This representation is called rational because it is valid for any field K.

As we will see later, two endomorphisms f, g (resp. two matrices $A, B \in M_n(K)$) are similar if and only if they have the same similarity invariants, hence the interest of the result.

Before moving to the proof let us recall some facts about the transposed map of an endomorphism. If $f : E \to E$ is an endomorphism of E, the transposed map ${}^tf$ is defined by: ${}^tf : E^* \to E^* : \varphi \to \varphi \circ f$. If $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$, then ${}^tA = \mathrm{Mat}({}^tf; \mathcal{B}^*, \mathcal{B}^*)$, where $\mathcal{B}^*$ is the dual basis of $\mathcal{B}$. In the following lemma we collect a few facts about transposed matrices some of which have already been seen in previous chapters.

Lemma 30.3. *Let $A, B \in M_n(K)$ be two $m \times n$ matrices. Then*

(1) ${}^t(AB) = {}^tB\,{}^tA$;

(2) *if A is invertible then ${}^t(A^{-1}) = ({}^tA)^{-1}$;*

(3) *we have ${}^t(A^m) = ({}^tA)^m$, $\forall m \geq 1$;*

(4) *if $P(x) \in K[x]$, ${}^t(P(A)) = P({}^tA)$;*

(5) *A and ${}^tA$ have the same characteristic and minimal polynomial.*

Proof. Exercise 30.10. □

Remark 30.4. The previous lemma can also be proved using the definition of ${}^tf$ and the fact that if $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$, then ${}^tA = \mathrm{Mat}({}^tf; \mathcal{B}^*, \mathcal{B}^*)$.

Note in particular the following translation of (3): $({}^tf)^i(\varphi) = \varphi \circ f^i$. We will see later that A and ${}^tA$ are similar.

Lemma 30.5. *Let E be a K-vector space, $f \in \mathrm{End}(E)$ and $v \in E$ be such that $\mathrm{Ann}(v) = (M_f)$. Then there exists a stable subvector space $W \subset E$ such that $E = \langle v \rangle_f \oplus W$.*

Proof. Let $F := \langle v \rangle_f$ be of dimension r, then $F = \langle v, f(v), \ldots, f^{r-1}(v) \rangle$ where the vectors $e_1 = v, e_2 = f(e_1), \ldots, e_r = f^{r-1}(e_1)$, are linearly independent. Recall that $r = \deg M_f$ (Remark 29.1). Complete to a basis $\mathcal{B} = (e_1, \ldots, e_r, e_{r+1}, \ldots, e_n)$ of E ($n = \dim(E)$).

Let $\varphi \in E^*$ be such that $\varphi(e_1) = \cdots = \varphi(e_{r-1}) = 0$ and $\varphi(e_r) = 1$ (for example $\varphi = e_r^*$), and $U \subset E^*$, $U = \langle \varphi \rangle_{{}^tf}$. So U is the cyclic subspace generated by

$$\varphi, {}^tf(\varphi), ({}^tf)^2(\varphi), \ldots, ({}^tf)^i(\varphi), \ldots$$

i.e. (Remark 30.4) $U = \langle \varphi \circ f^i \rangle$.

Now, let $W = U^\circ$. So $W = \{w \in E \mid (\varphi \circ f^i)(w) = 0, \forall i \geq 0\}$. Note that W is stable: if $w \in W$, $(\varphi \circ f^i)(f(w)) = (\varphi \circ f^{i+1})(w) = 0$, so $f(w) \in W$. Moreover, $W \cap F = \{0\}$. Let $w \in W \cap F$, then $w = \alpha_1 e_1 + \cdots + \alpha_r e_r$ ($w \in F$). Then $\varphi(w) = 0$ ($w \in W$), but $\varphi(w) = \alpha_r$. So $w = \alpha_1 e_1 + \cdots + \alpha_{r-1} e_{r-1}$. We have $(\varphi \circ f)(w) = 0 = \varphi(\alpha_1 e_2 + \cdots + \alpha_{r-1} e_r) = \alpha_{r-1}$. Proceeding in this way, up to $(\varphi \circ f^{r-1})(w) = 0 = \alpha_1$, we see that $w = 0$.

Furthermore $E = F \oplus W$. We have $\dim U = \deg m$, where m is the minimal polynomial of ${}^t f_{|\langle\varphi\rangle_{t_f}}$ (Remark 29.1). As usual $m \mid M_{tf}$. But ${}^t f$ and f have the same minimal polynomial (Lemma 30.3). So $\dim(U) \leq \deg M_f = \dim F$. Therefore $\dim W \geq n - \dim F$. If the inequality is strict we have, by Grassmann, $F \cap W \neq \{0\}$. So $\dim W = n - \dim F$ and $E = F \oplus W$. □

Now, we can prove the main theorem.

Proof. (of Theorem 30.1) We proceed by induction on $\dim E$. If $\dim E = 1$, there is nothing to prove. Suppose the theorem proved in dimension smaller that $n = \dim E$.

By the killer lemma (Lemma 29.8) there exists $v \in E$ such that $\mathrm{Ann}(v) = (M_f)$. If $\langle v\rangle_f = E$, E_f is cyclic, the matrix of f in the cyclic basis $(v, f(v), \ldots, f^{n-1}(v))$ is $\mathcal{C}(M_f)$. Furthermore $M_f = P_f$. Then conditions (i), (ii), (iii) are satisfied.

If $\langle v\rangle_f \neq E$, by Lemma 30.5 there exists a stable subspace W such that $E = F_1 \oplus W$ ($F_1 := \langle v\rangle_f$). By induction hypothesis W is a direct sum of cyclic subspaces: $W = F_2 \oplus \cdots \oplus F_r$ and if $P_i = M_{f_{|F_i}}$, $i = 2, \ldots, r$ then $P_r \mid P_{r-1} \mid \cdots \mid P_3 \mid P_2$. Furthermore, if $\mathcal{B}'$ is the basis of W given by the union of the cyclic bases of the various $F_i = \langle v_i\rangle_{f_{|W}}$, then $\mathrm{Mat}(f_{|W}; \mathcal{B}', \mathcal{B}')$ is the diagonal block matrix, whose blocks are the companion matrices of the polynomials P_i.

We have $P_2 = M_{f_{|W}}$. Since $M_f(w) = 0, \forall w \in E$ and therefore $\forall w \in W$, $P_2 = M_{f_{|W}} \mid P_1 = M_f$. Finally, the matrix of f in the basis $\mathcal{B}$, obtained by joining the cyclic basis of $\langle v\rangle_f$ to the basis $\mathcal{B}'$, is the diagonal block matrix, whose blocks are the matrices $\mathcal{C}(P_i)$. This shows the existence of the decomposition.

Let us now move to the uniqueness of the similarity invariants. Suppose we have two decompositions in cyclic subspaces

$$E = F_1 \oplus \cdots \oplus F_r = G_1 \oplus \cdots \oplus G_s$$

with $F_i = \langle v_i\rangle_f$, $G_j = \langle w_j\rangle_f$. Let $P_1, \ldots, P_r$, $Q_1, \ldots, Q_s$ be the associated polynomials ($P_i = M_{f_{|F_i}}, Q_j = M_{f_{|G_j}}$).

Suppose $(P_1, \ldots, P_r) \neq (Q_1, \ldots, Q_s)$. In any case $P_1 = Q_1 = M_f$. Furthermore, $\sum_i \deg(P_i) = \sum_j \deg(Q_j) = \dim E$. So there exists (even if $r \neq s$) an index $j \geq 2$ such that $P_i = Q_i$ if $i < j$ and $P_j \neq Q_j$.

We have $P_j(E) = P_j(F_1) \oplus \cdots \oplus P_j(F_r)$. In fact $P_j(F_i) \subset F_i$ (F_i is stable) and therefore the sum of the images is still direct. On the other hand, $P_j(F_i) = 0$ if $i \geq j$, because $P_j = M_{f_{|F_j}}$ and $M_{f_{|F_i}} = P_i \mid P_j$ if $i > j$. In conclusion:

$$P_j(E) = P_j(F_1) \oplus \cdots \oplus P_j(F_{j-1}).$$

On the other hand, we also have

$$P_j(E) = P_j(G_1) \oplus \cdots \oplus P_j(G_s).$$

The matrices of $f_{|F_i}$ and $f_{|G_k}$ in the cyclic basis are $\mathcal{C}(P_i)$, $\mathcal{C}(Q_k)$. So if $i < j$ they are equal (because $Q_i = P_i$). If $P_j(x) = x^t + a_{t-1}x^{t-1} + \cdots + a_0$ and $A_i = \mathcal{C}(P_i) = \mathcal{C}(Q_i)$ $(i < j)$ the matrix of $P_j(f_{|F_i})$ in the cyclic basis of F_i is $M_i := A_i^t + a_{t-1}A_i^{t-1} + \cdots + a_0 I_t$. Likewise the matrix of $P_j(f_{|G_i})$ in the cyclic basis of G_i is M_i.

It follows that $\dim P_j(F_i) = \dim P_j(G_i) = \mathrm{rk}\,(M_i)$, if $i < j$. This implies $P_j(G_j) = \cdots = P_j(G_s) = \{0\}$. In particular P_j annihilates G_j. So $Q_j = M_{f_{|G_j}}$ divides P_j. Exchanging P_j with Q_j, we get that $P_j \mid Q_j$ and hence $P_j = Q_j$. A contradiction. So $r = s$ and $P_i = Q_i$, $\forall i$.

If A is a diagonal block matrix, with blocks A_i, then clearly the characteristic polynomial of A is the product of the characteristic polynomials of the A_i. Since the characteristic polynomial of $\mathcal{C}(P)$ is P (Lemma 28.11), we have $P_f = P_1 \cdots P_r$. □

30.2. Frobenius form and similarity

The previous results justify the following definition.

Definition 30.6. Let K be a field and $A \in M_n(K)$. The matrix A is in rational canonical form (or is a Frobenius reduction) if

$$A = \begin{pmatrix} \mathcal{C}(P_1) & & \\ & \ddots & \\ & & \mathcal{C}(P_r) \end{pmatrix}$$

with $P_1 = M_A$ and $P_r \mid \cdots \mid P_2 \mid P_1$. The polynomials $P_1, \ldots, P_r$ are the similarity invariants of A.

Conversely, to each sequence of monic polynomials $(Q_1, \ldots, Q_s)$ in $K[x]$ with $Q_s \mid \cdots \mid Q_2 \mid Q_1$ we associate the Frobenius reduction

$$\mathrm{Can}(Q_1, \ldots, Q_s) = \begin{pmatrix} \mathcal{C}(Q_1) & & \\ & \ddots & \\ & & \mathcal{C}(Q_s) \end{pmatrix}.$$

Remark 30.7. In the previous definition we chose an order on the blocks: the first block at the top left is $\mathcal{C}(P_1)$ where $P_1 = M_A$. So $\mathrm{Can}(Q_1, \ldots, Q_s)$ is uniquely determined by $(Q_1, \ldots, Q_s)$. Obviously, each Q_i has positive degree (otherwise $M_{f_{|F_i}} = 1$ i.e. Id annihilates F_i i.e. $F_i = \{0\}$).

A diagonal block matrix with companion block matrices of polynomials $P_1, \ldots, P_r$ is not necessarily a Frobenius reduced matrix: we must have $P_r \mid \cdots \mid P_1$ and the blocks must be in the right order.

Proposition 30.8. *Let $M, N \in M_n(K)$ be two $n \times n$ matrices. Then*

(1) *the matrices M and N are similar if and only if they have the same similarity invariants;*

(2) *two Frobenius reduced matrices are similar if and only if they are equal;*

(3) *each matrix $M \in M_n(K)$ is similar to a unique Frobenius reduced matrix (the one constructed with its similarity invariants). This Frobenius reduction will be denoted by* $\mathrm{Can}(A)$ *(canonical or rational form of A).*

Proof. 1) If $M \approx N$ (M, N similar) then there exist a K-vector space E, of dimension n, an endomorphism $f \in \mathrm{End}(E)$ and bases $\mathcal{B}, \mathcal{B}'$ of E such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = M$ and $\mathrm{Mat}(f; \mathcal{B}', \mathcal{B}') = N$. Therefore (Theorem 30.1) the similarity invariants of M (resp. N) are those of f.

If $P_1, \ldots, P_r$ are the similarity invariants of M, N then

$$M \approx \begin{pmatrix} \mathcal{C}(P_1) & & \\ & \ddots & \\ & & \mathcal{C}(P_r) \end{pmatrix} \approx N.$$

So $M \approx N$.

2) Follows from 1).

3) Follows from Theorem 30.1 and from 2). □

From the above we finally obtain an exhaustive description of the similarity classes of $M_n(K)$ with K an arbitrary field.

Corollary 30.9. *The similarity classes of matrices in $M_n(K)$ are in one-to-one correspondence with the sequences of non constant monic polynomials $(P_1, \ldots, P_r)$ such that $P_r \mid \cdots \mid P_1$ and $\sum_i \deg(P_i) = n$.*

Exercises

Exercise 30.10. Prove Lemma 30.3.

Exercise 30.11. 1) Determine the similarity invariants of αI_3.
2) Let $D(a_1, \ldots, a_n)$ be the diagonal matrix of order n with $a_1, \ldots, a_n$ on the diagonal ($a_i = a_{ii}$). Determine the similarity invariants of $D(a_1, a_2, a_3)$ where $a_i \neq a_j$ if $i \neq j$.
3) Determine the similarity invariants and the canonical matrix of $D(a, a, b)$ (a, b distinct) and $D(a, a, a, b, b, c)$ (a, b, c distinct).

Exercise 30.12. Let E be a K-vector space and $f \in \mathrm{End}(E)$. Show that E_f is cyclic if and only if $P_f = M_f$.

Exercise 30.13. Let E be a K-vector space and $f \in \mathrm{End}(E)$. Show that P_f and M_f have the same irreducible factors.

Chapter 31

Canonical form: applications

31.1. Similarity of matrices of order at most three

One may wonder whether there is a simple criterion to determine if two matrices $A, B \in M_3(\mathbb{R})$ are similar or not. Of course, one can check if there exists an invertible matrix P such as $PA = BP$, but this is not always easy.

Proposition 31.1. *Let K be a field and $A, B \in M_n(K)$ with $2 \leq n \leq 3$. The following are equivalent:*

(1) *$A \approx B$ (A and B are similar);*

(2) *$P_A = P_B$ and $M_A = M_B$;*

(3) *$P_A = P_B$ and for each eigenvalue λ the eigenspaces associated to λ have the same dimension.*

Proof. We will do the proof for the case $n = 3$ leaving the proof for the case $n = 2$ (even easier) to the reader.

It is clear that (1) implies (2). We show that (2) implies (1). If $\deg(M_A) = 3$, then $P_A = M_A$ and the similarity invariants of A are $(P_1 = M_A)$, but we also have $P_B = M_B$. Therefore, A and B have the same invariants and hence they are similar.

If M_A has degree two the invariants of A are $(P_1 = M_A, P_2)$, with P_2 of degree one. Since $P_2 \mid P_1$ we are in the case $P_1 = (x-a)(x-b)$, $P_2 = x - a$ and $P_A = (x-a)^2(x-b)$ (possibly $a = b$). Since $P_A = P_B, M_A = M_B$, the invariants of B are $(Q_1 = M_B = (x-a)(x-b), Q_2)$, with $Q_2 \mid Q_1$. Furthermore, since $P_B = Q_1 Q_2$ also A and B have the same invariants. If $M_A = (x - \lambda)$ then $A = \lambda I_3$ and the same is true for B.

We showed that (1) $\Leftrightarrow$ (2). Now, we show that (1) and (2) $\Leftrightarrow$ (3).

(1) $\Rightarrow$ (3) If $A \approx B$, A and B represent the same endomorphism f in different bases and therefore the eigenspaces related to an eigenvalue λ have the same dimension that is $\dim E_A(\lambda) = \dim E_B(\lambda) = \dim E_f(\lambda)$.

(3) $\Rightarrow$ (2) If $P_A = P_B$ is irreducible then $M_A = P_A = M_B$. If $P := P_A = P_B = (x-\lambda)Q(x)$, with $Q(x)$ irreducible, then $M_A = P = M_B$. In fact, the minimal polynomial and the characteristic polynomial have the same irreducible factors (Exercise 30.13). Or look at the similarity invariants: if $M_A \neq P$, the invariants of A are $(P_1 = M_A, P_2)$, with $P = P_1P_2$ and $P_2 \mid P_1$, a contradiction.

We can therefore assume that P has all its roots in K. In this case also M_A, M_B all have their roots in K which are the same as those of P (Corollary 27.6). If the roots are distinct then A, B are diagonalizable and $M_A = P = M_B$.

Assume $P(x) = (x-\alpha)(x-\beta)^2$, with $\alpha \neq \beta$. If $\dim(E_A(\beta)) = \dim(E_A(\beta)) = 1$, A, B are not diagonalizable and M_A, M_B have a double root, then $M_A = P = M_B$. If instead the dimension of the eigenspace is two, A, B are diagonalizable and M_A, M_B do not have multiple roots, then $M_A = M_B = (x-\alpha)(x-\beta)$. It remains to deal with the case $P = P_A = P_B = (x-\lambda)^3$ (Exercise 31.20). □

Remark 31.2. Note this curious fact: in dimension three if M_f has degree two, then M_f can not be irreducible, therefore it is reducible and P_f has all its roots in K and one of them is at least double.

Remark 31.3. The previous result no longer holds if $n \geq 4$ (Exercise 31.21).

31.2. Each matrix $A \in M_n(K)$ is similar to its transpose

Lemma 31.4. *Let E be a K-vector space and $f \in \mathrm{End}(E)$. If E_f is cyclic $E_f = \langle v \rangle_f$, then $\mathcal{B} = (e_1 = v, e_2 = f(v), \ldots, e_n = f^{n-1}(v))$, the cyclic basis of $\langle v \rangle_f$, is a basis of E.*

(1) *Under these conditions $E^*_{{}^tf}$ is also cyclic: $E_{{}^tf} = \langle e_n^* \rangle_{{}^tf}$.*

(2) *If $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ for some basis $\mathcal{B}$ of E, then $A \approx {}^tA$.*

Proof. (1) Let $\mathcal{B}^* = (e_n^*, e_n^* \circ f, \ldots, e_n^* \circ f^{n-1}) = (({}^tf)^i(e_n^*)))_{0 \leq i \leq n-1}$. We show that $\mathcal{B}^*$ is a basis of E^*.

Let $\alpha_0 e_n^* + \alpha_1(e_n^* \circ f) + \cdots + \alpha_{n-1}(e_n^* \circ f^{n-1}) = 0$. Evaluating at $e_1 = v$, keeping in mind that $f^i(e_1) = e_{i+1}$ if $i \leq n-1$, we get $\alpha_{n-1} = 0$.

Suppose $\alpha_0 e_n^* + \alpha_1(e_n^* \circ f) + \cdots + \alpha_k(e_n^* \circ f^k) = 0$, with $k < n-1$. Evaluating at $e_{n-k} = f^{n-k-1}(e_1)$, we get $\alpha_k = 0$ (if $i < k$, $f^i(e_{n-k}) = f^{n-k-1+i}(e_1) = e_{n-k+i}$). We conclude that $\alpha_i = 0, \forall i$ and therefore $\mathcal{B}^*$ is a basis of E^*.
This shows that $E^*_{{}^tf} = \langle e_n^* \rangle_{{}^tf}$.

(2) The similarity invariants of f, and therefore of A, are ($P_1 = M_f = P_f$) because f is cyclic (Proposition 28.12). In the same way the invariants of ${}^tf$, and therefore of ${}^tA$, are ($Q_1 = M_{{}^tf} = P_{{}^tf}$). Since $P_f = P_{{}^tf}$ (Lemma 30.3), A and ${}^tA$ have the same invariants and are therefore similar. □

Proposition 31.5. *Let $A \in M_n(K)$, with K any field. Then A is similar to ${}^tA$.*

Proof. Let $f \in \mathrm{End}(E)$ be such that $A = \mathrm{Mat}(f; \mathcal{B}', \mathcal{B}')$ where $\mathcal{B}'$ is any basis of E. By Theorem 30.1 there exists a decomposition $E = F_1 \oplus \cdots \oplus F_r$ of E into cyclic subspaces $F_i = \langle v_i \rangle_f$. If $t := \dim F_i$ and $\mathcal{B}_i = (e_i = f^{i-1}(v_i))$ is the associated cyclic basis, then (Lemma 31.4) F_i^* is cyclic. Set $\mathcal{B}_i = (e_t^* \circ f^i), 0 \leq i \leq t-1$. In particular, we have an isomorphism $\Phi_i : F_i \to F_i^* : e_i \to e_t^* \circ f^{i-1}$. Putting these isomorphisms together we obtain an isomorphism $\Phi : E \to E^*$ which allows us to write $E^* = F_1^* \oplus \cdots \oplus F_r^*$ (the problem is that if $F \subset E$ is a proper subspace, F^* is not naturally a subspace of E^*).

The matrix of f in the basis $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ is the block diagonal matrix with block companion matrices $\mathcal{C}(P_{f_i})$, where $f_i = f_{|F_i}$. The matrix of ${}^tf$ in the basis $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ is the block diagonal matrix with companion blocks $\mathcal{C}(P_{{}^tf_i})$. Now, $P_{f_i} = P_{{}^tf_i}$ implies $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = \mathrm{Mat}({}^tf; \mathcal{B}, \mathcal{B}) := M$, and since $A \approx M$, ${}^tA \approx M$ we have $A \approx {}^tA$. □

31.3. Similarity invariants and field extensions

Let $A \in M_n(\mathbb{R})$, since $\mathbb{R} \subset \mathbb{C}$, we can consider A as an element of $M_n(\mathbb{C})$. Since $P_A(x) = \det(xI_n - A)$ depends only on the coefficients of A (determinant formula), the characteristic polynomial of A considered as an element of $M_n(\mathbb{R})$ is equal to the characteristic polynomial of A considered as an element of $M_n(\mathbb{C})$. What happens with the minimum polynomial? It is not clear a priori that we have $M_{A|\mathbb{R}} = M_{B|\mathbb{C}}$.

Let us consider the general case where $K \subset K'$ is a field extension. If $A \in M_n(K)$ we will indicate with $M_{A|K}$ the minimal polynomial of A as an element of $M_n(K)$ (same notation for the characteristic polynomial).

Proposition 31.6. *Let $K \subset K'$ be a field extension and $A \in M_n(K)$.*

(1) *The similarity invariants of A as an element of $M_n(K)$ are also the similarity invariants of A as an element of $M_n(K')$.*

(2) *$P_{A|K} = P_{A|K'}$ and $M_{A|K} = M_{A|K'}$.*

(3) *Let $A, B \in M_n(K)$. The matrices A, B are similar as elements of $M_n(K')$ if and only if they are similar as elements of $M_n(K)$.*

Proof. (1) Comes from the uniqueness of the similarity invariants (Theorem 30.1). (2) Follows from (1): if $(P_1, \ldots, P_r)$ are the similarity invariants of A, then $P_1 =$

$M_{A|K} = M_{A|K'}$ and $P_1 \cdots P_r = P_{A|K} = P_{A|K'}$.
(3) Follows from (1). □

31.4. Nilpotent matrices

As we know an endomorphism $f \in \mathrm{End}(E)$ (respectively a matrix $A \in M_n(K)$) is *nilpotent* if there exists t such that $f^t = 0$ ($A^t = 0$).

If λ is an eigenvalue of f then $f(v) = \lambda v, v \neq 0$, and $f^t(v) = \lambda^t v = 0$. It follows that $\lambda = 0$. So all eigenvalues of f are zero, and $P_f(x) = x^n$ ($n = \dim E$). Hence, $M_f(x) = x^m$, with $m \leq n$ and $M_f(f) = f^m = 0$. In particular, $\min\{i \mid f^i = 0\} \leq n$.

Clearly, the similarity invariants of f will be of the form ($P_1 = x^{m_1}, P_2 = x^{m_2}, \ldots, P_r = x^{m_r}$), with $m_1 \geq m_2 \geq \cdots \geq m_r \geq 1$ and $m_1 + \cdots + m_r = n$.

Conversely, it is clear that an endomorphism with similarity invariants of this type is nilpotent (because $P_1 = M_f$, therefore $f^{m_1} = 0$).

Definition 31.7. A partition of the integer n is a sequence of integers $(m_1, \ldots, m_r)$ with $m_1 \geq \cdots \geq m_r \geq 1$ and $m_1 + \cdots + m_r = n$. We denote by $\mathcal{P}(n)$ the set of partitions of n and $p(n) := \#\mathcal{P}(n)$.

Example 31.8. We have $5 = 4+1 = 3+2 = 3+1+1 = 2+2+1 = 2+1+1+1 = 1+1+1+1+1$. So $p(5) = 7$ ((n) is a partition of n). Obviously as n increases things get more complicated, for example $p(243) = 133\,978\,259\,344\,888$. The study of the numbers $p(n)$ and their arithmetic properties is an important topic in number theory [**HW79**, Chapter XIX].

Summarizing what has been done so far:

Proposition 31.9. *The set of similarity classes of nilpotent matrices of order n is in one-to-one correspondence with $\mathcal{P}(n)$, the set of partitions of n.*

Remark 31.10. In Proposition 31.9 the zero endomorphism is included, it corresponds to the partition $1+1+\cdots+1$ (which corresponds to the invariant $(x, \ldots, x)$). Therefore $M_f = x$, i.e. $f = 0$).

Let $A \in M_n(K)$ with $A^{n-1} \neq 0$ and $A^n = 0$. Then $M_A(x) = x^n = P_A(x)$, the similarity invariants are (x^n). In particular A is similar to the companion matrix of x^n, that is

$$A \approx \mathcal{C}(x^n) = \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & 0 \\ & & 1 & 0 \end{pmatrix}$$

Definition 31.11. A *Jordan block* of order r relative to λ is a matrix $J(\lambda, r) \in M_r(K)$ where

$$J(\lambda; r) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

with λ on the diagonal, 1 on the superdiagonal, and all other coefficients zero.

Going back to nilpotent matrices:

Lemma 31.12. *Let $A \in M_n(K)$, with $A^{n-1} \neq 0, A^n = 0$. Then $A \approx J(0; n)$.*

Proof. We have $A \approx \mathcal{C}(x^n) = {}^tJ(0; n)$. Using Proposition 31.5 we get $A \approx J(0; n)$. □

More generally:

Theorem 31.13 (Jordan form of nilpotent matrices)**.** *Let $A \in M_n(K)$ be a nilpotent matrix with similarity invariants $(x^{m_1}, \ldots, x^{m_r})$, $m_1 \geq \cdots \geq m_r \geq 1$, $m_1 + \cdots + m_r = n$.*

Then A is similar to $D(J(0; m_1), J(0; m_2), \ldots, J(0; m_r))$, the block diagonal matrix with blocks $J(0; m_i)$. The matrix $D(J(0; m_1), J(0; m_2), \ldots, J(0; m_r))$ is the Jordan form of A.

Proof. By Theorem 30.1 A is similar to the block-diagonal matrix, D_0 with $\mathcal{C}(x^{m_i})$ blocks. We have ${}^t\mathcal{C}(x^{m_i}) = J(0; m_i)$. By Proposition 31.5 $\mathcal{C}(x^{m_i}) \approx J(0; m_i)$.

We conclude that $D_0 \approx D(J(0; m_1), J(0; m_2), \ldots, J(0; m_r))$. In fact, if $A = D(A_1, .., A_r)$, $B = D(B_1, \ldots, B_r)$ are two diagonal block matrices, with blocks of the same order ($A_i, B_i \in M_{n_i}(K)$), then $A_i \approx B_i, \forall i \Rightarrow A \approx B$. This follows from the fact that the product of two such matrices is a block matrix of the following form: $AB = D(A_1B_1, \ldots, A_rB_r)$. □

Remark 31.14. Let $J = D(J(0; m_1), \ldots, J(0; m_r))$. Since a Jordan block $J(0; m)$ clearly has rank $m - 1$, the rank of J is $\sum_{i=1}^{r}(m_i - 1) = n - r$ (consider the row vectors), i.e. $\dim \mathrm{Ker}(J) = r$. We conclude that the number of blocks in the Jordan form of A is equal to $\dim \mathrm{Ker}(A)$.

31.5. Jordan form

Let $f \in \mathrm{End}(E)$, E a K-vector space. Under the hypothesis that $P_f(x)$ has all its roots in K, we show the existence of a basis $\mathcal{B}$ such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = J$, where J is a block diagonal matrix, with Jordan blocks $J(\lambda_i, m_i)$ where the λ_i are the

eigenvalues of f (the same λ_i may appear multiple times). The matrix J is uniquely determined by the similarity class of f and is called the *Jordan form* of f (or of A if $A = \mathrm{Mat}(f; C, C)$, C a basis of E).

Definition 31.15. A *Jordan matrix* is a block-diagonal matrix, J, with Jordan blocks. So $J = D(J(\alpha_1, m_1), \ldots, J(\alpha_r, m_r))$ (note the α_i are not necessarily distinct).

Theorem 31.16 (Jordan form)**.** *Let E be a K-vector space and $f \in \mathrm{End}(f)$. Assume that $P_f(x)$ has all its roots in K. Then there exists a basis $\mathcal{B}$ of E such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = J$ is a Jordan matrix. The matrix J is uniquely determined by the similarity class of f.*

Proof. By hypothesis we have $P_f(X) = \prod_{i=1}^{r}(x - \lambda_i)^{\alpha_i}$ ($\lambda_i \neq \lambda_j$ if $i \neq j$) and therefore (Corollary 27.6) $M_f(x) = \prod_{i=1}^{r}(x-\lambda_i)^{\beta_i}$, with $1 \leq \beta_i \leq \alpha_i$ and $\alpha_1 + \cdots + \alpha_r = n := \dim E$. By Theorem 29.3 we have the decomposition into characteristic subspaces: $E = E_1 \oplus \cdots \oplus E_r$, where $E_i = \mathrm{Ker}(\lambda_i Id - f)^{\beta_i}$. Furthermore $\dim E_i = \alpha_i$ (Theorem 29.6). Since E_i is stable $f_i = f_{|E_i} \in \mathrm{End}(E_i)$ and therefore also $g_i = f_i - \lambda_i Id_{E_i} \in \mathrm{End}(E_i)$.

The endomorphism g_i is obviously nilpotent ($E_i = \mathrm{Ker}(g_i^{\beta_i})$). Then by Theorem 31.13 there exists a basis $\mathcal{B}_i$ of E_i such that

$$\mathrm{Mat}(g_i; \mathcal{B}_i, \mathcal{B}_i) = D(J(0; m_{j_1}), \ldots, J(0; m_{j_i}))$$

where the integers $m_{j_1}, \ldots, m_{j_i}$ are determined by the similarity invariants of g_i. Since $f_i = g_i + \lambda_i Id$, we have $\mathrm{Mat}(f_i; \mathcal{B}_i, \mathcal{B}_i) = D(J(\lambda_i; m_{j_1}), \ldots, J(\lambda_i; m_{j_i})) = J_i$, where J_i is a Jordan matrix. If $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$, then $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = J = D(J_1, \ldots, J_r)$ is a Jordan matrix .

From what we have seen on nilpotent matrices it is clear that J depends only on the similarity class of $M = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$ ($\mathcal{B}$ any basis). Note that if $B = A - \lambda I_n$, $B' = A' - \lambda I_n$ we have $B \approx B'$ if and only if $A \approx A'$. □

Remark 31.17. So if $A \in M_n(K)$ is such that $P_A(x)$ has all its roots in K, then $A \approx J = D(J(\alpha_1, m_1), \ldots, J(\alpha_r, m_r))$ where the α_i are not necessarily distinct.

From the previous proof we see that every α_i is an eigenvalue and that every eigenvalue of A appears among the α_i. Furthermore, if λ_k is an eigenvalue then $\#\{i \mid \alpha_i = \lambda_k\} = \dim E_A(\lambda_k)$, where $E_A(\lambda_k)$ indicates the eigenspace relative to λ_k. Said differently, the number of blocks in which λ_k appears is equal to $\dim E_A(\lambda_k)$. This follows from Remark 31.14 and the previous proof.

Example 31.18. Let

$$B = \begin{pmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{pmatrix} \in M_3(\mathbb{R}).$$

Then $P_B(x) = (x-2)^3$. There is a Jordan form for B. Since 2 is the only eigenvalue it is just a matter of figuring out how many blocks there are. We have $\mathrm{rk}(2I_3 - B) = 1$, i.e. $\dim E_B(2) = 2$. So there are two blocks and, necessarily,

$$B \approx \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

31.6. How to find the Jordan form

Let us start with a simple but important observation. Let $J = J(0,n)$ be a Jordan matrix of order n for the eigenvalue 0. Then J is an $n \times n$ matrix with all 0 on the diagonal, all 1 above the diagonal, and the other coefficients are zero. If J corresponds to the endomorphism f, f is nilpotent with $f^n = 0$, $f^{(n-1)} \neq 0$. We have $\mathrm{rk}(J) = n-1$ and $\dim \mathrm{Ker}(f) = 1$.

Then if $d_i = \dim \mathrm{Ker}(f^i)$ we have $d_i = i$ if $1 \leq i \leq n$ and $d_k = n$ if $k \geq n$. In fact we have $f(e_1) = 0, f(e_i) = e_{i-1}$. So $f^2(e_2) = f(e_1) = 0$, if $i > 2$, $f^2(e_i) = f(e_{i-1}) = e_{i-2} \neq 0$. More generally we can write $f^k(e_i) = e_{i-k}$ with the convention $e_{i-k} = 0$ if $i - k \leq 0$. Hence $d_k = k$.

Let us now consider the general nilpotent case $J = D(J(0,m_1), \ldots, J(0,m_r))$ with $m_1 \geq \cdots \geq m_r$, $m_1 + \cdots + m_r = n$. The matrix J has rank $\sum^r (m_i - 1) = n - r$, so if $d_i = \dim \mathrm{Ker}(J^i)$, we have $d_1 = r$, that is, d_1 gives us the number of blocks. When we multiply by J, as we saw before, the dimension of the kernel of each non-zero block increases by one. So, for $i \leq m_r$, $d_i = ir$, if there are t blocks of dimension m_r (i.e. if $m_r = m_{r-1} = \cdots = m_{r-(t-1)}$), we will have $d_{m_r+1} = rm_r + (r-t)$, there are only $r - t$ non zero blocks each increasing the dimension of the kernel by one. The zero blocks do not increase the dimension of the kernel.

Set $d_0 = 0$ and $\delta_i = d_i - d_{i-1}$. We have $\delta_i = r$ for $i \leq m_r$ and $\delta_{m_r+1} = r - t$. Continuing in this way we see that δ_i is the number of blocks of dimension at least i. Furthermore, $m_i = \#\{t \mid \delta_t \geq i\}$. We also have that the number of blocks of dimension j is equal to $\delta_j - \delta_{j+1}$.

In conclusion, the sequence $\dim \mathrm{Ker}(J^i)$ determines the Jordan form (this proves the uniqueness, modulo the order of the blocks, of the Jordan form).

The previous considerations apply to the general case with multiple eigenvalues that are not necessarily zero, in this case we must consider the matrices $J - \lambda_i I_n$

for each eigenvalue λ_i and observe that the blocks $J(\lambda_k, m_i) - \lambda_i I_n$ are invertible if $\lambda_k \neq \lambda_i$.

Example 31.19. Consider the matrix

$$A = \begin{pmatrix} 5 & 0 & 4 & -2 & -3 \\ -2 & 3 & -3 & 2 & 4 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 1 & 0 & 2 & -1 & 1 \end{pmatrix}.$$

It turns out that $P_A(x) = (x-3)^5$. Furthermore $\dim E_A(3) = 2$. The matrix is not diagonalizable but has a Jordan form. We know that the Jordan form has two blocks, so there are two possibilities: $(4,1)$ or $(3,2)$.

The matrix $B = A - 3I_5$ is nilpotent and if $A \approx J$ then $B \approx J - 3I_5$. So it is enough to find the Jordan form of B.

The idea is to calculate B^3. If $B^3 \neq 0$ then we are in case $(4,1)$ otherwise we are in case $(3,2)$. Actually, one just needs to calculate B^2.

We see B^2 has rank 1, therefore a kernel of dimension 4. So $d_1 = 2, d_2 = 4$. In particular $\delta_2 = 2$ i.e. there are two blocks of dimension at least 2 and we are in case $(3,2)$.

Exercises

Exercise 31.20. 1) Prove Proposition 31.1 in the case $n = 2$.
2) Conclude the proof of Proposition 31.1.

Exercise 31.21. Show that there exist two matrices 4×4 with coefficients in K which have the same characteristic polynomial and the same minimal polynomial but which are not similar.

Exercise 31.22. Let $K \subset K'$ be a field extension, E a K-vector space, F a K'-vector space, and $\mathcal{B} = (e_1, \ldots, e_n)$ a basis of E. Assume that there exists a K-linear map $\varphi : E \to F$ (note that F is also a K-vector space) such that $\mathcal{C} = (\varphi(e_1), \ldots, \varphi(e_n))$ is a basis of the K'-vector space F (an example of this situation is the following $\varphi : \mathbb{R}^n \to \mathbb{C}^n$ which sends the canonical basis of $\mathbb{R}^n$ to the canonical basis of $\mathbb{C}^n$.
1) Compute $\dim_K(F)$ and $\dim_K(\mathrm{Ker}(\varphi))$.
2) Is $\mathrm{Im}(\varphi) \subset F$ a K'-subvector space of F?
3) Let $v_1, \ldots, v_m$ be m linearly independent vectors of E. Show that $\varphi(v_1), \ldots, \varphi(v_m)$ are linearly independent in the K'-vector space F.
4) Let $f \in \mathrm{End}(E)$ be such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = A$. We define an endomorphism, $\tilde{f}$, of the K'-vector space F by $\tilde{f}(\varphi(e_i)) = \varphi(f(e_i))$. Determine $\mathrm{Mat}(\tilde{f}; \mathcal{C}, \mathcal{C})$.

Exercise 31.23. Let $K \subset K'$ be a field extension and $A \in M_n(K)$ with minimal polynomial $m_A(x) \in K[x]$ of degree d. Via the natural inclusion $M_n(K) \hookrightarrow M_n(K')$ we can see A as an element of $M_n(K')$. Let $M_A(x) \in K'[x]$ be the minimal polynomial of A on K'. We have that $M_A(x) \mid m_A(x)$ in $K'[x]$ and, a priori, we could have $M_A(x) \neq m_A(x)$. Using Exercise 31.22 prove that $M_A(x) = m_A(x)$.

Exercise 31.24. We resume the situation of Exercise 31.22 with $\mathbb{R} \subset \mathbb{C}$ and $\varphi : \mathbb{R}^n \to \mathbb{C}^n$ which sends the canonical basis to the canonical basis.
1) Each vector $w \in \mathbb{C}^n$ is written (uniquely) as $w = u + iv$ with $u, v \in \mathbb{R}^n$. In the situation of item 4) of Exercise 31.22, show that $\overline{f}(w) = f(u) + if(v)$ (use $A = \mathrm{Mat}(f; \mathcal{B}, \mathcal{B})$).
2) Let λ be a non-real eigenvalue ($\lambda = a + ib, a, b \in \mathbb{R}, b \neq 0$)) of the matrix A and let $w = u + iv$ be an associated eigenvector. Observe that $Aw = \lambda w = Au + iAv$. Show that the subspace $\langle u, v \rangle \subset \mathbb{R}^n$ is stable, of dimension two, under A (i.e. under f). In conclusion, a complex, non-real root of $P_f(x)$ provides a stable subspace of f, of dimension two.
3) Give an example of an endomorphism f of $\mathbb{R}^3$, non-diagonalizable and such that $\mathbb{R}^3 = F \oplus G$, with F, G stable.

Exercise 31.25. Find the Jordan form and the invariant factors of

$$A = \begin{pmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{pmatrix} \in M_3(\mathbb{R}).$$

Exercise 31.26. Consider the matrix

$$A = \begin{pmatrix} 4 & 3 & -2 \\ -3 & -1 & 3 \\ 2 & 3 & 0 \end{pmatrix} \in M_3(\mathbb{R}).$$

1) State whether A is diagonalizable and find bases of the eigenspaces.
2) If v_2 is a vector such that $Av_2 = 2v_2$, look for a vector v_3 such that $(A-2I_3)v_3 = v_2$ (this involves solving a linear system, explain why the system is solvable).
3) Find an invertible matrix P such that $A = PJP^{-1}$, where J is the Jordan form of A.

Exercise 31.27. 1) We take the matrix A from Example 31.19, and set

$$A = \mathrm{Mat}(f; \mathcal{C}, \mathcal{C})$$

where $\mathcal{C}$ is the canonical basis of $\mathbb{R}^5$. Reasoning as in Exercise 31.26, find a basis $\mathcal{B}$ such that $\mathrm{Mat}(f; \mathcal{B}, \mathcal{B}) = J$, where J is the Jordan form of A.
2) Describe a general method for finding the Jordan form.

Exercise 31.28. Consider the matrix

$$A = \begin{pmatrix} 1 & -1 & 2 & 1 \\ -1 & 4 & -7 & -1 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in M_4(\mathbb{R}).$$

State whether A has a Jordan form and, if the answer is positive, determine it.

Exercise 31.29. 1) Show that a Jordan block $J(\lambda, n)$ is similar to its transpose (do a clever base change).
2) Let $A \in M_n(K)$ be such that $P_A(x)$ has all its roots in K. Show that A is similar to its transpose.

Exercise 31.30. The aim of this exercise is to show that if K is algebraically closed, with $\mathrm{char}(K) \neq 2$, then every matrix of $M_n(K)$ is similar to a symmetric matrix.
1) Show that if $S \in M_n(K)$ is a symmetric matrix of rank n, then there exists an invertible matrix P such that $S = {}^tPP$.
2) Show with a counterexample that 1) is not verified if $\mathrm{char}(K) = 2$ (consider an antidiagonal symmetric matrix 2×2).

3) If $J = J(\lambda, n)$ is a Jordan block, using Exercise 31.29, show that there exists a symmetric matrix S of rank n such that

$$S^{-1}JS = {}^tJ.$$

4) By 1) $S = {}^tPP$. Show that (with the notations of 2)) $P\,{}^tJP^{-1}$ is symmetric.
5) Conclude that every $A \in M_n(K)$, K algebraically closed with $\mathrm{char}(K) \neq 2$, is similar to a symmetric matrix.
6) So every matrix $A \in M_n(\mathbb{C})$ is similar to a symmetric matrix. Therefore, as we have seen, there exist symmetric matrices $S \in M_n(\mathbb{C})$ that are not diagonalizable.

Part 6

Projective Geometry

Projective geometry has its origins in the early Italian Renaissance, in particular the architectural drawings of *Filippo Brunelleschi* (1377–1446) and *Leon Battista Alberti* (1404–72), who invented the method of perspective. The eye of the painter is connected to points on the landscape by the sight lines. The intersection of these sight lines with a vertical picture plane generates the drawing. Thus, the reality plane is projected onto the picture plane, hence the name projective geometry.

The French mathematicians *Girard Desargues* (1591–1661) and *Blaise Pascal* (1623–62) took the first steps by studying the properties of figures that were preserved under perspective mappings. The importance of projective geometry, however, became clear only after 1800 in the works of several other French mathematicians, notably *Jean-Victor Poncelet* (1788–1867). Ignoring geometric measurements such as distances and angles, projective geometry provides a clearer understanding of some properties of geometric objects.

Euclidean geometry is contained within projective geometry, with the latter having a simpler foundation. General results in Euclidean geometry may be derived in a more transparent manner, and different but similar theorems of Euclidean geometry may be handled collectively within the framework of projective geometry. For instance, parallel and non parallel lines need not be treated as separate cases, rather the projective plane is endowed of a line at infinity along which lines, that would be parallel in the affine plane, meet.

Chapter 32

Projective spaces

Definition 32.1. Let V be a K-vector space. The *projective space* $\mathbb{P}(V)$ is the set of 1-dimensional subvector spaces of V. If the vector space V has dimension $n+1$ then $\mathbb{P}^n := \mathbb{P}(V)$ has dimension n.

Take $K = \mathbb{R}$. A line through the origin in $\mathbb{R}^{n+1}$ intersects the sphere

$$S^n = \left\{ x \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i^2 = 1 \right\}$$

in two antipodal points. Hence, $\mathbb{P}(\mathbb{R}^{n+1})$ can be viewed as S^n with the antipodal points identified.

Note that any line through the origin intersects the upper hemisphere in a point. So, in order to construct $\mathbb{P}(\mathbb{R}^{n+1})$ one could remove the lower hemisphere and identify the antipodal points on the equator. For instance, for $n = 1$ one gets an identification between $\mathbb{P}(\mathbb{R}^2)$ and S^1.

We now introduce the construction of $\mathbb{P}(V)$ as a quotient of V. On V we consider the following equivalence relation:

$$v \sim_\lambda w \iff v = \lambda w \text{ for some } \lambda \in K^*.$$

The quotient

$$\mathbb{P}(V) = \frac{V \setminus \{0\}}{\sim_\lambda}$$

is exactly the projective space of 1-dimensional subvector spaces of V. We will write $[v] \in \mathbb{P}(V)$ for the point of $\mathbb{P}(V)$ corresponding to the vector $v \in V$.

Let $(e_0, \ldots, e_n)$ be a basis of V. We can write a vector $v = \sum_{i=0}^n x_i e_i$ on the chosen basis, and we will indicate by $[v] = [x_0 : \cdots : x_n]$ the so called *homogeneous coordinates* of the point $[v] \in \mathbb{P}(V)$.

On the the subset

$$\mathcal{U}_0 = \{[x_0 : \cdots : x_n] \mid x_0 \neq 0\} \subset \mathbb{P}(V)$$

we can write

$$[x_0 : \cdots : x_n] = \left[1 : \frac{x_1}{x_0} : \cdots : \frac{x_n}{x_0}\right]$$

yielding an identification

$$\mathcal{U}_0 \cong K^n.$$

The complementary set

$$\mathcal{U}_0^c = \{[0 : x_1 : \cdots : x_n]\} \subset \mathbb{P}(V)$$

is a projective space $\mathbb{P}^{n-1}$ with homogeneous coordinates $[x_1 : \cdots : x_n]$. Therefore,

$$\mathbb{P}^n = \mathcal{U}_0 \cup \mathcal{U}_0^c = K^n \cup \mathbb{P}^{n-1}.$$

Actually on $\mathbb{P}^n$ we have $n+1$ *affine charts*

$$\mathcal{U}_i = \{[x_0 : \cdots : x_n] \mid x_i \neq 0\} \subset \mathbb{P}^n$$

for $i = 0, \ldots, n$. The set

$$\mathcal{U}_i^c = \{[x_0 : \cdots : x_{i-1} : 0 : x_{i+1} : \cdots : x_n]\} \subset \mathbb{P}^n$$

is called the *hyperplane at infinity* of the chart $\mathcal{U}_i$.

Definition 32.2. A *linear subspace* $\mathbb{P}(W)$ of the projective space $\mathbb{P}(V)$ is the set of 1-dimensional subvector spaces of a subvector space $W \subset V$.

Note that $\mathbb{P}(W)$ is a projective space of dimension $\dim(W)-1$. If $\dim \mathbb{P}(W) = 1$ we will say that $\mathbb{P}(W)$ is a *projective line* and that it is a *projective plane* when $\dim \mathbb{P}(W) = 2$.

Proposition 32.3. *Let $[v], [w] \in \mathbb{P}(V)$ be two distinct points. There exists a unique line $\langle [v], [w] \rangle \subset \mathbb{P}(V)$ passing through $[v]$ and $[w]$.*

Proof. Since $[v] \neq [w]$ the vectors $v, w \in V$ are independent, and hence v, w generate a 2-dimensional subvector space $\langle v, w \rangle \subset V$. The projectivization $\mathbb{P}(\langle v, w \rangle)$ is a line through $[v]$ and $[w]$.

Now, let $\mathbb{P}(W_1), \mathbb{P}(W_2)$ be two lines in $\mathbb{P}(V)$ through $[v]$ and $[w]$. Then W_1, W_2 are two 2-dimensional subvector spaces containing the pair of independent vectors v, w. Hence, $W_1 = W_2$ and $\mathbb{P}(W_1) = \mathbb{P}(W_2)$. □

The following result is probably the most representative difference between the affine plane and the projective plane.

Proposition 32.4. *Let V be a K-vector space of dimension three. Consider two distinct lines $\mathbb{P}(W_1), \mathbb{P}(W_2) \subset \mathbb{P}(V) = \mathbb{P}^2$. Then $\mathbb{P}(W_1), \mathbb{P}(W_2)$ intersect in a unique point.*

Proof. The subvector spaces $W_1, W_2 \subset V$ are two distinct planes. Since V has dimension three $W_1 \cap W_2$ is a vector line $L = \langle v \rangle$ with $v \in V \setminus \{0\}$. Hence $[v] \in \mathbb{P}(W_1) \cap \mathbb{P}(W_2)$.

Now, assume that $\mathbb{P}(W_1) \cap \mathbb{P}(W_2)$ contains another point $[w] \neq [v]$. Then $\langle v, w \rangle \subset W_1 \cap W_2$. Hence $W_1 = W_2$ and $\mathbb{P}(W_1) = \mathbb{P}(W_2)$, a contradiction. □

Proposition 32.5. *Two linear subspaces $\mathbb{P}(W_1), \mathbb{P}(W_2) \subset \mathbb{P}(V) = \mathbb{P}^n$ intersect in linear subspace of dimension at least $\dim \mathbb{P}(W_1) + \dim \mathbb{P}(W_2) - n$.*

Proof. By the Grassmann formula $W_1, W_2 \subset V$ intersect in a subvector space of dimension at least $\dim W_1 + \dim W_2 - (n+1)$. Therefore, $\mathbb{P}(W_1) \cap \mathbb{P}(W_2) = \mathbb{P}(W_1 \cap W_2)$ has dimension at least $\dim \mathbb{P}(W_1) + \dim \mathbb{P}(W_2) - n$. □

32.1. Projectivities

Let V, W be two K-vector spaces and $F : V \to W$ a linear map. Assume that the kernel of F is trivial. Then a 1-dimensional subvector space $L \subset V$ gets mapped to a 1-dimensional subvector space $F(L) \subset W$, and hence F yields a well-defined map

$$f : \mathbb{P}(V) \to \mathbb{P}(W)$$

called a *projectivity* from $\mathbb{P}(V)$ to $\mathbb{P}(W)$.

Proposition 32.6. *Two linear maps $F, G : V \to W$ define the same projectivity if and only if $F = \lambda G$ for some $\lambda \in K^*$.*

Proof. Assume that $G = \lambda F$ for some $\lambda \in K^*$. Then

$$[F(v)] = [\lambda G(v)] = [G(v)]$$

for any $v \in V$, and hence F, G define the same projectivity.

Now, assume that F, G define the same projectivity $f : \mathbb{P}(V) \to \mathbb{P}(W)$ and fix a basis $(e_0, \ldots, e_n)$ of V. Since

$$[F(e_i)] = f([e_i]) = [G(e_i)]$$

we have $G(e_i) = \lambda_i F(e_i)$ for some $\lambda_i \in K^*$, and

$$G(\sum_{i=0}^{n} e_i) = \lambda F(\sum_{i=0}^{n} e_i)$$

for some $\lambda \in K^*$. Therefore,

$$\sum_{i=0}^{n} \lambda F(e_i) = \lambda F(\sum_{i=0}^{n} e_i) = G(\sum_{i=0}^{n} e_i) = \sum_{i=0}^{n} G(e_i) = \sum_{i=0}^{n} \lambda_i F(e_i)$$

and hence

$$\sum_{i=0}^{n} (\lambda - \lambda_i) F(e_i).$$

Since F is injective the $F(e_i)$ are independent and hence $\lambda_i = \lambda$ for any $i = 0, \ldots, n$.

Finally, $G(e_i) = \lambda F(e_i)$ for any $i = 0, \ldots, n$ and since $(e_0, \ldots, e_n)$ is a basis of V we get that $G = \lambda F$. □

Example 32.7. Consider two distinct projective planes $H_1, H_2 \subset \mathbb{P}^3$ and a point $[v] \in \mathbb{P}^3 \setminus H_1 \cup H_2$. For any point $[v_1] \in H_1$ the projective line $L \subset \mathbb{P}^3$ through $[v_1]$ and $[v]$ intersects H_2 in a unique point $[F(v_1)]$. This yields a projectivity

$$f : H_1 \to H_2 : \ [v_1] \to [F(v_1)].$$

Fix homogeneous coordinates $[x_0, \ldots, x_3]$ on $\mathbb{P}^3$. We may assume that $H_1 = \{x_1 = 0\}$, $H_2 = \{x_2 = 0\}$ and $[v] = [0 : 1 : 1 : 0]$. Take a point $[a] = [a_0 : 0 : a_2 : a_3] \in H_1$. The line through $[v]$ and $[a]$ is given by

$$L = \{a_0 x_1 - a_0 x_2 + a_2 x_0 = a_2 x_3 + a_3 x_1 - a_3 x_2 = 0\}$$

and it intersects the plane H_2 in the point

$$[a_0 : -a_2 : 0 : a_3].$$

Therefore, in homogeneous coordinates the map f is given by

$$f : H_1 \to H_2 : \ [a_0 : 0 : a_2 : a_3] \to [a_0 : -a_2 : 0 : a_3].$$

The following is the projective analogue of linear independence.

Definition 32.8. Let $p_1, \ldots, p_{n+2} \in \mathbb{P}(V)$ be $n+2$ points in an n-dimensional projective space. The p_i are in *general position* if each subset $\{p_{i_1}, \ldots, p_{i_{n+1}}\} \subset \{p_1, \ldots, p_{n+2}\}$ of $n+1$ points has representative vector v_{i_j} with $p_{i_j} = [v_{i_j}]$ such that $v_{i_1}, \ldots, v_{i_{n+1}}$ are linearly independent.

For instance, any three distinct points of $\mathbb{P}^1$ are in general position.

Theorem 32.9. *Let $p_1, \ldots, p_{n+2} \in \mathbb{P}(V)$ and $q_1, \ldots, q_{n+2} \in \mathbb{P}(W)$ be two sets of $n+2$ points in general position. Then there exists a unique projectivity $f : \mathbb{P}(V) \to \mathbb{P}(W)$ such that $q_i = f(p_i)$ for $i = 1, \ldots, n+2$.*

Proof. Let $v_1, \ldots, v_{n+2} \in V$ be representative vectors for $p_1, \ldots, p_{n+2} \in \mathbb{P}(V)$. Since the p_i are in general position $v_1, \ldots, v_{n+1}$ form a basis of V, and we can write $v_{n+2} = \sum_{i=1}^{n+1} \lambda_i v_i$. Again since the p_i are in general position we must have

$\lambda_i \neq 0$ for any $i = 1, \ldots, n+1$ otherwise we would have a linear relation among $n+1$ of the v_i. Hence $[\lambda_i v_i] = [v_i] = p_i$. So we could choose $\lambda_i = 1$ and write $v_{n+2} = \sum_{i=1}^{n+1} \lambda_i v_i = \sum_{i=1}^{n+1} v_i$.

Similarly, we can choose representative vectors $w_1, \ldots, w_{n+2} \in W$ for the points $q_1, \ldots, q_{n+2} \in \mathbb{P}(W)$ such that $w_{n+2} = \sum_{i=1}^{n+1} w_i$.

Now, there exists a unique linear map $F : V \to W$ such that $F(v_i) = w_i$ for $i = 1, \ldots, n+1$, and since $w_1, \ldots, w_{n+1}$ form a basis of W the map F is invertible. Moreover

$$F(v_{n+2}) = F(\sum_{i=1}^{n+1} v_i) = \sum_{i=1}^{n+1} F(v_i) = \sum_{i=1}^{n+1} w_i = w_{n+2}$$

and hence $f(p_i) = q_i$ for $i = 1, \ldots, n+2$, where $f : \mathbb{P}(V) \to \mathbb{P}(W)$ is the projectivity induced by $F : V \to W$.

Now, let $g : \mathbb{P}(V) \to \mathbb{P}(W)$ be another projectivity, induced by a linear map $G : V \to W$, such that $g(p_i) = q_i$ for $i = 1, \ldots, n+2$. Then $G(v_i) = \alpha_i w_i$ for $i = 1, \ldots, n+2$, and

$$G(v_{n+2}) = \alpha_{n+2} w_{n+2} = G(\sum_{i=1}^{n+1} v_i) = \sum_{i=1}^{n+1} G(v_i) = \sum_{i=1}^{n+1} \alpha_i w_i.$$

Therefore,

$$\sum_{i=1}^{n+1} \alpha_i w_i = \alpha_{n+2} w_{n+2} = \alpha_{n+2} \sum_{i=1}^{n+1} w_i = \sum_{i=1}^{n+1} \alpha_{n+2} w_i$$

that is $\sum_{i=1}^{n+1} (\alpha_i - \alpha_{n+2}) w_i = 0$, and since $w_1, \ldots, w_{n+2}$ are linearly independent we get $\alpha_i = \alpha_{n+2}$ for $i = 1, \ldots, n+1$. So

$$G(v_i) = \alpha_i w_i = \alpha_i F(v_i) = \alpha_{n+2} F(v_i)$$

for $i = 1, \ldots, n+2$. Finally, $G = \alpha_{n+2} F$ and hence $g = f$. □

Remark 32.10. Let $p_1, \ldots, p_{n+2} \in \mathbb{P}^n$ be $n+2$ points in general position. Theorem 32.9 says that there exists a unique projectivity $\mathbb{P}^n \to \mathbb{P}^n$ moving $p_1, \ldots, p_{n+2}$ to the *fundamental points* $[1 : 0 : \cdots : 0], \ldots, [0 : \cdots : 0 : 1], \ldots, [1 : \cdots : 1] \in \mathbb{P}^n$.

32.2. Dual projective space

The dual of the n-dimensional projective space $\mathbb{P}(V)$ is the n-dimensional projective space $\mathbb{P}^n = \mathbb{P}(V^*)$ where V^* is the dual K-vector space of V. A point of $\mathbb{P}(V^*)$ corresponds to a linear form on V, modulo scalar multiplication, which in turn defines a hyperplane in V and hence in $\mathbb{P}(V)$. Therefore, the points of $\mathbb{P}(V^*)$ are in 1-to-1 correspondence with the hyperplanes of $\mathbb{P}(V)$.

Proposition 32.11. *The points of a subspace $\mathbb{P}(U) \subset \mathbb{P}(V^*)$ of dimension r correspond to the hyperplanes of $\mathbb{P}(V)$ containing a fixed linear subspace $\mathbb{P}(W) \subset \mathbb{P}(V)$ of dimension $n-r-1$.*

Proof. The subvector space $U \subset V^*$ consists of the linear forms $V \to K$ vanishing on a subvector space $W \subset V$. Furthermore, W has dimension $\dim(V) - \dim(U) = n+1-(r+1) = n-r$. Such linear forms in turn are in bijection with the hyperplanes of $\mathbb{P}(V)$ containing $\mathbb{P}(W)$ which is a linear subspace of dimension $\dim(W) - 1 = n-r-1$. □

Remark 32.12. For instance, lines in $\mathbb{P}^2$ through a fixed point $p \in \mathbb{P}^n$ are in 1-to-1 correspondence with points of $\mathbb{P}^{2*}$ contained in a line. Therefore, the lines in $\mathbb{P}^2$ through a fixed point are in bijection with the points of $\mathbb{P}^1$.

Remark 32.13. Duality allows us to associate to a statement its dual statement whose proof could be easier than that of the original statement. Here is a basic example: Consider the statement

$$\mathcal{S}: \text{ two distinct lines in } \mathbb{P}^2 \text{ intersect in a unique point.}$$

The dual statement is

$$\mathcal{S}^*: \text{ two distinct points in } \mathbb{P}^{2*} \text{ lie on a unique line.}$$

Exercises

Exercise 32.14. Compute the projectivity $f : \mathbb{P}^2 \to \mathbb{P}^2$ of the projective plane $\mathbb{P}^2$ such that $f([1:0:0]) = [0:1:0]$, $f([0:1:0]) = [1:1:1]$, $f([0:0:1]) = [1:0:0]$, $f([1:1:1]) = [0:0:1]$.

Exercise 32.15. Compute the duals in $\mathbb{P}^{3*}$ of the point $p = [1:2:-1:3] \in \mathbb{P}^3$, the line $L = \{x_0 - x_3 = x_1 + x_2 = 0\} \subset \mathbb{P}^3$ and the plane $H = \{x_0 - x_1 + x_2 - x_3 = 0\} \subset \mathbb{P}^3$, where $[x_0 : \cdots : x_3]$ are homogeneous coordinates on $\mathbb{P}^3$.

Exercise 32.16. Let $f : \mathbb{P}(V) \to \mathbb{P}(V)$ be the projectivity induced by the invertible linear map $F : V \to V$. Show that

(1) if $v \in V$ is an eigenvector of F then $[v] \in \mathbb{P}(V)$ is a fixed point of f;

(2) if V is an $\mathbb{R}$-vector space of dimension two then any projectivity $f : \mathbb{P}(V) \to \mathbb{P}(V)$ has a fixed point.

Exercise 32.17. Let $L, R \subset \mathbb{P}^3$ be two disjoint lines and $p \in \mathbb{P}^3 \setminus \{L \cup R\}$ a point. Show that there is a unique line L_p such that

$$p \in L_p \text{ and } L_p \text{ intersects both } L \text{ and } R.$$

Exercise 32.18 (Linear projections)**.** Let $H \subset \mathbb{P}^n$ be a linear subspace of dimension m. Fix a linear subspace $\Lambda \subset \mathbb{P}^n$ of dimension $n - m - 1$ such that $H \cap \Lambda = \emptyset$. Given a point $p \in \mathbb{P}^n \setminus H$ let $\langle p, H \rangle$ be the subspace generated by p and H.

(1) Show that $\langle p, H \rangle$ intersects Λ in a point $\pi_H(p)$.

(2) The map

$$\pi_H : \mathbb{P}^n \setminus H \to \Lambda : \ p \to \pi_H(p)$$

is the linear projection from H. Show that π_H is not defined at the points of H.

(3) Let $H' \subset \mathbb{P}^n$ be a linear subspace of dimension h, not entirely contained in H, and intersecting H in a linear subspace of dimension s. Show that $\pi_H(H')$ is a linear subspace of Λ of dimension $h - s - 1$.

Chapter 33

Quadric hypersurfaces

A hypersurface X_d of degree d in $\mathbb{P}^n$ is the zero locus of a homogeneous polynomial $f \in K[x_0, \ldots, x_n]_d$ of degree d that is

$$X_d = \{[x_0 : \cdots : x_n] \in \mathbb{P}^n \mid f(x_0, \ldots, x_n) = 0,\ f \in K[x_0, \ldots, x_n]_d\}.$$

A hyperplane is a hypersurface of degree $d = 1$, a *quadric* is a hypersurface of degree $d = 2$. A quadric in $\mathbb{P}^2$ is called a *conic*.

We will always assume that the characteristic of the base field K is different from two so that there is a bijective correspondence between quadric hypersurfaces and symmetric bilinear forms. A quadric is said to be *smooth* or *non singular* if the associated symmetric bilinear form is non degenerated.

We know that, up to a change of basis, a polynomial $f \in K[x_0, \ldots, x_n]_2$ can be written as

- $f = \sum_{i=0}^{r} x_i^2$ if $K = \mathbb{C}$;
- $f = \sum_{i=0}^{p} x_i^2 - \sum_{j=p+1}^{q} x_j^2$ if $K = \mathbb{R}$.

The number of squares appearing in the expression of f is the *rank* of X_2, and the ordered pair made of the numbers of squares with coefficient 1 and those with coefficient -1 is the *signature* of X_2.

Example 33.1. A quadric $X_2 \subset \mathbb{P}^1$ is either empty, when the polynomial f does not have roots in K, or a single point with multiplicity two or a pair of points.

Example 33.2. Consider the conic

$$X_2 = \{f = x_0x_1 + x_1x_2 + x_0x_2 = 0\} \subset \mathbb{P}^2$$

with $K = \mathbb{R}$. Setting $y_0 = \frac{x_0+x_1}{2}, y_1 = \frac{x_0-x_1}{2}$ and then $z_0 = y_0+x_2, z_1 = y_1, z_2 = y_2$ we get $f = z_0^2 - z_1^2 - z_2^2$. So that X_2 is smooth (it has rank three) and has signature $(1, 2)$.

Example 33.3. Consider the conic

$$X_2 = \{x_0x_1 - x_2^2 = 0\} \subset \mathbb{P}^2.$$

In the affine charts $\mathcal{U}_0 = \{x_0 \neq 0\}$ and $\mathcal{U}_1 = \{x_1 \neq 0\}$ the conic X_2 looks like a parabola while in the chart $\mathcal{U}_2 = \{x_2 \neq 0\}$ it looks like a hyperbola.

Definition 33.4. Let $X_2 = \{f = 0\} \subset \mathbb{P}^n$ be a quadric. The *singular locus* $\mathrm{Sing}(X_2)$ of X_2 is the locus of points of X_2 where all the partial derivatives of f vanish. A point in $X_2 \setminus \mathrm{Sing}(X_2)$ is called a smooth point of X_2.

Given a point $p \in X_2$ the *tangent space* of X_2 in p is defined as

$$T_pX_2 = \left\{ x_0\frac{\partial f}{\partial x_0}(p) + \cdots + x_0\frac{\partial f}{\partial x_n}(p) = 0 \right\} \subset \mathbb{P}^n.$$

In particular, if $p \in \mathrm{Sing}(X_2)$ then $T_pX_2 = \mathbb{P}^n$ while if p is a smooth point of X_2 then T_pX_2 is a hyperplane in $\mathbb{P}^n$.

Let $K = \mathbb{C}$. Then we may write $f = \sum_{i=0}^{r} x_i^2$. Let us distinguish three cases depending on r:

$r = 0$ In this case $X_2 = \{x_0^2 = 0\} \subset \mathbb{P}^n$ has rank one and it is the hyperplane $\{x_0 = 0\}$ counted with multiplicity two. We say that X_2 is irreducible (just one piece) but non reduced (with multiplicity two).

$r = 1$ In this case $X_2 = \{x_0^2+x_1^2 = (x_0+ix_1)(x_0-ix_1) = 0\} \subset \mathbb{P}^n$ has rank two and it is the union of the hyperplanes $H_+ = \{x_0+ix_1 = 0\}$ and $H_- = \{x_0-ix_1 = 0\}$. We say that X_2 is reducible (two pieces) and reduced (both with multiplicity one).

$r \geq 3$ The singular locus of X_2 is given by

$$\mathrm{Sing}(X_2) = \{x_0 = \cdots = x_r = 0\} \subset \mathbb{P}^n$$

which is a linear subspace of dimension $n-r-1$. The quadric X_2 is irreducible. We say that X_2 is a cone of vertex $\mathrm{Sing}(X_2)$. In particular, X_2 is smooth if and only if $r = n$.

33.1. Conics

Assume that $\mathrm{char}(K) \neq 2$ and fix homogeneous coordinates $[x_0 : x_1 : x_2]$ on $\mathbb{P}^2$. To a conic

$$X_2 = \left\{ Z_0x_0^2 + Z_1x_0x_1 + Z_2x_0x_2 + Z_3x_1^2 + Z_4x_1x_2 + Z_5x_2^2 = 0 \right\} \subset \mathbb{P}^2$$

we can associate the symmetric matrix

$$M_{X_2} = \begin{pmatrix} Z_0 & \frac{Z_1}{2} & \frac{Z_2}{2} \\ \frac{Z_1}{2} & Z_3 & \frac{Z_4}{2} \\ \frac{Z_2}{2} & \frac{Z_4}{2} & Z_5 \end{pmatrix}.$$

The points of the degree three hypersurface

$$X_3 = \{\det(M_{X_2}) = 0\} \subset \mathbb{P}^5$$

where $[Z_0 : \cdots : Z_5]$ are the homogeneous coordinates on $\mathbb{P}^5$, correspond to singular conics.

A *pencil* of conics is a line in $\mathbb{P}^5$. Now, take two smooth conics $C, C' \in \mathbb{P}^5$ and let $L = \langle C, C' \rangle \subset \mathbb{P}^5$ be the pencil generated by them. Note that L is not contained in X_3 since C, C' belongs to L. Hence, L intersects X_3 in three points counted with multiplicity that is in a pencil generated by two smooth conics there are three (counted with multiplicity) singular conics.

33.2. Projective spaces parametrizing quadrics

Assume that $\mathrm{char}(K) \neq 2$ and fix homogeneous coordinates $[x_0 : \cdots : x_n]$ on $\mathbb{P}^n$. A quadric is a hypersurface of the form

$$X_2 = \left\{ Z_0 x_0^2 + Z_1 x_0 x_1 + Z_2 x_0 x_2 + \cdots + Z_{N_2} x_n^2 = 0 \right\} \subset \mathbb{P}^n$$

where $[Z_0 : \cdots : Z_{N_2}]$ are homogeneous coordinates on the projective space $\mathbb{P}^{N_2}$, with $N_2 = \binom{n+2}{2} - 1$, whose points parametrize homogeneous polynomials of degree two in $x_0, \ldots, x_n$ modulo scalar that is quadric hypersurfaces in $\mathbb{P}^n$. To X_2 we can associated the symmetric matrix

$$M_{X_2} = \begin{pmatrix} Z_0 & \cdots & \frac{Z_n}{2} \\ \vdots & \ddots & \vdots \\ \frac{Z_n}{2} & \cdots & Z_{N_2} \end{pmatrix}.$$

Let $R_s \subset \mathbb{P}^{N_2}$ be the locus of points at which all the $(s+1) \times (s+1)$ minors of M_{X_2} vanish. Then the points of R_s correspond to quadrics of rank at most s and we have a chain of inclusions

$$R_1 \subset R_2 \subset \cdots \subset R_{n+1} \subset \mathbb{P}^{N_2}.$$

In particular, the points of R_1 corresponds to quadrics of rank one that is double hyperplanes, the points of R_{n+1} corresponds to quadrics of rank n that is quadric cones with a point as vertex, and the points of $\mathbb{P}^{N_2} \setminus R_{n+1}$ correspond to smooth quadrics. Note that R_{n+1} is the degree $n+1$ hypersurface defined by the vanishing of the determinant of M_{X_2}.

33.3. Parametrizations of quadrics

Let $X_2 \subset \mathbb{P}^n$ be a quadric hypersurface of rank greater than or equal to three, and $p \in X_2$ a smooth point with homogeneous coordinates in the base field K. Fix a hyperplane $H \subset \mathbb{P}^n$ not passing through p. Given a point $x \in H$ we denote by $L_x \subset \mathbb{P}^n$ the line spanned by x and p. For the intersection $X_2 \cap L_x$ we have three possibilities:

- L_x intersects X_2 in p plus another point $\overline{x} \neq p$;
- L_x intersects X_2 in p with multiplicity two;
- L_x is contained in X_2.

Let $\overline{P}_x = P_{|L_x}$ be the restriction of the polynomials $P = P(x_0, \ldots, x_n)$, defining X_2, to the line L_x. Then $\overline{P}_x$ is a homogeneous polynomial of degree two in two variables whose coefficients depend on the homogeneous coordinates of the point x.

Therefore, L_x intersects X_2 in p with multiplicity two if and only if the discriminant $\Delta_{\overline{P}_x}$ of $\overline{P}_x$ vanishes, and L_x is contained in X_2 if and only if all the coefficients of $\overline{P}_x$ vanish. Let us denote by $C_1 \subset H$ the locus of points $x \in H$ for which $\Delta_{\overline{P}_x} = 0$ and $C_0 \subset H$ the locus of points $x \in H$ such that all the coefficients of $\overline{P}_x$ vanish. Then $C_0 \subset C_1 \subset H$. Set $\mathcal{U} = H \setminus C_1$. Note that since $x \in X_2$ is a smooth point $\mathcal{U}$ is non empty.

Furthermore, let $\mathcal{V} \subset X_2$ be the set of points $\overline{x} \in X_2$ such that the line $\langle p, \overline{x} \rangle$ intersects X_2 in two distinct points.

Since for any point $x \in \mathcal{U}$ the line L_x intersects X_2 in p plus another point $\overline{x} \neq p$ we get a map

$$\varphi_p : \mathcal{U} \subset H \to \mathcal{V} \subset X_2 : x \to \overline{x}$$

whose inverse is the restriction to X_2 of the linear projection $\pi_p : \mathbb{P}^n \setminus \{p\} \to H$ from p. Hence, we have a bijective correspondence between the points of $\mathcal{U}$ and those of $\mathcal{V}$. The map φ_p is a *rational parametrization* of X_2.

Now, let us focus on smooth conics $X_2 \subset \mathbb{P}^n$. Assume that X_2 has a point with homogeneous coordinates in the base field K, this is always the case if K for instance is algebraically closed), but for example the conic

$$X_2 = \{x_0^2 + x_1^2 + x_2^2 = 0\} \subset \mathbb{P}^2$$

does not have any point defined over the field of real numbers $\mathbb{R}$. In this situation $C_0 = \emptyset$. Indeed, if $L_x \subset X_2$ then $X_2 = L_x \cup L$ would be the union of two lines and hence it would have rank at most two.

Furthermore, there is just one line intersecting X_2 with multiplicity two in p, namely the tangent line T_pX_2 of X_2 in p. Hence, C_1 consists of a single point

$\overline{p} \in H$. Therefore, since the limit of $\pi_p(x)$ for $x \to p$ is exactly $\overline{p}$ we can extend the map $\pi_{p|\mathcal{V}}$ to a bijection defined on the whole of X_2 by mapping $p \to \overline{p}$.

Finally, we proved that a smooth conic X_2 over a field K with a point defined over K is in bijection with the projective line $\mathbb{P}^1$ over K.

33.4. Dual quadrics

Let $X_2 \subset \mathbb{P}^n$ be a smooth quadric hypersurface, and consider the set

$$\begin{array}{ccccc} & & \mathcal{I} = \{(x, T_x X_2),\ x \in X_2\} \subset X_2 \times \mathbb{P}^{n*} & & \\ & \overset{\pi_1}{\swarrow} & & \overset{\pi_2}{\searrow} & \\ X_2 & & & & \mathbb{P}^{n*} \end{array}$$

with the projection π_1, π_2 onto the factors. Set $X_2^* = \pi_2(\mathcal{I})$. As usual consider the matrix

$$M_{X_2} = \begin{pmatrix} Z_0 & \dots & \frac{Z_n}{2} \\ \vdots & \ddots & \vdots \\ \frac{Z_n}{2} & \dots & Z_{N_2} \end{pmatrix}.$$

of X_2. Let us focus on the case $n = 2$. The inverse of M_{X_2} is given by

$$M_{X_2}^{-1} = \frac{1}{\det(M_{X_2})} \begin{pmatrix} 4Z_3Z_5 - Z_4^2 & Z_2Z_4 - 2Z_1Z_5 & Z_1Z_4 - 2Z_2Z_3 \\ Z_2Z_4 - 2Z_1Z_5 & 4Z_0Z_5 - Z_2^2 & Z_1Z_2 - 2Z_0Z_4 \\ Z_1Z_4 - 2Z_2Z_3 & Z_1Z_2 - 2Z_0Z_4 & 4Z_0Z_3 - Z_1^2 \end{pmatrix}$$

where

$$\det(M_{X_2}) = 4Z_0Z_3Z_5 - Z_0Z_4^2 - Z_1^2Z_5 + Z_1Z_2Z_4 - Z_2^2Z_3.$$

For $x \in X_2$ we have

$$T_x X_2 = [2Z_0x_0 + Z_1x_1 + Z_2x_2 : Z_1x_0 + 2Z_3x_1 + Z_4x_2 : x_0Z_2 + Z_4x_1 + 2Z_5x_2] \in \mathbb{P}^{2*}.$$

and

$$T_x X_2 \cdot M_{X_2}^{-1} \cdot T_x X_2^t = 4(Z_0x_0^2 + Z_1x_0x_1 + Z_2x_0x_2 + Z_3x_1^2 + Z_4x_1x_2 + Z_5x_2^2).$$

Since $x \in X_2$ we get that $T_x X_2 \cdot M_{X_2}^{-1} \cdot T_x X_2^t = 0$ that is

$$X_2^* = \{y \in \mathbb{P}^{n*} \mid y \cdot M_{X_2}^{-1} \cdot y^t = 0\} \subset \mathbb{P}^{2*}.$$

The conic $X_2^* \subset \mathbb{P}^{2*}$ is the *dual conic* of $X_2 \subset \mathbb{P}^2$.

More generally the dual quadric $X_2^* \subset \mathbb{P}^{n*}$ of the quadric

$$X_2 = \{x \cdot M_{X_2} \cdot x^t = 0\} \subset \mathbb{P}^n$$

is the quadric

$$X_2^* = \{y \cdot M_{X_2}^{-1} \cdot y^t = 0\} \subset \mathbb{P}^{n*}$$

associated to the matrix $M_{X_2}^{-1}$.

Exercises

Exercise 33.5. Consider five points $p_1, \dots, p_5 \in \mathbb{P}^2$ such that no three are collinear. Prove that there exists a unique smooth conic passing through all the p_i.

Exercise 33.6. Let $X_2 \subset \mathbb{P}^2$ be a smooth conic and $p \in \mathbb{P}^2 \setminus X_2$ a point. Prove that there are two lines passing through p that are tangent to X_2.

Exercise 33.7. Let $\Gamma_f \subset \mathbb{P}^1 \times \mathbb{P}^1$ be the graph of a projectivity $f : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^1$. Consider the embedding

$$s : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3 : ([u_0 : u_1], [v_0 : v_1]) \to [u_0v_0 : u_0v_1 : u_1v_0 : u_1v_1].$$

Prove that $s(\Gamma_f) \subset \mathbb{P}^3$ is the intersection of a quadric surface and a plane.

Exercise 33.8. Compute the dual conic $X_2^* \subset \mathbb{P}^{2*}$ of the conic

$$X_2 = \{x_0x_2 - x_1^2 = 0\} \subset \mathbb{P}^2$$

and the tangent lines of X_2 passing through the point $p = [0 : 1 : 0]$.

Exercise 33.9. Let $p_1, p_2, p_3 \in \mathbb{P}^2$ be three non collinear points. Prove that the conics of $\mathbb{P}^2$ passing through the p_i form a plane Π in the projective space $\mathbb{P}^5$ parametrizing plane conics.

Moreover, prove that the conics through the p_i and tangent to a fixed line form a conic in Π.

Exercise 33.10. Let $L_1, L_2, L_3 \subset \mathbb{P}^3$ be three pairwise skew lines. Prove that through any point $p \in L_1$ there is a unique line L_p passing through p and intersecting L_2 and L_3 (Exercise 32.17). Consider the set

$$S = \bigcup_{p \in L_1} L_p \subset \mathbb{P}^3.$$

Prove that $S \subset \mathbb{P}^3$ is a smooth quadric surface and that $S \subset \mathbb{P}^3$ is the unique quadric surface containing $L_1, L_2, L_3 \subset \mathbb{P}^3$.

Chapter 34

Grassmannians

We have seen in Section 12.1 the notion of multilinear form. In this chapter will see how this concept in crucial in the construction of a geometric object parametrizing lines in $\mathbb{P}^3$.

Definition 34.1. Let V be a K-vector space of dimension $n+1$. The p-th exterior power $\bigwedge^p V$ is the dual space of the vector space of alternating p-linear forms on V. The elements of $\bigwedge^p V$ are called p-vectors.

The exterior product $v_1 \wedge \cdots \wedge v_p \in \bigwedge^p V$ of the vectors $v_1, \ldots, v_p \in V$ is the linear form on the vector space of alternating p-linear forms on V given by

$$(v_1 \wedge \cdots \wedge v_p)(F) = F(v_1, \ldots, v_p)$$

where F is an alternating p-linear on V.

As we have seen in Section 12.1:

- $v_1 \wedge \cdots \wedge v_p$ is linear with respect to each v_i;
- switching two of the v_i changes the sign of the exterior product;
- if two of the v_i are proportional then $v_1 \wedge \cdots \wedge v_p = 0$.

Given a basis $(e_0, \ldots, e_n)$ the p-vectors $e_{i_1} \wedge \cdots \wedge e_{i_p}$ with $i_1 < \cdots < i_p$ form a basis of $\bigwedge^p V$ which has therefore dimension $\binom{n+1}{p}$. Now, given

$$v = \sum_{i_1 < \cdots < i_p} a_{i_1, \ldots, i_p} v_1 \wedge \cdots \wedge v_p \in \bigwedge^p V$$

and

$$w = \sum_{j_1 < \cdots < j_q} b_{j_1, \ldots, j_q} w_1 \wedge \cdots \wedge w_q \in \bigwedge^q V$$

we define $v \wedge w \in \bigwedge^{p+q} V$ by

$$(v \wedge w)(F) = \sum_{i_1<\cdots<i_p;\, j_1<\cdots<j_q} a_{i_1,\ldots,i_p} b_{j_1,\ldots,j_q} F(v_1, \ldots, v_p, w_1, \ldots, w_q)$$

where F is an alternating $(p+q)$-linear on V. Note that $v \wedge w$ only depends on v, w and not on how they are written on a basis of V. Furthermore

- $u \wedge (v + w) = u \wedge v + u \wedge w$ for all $u \in \bigwedge^p V; v, w \in \bigwedge^q V$;
- $(u \wedge v) \wedge w = u \wedge (v \wedge w)$ for all $u \in \bigwedge^p V, v \in \bigwedge^q V, w \in \bigwedge^r V$;
- $u \wedge v = (-1)^{pq}(v \wedge u)$ for all $u \in \bigwedge^p V, v \in \bigwedge^q V$.

Lemma 34.2. *The exterior product $v_1 \wedge \cdots \wedge v_p \in \bigwedge^p V$ vanishes if and only if the $v_i \in V$ are linearly dependent.*

Proof. Assume that one the v_i, say v_1, is linear combination of the other v_i. Then $v_1 = \sum_{i=2}^p \lambda_i v_i$ and hence

$$v_1 \wedge \cdots \wedge v_p = \left(\sum_{i=2}^p \lambda_i v_i \right) \wedge v_2 \wedge \cdots \wedge v_p = 0$$

since each term has a repeated v_i. If $v_1, \ldots, v_p \in V$ are linearly independent then they can be completed to a basis of V. Hence, $v_1 \wedge \cdots \wedge v_p$ is a basis vector for $\bigwedge^p V$ and in particular $v_1 \wedge \cdots \wedge v_p \neq 0$. □

34.1. Grassmannians of lines and 2-vectors

We want to characterize 2-vectors $v \in \bigwedge^2 V$ that are *decomposable* i.e. that can written as $v = v_1 \wedge v_2$ with $v_1, v_2 \in V$.

Proposition 34.3. *A 2-vector $v \in \bigwedge^2 V$ is decomposable if and only if $v \wedge v = 0$ in $\bigwedge^4 V$.*

Proof. Assume that $v \wedge v = 0$. If $\dim(V) \leq 1$ then $\dim(\bigwedge^2 V) = 0$. Furthermore, if $\dim(V) = 2$ then $\dim(\bigwedge^2 V) = 1$ and $v_1 \wedge v_2 \neq 0$ if and only if v_1, v_2 form a basis of V, so any $v \in \bigwedge^2 V$ is decomposable. Consider the case $\dim(V) = 3$. For a fixed $v \in \bigwedge^2 V$ we define the map

$$A : V \to \bigwedge^3 V : w \to v \wedge w.$$

Since $\dim(\bigwedge^3 V) = 1$ we have that the kernel of A has dimension at least two. Choose linearly independent vectors $v_1, v_2 \in \mathrm{Ker}(A)$ and complete to a basis (v_1, v_2, v_3) of V. We can write

$$v = \lambda_1 v_2 \wedge v_3 + \lambda_2 v_3 \wedge v_1 + \lambda_3 v_1 \wedge v_2.$$

Now, since $v_1, v_2 \in \mathrm{Ker}(A)$ we get

$$v \wedge v_1 = \lambda_1 v_2 \wedge v_3 \wedge v_1 = 0;$$
$$v \wedge v_2 = \lambda_2 v_3 \wedge v_1 \wedge v_2 = 0.$$

Hence, $\lambda_1 = \lambda_2 = 0$ and $v = \lambda_3 v_1 \wedge v_2$ is decomposable.

Assume that the result holds for $\dim(V) \leq n$. Fix a basis $(e_0, \ldots, e_n)$ of V and write

$$v = \sum_{0 \leq i < j \leq n} a_{i,j} e_i \wedge e_j = \sum_{i=0}^{n-1} a_{i,n} e_i \wedge v_n + \sum_{0 \leq i < j \leq n-1} a_{i,j} e_i \wedge e_j = u \wedge e_n + v'$$

where $u \in U = \langle e_0, \ldots, e_{n-1} \rangle$ and $v' \in \bigwedge^2 U$. Note that $\dim(U) = n$ so that we can proceed by induction on the dimension. We have

$$v \wedge v = 2u \wedge v' \wedge e_n + v' \wedge v' = 0$$

and since e_n does not appear neither in the expansion of $u \wedge v'$ nor in that of $v' \wedge v'$ we get that $u \wedge v' = v' \wedge v' = 0$.

By induction $v' \wedge v' = 0$ yields that $v' = v_1 \wedge v_2$ with $v_1, v_2 \in V$, and so $u \wedge v_1 \wedge v_2 = 0$. Therefore, Lemma 34.2 yields that there is a linear relation

$$\lambda u + \lambda_1 v_1 + \lambda_2 v_2 = 0.$$

If $\lambda = 0$ then v_1, v_2 are linearly dependent and $v' = v_1 \wedge v_2 = 0$. So $v = u \wedge e_n$ is decomposable. If $\lambda \neq 0$ then

$$u = \alpha_1 v_1 + \alpha_2 v_2$$

with $\alpha_i = \frac{\lambda_i}{\lambda}$. So

$$v = (\alpha_1 v_1 + \alpha_2 v_2) \wedge e_n + v_1 \wedge v_2$$

in the subspace $\langle v_1, v_2, e_n \rangle$, and we conclude that v is decomposable by the 3-dimensional case in the first part of the proof.

Conversely, if $v = v_1 \wedge v_2$ for $v_1, v_2 \in V$ then

$$v \wedge v = v_1 \wedge v_2 \wedge v_1 \wedge v_2 = 0$$

since v_1 (or equivalently v_2) is repeated twice in the expression of v. □

34.2. The Grassmannian of lines in $\mathbb{P}^3$

The first interesting case is $p = 2, n + 1 = 4$. We construct a geometric object parametrizing vector planes in V or equivalently projective lines in $\mathbb{P}(V) = \mathbb{P}^3$.

Fix a basis $(e_0, \ldots, e_3)$ of V. Then $(e_0 \wedge e_1, e_0 \wedge e_2, e_0 \wedge e_3, e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3)$ is a basis of $\bigwedge^2 V$ and $(e_0 \wedge e_1 \wedge e_2 \wedge e_3)$ is a basis of $\bigwedge^4 V$.

Now, if $\Pi \subset V$ is a vector plane we can choose a basis (v, w) of Π and form the 2-vector $v \wedge w$. If (v', w') is another basis of Π then $v' \wedge w' = \lambda(v \wedge w)$ where

λ is the determinant of the change of basis matrix. Hence, we can associated to Π a well defined point $[v \wedge w] \in \mathbb{P}(\bigwedge^2 V)$.

Conversely, take $v \in \bigwedge^2 V$ and write

$$v = Z_0 e_0 \wedge e_1 + Z_1 e_0 \wedge e_2 + Z_2 e_0 \wedge e_3 + Z_3 e_1 \wedge e_2 + Z_4 e_1 \wedge e_3 + Z_5 e_2 \wedge e_3.$$

By Proposition 34.3 we have that $v \in \bigwedge^2 V$ is decomposable if and only if $v \wedge v = 0$. Now

$$\begin{aligned} v \wedge v &= Z_0 Z_5 e_0 \wedge e_1 \wedge e_2 \wedge e_3 + Z_1 Z_4 e_0 \wedge e_2 \wedge e_1 \wedge e_3 + Z_2 Z_3 e_0 \wedge e_3 \wedge e_1 \wedge e_2 + \\ &\quad Z_3 Z_2 e_1 \wedge e_2 \wedge e_0 \wedge e_3 + Z_4 Z_1 e_1 \wedge e_3 \wedge e_0 \wedge e_2 + Z_5 Z_0 e_2 \wedge e_3 \wedge e_0 \wedge e_1 \\ &= 2(Z_0 Z_5 - Z_1 Z_4 + Z_2 Z_3) e_0 \wedge e_1 \wedge e_2 \wedge e_3. \end{aligned}$$

Therefore $v \in \bigwedge^2 V$ is decomposable if and only if $[v] = [Z_0 : \cdots : Z_5] \in \mathbb{P}(\bigwedge^2 V) = \mathbb{P}^5$ belongs to the quadric hypersurface

$$\mathbb{G}(1,3) = \{Z_0 Z_5 - Z_1 Z_4 + Z_2 Z_3 = 0\} \subset \mathbb{P}^5 \tag{34.1}$$

which is the Grassmannian parametrizing lines in $\mathbb{P}^3$.

Proposition 34.4. *Let $L_1, L_2 \subset \mathbb{P}^3$ be two lines and $[L_1], [L_2] \in \mathbb{G}(1,3)$ the corresponding points. Then $L_1 \cap L_2 \neq \emptyset$ if and only if the line $\langle [L_1], [L_2] \rangle$ is contained in $\mathbb{G}(1,3)$.*

Proof. Let $U_1, U_2 \subset V$ be the two planes in V, where $\mathbb{P}^3 = \mathbb{P}(V)$, such that $\mathbb{P}(U_i) = L_i$. Assume that $L_1 \cap L_2 = \{p\}$ and $v \in V$ such that $[v] = p$. Let (u, u_1) and (u, u_2) be bases of U_1 and U_2. The line in $\mathbb{P}^5 = \mathbb{P}(\bigwedge^2 V)$ spanned by $[L_1], [L_2]$ is then $\mathbb{P}(W)$ where $W \subset \bigwedge^2 V$ is spanned by $u \wedge u_1$ and $u \wedge u_2$. Any 2-vector in W is of the form

$$\lambda_1 u \wedge u_1 + \lambda_2 u \wedge u_2 = u \wedge (\lambda_1 u_1 + \lambda_2 u_2)$$

and hence it is decomposable. So $[\lambda_1 u \wedge u_1 + \lambda_2 u \wedge u_2] \in \mathbb{G}(1,3)$.

Now, assume that $L_1 \cap L_2 = \emptyset$ that is $U_1 \cap U_2 = \{0\}$. Let $(u_1, v_1), (u_2, v_2)$ be bases of U_1, U_2. Then $V = U_1 \oplus U_2$ and (u_1, v_1, u_2, v_2) is a basis of V. So $u_1 \wedge v_1 \wedge u_2 \wedge v_2 \neq 0$. A point on the line spanned by $[L_1], [L_2]$ is represented by a vector of the form $v = \lambda_1 u_1 \wedge v_1 + \lambda_2 u_2 \wedge v_2$, and

$$v \wedge v = 2\lambda_1 \lambda_2 u_1 \wedge v_1 \wedge u_2 \wedge v_2.$$

So $v \wedge v = 0$ if and only if either $\lambda_1 = 0$ or $\lambda_2 = 0$. Therefore, the line spanned by $[L_1], [L_2]$ intersects $\mathbb{G}(1,3)$ only at the points $[L_1], [L_2]$. □

Proposition 34.5. *The lines in $\mathbb{P}^3$ passing through a fixed point correspond to the points of a plane contained in $\mathbb{G}(1,3)$.*

Proof. The line $\mathbb{P}(U)$ passes through a fixed point $[v] \in \mathbb{P}^3$ if and only if $v \in U$. Let (v, v_1, v_2, v_3) be a basis of V. A decomposable 2-vector of the form $v \wedge u$ can be written as

$$v \wedge (\lambda v + \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3) = \sum_{i=1}^{3} \lambda_i v \wedge v_i.$$

Hence, a line $\mathbb{P}(U)$ that passes through $[v] \in \mathbb{P}^3$ yields a point of the plane spanned by $v \wedge v_1, v \wedge v_2, v \wedge v_3$ in $\mathbb{G}(1,3)$ and any point in such plane corresponds to a line through $[v] \in \mathbb{P}^3$. □

Proposition 34.6. *The lines in $\mathbb{P}^3$ contained in a plane correspond to the points of a plane contained in $\mathbb{G}(1,3)$.*

Proof. The lines in $\mathbb{P}^3$ contained in a plane Π correspond to the lines in $\mathbb{P}^{3*}$ passing through the point Π^* which by Proposition 34.5 form a plane in the Grassmannian of lines in $\mathbb{P}^{3*}$. Finally, by duality the lines in $\mathbb{P}^3$ contained in Π form a plane in $\mathbb{G}(1,3)$. □

Theorem 34.7. *Let $H \subset \mathbb{G}(1,3)$ be a plane. Then H parametrizes either the lines through a fixed point or the lines contained in a fixed plane.*

Proof. Let $[L_1], [L_2], [L_3] \in H$ be three non collinear points. The lines joining two of the $[L_i]$ is contained in H and hence in $\mathbb{G}(1,3)$. So Proposition 34.4 yields that $L_i \cap L_j \neq \emptyset$ for all $i, j = 1, 2, 3$. We distinguish two cases:

- the L_i intersect in a point p. Let $H_p \subset \mathbb{G}(1,3)$ be the plane parametrizing the lines through p. Then H_p and H intersect in the three non collinear points $[L_1], [L_2], [L_3]$, and so $H = H_p$.
- The L_i intersect in three distinct points. Let v_1, v_2, v_3 be representative vectors for these three points. Then $v_2 \wedge v_3, v_1 \wedge v_3, v_1 \wedge v_2$ are representative vectors for $[L_1], [L_2], [L_3]$. A point of the plane H can then be written as

  $$\lambda_1 v_2 \wedge v_3 + \lambda_2 v_1 \wedge v_3 + \lambda_3 v_1 \wedge v_2.$$

  So $L_1, L_2, L_3 \subset \Pi$ where Π is the plane generated by $[v_1], [v_2], [v_3]$. Let $H_v \subset \mathbb{G}(1,3)$ be the plane parametrizing the lines contained in Π. Since H_v and H intersect in the three non collinear points $[L_1], [L_2], [L_3]$ we get that $H = H_v$.

We conclude that H is a plane either of the type in Proposition 34.5 or of the type in Proposition 34.6. □

34.3. The Plücker embedding of $\mathbb{G}(k,n)$

We will denote by $\mathbb{G}(k,n)$ the set of k-dimensional linear subspaces of $\mathbb{P}^n = \mathbb{P}(V)$:

$$\mathbb{G}(k,n) = \{\mathbb{P}(H) \subset \mathbb{P}^n \text{ where } H \subset V \text{ is a subvector space and } \dim(H) = k+1\}.$$

The geometric object $\mathbb{G}(k,n)$ is the Grassmannian parametrizing the k-dimensional linear subspaces of $\mathbb{P}^n$. For instance, $\mathbb{G}(0,n)=\mathbb{P}^n$, $\mathbb{G}(n-1,n)=\mathbb{P}^{n*}$ and $\mathbb{G}(1,3)$ is the Grassmannian of lines in $\mathbb{P}^3$ constructed in Section 34.2.

Remark 34.8. Recall that by duality there is a bijection between the k-dimensional linear subspaces of $\mathbb{P}^n$ and the $(n-k-1)$-dimensional linear subspaces of $\mathbb{P}^{n-k-1}$. Therefore, the is a bijective correspondence between $\mathbb{G}(k,n)$ and $\mathbb{G}(n-k-1,n)$.

Now, let $(e_0,\dots,e_n)$ be a basis of V, $H\subset V$ a subvector space of dimension $k+1$, and $(v_0,\dots,v_k)$ a basis of H. Write $v_i=\sum_{j=0}^n x_{ij}e_j$ and consider the matrix

$$M_H=\begin{pmatrix} x_{00} & x_{01} & \dots & x_{0n}\\ \vdots & \vdots & \ddots & \vdots\\ x_{k0} & x_{k1} & \dots & x_{kn}\end{pmatrix}.$$

Note that the coordinates of the $(k+1)$-vector $v_0\wedge\cdots\wedge v_k\in\bigwedge^{k+1}V$, in the basis $(e_{i_0}\wedge\cdots\wedge e_{i_k})_{0\le i_0<\cdots<i_n\le n}$, are given by the determinants of the $(k+1)\times(k+1)$ submatrices $M_{H;i_0,\dots,i_k}$ obtained extracting from M_H the columns indexed by $i_0,\dots,i_k$:

$$v_0\wedge\cdots\wedge v_k=\sum_{0\le i_0<\cdots<i_n\le n}\det(M_{H;i_0,\dots,i_k})e_{i_0}\wedge\cdots\wedge e_{i_k}.$$

Furthermore, if $(v_0',\dots,v_k')$ is another basis of H we have that $v_0'\wedge\cdots\wedge v_k'=\det(N)v_0\wedge\cdots\wedge v_k$ where N is the base change matrix. Therefore, the is a well defined map

$$pl_{k,n}:\mathbb{G}(k,n)\to\mathbb{P}(\bigwedge^{k+1}V)=\mathbb{P}^{N_{k,n}}:H\to[v_0\wedge\cdots\wedge v_k]$$

where $N_{k,n}=\binom{n+1}{k+1}-1$. The map $pl_{k,n}$ is called the *Plücker embedding* of $\mathbb{G}(k,n)$ and allows us to realize $\mathbb{G}(k,n)$ as an object living inside $\mathbb{P}^{N_{k,n}}$.

Example 34.9. Let $\mathbb{P}(H)\subset\mathbb{P}^3$ be a line, and (v_0,v_1) a basis of H. Write $v_0=x_{00}e_0+x_{01}e_1+x_{02}e_2+x_{03}e_3$, $v_1=x_{10}e_0+x_{11}e_1+x_{12}e_2+x_{13}e_3$ and consider the matrix

$$M_H=\begin{pmatrix} x_{00} & x_{01} & x_{02} & x_{03}\\ x_{10} & x_{11} & x_{12} & x_{13}\end{pmatrix}.$$

Then the Plücker embedding is the map

$$pl_{1,3}:\mathbb{G}(1,3)\to\mathbb{P}(\bigwedge^{2}V)=\mathbb{P}^5$$

mapping H to the point $[v_0\wedge v_1]=[x_{00}x_{11}-x_{01}x_{10}:x_{00}x_{12}-x_{02}x_{10}:x_{00}x_{13}-x_{03}x_{10}:x_{01}x_{12}-x_{02}x_{11}:x_{01}x_{13}-x_{03}x_{11}:x_{02}x_{13}-x_{03}x_{12}]\in\mathbb{P}^5$. Note that, if $[Z_0:\cdots:Z_5]$ are homogeneous coordinates on $\mathbb{P}^5$ then the coordinates of $[v_0\wedge v_1]$

satisfy the quadratic relation $Z_0Z_5 - Z_1Z_4 + Z_2Z_3 = 0$ which is exactly the equation (34.1) defining $\mathbb{G}(1,3) \subset \mathbb{P}^5$.

Exercises

Exercise 34.10. Show that $\mathbb{G}(1,3) \subset \mathbb{P}^5$ is smooth.

Exercise 34.11. Let $L \subset \mathbb{P}^3$ be a line. Show that the lines in $\mathbb{P}^3$ intersecting L correspond to the points in the intersection of $\mathbb{G}(1,3)$ with a hyperplane in $\mathbb{P}^5$.

Exercise 34.12. Let $\Pi \subset \mathbb{P}^3$ be a plane and $p \in \Pi$ a point. Show that the lines contained in Π and passing through p correspond to the points of a line in $\mathbb{G}(1,3)$.

Exercise 34.13. Consider the quadric surface $Q = \{x_0x_3 - x_1x_2 = 0\} \subset \mathbb{P}^3$.

- Show that Q is smooth.
- Show that Q contains two families $\mathcal{F}_\alpha, \mathcal{F}_\beta$ of lines such that two lines of the same family do not intersect and two lines of different families intersect in a point.
- Show that the lines of the families $\mathcal{F}_\alpha, \mathcal{F}_\beta$ correspond to the points of two smooth conics $C_\alpha, C_\beta \subset \mathbb{G}(1,3)$, and that the planes Π_α, Π_β generated by C_α, C_β do not intersect.

Exercise 34.14. Write down the Plücker embedding of $\mathbb{G}(1,4)$, the Grassmannian of lines in $\mathbb{P}^4$, in $\mathbb{P}^9$ and work out five homogeneous polynomials of degree two, in the homogeneous coordinates of $\mathbb{P}^9$, vanishing on $\mathbb{G}(1,4)$.

Exercise 34.15. Prove that through any two points $[L_1], [L_2] \in \mathbb{G}(1,n) \subset \mathbb{P}^{N_1}$ there passes a conic contained in $\mathbb{G}(1,n)$.

Bibliography

[DF04] D. S. Dummit and R. M. Foote, *Abstract algebra*, third ed., John Wiley & Sons, Inc., Hoboken, NJ, 2004. MR 2286236

[Hal74] P. R. Halmos, *Naive set theory*, Undergraduate Texts in Mathematics, Springer-Verlag, New York-Heidelberg, 1974, Reprint of the 1960 edition. MR 453532

[HW79] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, fifth ed., The Clarendon Press, Oxford University Press, New York, 1979. MR 568909

[Lan87] S. Lang, *Linear algebra*, third ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1987. MR 874113

[Ser00] E. Sernesi, *Geometria 1*, Programma di mat. fisica elettronica, Bollati Boringhieri, 2000.