A Course in Algebra

E. B. Vinberg

Graduate Studies
in Mathematics
Volume 56

American Mathematical Society

A Course in Algebra

A Course in Algebra

E. B. Vinberg

Graduate Studies
in Mathematics
Volume 56

American Mathematical Society
Providence, Rhode Island

Э. Б. Винберг

КУРС АЛГЕБРЫ

"Факториал Пресс", Москва, 1999, 2001

Translated from the Russian by Alexander Retakh
This work was originally published in Russian by Factorial Press under the title, Kurs Algebry © 2001. The present translation was created under license for the American Mathematical Society and is published by permission.

2000 *Mathematics Subject Classification.* Primary 13-01, 15-01, 16-01, 20-01.

For additional information and updates on this book, visit
www.ams.org/bookpages/gsm-56

Library of Congress Cataloging-in-Publication Data

Vinberg, E. B. (Ernest Borisovich)
[Kurs algebry. English]
A course in algebra / E. B. Vinberg.
p. cm. — (Graduate studies in mathematics, ISSN 1065-7339 ; v. 56)
Includes bibliographical references and index.
ISBN 0-8218-3318-9 (acid-free paper) ISBN 0-8218-3413-4 (softcover)
1. Algebra. I. Title. II. Series.

QA154.3.V56 2003
512—dc21 2002033011

10 9 8 7 6 5 4 3 2 1 08 07 06 05 04 03

Contents

Preface

My motivation for writing this book came from teaching a two year course in algebra at the Mathematical College of the Independent University of Moscow in 1992–1994. The students' enthusiasm and a relatively small class allowed me to keep the level of presentation higher than it is usually done at the Mechanics and Mathematics Department of Moscow State University, and to touch on a number of subjects beyond a regular university course. However, in writing this book I used my experience in teaching at Moscow State University, and so the final version of the book is only partially related to the course given at the Independent University.

Chapters 1–7 and part of Chapter 8 more or less correspond to the first year algebra course at the Mechanics and Mathematics Department of Moscow State University. The remaining chapters cover, and, in fact, substantially exceed the second year algebra course. These chapters are intended mainly for students specializing in algebra.

Note that Chapter 7 is devoted to geometry of Euclidean, affine, and projective spaces. However, this chapter should not be viewed as an exposition of geometry; rather, it describes the algebraic approach to geometry.

In the first four chapters I tried to make the presentation sufficiently detailed to be suitable for a reader such as a mathematics freshman at Moscow State University. (However, the language of sets and maps is used from the very beginning without any explanations.) In later chapters I allowed myself to skip details that can be easily reconstructed, since I believe that a reader should gradually acquire mathematical culture.

There are almost no technically difficult proofs in this book. Following my point view on mathematics, I tried to replace calculations and difficult

deductions with conceptual proofs. Some readers may find this style hard, but the efforts spent in absorbing new ideas will pay off when the students start solving problems not discussed in this book.

For the English edition, the bibliography at the end of the book was revised. It is certainly not complete and, to some extent, arbitrary, but I believe the reader might find it helpful.

I am grateful to all current and former members of the Chair of Higher Algebra at the Mechanics and Mathematics Department of Moscow State University who helped me to shape my approach to teaching algebra.

In the English translation a number of misprints and inaccuracies were corrected and some explanations added.

Moscow, November 2002 E. B. Vinberg

Chapter 1

Algebraic Structures

When you are introduced to people, at first you only learn their names and faces. Meeting them later, you begin to know them better, maybe even become friends with them.

In the first chapter, you will be only introduced to most of the algebraic structures considered in this book. A deeper understanding of them should come later, through reading and problem-solving.

1.1. Introduction

If it is at all possible to define the subject of algebra precisely, then this is the study of algebraic structures: sets on which operations are defined. By an operation on a set M, we mean a map

$$M \times M \to M,$$

i.e., a rule that assigns to every two elements of M some element of the same set M. These elements can be numbers or objects of a different kind.

The following number sets are well-known important examples of algebraic structures. They have the operations of addition and multiplication:

$\mathbb{N}$, the set of all natural numbers,

$\mathbb{Z}$, the set of all integers,

$\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$, the set of all nonnegative integers,

$\mathbb{Q}$, the set of all rational numbers,

$\mathbb{R}$, the set of all real numbers,

$\mathbb{R}_+$, the set of all nonnegative real numbers.

Remark that the operations of addition and multiplication are not defined on every number set. For example, multiplication is not defined on the set of negative numbers because the product of two negative numbers is positive. On the set of irrational numbers, neither multiplication nor addition is defined, since the sum and the product of two irrational numbers can be rational.

Here are some examples of algebraic structures whose elements are not numbers:

Example 1.1. Let M, N, P be sets and let

$$f : N \to M, \qquad g : P \to N$$

be maps between them. The *product* or *composition* of f and g is the map

$$fg : P \to M,$$

defined as

$$(fg)(a) = f(g(a)) \qquad \forall a \in P,$$

i.e., the result of successive application of, first, g and, then, f. In particular, when $M = N = P$, we obtain an operation on the set of all maps from M to itself. This operation provides many important examples of algebraic structures that are called groups. For example, according to the axioms of Euclidean geometry, the product of two motions of the plane is again a motion. When we consider the operation of multiplication on the set of all such motions, we obtain the algebraic structure called the group of motions of the plane.

Example 1.2. The set of all vectors in the three-dimensional space with the operations of addition and cross product is an example of an algebraic structure with two operations. Notice, however, that the inner product is not an operation as defined above. Indeed, its result does not belong to the same set as the original vectors. More general operations such as the inner product are also considered in algebra but we will not concern ourselves with them for now.

All the above examples are natural in the sense that they arose from the studies of the real world or the internal progress in mathematics. But actually, it is possible to consider any operation on any set. For instance, one can consider the set $\mathbb{Z}_+$ with the operation that assigns to a pair of numbers the number of coinciding digits in their decimal representations. However, only few algebraic structures are of real interest.

Also, an algebraist is interested only in such properties of algebraic structures and their elements that can be stated in terms of their particular operations. This view is formally expressed in the concept of isomorphism.

Definition 1.3. Let M be a set with an operation $\circ$ and N a set with an operation $*$. Algebraic structures $(M, \circ)$ and $(N, *)$ are called *isomorphic* if there exists a bijective map

$$f : M \to N$$

such that

$$f(a \circ b) = f(a) * f(b)$$

for any $a, b \in M$. We denote this fact as $(M, \circ) \simeq (N, *)$. The map f is called an *isomorphism* between $(M, \circ)$ and $(N, *)$.

In a similar way we define an isomorphism between algebraic structures with two or more operations.

Example 1.4. The map

$$a \mapsto 2^a$$

is an isomorphism between the set of all real numbers with the operation of addition and the set of positive numbers with the operation of multiplication. Indeed,

$$2^{a+b} = 2^a 2^b.$$

Instead of base 2 we can consider any other positive base different from 1. This shows that there might exist many different isomorphisms between two isomorphic algebraic structures.

Example 1.5. Let M be the set of parallel translations of the plane along a fixed line. For a real number a, denote by t_a the translation from M defined by the vector of length $|a|$ with direction defined by the sign of a. (When $a = 0$, t_a is the zero translation.) It is easy to see that

$$t_{a+b} = t_a \circ t_b,$$

where $\circ$ denotes the product (composition) of parallel translations. Hence, the map $a \to t_a$ is an isomorphism between the algebraic structures $(\mathbb{R}, +)$ and $(M, \circ)$.

It is clear that if two algebraic structures are isomorphic, any statement made only in terms of operations is valid for one of this structures if and only if it is valid for the other.

For example, an operation $\circ$ on a set M is called *commutative* if

$$a \circ b = b \circ a$$

for any $a, b \in M$. If $(M, \circ)$ is isomorphic to $(N, *)$ and the operation $\circ$ on M is commutative, then the operation $*$ on N is commutative too.

So, it does not matter which of the isomorphic structures to study: all of them are models of the same object. However, the choice of a model might be important for the solution of a specific problem. Sometimes a particular

model turns out to be more useful. For instance, if a model has geometric nature, one can study it with geometric methods.

1.2. Abelian Groups

Addition of real numbers has the following properties:

(A1) $a + b = b + a$ (commutativity);

(A2) $(a + b) + c = a + (b + c)$ (associativity);

(A3) $a + 0 = a$;

(A4) $a + (-a) = 0$.

One can logically deduce other properties from these ones. For instance, they imply that there exists an operation inverse to addition, i.e., subtraction. This means that for any a, b, the equation

$$x + a = b$$

has a unique solution. Let us prove this. If c is a solution of this equation (i.e.,$c + a = b$), then

$$(c + a) + (-a) = b + (-a).$$

From (A2)–(A4), we obtain

$$(c + a) + (-a) = c + (a + (-a)) = c + 0 = c.$$

Therefore,

$$c = b + (-a).$$

This shows that if a solution exists, it is unique and equals $b + (-a)$. Conversely, after substituting $x = b + (-a)$ in the original equation, we see that $b + (-a)$ is indeed a solution:

$$(b + (-a)) + a = b + ((-a) + a) = b + (a + (-a)) = b + 0 = b.$$

Multiplication of real numbers has properties similar to (A1)–(A4):

(M1) $ab = ba$ (commutativity);

(M2) $(ab)c = a(bc)$ (associativity);

(M3) $a1 = a$;

(M4) $aa^{-1} = 1$ for $a \neq 0$.

Properties (M1)–(M4) differ from (A1)–(A4) almost exclusively in notation. The only small difference is that in (M4) we assume that $a \neq 0$, while in (A4) we make no assumptions about a. We deduced from (A1)–(A4) the existence of the operation of subtraction. Translated into the language of

multiplication, this deduction implies the existence of the operation of division, inverse to that of multiplication. More precisely, we can deduce from (M1)–(M4) that for any $a \neq 0$ and any b, the equation $xa = b$ has a unique solution $x = ba^{-1}$.

This discussion is put here not to teach you something new about real numbers but to communicate an idea which is very important in algebra. It is the idea of axiomatic approach. This means the simultaneous study of whole classes of algebraic structures defined by various axioms, that is, particular properties of operations on these structures. It does not matter how these operations are defined in each specific case. As long as the axioms are satisfied, each theorem deduced from these axioms is true.

Of course, only few systems of axioms are really interesting. It is impossible to come up off the top of one's head with a system of axioms that will lead to a reasonable theory. All systems of axioms considered in modern algebra have a long history and are products of analysis of algebraic structures that arose in a natural way. Such are the systems of axioms of a group, ring, field, vector space, and other structures that you will encounter in this book.

Properties (A1)–(A4), as well as (M1)–(M4), are in fact the system of axioms of an abelian group. Before we state these explicitly, a few words about terminology. Names and notation of operations on algebraic structures carry no particular meaning; however, most often they are called addition or multiplication and are also denoted appropriately. This allows us to use the well-developed terminology and notational system for operations with real numbers. This also reminds us of sometimes helpful similarities between the numbers and the structure we are considering.

First, we define an abelian group using the language of addition.

Definition 1.6. An additive *abelian group* is a set A with an operation of addition that has the following properties:

(i) $a + b = b + a$ for any $a, b \in A$ (*commutativity*);

(ii) $(a + b) + c = a + (b + c)$ for any $a, b, c \in A$ (*associativity*);

(iii) there exists an element $0 \in A$ (*zero*) such that $a + 0 = a$ for any $a \in A$;

(iv) for any element $a \in A$, there exists an element $-a \in A$ (an *opposite*) such that $a + (-a) = 0$.

We will deduce a few simple properties from these axioms.

(i) Zero is unique. Indeed, let both 0_1 and 0_2 be zeros. Then

$$0_1 = 0_1 + 0_2 = 0_2.$$

(ii) The opposite is unique. Indeed, let both $(-a)_1$ and $(-a)_2$ be opposites of a. Then

$$(-a)_1 = (-a)_1 + (a + (-a)_2) = ((-a)_1 + a) + (-a)_2 = (-a)_2.$$

(iii) For any a, b, the equation $x + a = b$ has a unique solution equal to $b + (-a)$. For the proof, see the beginning of this section. This solution is called the *difference* of b and a and is denoted $b - a$.

It is not difficult to deduce from the associativity property that the sum of an arbitrary number of elements (not just three) does not depend on where parentheses are put in the expression for the sum (try to prove this). For this reason, parentheses are usually omitted altogether.

Example 1.7. The sets $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are abelian groups with respect to the ordinary addition.

Example 1.8. The set of vectors (on the plane or in space) is an abelian group with respect to the standard addition of vectors.

Example 1.9. A *row* of length n is a sequence of n numbers. The set of all rows of length n with entries from $\mathbb{R}$ is denoted $\mathbb{R}^n$. Define addition of rows by the rule

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

Obviously, $\mathbb{R}^n$ is an abelian group with respect to the above operation. Its zero is the zero row

$$0 = (0, 0, \ldots, 0).$$

Example 1.10. The set of all functions defined on a given subset of the real line is an abelian group with respect to the standard addition of functions.

Now, we define abelian groups using the language of multiplication.

Definition 1.6′. A multiplicative *abelian group* is a set A with an operation of multiplication that has the following properties:

(i) $ab = ba$ for any $a, b \in A$ (*commutativity*);

(ii) $(ab)c = a(bc)$ for any $a, b, c \in A$ (*associativity*);

(iii) there exists an element $e \in A$ (*identity*) such that $ae = a$ for any $a \in A$;

(iv) for any element $a \in A$, there exists an element $a^{-1} \in A$ (an *inverse*) such that $aa^{-1} = e$.

The identity of a multiplicative abelian group is sometimes denoted 1.

The simplest corollaries of the axioms of an abelian group that we first obtained in the additive language are translated into the multiplicative language as follows:

(i) The identity is unique.

(ii) The inverse is unique.

(iii) For any a, b, the equation $xa = b$ has a unique solution equal to ba^{-1}. It is called the *ratio* of b and a and is denoted $\frac{b}{a}$ (or b/a).

Example 1.11. Sets $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ and $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ are abelian groups with respect to the ordinary multiplication.

Later we define the general notion of a group (not necessarily abelian). There the operation is not required to be commutative.

1.3. Rings and Fields

Unlike groups, fields and rings are algebraic structures with two operations, which are usually called addition and multiplication. Their axioms are suggested by the properties of operations over real numbers, just like the axioms of an abelian group. The axioms of a ring is a reasonably chosen set of conditions on these operations. Such a choice allows us to consider other important examples of algebraic structures satisfying these axioms. Of these, we have already mentioned the set of vectors of the Euclidean space with the operations of addition and cross product.

Definition 1.12. A *ring* is a set K with the operations of addition and multiplication that have the following properties:

(i) K is an abelian group with respect to addition (this group is called the *additive group* of K);

(ii) $a(b+c) = ab+bc$ and $(a+b)c = ac+bc$ for any $a, b, c \in K$ (*distributive laws*).

We deduce here several corollaries of the axioms of a ring which are not among the corollaries of the axioms of an abelian group given already in Section 1.2.

(i) $a0 = 0a = 0$ for any $a \in K$. Indeed, let $a0 = b$. Then

$$b + b = a0 + a0 = a(0 + 0) = a0 = b$$

implying

$$b = b - b = 0.$$

The proof of $0a = 0$ is similar.

(ii) $a(-b) = (-a)b = -ab$ for any $a, b \in K$. Indeed,

$$ab + a(-b) = a(b + (-b)) = a0 = 0$$

and, similarly, $ab + (-a)b = 0$.

(iii) $a(b-c) = ab - ac$ and $(a-b)c = ac - bc$ for any $a, b, c \in K$. Indeed,

$$a(b-c) + ac = a(b-c+c) = ab$$

and, similarly, $(a-b)c + bc = ac$.

A ring K is called *commutative* if its multiplication is commutative, i.e., if

$$ab = ba \qquad \forall a, b,$$

and *associative* if its multiplication is associative, i.e., if

$$(ab)c = a(bc) \qquad \forall a, b, c.$$

An element 1 of a ring is called a *unity* if

$$a1 = 1a = a \qquad \forall a.$$

Just as for the identity of a multiplicative abelian group, it can be shown that a ring cannot possess more than one unity (but there might be none).

Remark 1.13. If $1 = 0$, then for any a

$$a = a1 = a0 = 0,$$

i.e., the ring consists of zero only. Thus, if a ring has more than one element, $1 \neq 0$.

Remark 1.14. If a ring is commutative, one of the distributive laws implies the other. The same is true for the defining identities of unity.

Example 1.15. The number sets $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are commutative associative rings with unities with respect to the ordinary addition and multiplication.

Example 1.16. The set $2\mathbb{Z}$ of even integers is a commutative associative ring without unity.

Example 1.17. The set of all functions defined on a given subset of the real line is a commutative associative ring with unity with respect to the ordinary addition and multiplication of functions.

Example 1.18. The set of vectors of the Euclidean space endowed with operations of addition and cross product is a noncommutative and nonassociative ring. Yet it satisfies the following identities which, in some sense, replace commutativity and associativity:

$a \times b + b \times a = 0$ (*anticommutativity*),

$(a \times b) \times c + (b \times c) \times a + (c \times a) \times b = 0$ (*the Jacobi identity*).

Anticommutativity follows directly from the definition of the cross product. As for the proof of the Jacobi identity, see Example 1.75.

Exercise 1.19. Let X be a set and 2^X the set of all its subsets. Prove that 2^X is a ring with respect to the operations of symmetric difference

$$M \triangle N = (M \setminus N) \cup (N \setminus M)$$

and intersection, taken for addition and multiplication, respectively. Prove that this ring is commutative and associative.

An element a^{-1} of a ring with unity is called an *inverse* of an element a if

$$aa^{-1} = a^{-1}a = 1.$$

(In a commutative ring, it suffices to require $aa^{-1} = 1$.) As in the case of a multiplicative abelian group, it can be proved that in an associative ring with unity no element can have more than one inverse (but may have none). An element that has an inverse is called *invertible*.

Definition 1.20. A *field* is a commutative associative ring with unity where every nonzero element is invertible.

Remark 1.21. A ring that consists of only zero is not regarded as a field.

Examples of fields are the field of rational numbers $\mathbb{Q}$ and the filed of real numbers $\mathbb{R}$. The ring $\mathbb{Z}$ is not a field; its only invertible elements are ± 1.

Exercise 1.22. Prove that there exists a field that consists of two elements. (*Hint*: it is clear that one of these elements must be zero and the other unity.)

Every field has the following important property:

$$ab = 0 \quad \Longrightarrow \quad \{a = 0 \text{ or } b = 0\}.$$

Indeed, if $a \neq 0$, then when we multiply both sides of $ab = 0$ by a^{-1}, we obtain $b = 0$.

There exist other rings that have the above property, for example, the ring $\mathbb{Z}$. They are called *rings without zero divisors*. In a ring without zero divisors, cancellations are possible:

$$\{ac = bc \text{ (or } ca = cb) \text{ and } c \neq 0\} \quad \Longrightarrow \quad a = b.$$

Indeed, the equality $ac = bc$ can be rewritten as $(a - b)c = 0$. Hence, if $c \neq 0$, we obtain $a - b = 0$, i.e., $a = b$.

Here is an example of a commutative associative ring with zero divisors:

Example 1.23. In the ring of functions on a subset X of the real line (see Example 1.17), there exist zero divisors if and only if X consists of more than one point. Indeed, split X into two nonempty sets X_1 and X_2 (i.e., $X_1 \cup X_2 = X, X_1 \cap X_2 = \varnothing$) and let

$$f_i(x) = \begin{cases} 1 & \text{for } x \in X_i, \\ 0 & \text{for } x \notin X_i, \end{cases} \qquad \text{for } i = 1, 2.$$

Then $f_1, f_2 \neq 0$ but $f_1 f_2 = 0$. When X consists of just one point, the ring of real-valued functions on X is isomorphic to $\mathbb{R}$.

Since there are no zero divisors in a field, the product of two nonzero elements is nonzero. Nonzero elements of a field K form an abelian group with respect to multiplication. It is called the *multiplicative group of* K and is denoted K^*.

1.4. Subgroups, Subrings, and Subfields

Consider a set M with an operation $\circ$ and a subset N of M. The subset N is said to be *closed with respect to the operation* $\circ$ if

$$a, b \in N \implies a \circ b \in N.$$

If this happens, the operation $\circ$ is defined on N, thus making it an algebraic structure. If the operation $\circ$ on M has some property expressed as an identity (e.g., commutativity or associativity), then it obviously has the same property as an operation on N. However, some other properties of $\circ$ might not be inherited in N.

For instance, a subset of an abelian group closed with respect to addition might not be an abelian group, since it does not necessarily contain zero or opposites of all elements. Consider an example: the subset $\mathbb{Z}_+$ of the abelian group $\mathbb{Z}$. It is closed with respect to addition but it is not an abelian group (not even a group) because it does not contain opposites of any of its elements except for zero.

Definition 1.24. A subset B of an abelian group A is called a *subgroup* if

(i) B is closed with respect to addition;

(ii) $a \in B \implies -a \in B$;

(iii) $0 \in B$.

Remark 1.25. It is easy to see that when B is not empty, the first two conditions imply the third. Therefore, instead of the third condition, we may require B to be nonempty.

Obviously, every subgroup of an additive abelian group is also an abelian group with respect to the same operation.

Example 1.26. The additive group $\mathbb{R}$ contains the following chain of subgroups:

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Example 1.27. The set of vectors in space parallel to a given line or plane is a subgroup.

There are two "trivial" subgroups in every abelian group: the group itself and the subgroup that consists only of zero.

Exercise 1.28. Prove that every subgroup of $\mathbb{Z}$ has the form $n\mathbb{Z}$, where $n \in \mathbb{Z}_+$. (A solution of this exercise can be found in Section 4.3.)

Here is the multiplicative version of Definition 1.24.

Definition 1.24′. A subset B of a multiplicative abelian group A is called a *subgroup* if

(i) B is closed with respect to multiplication;

(ii) $a \in B \implies a^{-1} \in B$;

(iii) $e \in B$.

Example 1.29. The group $\mathbb{R}^*$ contains the following chain of subgroups:

$$\{\pm 1\} \subset \mathbb{Q}^* \subset \mathbb{R}^*.$$

Discussion at the beginning of this section can also be extended to the case of algebraic structures with several operations. In this way we arrive at the definitions of a subring and a subfield.

Definition 1.30. A subset L of a ring K is called a *subring* if

(i) L is a subgroup of the additive group of the ring K;

(ii) L is closed with respect to multiplication.

Clearly, every subring is itself a ring with respect to the same operations. It also inherits properties such as commutativity and associativity from the larger ring.

Example 1.31. The chain of subgroups of the additive group $\mathbb{R}$ from Example 1.26 is also a chain of subrings.

Example 1.32. For any $n \in \mathbb{Z}_+$, the set $n\mathbb{Z}$ is a subring of $\mathbb{Z}$.

Exercise 1.33. Prove that all finite subsets of a set X form a subring of the ring 2^X constructed in Exercise 1.19.

Definition 1.34. A subset L of a field K is called a *subfield* if

(i) L is a subring of K;

(ii) $a \in L, a \neq 0 \implies a^{-1} \in L$;

(iii) $1 \in L$.

Clearly, every subfield is a field with respect to the same operations.

Example 1.35. The field $\mathbb{Q}$ is a subfield of $\mathbb{R}$.

Exercise 1.36. Prove that a subset L of a field K is a subfield if and only if

(i) L is closed with respect to subtraction and division;

(ii) $L \ni 0, 1$.

Exercise 1.37. Prove that the field $\mathbb{Q}$ does not contain any nontrivial subfields (i.e., subfields other than itself).

1.5. The Field of Complex Numbers

It is impossible to define division on the ring of integers, and this makes it necessary to extend integers to the field of rational numbers. Similarly, since it is impossible to extract square roots from negative numbers in the field of real numbers, we are forced to extend this field to a bigger one called the field of complex numbers.

To understand better what the field of complex numbers is, we should first try to understand the nature of the field of real numbers. Its rigorous construction is usually covered in a real analysis course. We will not consider all the details here; however, we remark that there are several such constructions (e.g., infinite decimal fractions, Dedekind cuts of rationals, etc). Formally speaking, each method leads to a different field. Which one is the "real" field of reals? The answer is that they are all isomorphic, and we should simply view them as different models of the same object called the field of real numbers.

In such a situation, the most satisfactory approach is always the axiomatic one: first, we formulate as axioms the properties that our object must have, and then prove that these properties define it uniquely up to isomorphism. Finally, we prove the existence by constructing a model. In the case of the field of real numbers such a choice of axioms (in addition to the axioms of a field) may be the order axioms, the Archimedean postulate, and the axiom of continuity.

Remark 1.38. It is not difficult to prove that not only every two models of the field of real numbers are just isomorphic, but that the isomorphism between them is unique. (The proof comes down to demonstrating that an isomorphism from $\mathbb{R}$ to itself is the identity map. This demonstration, in turn, relies on the observation that nonnegative numbers in $\mathbb{R}$ are squares, and hence must be mapped to nonnegative numbers under any isomorphism.) This means that every element of $\mathbb{R}$ is specific, i.e., in any model we can identify numbers 10, $\sqrt{2}$, π, etc.

Now we can state the axiomatic definition of the field of complex numbers.

Definition 1.39. The *field of complex numbers* is a field $\mathbb{C}$ such that

(i) it contains the field of real numbers $\mathbb{R}$ as a subfield;

(ii) it contains an element $\imath$ such that $\imath^2 = -1$;

(iii) it is minimal among the fields with properties (i) and (ii), i.e., if K is a subfield of $\mathbb{C}$ containing $\mathbb{R}$ and $\imath$, then $K = \mathbb{C}$.

Remark 1.40. Equality $x^2 + 1 = (x - \imath)(x + \imath)$ implies that the equation $x^2 = -1$ has exactly two solutions in $\mathbb{C}$: $\imath$ and $-\imath$. If a subfield of $\mathbb{C}$ contains one of these solutions, it must contain the other.

Theorem 1.41. *The field of complex numbers exists and is unique up to an isomorphism that maps all real numbers to themselves. Every complex number can be uniquely written as $a + b\imath$, where $a, b \in \mathbb{R}$ and $\imath$ is a (fixed) element such that $\imath^2 = -1$.*

Proof. (i) Let $\mathbb{C}$ be a field of complex numbers (we assume for now that it exists). Consider its subset

$$K = \{a + b\imath : a, b \in \mathbb{R}\}.$$

Properties of field operations and the identity $\imath^2 = -1$ imply

$$(a_1 + b_1\imath) + (a_2 + b_2\imath) = (a_1 + a_2) + (b_1 + b_2)\imath, \tag{1.1}$$

$$(a_1 + b_1\imath)(a_2 + b_2\imath) = (a_1a_2 - b_1b_2) + (a_1b_2 + b_1a_2)\imath. \tag{1.2}$$

By solving appropriate equations, we also obtain

$$-(a + b\imath) = (-a) + (-b)\imath, \tag{1.3}$$

$$(a + b\imath)^{-1} = \frac{a}{a^2 + b^2} + \left(-\frac{b}{a^2 + b^2}\right)\imath, \qquad \text{whenever } a^2 + b^2 \neq 0. \tag{1.4}$$

Equations (1.1)–(1.4) show that K is a subfield of $\mathbb{C}$. Since K obviously contains $\imath$ and $\mathbb{R}$, $K = \mathbb{C}$.

Therefore, every element of $\mathbb{C}$ can be presented in the form $a+bi$, where $a, b \in \mathbb{R}$. We have to show that such a presentation is unique. Let $a_1+b_1i = a_2+b_2i$, $a_1, b_1, a_2, b_2 \in \mathbb{R}$. Then

$$a_1 - a_2 = (b_2 - b_1)i.$$

Taking squares of both sides, we get

$$(a_1 - a_2)^2 = -(b_2 - b_1)^2,$$

which implies

$$a_1 - a_2 = b_2 - b_1 = 0,$$

as required.

Now, let $\mathbb{C}'$ be another field of complex numbers and $i' \in \mathbb{C}'$ an element such that $(i')^2 = -1$. As equations (1.1) and (1.2) remain valid when i is replaced with i', the map

$$f : \mathbb{C} \to \mathbb{C}', \quad a + bi \mapsto a + bi' \qquad (a, b \in \mathbb{R}),$$

is an isomorphism from $\mathbb{C}$ to $\mathbb{C}'$.

(ii) The above discussion suggests a way of proving the existence of the field of complex numbers. Consider the set $\mathbb{C}$ of pairs (a, b) with $a, b \in \mathbb{R}$. Define addition and multiplication according to the following formulas suggested by (1.1) and (1.2):

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2),$$
$$(a_1, b_1)(a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + b_1a_2).$$

Obviously, $\mathbb{C}$ is an abelian group with respect to addition (see Example 1.9). Moreover, its multiplication is commutative and satisfies the distributive law. A direct calculation shows that it is also associative. Therefore, $\mathbb{C}$ is a commutative associative ring.

Since

$$(a, b)(1, 0) = (a, b),$$

the element $(1, 0)$ is the unity of $\mathbb{C}$. Equation (1.4) suggests a form of the inverse of (a, b), whenever $a^2 + b^2 \neq 0$. Indeed, we can check directly that

$$(a, b)\left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2}\right) = (1, 0).$$

Therefore, $\mathbb{C}$ is a field.

Furthermore,

$$(a_1, 0) + (a_2, 0) = (a_1 + a_2, 0),$$
$$(a_1, 0)(a_2, 0) = (a_1a_2, 0),$$

i.e., operations on pairs of the form $(a,0)$ arise from corresponding operations on the pairs' first components. We identify each pair $(a,0)$ with a and can claim now that $\mathbb{C}$ contains $\mathbb{R}$ as a subfield.

Put $\imath = (0,1)$. Then

$$\imath^2 = (-1,0) = -1,$$
$$a + b\imath = (a,b) \qquad \text{for } a,b \in \mathbb{R}.$$

Therefore, every element of $\mathbb{C}$ can be (uniquely) presented in the form $a+b\imath$, where $a,b \in \mathbb{R}$. It follows that if a subfield K of $\mathbb{C}$ contains $\imath$ and $\mathbb{R}$, then $K = \mathbb{C}$. We conclude that $\mathbb{C}$ is indeed the field of complex numbers. □

The decomposition of a complex number $c \in \mathbb{C}$ as $a + b\imath$, $a,b \in \mathbb{R}$, is called its *algebraic form*; a is called the *real part* of c and b the *imaginary part* of c. Notation: $a = \Re c, b = \Im c$. Complex numbers that are not real are called *imaginary*; those of the form $b\imath$, $b \in \mathbb{R}$, are called *purely imaginary*.

If we substitute $\mathbb{C}$ for $\mathbb{C}'$ and $-\imath$ for $\imath'$ in the first part of the proof of Theorem 1.41, we see that the map

$$c = a + b\imath \mapsto \bar{c} = a - b\imath \qquad (a,b \in \mathbb{R})$$

is an isomorphism of $\mathbb{C}$ into itself. This map is called *complex conjugation*. In general, an isomorphism of an algebraic structure into itself is called an *automorphism*. Thus, complex conjugation $c \mapsto \bar{c}$ is an automorphism of the field of complex numbers. It is clear that $\bar{\bar{c}} = c$.

Real numbers are precisely the complex numbers equal to their conjugates. It follows that for any $c \in \mathbb{C}$, $c + \bar{c}$ and $c\bar{c}$ are real. Indeed,

$$\overline{c+\bar{c}} = \bar{c} + c = c + \bar{c}, \qquad \overline{c\bar{c}} = \bar{c}c = c\bar{c}.$$

It is easy to see that if $c = a + b\imath$, $a,b \in \mathbb{R}$, then

$$c + \bar{c} = 2a, \qquad c\bar{c} = a^2 + b^2. \tag{1.5}$$

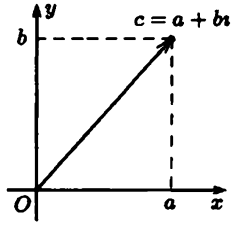


Figure 1.1

We can depict complex numbers as points or vectors on the plane. Namely, a number $c = a + bi$ is represented as a point or vector with Cartesian coordinates (a, b) (Figure 1.1). In some cases it is easier to depict complex numbers as points, in other cases, as vectors. In the vector form, addition of complex numbers corresponds to the standard vector addition.

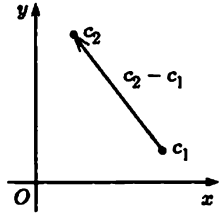


Figure 1.2

Notice also that the difference of complex numbers c_1 and c_2 is represented by the vector connecting points that represent c_1 and c_2 (Figure 1.2).

It is sometimes more convenient to use polar instead of Cartesian coordinates. This leads to the introduction of the following concepts:

The *absolute value* of a complex number $c = a + bi$ is the length of the vector representing c. The absolute value of c is denoted $|c|$. Clearly,

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

The *argument* of a complex number is the angle formed by the corresponding vector with the polar axis. The argument is determined up to 2π and is not defined for 0. The notation is $\arg c$.

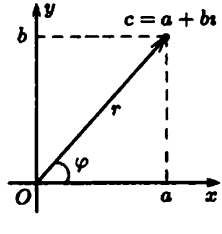


Figure 1.3

Let r and φ be the absolute value and argument of a complex number c (Figure 1.3). Obviously,

$$a = r\cos\varphi, \qquad b = r\sin\varphi,$$

hence,

$$c = r(\cos\varphi + \imath\sin\varphi).$$

The above representation of a complex number c is called its *trigonometric form*. Since the trigonometric form of a complex number is determined up to addition to φ a multiple of 2π,

$$r_1(\cos\varphi_1 + \imath\sin\varphi_1) = r_2(\cos\varphi_2 + \imath\sin\varphi_2)$$
$$\iff \{r_1 = r_2,\ \varphi_1 = \varphi_2 + 2\pi k,\ k \in \mathbb{Z}\}, \qquad \text{for } r_1, r_2 > 0.$$

Expressing complex numbers in trigonometric form is useful for performing on them such operations as multiplication, division, raising to a power, and root extraction. Namely, formulas for the sine and cosine of a sum of two angles imply

$$r_1(\cos\varphi_1 + \imath\sin\varphi_1)\cdot r_2(\cos\varphi_2 + \imath\sin\varphi_2)$$
$$= r_1r_2(\cos(\varphi_1 + \varphi_2) + \imath\sin(\varphi_1 + \varphi_2)),$$

i.e., when multiplying complex numbers, we multiply their absolute values and add their arguments. Thus, we have the following formulas for the ratio of complex numbers:

$$\frac{r_1(\cos\varphi_1 + \imath\sin\varphi_1)}{r_2(\cos\varphi_2 + \imath\sin\varphi_2)} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + \imath\sin(\varphi_1 - \varphi_2)),$$

and for a positive integer power of a complex number:

$$[r(\cos\varphi + \imath\sin\varphi)]^n = r^n(\cos n\varphi + \imath\sin n\varphi) \qquad \text{(De Moivre's formula)}.$$

To *extract a root* of the nth degree from a complex number $c = r(\cos\varphi + \imath\sin\varphi)$ is to solve the equation $z^n = c$. Put $|z| = s$, $\arg z = \psi$. Then $s^n = r$, $n\psi = \varphi + 2\pi k$, $k \in \mathbb{Z}$. Hence,

$$s = \sqrt[n]{r} \quad \text{(positive root)}, \qquad \psi = \frac{\varphi + 2\pi k}{n}.$$

Combining these formulas, we get

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

With the above formula, we obtain the same answer for different values of k if and only if they are congruent modulo n. It follows that for $c \neq 0$, the equation $z^n = c$ has exactly n solutions corresponding to, say, $k = 0, 1, \ldots, n-1$. Represented geometrically, these numbers lie at the vertices of a *regular n-gon* with its center at the origin (Figure 1.4 illustrates the case $n = 8$).

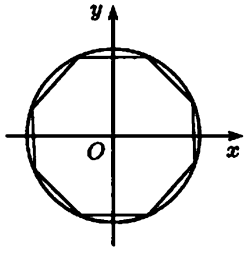


Figure 1.4

1.6. Rings of Residue Classes

Extending the ring of integers, we obtain the chain

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

We will see later that this chain can be enlarged (and, in particular, continued to the right). Rings of residue classes are also constructed from the integers; however, the approach is entirely different. It is a standard mathematical method of "sticking together": forming a quotient by an equivalence relation.

Consider a set M. Any subset $R \subset M \times M$ is called a *relation* on the set M. If $(a, b) \in R$, we say that a and b are related and denote this aRb.

Here are some examples of relations:

Example 1.42. M is the set of all people; aRb if a knows b.

Example 1.43. Same M; aRb if a is a friend of b's.

Example 1.44. Same M; aRb if a and b live in the same building.

Example 1.45. $M = \mathbb{R}$; aRb if $a \leq b$.

Example 1.46. M is the set of circles on the plane; aRb if circles a and b are congruent, i.e., if there exists a motion that identifies one with the other.

An *equivalence relation* R is a relation which is

(i) reflexive: aRa;

(ii) symmetric: $aRb \implies bRa$;

(iii) transitive: aRb and $bRc \implies aRc$.

Among the relations in Examples 1.42–1.46 only the third and fifth are equivalence relations: the first and fourth are not symmetric, whereas the second is symmetric but not transitive.

Equivalence relation is usually written as $a \overset{R}{\sim} b$ or simply $a \sim b$.

Let R be an equivalence relation on a set M. For any $a \in M$, put

$$R(a) = \{b \in M \,:\, a \overset{R}{\sim} b\}.$$

Properties of equivalence relations clearly imply that $a \in R(a)$ and

$$R(a) \cap R(b) \neq \varnothing \quad \Longrightarrow \quad R(a) = R(b).$$

Therefore, the subsets $R(a)$ form a partition of M, i.e., their union covers M and the intersection of each pair is empty. These subsets are called *equivalence classes under* R. Two elements of M are equivalent if and only if they belong to the same class.

The set of equivalence classes under R is called the *quotient set of* M *by* R and is denoted M/R. The map

$$M \to M/R, \qquad a \mapsto R(a)$$

is called the *quotient map*.

For instance, in Example 1.44 an equivalence class is the set of tenants of a building. We can identify the quotient set with the set of all buildings; then the quotient map is a map that assigns to each person a building where this person lives. In Example 1.46, equivalence classes are the sets of circles with the same radius, the quotient set can be identified with the set of positive real numbers, and the quotient map assigns to each circle its radius.

Consider a set M with an operation $(x, y) \mapsto x * y$. An equivalence relation R *agrees* with the operation $*$ if

$$\{a \overset{R}{\sim} a', \quad b \overset{R}{\sim} b'\} \quad \Longrightarrow \quad a * b \overset{R}{\sim} a' * b'.$$

In this case, we can define operation $*$ on the quotient set M/R by the rule

$$R(a) * R(b) = R(a * b). \tag{1.6}$$

Verbally the above definition goes as follows: to perform an operation on two equivalence classes, one should choose an arbitrary representative in each of them, perform this operation on the representatives, and take the class where the result of this operation lies. This class does not depend on the initial choice of representatives because the relation agrees with the operation.

Obviously, the operation on M/R inherits all properties of the operation on M that are expressed in the form of identities, for instance, commutativity or associativity. This is also true for the existence of zero (identity) and inverse elements. More precisely, if we call an operation on M an addition and M contains a zero element 0 for this operation, then $R(0)$ is a zero element in M/R. If $-a$ is an opposite of a, then $R(-a)$ is an opposite of $R(a)$ as an element of M/R.

Now we are ready to construct rings of residue classes. Fix a natural number n. Consider the following equivalence relation, called *congruence modulo n*, on the set $\mathbb{Z}$: a is congruent to b modulo n if $a-b$ is divisible by n or, equivalently, if a and b have the same remainder (residue) when divided by n. (Notation: $a \equiv b \pmod{n}$.)

Clearly, this is an equivalence relation; moreover, equivalence classes can be enumerated by numbers $0, 1, \ldots, n-1$, so that the rth class consists of all integers whose remainder after division by n is r.

The equivalence class that contains an integer a is called the *residue class* of a *(modulo n)* and is denoted $[a]_n$ or, when it is clear what n is, simply $[a]$.

The quotient set of $\mathbb{Z}$ by the relation of congruence modulo n is denoted $\mathbb{Z}_n$. We may write

$$\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}$$

but should keep in mind that any element of $\mathbb{Z}_n$ can be denoted differently. For instance, element $[1]_n$ might as well be denoted by $[2n+1]_n$, $[-(n-1)]_n$, etc.

We must prove now that the relation of congruence modulo n agrees with the operations of addition and multiplication. Let

$$a \equiv a' \pmod{n}, \qquad b \equiv b' \pmod{n}.$$

Then

$$a+b \equiv a'+b \equiv a'+b' \pmod{n}$$

and, similarly,

$$ab \equiv a'b \equiv a'b' \pmod{n}.$$

Therefore, we can define operations of addition and multiplication on the set $\mathbb{Z}_n$ by the following formulas:

$$[a]_n + [b]_n = [a+b]_n, \qquad [a]_n\,[b]_n = [ab]_n$$

(this is valid for all $a, b \in \mathbb{Z}$). Thus $\mathbb{Z}_n$ becomes a commutative associative ring with unity. It is called the *ring of residue classes* (or, sometimes, simply the *residue ring*) modulo n.

Example 1.47. Here are the addition and multiplication tables for the ring $\mathbb{Z}_5$. For simplicity, we omit brackets when writing elements of this ring.

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

×	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Observe, in particular, that elements 2 and 3 are inverses of each other and that 4 is the inverse of itself.

Example 1.48. Let us calculate $[2]^{100}$ in the ring $\mathbb{Z}_{125}$:

$$[2]^7 = [128] = [3], \qquad [2]^{35} = ([2]^7)^5 = [3]^5 = [243] = [-7],$$
$$[2]^{50} = [2]^{35}([2]^7)^2[2] = [-7][3]^2[2] = [-126] = [-1],$$
$$[2]^{100} = ([2]^{50})^2 = [1].$$

This means that

$$2^{100} \equiv 1 \pmod{125}.$$

Given that 2^{100} is divisible by 8, we deduce that

$$2^{100} \equiv 376 \pmod{1000},$$

i.e., that the decimal representation of 2^{100} ends with 376.

The ring $\mathbb{Z}_n$ has all properties of a field except, perhaps, the property of having inverses for all its nonzero elements. Clearly, $\mathbb{Z}_2$ is the field of two elements discussed in Exercise 1.22. The above multiplication table of the ring $\mathbb{Z}_5$ shows that $\mathbb{Z}_5$ is a field as well. On the other hand, $\mathbb{Z}_4$ is not a field since its element [2] is not invertible.

Theorem 1.49. *The ring $\mathbb{Z}_n$ is a field if and only if n is a prime number.*

Proof. (i) Let n be a composite number, i.e., $n = kl$ for $1 \le k, l \le n$. Then $[k]_n, [l]_n \neq 0$, yet

$$[k]_n[l]_n = [kl]_n = [n]_n = 0.$$

Therefore, $\mathbb{Z}_n$ contains zero divisors, hence it is not a field.

(ii) Conversely, let n be prime and $[a]_n \neq 0$, i.e., let a be not divisible by n. We should look for the inverse of $[a]_n$ by multiplying it by each element of the ring. In the process, we obtain elements

$$[0]_n,\ [a]_n,\ [2a]_n,\ \ldots,\ [(n-1)a]_n. \tag{1.7}$$

They are all distinct. Indeed, if $[ka]_n = [la]_n$, $0 \le k < l \le n-1$, then $[(l-k)a]_n = 0$, i.e., $(l-k)a$ is divisible by n, which is impossible as n divides neither $l-k$ nor a. (This is where we use that n is prime.) Hence, sequence (1.7) contains all elements of $\mathbb{Z}_n$, and in particular, $[1]_n$. This means that $[a]_n$ is invertible. □

Exercise 1.50. Prove that for any n, $[k]_n$ is invertible in the ring $\mathbb{Z}_n$ if and only if n and k are relatively prime.

In fields of residue classes we encounter a new phenomenon that does not appear in number fields (subfields of the field of complex numbers). Namely, in the field $\mathbb{Z}_n$ (n prime), the following equality is valid:

$$\underbrace{1+1+\cdots+1}_{n}=0. \tag{1.8}$$

(Of course, this is also true in a ring $\mathbb{Z}_n$ for any n.) It leads to several specific features of algebraic transformations in this field (see below).

In general, let K be a field. The least natural n such that equality (1.8) is valid in K is called the *characteristic* of this field. If such n does not exist, we call K the field of zero characteristic. Thus, $\mathbb{Z}_n$ with n prime is a field of characteristic n, while number fields have zero characteristic. The characteristic of K is denoted $\operatorname{char} K$.

If $\operatorname{char} K = n$, then for any $a \in K$,

$$\underbrace{a+a+\cdots+a}_{n}=\underbrace{(1+1+\cdots+1)}_{n}a=0a=0.$$

When nonzero, the characteristic of a field is always prime. Indeed, assume that $\operatorname{char} K = n = kl$ $(1 < k, l < n)$. Then

$$\underbrace{1+1+\cdots+1}_{n}=\underbrace{(1+1+\cdots+1)}_{k}\underbrace{(1+1+\cdots+1)}_{l}=0,$$

hence either $\underbrace{1+1+\cdots+1}_{k}=0$ or $\underbrace{1+1+\cdots+1}_{l}=0$ contradicting the definition of characteristic.

Most of the formulas of elementary algebra are valid in every field, as their deductions use only properties of addition and multiplication that are either axioms of a field or their corollaries. Specific features of fields of positive characteristic emerge only in formulas that contain multiplication or division by natural numbers.

For example, consider the formula

$$(a+b)^2=a^2+2ab+b^2.$$

It is valid in any field if $2ab$ is understood as $ab+ab$. However, in a field of characteristic 2, it acquires a simpler form

$$(a+b)^2=a^2+b^2.$$

More generally, in a field of characteristic p the following is valid:

$$(a+b)^p=a^p+b^p.$$

Indeed, the binomial theorem gives

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} a^{p-k} b^k.$$

For $0 < k < p$,

$$\binom{p}{k} = \frac{p(p-1)\cdots(p-k+1)}{k!},$$

which is obviously divisible by p. Hence, all terms in the binomial formula but the first and the last, equal zero.

Exercise 1.51. Conclude from the aforesaid that every $a \in \mathbb{Z}_p$ satisfies the identity $a = a^p$. (This fact is called Fermat's little theorem; for another proof, see Section 2.5.)

Things get worse when we need to divide by a natural number, e.g., when we try to calculate ab from the above formula for the sum's square. To make this division meaningful in some sense, we may regard multiplication by a natural k as multiplication by the element $\underbrace{1+1+\cdots+1}_{k}$ of the given field. Then division by k would be seen as division by this element. However, if k is divisible by the field's characteristic, this element is zero and such division is impossible.

For instance, the formula for solutions of a quadratic equation involves a division by 2. So it is applicable (in the above sense) in every field of characteristic not equal to 2 but not in a field of characteristic 2.

Example 1.52. Let us solve the quadratic equation

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

in the field $\mathbb{Z}_{11}$. The standard formula yields

$$x_{1,2} = \frac{[-1] \pm \sqrt{[5]}}{[2]}.$$

Since $[5] = [16] = [4]^2$, we can assume that $\sqrt{[5]} = [4]$ (one of the values of the square root). Hence,

$$x_1 = \frac{[-1]+[4]}{[2]} = \frac{[3]}{[2]} = \frac{[14]}{[2]} = [7], \quad x_2 = \frac{[-1]-[4]}{[2]} = \frac{[-5]}{[2]} = \frac{[6]}{[2]} = [3].$$

1.7. Vector Spaces

In elementary geometry, we not only add vectors but also multiply them by numbers. Analyzing properties of these two operations, we come to the notion of a vector space.

Before actually defining it, we should note that here we step away from what we previously meant by an operation. Multiplication of a vector by a number is not an operation on two elements of the same set. It is an operation that assigns to each pair (number, vector) another vector. This is how things also are in the general definition of a vector space; however, there elements of an arbitrary (fixed) field replace real numbers.

Definition 1.53. A *vector* (or *linear*) *space* over a field K is a set V with operations of addition and multiplication by elements of the field with the following properties:

(i) V is an abelian group with respect to addition;

(ii) $\lambda(a+b) = \lambda a + \lambda b$ for any $\lambda \in K$, $a, b \in V$;

(iii) $(\lambda + \mu)a = \lambda a + \mu a$ for any $\lambda, \mu \in K$, $a \in V$;

(iv) $(\lambda\mu)a = \lambda(\mu a)$ for any $\lambda, \mu \in K$, $a \in V$;

(v) $1a = a$ for any $a \in V$.

Elements of a vector space are called *vectors*. We will abuse language sometimes and call elements of K numbers even when K is not a number field.

From now on, vectors in the sense of elementary geometry will be called *geometric vectors*. Their operations satisfy all axioms of a vector space; this is, in fact, the reason for the above definition. We will denote the space of vectors on the Euclidean plane (respectively, in the three-dimensional Euclidean space) as E^2 (respectively, E^3). Observe that this is a vector space over $\mathbf{R}$. Here are a few other important examples of vector spaces:

Example 1.54. The set K^n of rows of length n with entries from a field K is a vector space over K with respect to operations defined as follows:

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n),$$
$$\lambda(a_1, a_2, \ldots, a_n) = (\lambda a_1, \lambda a_2, \ldots, \lambda a_n).$$

Example 1.55. The set $F(X, K)$ of all functions on a set X with values in a field K is a vector space with respect to standard operations on functions, namely:

$$(f+g)(x) = f(x) + g(x), \qquad (\lambda f)(x) = \lambda f(x).$$

Example 1.56. Let K be a subfield of a field L. Then L can be regarded as a vector space over K with multiplication of elements of L by elements of K defined simply as multiplication in L. In particular, this makes $\mathbb{C}$ a vector space over $\mathbf{R}$.

We continue with several corollaries of the axioms of a vector space. All of them are proved similarly to analogous corollaries of the axioms of a ring (see Section 1.3). Symbol 0 stands for both zero of the field K and the zero vector, i.e., zero of the additive group V, but this should not be confusing.

(i) $\lambda 0 = 0$ for any $\lambda \in K$ (here 0 is the zero vector).

(ii) $\lambda(-a) = -\lambda a$ for any $\lambda \in K$, $a \in V$.

(iii) $\lambda(a - b) = \lambda a - \lambda b$ for any $\lambda \in K$, $a, b \in V$.

(iv) $0a = 0$ for any $a \in V$ (here 0 on the left is a number, and on the right, a vector).

(v) $(-1)a = -a$ for any $a \in V$.

(vi) $(\lambda - \mu)a = \lambda a - \mu a$ for any $\lambda, \mu \in K, a \in V$.

Definition 1.57. A subset U of a vector space V is called a *subspace* if

(i) U is a subgroup of the additive group V;

(ii) $a \in U \implies \lambda a \in U$ for any $\lambda \in K$.

Remark 1.58. The definition of a subgroup requires that

$$a \in U \implies -a \in U.$$

One can note that condition (ii) automatically implies this, as $-a = (-1)a$.

A subspace of a vector space is itself a vector space with respect to the same operations.

Example 1.59. In the space E^3, the set of vectors parallel to a given plane or line, is a subspace.

Example 1.60. In the space $F(X, \mathbb{R})$ of all functions on a given interval X of the real line, the set of continuous functions is a subspace.

Every vector space V contains two "trivial" subspaces: V itself and the zero subspace (which consists of the zero vector only). We will denote the latter simply by the symbol 0.

Definition 1.61. Vector spaces V and U over a field K are called *isomorphic* if there exists a bijective map

$$\varphi : V \to U$$

such that

(i) $\varphi(a + b) = \varphi(a) + \varphi(b)$ for any $a, b \in V$;

(ii) $\varphi(\lambda a) = \lambda \varphi(a)$ for any $\lambda \in K$, $a \in V$.

If so, the map φ is called an *isomorphism* between V and U.

We will see in Section 2.2 that it is quite easy to describe vector spaces up to isomorphism. In particular, we will mostly concern ourselves in this book with the so-called finite-dimensional vector spaces; and all of them are isomorphic to spaces K^n. The key notion for this theory is the notion of a basis.

An expression of the form

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n, \qquad \lambda_1, \lambda_2, \ldots, \lambda_n \in K,$$

is called a *linear combination* of vectors $a_1, a_2, \ldots, a_n \in V$. We say that a vector b can be *expressed as a linear combination* of vectors $a_1, a_2, \ldots, a_n$ if it equals a linear combination of them.

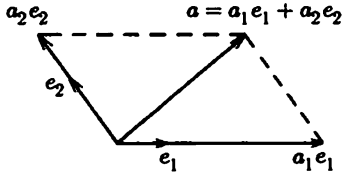


Figure 1.5

Definition 1.62. A system of vectors $\{e_1, e_2, \ldots, e_n\} \subset V$ is called a *basis* of V if every vector $a \in V$ can be uniquely expressed as a linear combination of $e_1, e_2, \ldots, e_n$. Coefficients of this expression are called *coordinates* of a in the basis $\{e_1, e_2, \ldots, e_n\}$.

Example 1.63. Recall that geometric vectors are called *collinear* if they are parallel to the same line and *coplanar* if they are parallel to the same plane. It is known from geometry that any two noncollinear vectors e_1, e_2 form a basis of E^2 (Figure 1.5). Similarly, any three noncoplanar vectors form a basis of E^3.

Example 1.64. Unit rows

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0), \\ e_2 &= (0, 1, \ldots, 0), \\ &\ldots\ldots\ldots\ldots \\ e_n &= (0, 0, \ldots, 1) \end{aligned}$$

form a basis of the space K^n. A row $a = (a_1, a_2, \ldots, a_n)$ has coordinates $a_1, a_2, \ldots, a_n$ in this basis. Of course, K^n has other bases as well.

Example 1.65. Take $\{1, i\}$ as a basis of $\mathbb{C}$ regarded as a vector space over $\mathbb{R}$ (see Example 1.56). The coordinates of a complex number in this basis are its real and imaginary parts.

Proposition 1.66. *Any vector space V over a field K with a basis consisting of n vectors is isomorphic to the space K^n.*

Proof. Let $\{e_1, e_2, \ldots, e_n\}$ be a basis of V. Consider the map

$$\varphi : V \to K^n$$

that assigns to each vector the row of its coordinates in the basis $\{e_1, e_2, \ldots, e_n\}$. Clearly, this map is bijective. Now, if

$$a = a_1e_1 + a_2e_2 + \cdots + a_ne_n, \qquad b = b_1e_1 + b_2e_2 + \cdots + b_ne_n,$$

then

$$a + b = (a_1 + b_1)e_1 + (a_2 + b_2)e_2 + \cdots + (a_n + b_n)e_n,$$
$$\lambda a = (\lambda a_1)e_1 + (\lambda a_2)e_2 + \cdots + (\lambda a_n)e_n.$$

Hence, φ is an isomorphism. □

Example 1.67. The space E^2 (respectively, E^3) is isomorphic to $\mathbb{R}^2$ (respectively, $\mathbb{R}^3$).

1.8. Algebras

Because their structure is so simple, vector spaces are not very interesting on their own. However, the notion of a vector space is a part of many algebraic (and not just algebraic) theories. For instance, by combining the notions of a vector space and a ring, we arrive at the important notion of an algebra.

Definition 1.68. An *algebra* is a set over a field K with operations of addition, multiplication, and multiplication by elements of K that have the following properties:

(i) A is a vector space with respect to addition and multiplication by elements of the field;

(ii) A is a ring with respect to addition and multiplication;

(iii) $(\lambda a)b = a(\lambda b) = \lambda(ab)$ for any $\lambda \in K$, $a, b \in A$.

Remark 1.69. So far we have used the word *algebra* to describe a particular branch of mathematics. In the above definition its meaning is, of course, different.

Example 1.70. Every field L having K as a subfield can be regarded as an algebra over K. In particular, $\mathbb{C}$ is an algebra over $\mathbb{R}$.

Example 1.71. The space E^3 is an algebra with respect to the operation of cross product.

Example 1.72. The set $F(X, K)$ of functions on a set X with values in a field K (see Example 1.55) is an algebra over K with respect to the standard operations of addition and multiplication of functions and multiplication of a function by a number. This algebra is commutative, associative, and has unity (which is the constant function equal to 1).

Exercise 1.73. Prove that the ring 2^X in Exercise 1.19 becomes an algebra over the field $\mathbb{Z}_2$ if we define multiplication by elements of this field by the following rules:

$$0M = \varnothing, \qquad 1M = M \qquad \forall M \in 2^X.$$

Assume that an algebra A regarded as a vector space over K has a basis $\{e_1, e_2, \ldots, e_n\}$. Let

$$a = a_1e_1 + a_2e_2 + \cdots + a_ne_n = \sum_{i=1}^{n} a_ie_i,$$

$$b = b_1e_1 + b_2e_2 + \cdots + b_ne_n = \sum_{i=1}^{n} b_ie_i$$

be two arbitrary elements of A. Then the distributive laws of multiplication imply

$$ab = \sum_{i=1}^{n} a_i(e_ib) = \sum_{i=1}^{n} a_i \left(\sum_{j=1}^{n} b_j(e_ie_j) \right) = \sum_{i,j=1}^{n} a_ib_j(e_ie_j).$$

This shows that multiplication in an algebra is completely determined by the products of basis vectors.

If multiplication of basis vectors is commutative, i.e., if

$$e_ie_j = e_je_i \qquad \forall i, j,$$

then multiplication in A is commutative in general. Indeed, in the above notation, for every $a, b \in A$, we have

$$ab = \sum_{i,j} a_ib_j(e_ie_j) = \sum_{i,j} b_ja_i(e_je_i) = ba.$$

Similarly, it is possible to prove that if multiplication of basis vectors is associative, i.e., if

$$(e_ie_j)e_k = e_i(e_je_k) \qquad \forall i, j, k,$$

then multiplication in A is associative in general.

On the other hand, if V is a vector space with a basis $\{e_1, e_2, \ldots, e_n\}$ and e_{ij}, $i, j = 1, 2, \ldots, n$, are arbitrarily chosen vectors from this space, we

can define multiplication on V by the following rule:

$$ab = \sum_{i,j} a_i b_j e_{ij},$$

thus making V an algebra.

Example 1.74. The field $\mathbb{C}$ regarded as an algebra over $\mathbb{R}$ has the following multiplication table for its basis vectors:

$\times$	1	i
1	1	i
i	i	-1

To check that multiplication in $\mathbb{C}$ is commutative and associative, it suffices to check that multiplication of 1 and i is commutative and associative, and this is trivial.

Example 1.75. For an orthonormal basis $\{i, j, k\}$ of E^3 (i.e., a basis that consists of orthogonal unit vectors), the multiplication table for the cross product looks like this:

$\times$	i	j	k
i	0	k	$-j$
j	$-k$	0	i
k	j	$-i$	0

This multiplication is anticommutative and satisfies the Jacobi identity (see Example 1.18). It is enough to check the latter identity for the basis vectors, which is not difficult (do this!).

Example 1.76. The *algebra of quaternions* $\mathbb{H}$ is given by its basis $\{1, i, j, k\}$ with the following multiplication table:

$\times$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

This algebra is associative (check this!) but not commutative. It contains the algebra of complex numbers as a subalgebra (see the definition in the next paragraph). Later we will see that just as in a field, every nonzero element of $\mathbb{H}$ is invertible. Thus, it is a "noncommutative field".

A subset of algebra is called a *subalgebra* if it is simultaneously a subring and a subspace. A map between algebras is called an *isomorphism* if it is simultaneously an isomorphism of vector spaces and of rings.

1.9. Matrix Algebras

An $m \times n$ *matrix* over a field K is a rectangular table of elements from K with m *rows* and n *columns*. Entries of a matrix are usually denoted by the same letter with two subscripts, the first being the row number and the second the column number:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Sometimes for brevity we will simply write $A = (a_{ij})$.

The sum of matrices $A = (a_{ij})$ and $B = (b_{ij})$ of the same size is the matrix

$$A + B = (a_{ij} + b_{ij}).$$

The product of a matrix $A = (a_{ij})$ and an element $\lambda \in K$ is the matrix

$$\lambda A = (\lambda a_{ij}).$$

With respect to these two operation all $m \times n$ matrices form a vector space that we will denote $K^{m \times n}$. In fact, it is no different from the space K^{mn} of rows of length mn. The special nature of matrices comes through in the definition of their multiplication.

The product of an $m \times n$ matrix $A = (a_{ij})$ and an $n \times p$ matrix $B = (b_{ij})$ is the $m \times p$ matrix $AB = (c_{ik})$ whose entries are determined by the following formula:

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}.$$

(We will explain the reason for this definition in Section 2.3.)

Observe that a product of two matrices is defined only when their sizes agree, namely, when the number of columns of the first matrix equals the number of rows of the second one.

Example 1.77.

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 5 \\ 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 \cdot 2 + 0 \cdot 0 + 2 \cdot 1 & 1 \cdot (-1) + 0 \cdot 5 + 2 \cdot 1 \\ 0 \cdot 2 + (-1) \cdot 0 + 3 \cdot 1 & 0 \cdot (-1) + (-1) \cdot 5 + 3 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 3 & -2 \end{pmatrix}.$$

Example 1.78.

$$\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}$$
$$=\begin{pmatrix}\cos\alpha\cos\beta-\sin\alpha\sin\beta & -\cos\alpha\sin\beta-\sin\alpha\cos\beta\\ \sin\alpha\cos\beta+\cos\alpha\sin\beta & -\sin\alpha\sin\beta+\cos\alpha\cos\beta\end{pmatrix}$$
$$=\begin{pmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix}.$$

Matrix multiplication is associative, meaning that

$$(AB)C=A(BC), \tag{1.9}$$

whenever sizes of matrices A, B, C agree so that all these products make sense.

Indeed, let

$$(AB)C=(u_{il}), \qquad A(BC)=(v_{il}).$$

Then

$$u_{il}=\sum_k\left(\sum_j a_{ij}b_{jk}\right)c_{kl}=\sum_{j,k}a_{ij}b_{jk}c_{kl},$$
$$v_{il}=\sum_j a_{ij}\left(\sum_k b_{jk}c_{kl}\right)=\sum_{j,k}a_{ij}b_{jk}c_{kl},$$

hence $u_{il}=v_{il}$.

An $n\times n$ matrix is called a *square matrix of order n*. A square matrix has two diagonals. The one leading from the upper left corner to the lower right corner is called the *main diagonal* or simply the *diagonal* and the other one the *secondary diagonal*. A square matrix is called *diagonal* if all its entries outside of the (main) diagonal are zero. Multiplication by a diagonal matrix looks especially simple:

$$\begin{pmatrix}a_1 & 0 & \dots & 0\\ 0 & a_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & a_n\end{pmatrix}\begin{pmatrix}b_{11} & b_{12} & \dots & b_{1p}\\ b_{21} & b_{22} & \dots & b_{2p}\\ \dots & \dots & \dots & \dots\\ b_{n1} & b_{n2} & \dots & b_{np}\end{pmatrix}=\begin{pmatrix}a_1b_{11} & a_1b_{12} & \dots & a_1b_{1p}\\ a_2b_{21} & a_2b_{22} & \dots & a_2b_{2p}\\ \dots & \dots & \dots & \dots\\ a_nb_{n1} & a_nb_{n2} & \dots & a_nb_{np}\end{pmatrix}$$

(every row of the second matrix is multiplied by the respective diagonal entry of the first matrix). Similarly,

$$\begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{m1} & a_{m2} & \dots & a_{mn}\end{pmatrix}\begin{pmatrix}b_1 & 0 & \dots & 0\\ 0 & b_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & b_n\end{pmatrix}=\begin{pmatrix}a_{11}b_1 & a_{12}b_2 & \dots & a_{1n}b_n\\ a_{21}b_1 & a_{22}b_2 & \dots & a_{2n}b_n\\ \dots & \dots & \dots & \dots\\ a_{m1}b_1 & a_{m2}b_2 & \dots & a_{mn}b_n\end{pmatrix}$$

(every column of the first matrix is multiplied by the respective diagonal entry of the second matrix). We will denote the diagonal matrix

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix}$$

by $\operatorname{diag}(a_1, a_2, \dots, a_n)$.

The diagonal matrix

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

is called the *identity matrix*. The above formulas imply that for any $m \times n$ matrix A,

$$AE = A, \qquad EA = A, \tag{1.10}$$

where in the first equality E stands for the identity matrix of order n and in the second, for the identity matrix of order m.

The following obvious properties relate matrix multiplication to other operations:

$$A(B + C) = AB + AC, \qquad (A + B)C = AC + BC, \tag{1.11}$$

$$(\lambda A)B = A(\lambda B) = \lambda(AB) \qquad \forall \lambda \in K. \tag{1.12}$$

(As in the statement concerning associativity, we assume here that the sizes of matrices agree so that all operations make sense.)

The sum and product of square matrices of the same order n are well defined; they are also square matrices of order n. Properties (1.9)–(1.12) show that all square matrices of order n form an associative algebra with unity. We denote it $\mathrm{L}_n(K)$.[1]

We notice below several "negative" properties of the algebra $\mathrm{L}_n(K)$ for $n \geq 2$. ($\mathrm{L}_1(K)$ is the field K itself.)

(i) The algebra $\mathrm{L}_n(K)$ is not commutative. The following example demonstrates this for $n = 2$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Similar examples can be provided for $n > 2$.

[1] This algebra is often denoted $\mathrm{M}_n(K)$. In our notation the letter "L" comes from "linear", and the reason for this choice is that matrices can be interpreted as linear maps (see Section 2.3).

(ii) The algebra $L_n(K)$ contains zero divisors. This follows, for instance, from the second equality above. Moreover, there exist nonzero matrices with zero squares, e.g.,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(iii) Not every nonzero element of $L_n(K)$ is invertible. This follows from the existence of zero divisors, since a zero divisor cannot be an invertible element (see the proof of absence of zero divisors in a field in Section 1.3). For instance, matrices $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ are not invertible in $L_2(K)$.

Exercise 1.79. A matrix E_{ij} that has 1 as the (i,j)th entry and zero in all other places is called a *matrix unit* (not to be confused with the identity matrix!). Matrix units E_{ij}, $i,j = 1,\ldots,n$, form a basis of the vector space $L_n(K)$. Write down the multiplication table of the algebra $L_n(K)$ in this basis.

Exercise 1.80. Matrices of the type λE, $\lambda \in K$, are called *scalar*. Clearly, any scalar matrix commutes with all other square matrices of the same order. Prove the converse: a square matrix that commutes with all other square matrices of the same order is scalar.

Exercise 1.81. Prove that in $L_2(\mathbb{R})$, matrices of the type

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

form a subalgebra isomorphic to the algebra of complex numbers.

Exercise 1.82. Prove that in the algebra $L_2(\mathbb{C})$ regarded as an algebra over $\mathbb{R}$, matrices of the type

$$\begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix}$$

form a subalgebra isomorphic to the algebra of quaternions (see Example 1.76).

For any matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

define the *transposed matrix*

$$A^{\mathsf{T}} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix},$$

whose columns are rows of A and vice versa. If we denote the (i,j)th entry of the transposed matrix by a_{ij}^{T}, then

$$a_{ij}^{\mathsf{T}} = a_{ji}.$$

Obviously,

$$(A^{\mathsf{T}})^{\mathsf{T}} = A,$$
$$(A+B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}},$$
$$(\lambda A)^{\mathsf{T}} = \lambda A^{\mathsf{T}} \qquad \forall \lambda \in K.$$

Also, it is true that

$$(AB)^{\mathsf{T}} = B^{\mathsf{T}} A^{\mathsf{T}}.$$

Indeed, let $AB = C = (c_{ik})$. Then

$$c_{ki}^{\mathsf{T}} = c_{ik} = \sum_j a_{ij} b_{jk} = \sum_j b_{kj}^{\mathsf{T}} a_{ji}^{\mathsf{T}},$$

implying that $C^{\mathsf{T}} = B^{\mathsf{T}} A^{\mathsf{T}}$.

Remark 1.83. Observe that all constructions in the last three sections would remain unchanged if we replaced K with a commutative associative ring with unity, for instance, the ring of integers or a ring of residue classes. The only difference lies in terminology: in this more general situation the term *module* is used instead of *vector space* (see Section 9.3).

Chapter 2

Elements of Linear Algebra

2.1. Systems of Linear Equations

Fix a field K. We are going to abuse the language slightly and call elements of K numbers. If it is difficult for you to think of a generic field, you can assume that $K = \mathbb{R}$; though, this case is not simpler than the generic one.

A *linear equation* with variables $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where *coefficients* $a_1, a_2, \ldots, a_n$ and the *free term* b belong to K. A linear equation is called *homogeneous* if $b = 0$.

A system of m linear equations with n variables has the following general form:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{2.1}$$

The matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

is called the *coefficient matrix* and the matrix

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

the *extended matrix* of system (2.1).

A system of equations is called *compatible* if it has at least one solution and *incompatible* otherwise. A compatible system can have one or more solutions. To solve a system of equations means to find all its solutions.

Observe that a solution of a system of equations with n variables is an ordered collection of n numbers, i.e., an element of K^n.

There exists a simple general method for solving systems of linear equations called *Gaussian elimination*. Its idea lies in reducing every system of linear equations to an equivalent system that has a simple form and whose solutions are easy to find. Recall that two systems of equations are called *equivalent* if their sets of solutions coincide, i.e., if every solution of the first system is a solution of the second and vice versa. Gaussian elimination is performed using special transformations called elementary.

Definition 2.1. An *elementary transformation* of a system of linear equations is a transformation of one of the following three types:

(i) adding an equation multiplied by a number to another equation;

(ii) interchanging two equations;

(iii) multiplying an equation by a nonzero number.

Notice that a transformation of the first type changes only one equation, the one to which the other, multiplied by a number, is being added.

Clearly, every solution of the original system of equations is a solution of the system obtained using an elementary transformation. On the other hand, the original system of equations can be reconstructed from the new one using an appropriate elementary transformation of the same type. For instance, if we add to the first equation the second one multiplied by c, we can get back by adding to the first equation of the new system the second equation (it is the same as in the original system) multiplied by $-c$. Thus, under any elementary transformation we obtain a system that is equivalent to the original one.

Since it is easier for us to work not with systems themselves but with their (extended) matrices, here is the corresponding definition for matrices:

Definition 2.1′. An *elementary row transformation* of a matrix is a transformation of one of the following three types:

(i) adding a row multiplied by a number to another row;

(ii) interchanging two rows;

(iii) multiplying a row by a nonzero number.

Obviously, every elementary transformation of a system of equations leads to a corresponding elementary row transformation of its extended matrix and its coefficient matrix.

We can show now that every matrix can be reduced to quite a simple form by elementary transformations.

Call the first nonzero element of a row $(a_1, a_2, \ldots, a_n)$ its *pivotal element*; if this element is a_k, then k is called the index of this pivotal element.

Definition 2.2. A matrix is in *step form* if

(i) the indices of pivotal elements of its nonzero rows form a strictly increasing sequence;

(ii) zero rows, if exist, are at the bottom.

That is, a matrix in step form looks as follows:

$$\begin{pmatrix} \boxed{a_{1j_1}} & \ldots\ldots\ldots\ldots\ldots\ldots \\ & \boxed{a_{2j_2}} \quad \ldots\ldots\ldots\ldots \\ & \qquad\qquad \ldots\ldots\ldots\ldots \\ 0 & \qquad\qquad\quad \boxed{a_{rj_r}} \quad \ldots \end{pmatrix}. \tag{2.2}$$

Here the entries $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ are nonzero and the entries to the left or below them are zero. Also, $j_1 < j_2 < \cdots < j_r$.

Theorem 2.3. *Every matrix can be reduced to step form by elementary transformations.*

Proof. If the given matrix is the zero one, it is already in step form. If it is nonzero, let j_1 be the index of its first nonzero column. By exchanging the rows, if necessary, we obtain a matrix where $a_{1j_1} \neq 0$. Then, we add to every row from the second down the first row multiplied by an appropriate number, so that all entries of the j_1th column, except the first one, become zero. We obtain a matrix of the form

$$\begin{pmatrix} 0 \ldots 0 & \boxed{a_{1j_1}} & \ldots\ldots \\ 0 & \multicolumn{2}{c}{\boxed{A_1}} \end{pmatrix}.$$

Applying the same procedure to the matrix A_1, we finally obtain a matrix of the form (2.2). □

Remark 2.4. In this proof we did not use elementary transformations of the third type. But they can be useful in solving particular systems.

Example 2.5. Reduce the following matrix to step form:

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

By subtracting from the 2nd, 3rd, and 4th rows the 1st row multiplied by 1, 2, and 2, respectively, we obtain the matrix

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & -3 & -3 & 3 & -6 \\ 0 & -4 & -4 & 3 & -3 \end{pmatrix}.$$

Then, by adding to the 3rd and 4th rows the 2nd row multiplied by 3 and 4, respectively, we get

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 5 \end{pmatrix}.$$

Finally, by exchanging the 3rd and 4th rows, we obtain a matrix in step form:

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & -1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Remark 2.6. The previous example was specially designed so that $j_1, \ldots, j_r$ would not be just the first r natural numbers. In some sense, this situation is an exception. For example, $j_1 \neq 1$ only when the first column of the original matrix is zero. Usually,

$$j_1 = 1, \qquad j_2 = 2, \qquad \ldots, \qquad j_r = r.$$

In such a case matrix (2.2) is called *trapezoidal*.

Now we apply the above theorem to solving systems of linear equations.

Definition 2.7. A system of linear equations is said to be in *step form* if its extended matrix is in step form.

Theorem 2.3 implies that every system of linear equations can be reduced to step form by elementary transformations. Thus, it is enough to learn how to solve systems already in step form.

We need to introduce a few terms. A square matrix $A = (a_{ij})$ is called *upper triangular* or simply *triangular* if $a_{ij} = 0$ for $i > j$ and *strictly triangular* if additionally $a_{ii} \neq 0$ for all i. A system of linear equations is called (*strictly*) *triangular* if its coefficient matrix is (strictly) triangular.

Remark 2.8. A square matrix $A = (a_{ij})$ is called *lower triangular* if $a_{ij} = 0$ for $i < j$.

Consider an arbitrary system of linear equations in step form. Denote the number of nonzero rows (number of steps) in this matrix by r and the number of nonzero rows of its extended matrix by $\tilde{r}$. Clearly, $\tilde{r} = r$ or $r + 1$.

Three cases are possible.

First case: $\tilde{r} = r + 1$. In this case the system contains an equation of the form

$$0x_1 + 0x_2 + \cdots + 0x_n = b,$$

where $b \neq 0$. Hence, it is incompatible.

Second case: $\tilde{r} = r = n$. In this case, after deleting zero equations (i.e., equations with all coefficients equal to zero), we obtain a strictly triangular system. We can uniquely determine x_n from the last equation, x_{n-1} from the next to last, and so on. Therefore, the system has a unique solution.

Third case: $\tilde{r} = r < n$. In this case, let $j_1, j_2, \ldots, j_r$ be the indices of pivotal coefficients of nonzero equations in the system. Call variables $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ *principal* and other variables *free*. Delete zero equations and carry terms with free variables to the right-hand side. In this way we obtain a strictly triangular system with respect to the principal variables. By solving it just like the one in the second case, we can express principal variables via the free ones. Together these expressions are called the *general solution* of the system. All solutions of the system are obtained from the general one by choosing some values for the free variables. Since these values can be chosen arbitrarily, the system has more than one solution and, when K is infinite, infinitely many solutions.

A compatible system of linear equations is called *determined* if it has a unique solution and *underdetermined* if it has more than one solution. As follows from the previous discussion, an underdetermined system has infinitely many solutions whenever K is infinite. Up to renumeration of variables, a general solution of such a system has the following form:

$$\text{(2.3)} \qquad \begin{cases} x_1 = c_{11}x_{r+1} + c_{12}x_{r+2} + \cdots + c_{1,n-r}x_n + d_1, \\ x_2 = c_{21}x_{r+1} + c_{22}x_{r+2} + \cdots + c_{2,n-r}x_n + d_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_r = c_{r1}x_{r+1} + c_{r2}x_{r+2} + \cdots + c_{r,n-r}x_n + d_r. \end{cases}$$

Example 2.9. The matrix from Example 2.5 is the extended matrix of the system

$$\begin{cases} x_1 + 2x_2 + x_3 = 2, \\ x_1 + 3x_2 + 2x_3 - x_4 = 4, \\ 2x_1 + x_2 - x_3 + 3x_4 = -2, \\ 2x_1 - 2x_3 + x_4 = 1. \end{cases}$$

Calculations in Example 2.5 show that this system is equivalent to the following system in step form:

$$\begin{cases} x_1 + 2x_2 + x_3 = 2, \\ x_2 + x_3 - x_4 = 2, \\ -x_4 = 5. \end{cases}$$

Taking variables x_1, x_2, and x_4 as principal and variable x_3 as free, we rewrite this system as

$$\begin{cases} x_1 + 2x_2 = -x_3 + 2, \\ x_2 - x_4 = -x_3 + 2, \\ -x_4 = 5. \end{cases}$$

Solving it with respect to x_1, x_2, and x_4, we find the general solution

$$\begin{cases} x_1 = x_3 + 8, \\ x_2 = -x_3 - 3, \\ x_4 = -5. \end{cases}$$

Remark 2.10. For consistency, we can think that for determined systems all variables are principal and no variable is free. Then the general solution is the unique solution of the system.

A strictly triangular matrix can be reduced to the identity matrix by elementary row transformations. To achieve this, we add the last row multiplied by an appropriate coefficient to all other rows. This coefficient is chosen here in such a way that all entries of the last column but the last one become zero. Then, similarly, we add the penultimate row to others so that all entries of the next to the last column (except for the diagonal entry) become zero, etc. Finally, we obtain a diagonal matrix. By multiplying its rows by appropriate numbers, we obtain the identity matrix. Using this method, we do not stop at the step form when solving a system of linear equations but continue with the transformations and reduce the coefficient matrix for the principal variables to the identity matrix. Then the general solution is easily obtained from the matrix we just got. This procedure is called the *reverse Gaussian elimination.*

Example 2.11. We continue reducing the matrix from Example 2.5. First, we delete the zero row. Then, by subtracting the third row from the second, we get

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 2 \\ 0 & 1 & 1 & 0 & -3 \\ 0 & 0 & 0 & -1 & 5 \end{pmatrix}.$$

We subtract the doubled second row from the first, multiply the third row by -1, and obtain

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 8 \\ 0 & 1 & 1 & 0 & -3 \\ 0 & 0 & 0 & 1 & -5 \end{pmatrix}.$$

Therefore, the system of equations from Example 2.9 is equivalent to

$$\begin{cases} x_1 \quad - x_3 \quad = 8, \\ \quad x_2 + x_3 \quad = -3, \\ \qquad\qquad x_4 = -5. \end{cases}$$

After carrying terms containing x_3 to the right-hand side, we obtain the general solution of this system (that we already found in Example 2.9).

A system of homogeneous linear equations is always compatible as it has the zero solution. If it is determined, then it has just the zero solution, and if it is underdetermined, it has at least one nonzero solution (even infinitely many if K is infinite). In the preceding notation, the latter case holds when $r < n$. Since $r < m$ always, we arrive at the following theorem, which is an important theoretical consequence of Gaussian elimination.

Theorem 2.12. *Every system of homogeneous linear equations for which the number of equations is less than the number of variables, has a nonzero solution.*

Underdetermined systems of linear equations differ by the *degree of indeterminacy*, which is naturally defined as the number of free variables in the general solution of a system. For instance, a line in three-dimensional space is given by a system of (two) linear equations with one free variable, and a plane by a system (of one equation) with two free variables. The same system of linear equations can admit different general solutions with different free variables, so it is natural to ask if the number of free variables always remains constant. A positive answer to this question relies on the concept of dimension introduced in the next section.

In the remaining part of this section, we will interpret Gaussian elimination in the language of matrix multiplication.

First of all, if X denotes the column of variables and B the column of free terms, system (2.1) can be rewritten in matrix form as follows:

$$AX = B. \tag{2.4}$$

Indeed, according to the rules of matrix multiplication, the matrix AX is a column of height m whose ith element equals

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n.$$

Setting this element equal to the ith element of the column B, we obtain exactly the ith equation of system (2.1).

Let U be a square matrix of order m. Multiplying both sides of equation (2.4) by U on the left, we obtain the following equation:

$$UAX = UB. \tag{2.5}$$

Obviously, every solution of (2.4) satisfies (2.5). Moreover, if U is invertible, multiplication by U^{-1} on the left changes (2.5) back into (2.4); hence these equations are equivalent.

Equation (2.5) corresponds to a system of linear equations with the coefficient matrix UA and the column UB of free variables. It is easy to see that the extended matrix of this system is $U\tilde{A}$.

Furthermore, a direct check shows that elementary transformations of a matrix A are equivalent to multiplying it on the left by the so-called *elementary matrices* of the following three types:

$$\begin{array}{c} \\ i \end{array} \begin{array}{c} \begin{array}{ccccccc} & & & & j & & \end{array} \\ \begin{pmatrix} 1 & & \vdots & & \vdots & & \\ & \ddots & \vdots & & \vdots & & \\ \cdots\cdots & & 1 & \cdots & c & \cdots\cdots & \\ & & & \ddots & \vdots & & \\ & & & & 1 & \cdots\cdots & \\ & & & & \vdots & \ddots & \\ & & & & \vdots & & 1 \end{pmatrix} \end{array} = E + cE_{ij},$$

$$\begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\begin{pmatrix} 1 & \vdots & & \vdots & \\ & \ddots & \vdots & & \vdots & \\ \cdots\cdots & 0 & \cdots & 1 & \cdots\cdots \\ & \vdots & \ddots & \vdots & \\ \cdots\cdots & 1 & \cdots & 0 & \cdots\cdots \\ & \vdots & & \vdots & \ddots & \\ & \vdots & & \vdots & & 1 \end{pmatrix} = P_{ij}, \quad \begin{array}{c} \\ i \\ \\ \\ \end{array}\begin{pmatrix} 1 & \vdots & \\ & \ddots & \vdots & \\ \cdots\cdots & \mathrm{c} & \cdots\cdots \\ & \vdots & \ddots & \\ & \vdots & & 1 \end{pmatrix} = Q_i(\mathrm{c}),$$

where $i \neq j$, $\mathrm{c} \neq 0$, and all entries in these matrices that are not written down explicitly are the same as in the identity matrix.

For example, multiplication of a matrix A by $E + cE_{ij}$, $i \neq j$, on the left adds the jth row of A multiplied by c to its ith row (and leaves other rows unchanged).

All elementary matrices are invertible; moreover, their inverses are elementary matrices corresponding to inverse elementary transformations:

$$(E + \mathrm{c}E_{ij})^{-1} = E - \mathrm{c}E_{ij}, \qquad P_{ij}^{-1} = P_{ij}, \qquad Q_i(\mathrm{c})^{-1} = Q_i(c^{-1}).$$

In the language of matrices, Gaussian elimination consists of successive multiplications on the left of equation (2.4) by elementary matrices with the goal of reducing A (and also the extended matrix $\tilde{A}$) to step form.

Remark 2.13. Use of other matrices instead of elementary ones provides us with different methods for solving systems of linear equations. Their theoretical underpinnings may not be so obvious, yet they are sometimes more useful for approximate calculations (when $K = \mathbb{R}$). For instance, such is the method of rotations; here matrices U are taken to be of the from

$$\begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\begin{pmatrix} 1 & \vdots & & \vdots & \\ & \ddots & \vdots & & \vdots & \\ \cdots\cdots & \cos\alpha & \cdots & -\sin\alpha & \cdots\cdots \\ & \vdots & \ddots & \vdots & \\ \cdots\cdots & \sin\alpha & \cdots & \cos\alpha & \cdots\cdots \\ & \vdots & & \vdots & \ddots & \\ & \vdots & & \vdots & & 1 \end{pmatrix}.$$

2.2. Basis and Dimension of a Vector Space

The concept of dimension is one of the most fundamental ideas in mathematics. In different branches of mathematics, it assumes various forms (as

does the concept of space itself). In this section we will define the dimension of a vector space and discuss questions related to this notion.

In Section 1.7, we introduced the notion of a basis of a vector space and proved that a vector space over a field K with a basis of n vectors is isomorphic to the space of rows K^n. The dimension of a vector space is defined as the number of vectors in its basis. However, before we state the definition precisely, two questions must be answered: which vector spaces possess a basis and whether a vector space may have two bases with different numbers of vectors.

To answer these questions, we need to introduce several new notions and prove a few statements that are also of independent interest.

Let V be a vector space over a field K.

A linear combination

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n, \qquad \lambda_1, \lambda_2, \ldots, \lambda_n \in K,$$

of vectors $a_1, a_2, \ldots, a_n \in V$ is called *trivial* if $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$ and *nontrivial* otherwise.

Definition 2.14. Vectors $a_1, a_2, \ldots, a_n$ are said to be *linearly dependent* if there exists a nontrivial linear combination of them that equals zero. Otherwise, they are said to be *linearly independent*.

Note that the notion of linear dependence (or independence) refers not to separate vectors but to their collections, or *systems*.

Remark 2.15. The notion of a system of vectors is different from that of a set of vectors. First, vectors in a system are assumed to be numbered. Second, some of them may be equal to each other. Thus, a system of n vectors is actually a mapping of the set $\{1, 2, \ldots, n\}$ into V. Notice, though, that the property of being dependent or independent does not depend on how vectors are numbered within the system.

Remark 2.16. The term "linear combination" actually has two meanings: the description of an action performed on given vectors (same as listing the coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n$) and the description of its result. In the statement "a nontrivial linear combination of these vectors equals zero," the word "nontrivial" is taken within the first meaning and "equals zero" within the second.

In other words, linear independence of vectors $a_1, a_2, \ldots, a_n$ means that equality

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n = 0$$

holds only when $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Example 2.17. A system that consists of exactly one vector is linearly dependent if and only if this vector is zero.

Example 2.18. A system that consists of two vectors is linearly dependent if and only if these vectors are proportional.

Example 2.19. Three geometric vectors (see Section 1.7) are linearly dependent if and only if they are coplanar (parallel to the same plane).

Clearly, if a system of vectors contains a linearly dependent subsystem, it is linearly dependent itself. For instance, any system of vectors that contains proportional vectors is linearly dependent.

Lemma 2.20. *Vectors $a_1, a_2, \ldots, a_n$, $n > 1$, are linearly dependent if and only if one of them is a linear combination of others.*

Proof. (i) If, for instance,

$$a_1 = \mu_2 a_2 + \cdots + \mu_n a_n,$$

then

$$a_1 - \mu_2 a_2 - \cdots - \mu_n a_n = 0,$$

which shows linear dependence of $a_1, a_2, \ldots, a_n$.

(ii) Conversely, let

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n = 0,$$

where not all coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n$ equal zero. Assume $\lambda_1 \neq 0$. Then

$$a_1 = -\frac{\lambda_2}{\lambda_1} a_2 - \cdots - \frac{\lambda_n}{\lambda_1} a_n,$$

i.e., a_1 can be expressed as a linear combination of $a_2, \ldots, a_n$. □

Remark 2.21. *Not every* vector in a linearly dependent system can be expressed as a linear combination of others. For example, let a be a nonzero vector. The system $\{a, 0\}$ is linearly dependent because

$$0a + 1 \cdot 0 = 0;$$

however, it is clear that a cannot be expressed via the zero vector.

Lemma 2.22. *Let vectors $a_1, a_2, \ldots, a_n$ be linearly independent. A vector b can be expressed as a linear combination of $a_1, a_2, \ldots, a_n$ if and only if vectors $a_1, a_2, \ldots, a_n, b$ are linearly dependent.*

Proof. If b can be expressed as a linear combination of $a_1, a_2, \ldots, a_n$, then $a_1, a_2, \ldots, a_n, b$ are linearly dependent by the previous lemma. Conversely, let

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n + \mu b = 0,$$

where not all coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n, \mu$ are equal to zero. We claim that $\mu \neq 0$. Indeed, otherwise $a_1, a_2, \ldots, a_n$ would be linearly dependent, thus contradicting the assumptions of the lemma. But then,

$$b = -\frac{\lambda_1}{\mu}a_1 - \frac{\lambda_2}{\mu}a_2 - \cdots - \frac{\lambda_n}{\mu}a_n.$$

□

Lemma 2.23. *Let b be a vector expressed as a linear combination of vectors $a_1, a_2, \ldots, a_n$. This expression is unique if and only if $a_1, a_2, \ldots, a_n$ are linearly independent.*

Proof. (i) Assume that b has two distinct expressions in terms of the vectors $a_1, a_2, \ldots, a_n$:

$$b = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n = \lambda_1' a_1 + \lambda_2' a_2 + \cdots + \lambda_n' a_n.$$

Then

$$(\lambda_1' - \lambda_1)a_1 + (\lambda_2' - \lambda_2)a_2 + \cdots + (\lambda_n' - \lambda_n)a_n = 0$$

is an expression of linear dependence of $a_1, a_2, \ldots, a_n$.

(ii) Conversely, let

$$\mu_1 a_1 + \mu_2 a_2 + \cdots + \mu_n a_n = 0$$

be an expression of linear dependence of $a_1, a_2, \ldots, a_n$. Then, if

$$b = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n,$$

we have

$$b = (\lambda_1 + \mu_1)a_1 + (\lambda_2 + \mu_2)a_2 + \cdots + (\lambda_n + \mu_n)a_n$$

as well, which is a different way of expressing b in terms of $a_1, a_2, \ldots, a_n$. □

Let $S \subset V$ be a subset. The collection of all possible (finite) linear combinations of vectors from S is called the *linear span* of S and is denoted $\langle S \rangle$. It is the smallest subspace of V containing S (check this!). We say that S spans V if $V = \langle S \rangle$.

Definition 2.24. A vector space is called *finite-dimensional* if it is spanned by a finite number of vectors, and *infinite-dimensional* otherwise.

Proposition 2.25. *If a vector space V is spanned by n vectors, then any $m > n$ vectors in V are linearly dependent.*

Proof. Let $V = \langle a_1, a_2, \ldots, a_n \rangle$ and $b_1, b_2, \ldots, b_m$, $m > n$, be some vectors in V. We can express them in terms of $a_1, a_2, \ldots, a_n$:

$$\begin{aligned} b_1 &= \mu_{11}a_1 + \mu_{12}a_2 + \cdots + \mu_{1n}a_n, \\ b_2 &= \mu_{21}a_1 + \mu_{22}a_2 + \cdots + \mu_{2n}a_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_m &= \mu_{m1}a_1 + \mu_{m2}a_2 + \cdots + \mu_{mn}a_n. \end{aligned}$$

For any $\lambda_1, \lambda_2, \ldots, \lambda_m \in K$, we have

$$\begin{aligned} \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_m b_m = &(\lambda_1\mu_{11} + \lambda_2\mu_{21} + \cdots + \lambda_m\mu_{m1})a_1 \\ &+(\lambda_1\mu_{12} + \lambda_2\mu_{22} + \cdots + \lambda_m\mu_{m2})a_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &+(\lambda_1\mu_{1n} + \lambda_2\mu_{2n} + \cdots + \lambda_m\mu_{mn})a_n. \end{aligned}$$

Consider the following system of n homogeneous linear equations with m variables:

$$\begin{cases} \mu_{11}x_1 + \mu_{21}x_2 + \cdots + \mu_{m1}x_m = 0, \\ \mu_{12}x_1 + \mu_{22}x_2 + \cdots + \mu_{m2}x_m = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mu_{1n}x_1 + \mu_{2n}x_2 + \cdots + \mu_{mn}x_m = 0. \end{cases}$$

If $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is a solution of this system, then

$$\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_m b_m = 0.$$

On the other hand, this system has a nonzero solution by Theorem 2.12. Therefore, vectors $b_1, b_2, \ldots, b_m$ are linearly dependent. □

In view of Lemma 2.23, Definition 1.62 of a basis of a vector space can be reformulated as follows:

Definition 2.26. A *basis* of a vector space V is a linearly independent system of vectors that spans V.

Theorem 2.27. *Every finite-dimensional subset V has a basis. More precisely, every finite subset S of V that spans V contains a basis of V.*

Proof. If S is linearly dependent, it contains a vector that can be expressed in terms of other vectors by Lemma 2.20. If we remove this vector from S, we obtain a set that still spans V but contains a smaller number of vectors. Continuing further, we will finally obtain a linearly independent set that spans V, i.e., a basis of V. □

Theorem 2.28. *All bases of a finite-dimensional space V contain the same number of vectors.*

This number is called the *dimension* of V and is denoted $\dim V$.

Proof. Assume V contains two bases with different numbers of vectors. Then according to Proposition 2.25, the one with the greater number of vectors is linearly dependent, contradicting the definition of basis. □

Remark 2.29. The zero vector space (that consists of the zero vector only) is regarded as having the "empty basis"; accordingly, its dimension is considered to be zero.

Example 2.30. The dimension of E^2 (respectively, E^3) is 2 (respectively, 3).

Example 2.31. It follows from Example 1.64 that K^n has dimension n.

Example 2.32. The field of complex numbers regarded as a vector space over $\mathbb{R}$ has dimension 2, and the algebra of quaternions (see Example 1.76) has dimension 4.

Example 2.33. If X is a finite set of n elements, then the vector space $F(X, K)$ of all functions on X with values in K (see Example 1.57) has dimension n. Indeed, consider the so-called δ-function δ_a ($a \in X$) defined as

$$\delta_a(x) = \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{if } x \neq a. \end{cases}$$

Clearly, any function $\varphi \in F(X, K)$ can be uniquely expressed in terms of δ-functions, namely

$$\varphi = \sum_{a \in X} \varphi(a)\delta_a.$$

Therefore, the functions δ_a, $a \in X$, form a basis of $F(X, K)$, and in this basis the coordinates of a function are its values.

If X is infinite, then for any n, $F(X, K)$ contains n linearly independent vectors, e.g., $\delta_{a_1}, \delta_{a_2}, \ldots, \delta_{a_n}$ for distinct $a_1, a_2, \ldots, a_n \in X$. Thus, in this case $F(X, K)$ is infinite-dimensional.

Example 2.34. The field $\mathbb{R}$ regarded as a vector space over $\mathbb{Q}$ is infinite-dimensional. Indeed, if it were finite-dimensional, every real number would be determined by its coefficients in some basis, i.e., by a finite collection of rational numbers. But then the set of real numbers would be countable and this is not so.

Exercise 2.35. Determine the number of vectors in an n-dimensional vector space over a finite field with q elements.

Exercise 2.36. Prove that the space of all continuous functions on an interval is infinite-dimensional.

Proposition 2.25 implies that a (finite or infinite) set S of vectors in a finite-dimensional vector space V contains a maximal linearly independent

subset, i.e., a linearly independent subset that becomes linearly dependent when any vector in S is added to it. Moreover, each linearly independent subset of S can be completed to a maximal linearly independent subset.

Proposition 2.37. *Any maximal linearly independent subset* $\{e_1, e_2, \ldots, e_k\}$ *of a set* S *is a basis of the linear span* $\langle S \rangle$ *of* S.

Proof. We need to show that every vector in $\langle S \rangle$ can be expressed as a linear combination of $e_1, e_2, \ldots, e_k$. By definition, every vector in $\langle S \rangle$ can be expressed as a linear combination of vectors from S. Hence, it suffices to show that every vector $a \in S$ can be expressed as a linear combination of $e_1, \ldots, e_k$. For $a \in \{e_1, \ldots, e_k\}$, this is obvious. For $a \notin \{e_1, \ldots, e_k\}$, this follows from Lemma 2.22. □

Applying the above considerations to $S = V$, we obtain the following theorem:

Theorem 2.38. *Any linearly independent system of vectors in a vector space* V *can be completed to a basis.*

In particular, any nonzero vector is contained in some basis and any n linearly independent vectors in an n-dimensional vector space already form a basis.

Exercise 2.39. Determine the number of bases of an n-dimensional space over a field of q elements.

Theorem 2.40. *Any subspace* U *of a finite-dimensional space* V *is also finite-dimensional, and* $\dim U \leq \dim V$. *Moreover, if* $U \neq V$, *then* $\dim U < \dim V$.

Proof. Let $\{e_1, e_2, \ldots, e_k\}$ be a maximal linearly independent system of vectors in U. By Proposition 2.37, $\{e_1, e_2, \ldots, e_k\}$ is a basis of U. Therefore, $\dim U = k$. The linearly independent system $\{e_1, e_2, \ldots, e_k\}$ can be completed to a basis of V. Thus, if $U \neq V$, $\dim V > k$. □

Exercise 2.41. Determine the number of k-dimensional subspaces of an n-dimensional vector space over a field of q elements.

The next theorem provides a complete description of all finite-dimensional vector spaces.

Theorem 2.42. *Finite-dimensional vector spaces over the same field are isomorphic if and only if their dimensions are the same.*

Proof. If $f : V \to U$ is an isomorphism of vector spaces and $\{e_1, e_2, \ldots, e_n\}$ is a basis of V, then $\{f(e_1), f(e_2), \ldots, f(e_n)\}$ is a basis of U, hence $\dim V =$

$\dim U$. Conversely, by Proposition 1.66, every n-dimensional vector space over a field K is isomorphic to K^n; therefore, all such spaces are isomorphic. □

Thus in any of our statements, we can replace an n-dimensional vector space over K with the space of rows K^n. The space K^n possesses a "distinguished" basis consisting of unit rows (see Example 1.64). On the other hand, if we fix a basis in an n-dimensional space V, then by assigning to each vector the row of its coordinates in this basis (as in the proof of Proposition 1.66), we establish a canonical isomorphism between V and K^n. This isomorphism maps basis vectors to unit rows. In this sense we can say that the space of rows is nothing but a finite-dimensional space with a fixed basis.

The set of all bases of an n-dimensional vector space V can be described in the following way. Fix a basis $\{e_1, \dots, e_n\}$. Any system $\{e'_1, \dots, e'_n\}$ of n vectors is given by a square matrix $C = (c_{ij})$ whose entries are defined by equalities

$$e'_j = \sum_i e_i c_{ij}, \qquad j = 1, \dots, n. \tag{2.6}$$

This matrix is called the *transition matrix* from the basis $\{e_1, \dots, e_n\}$ to the system $\{e'_1, \dots, e'_n\}$. According to its definition, the jth column of C is the column of coordinates of e'_j in the basis $\{e_1, \dots, e_n\}$. Thus, vectors $e'_1, \dots, e'_j$ are linearly independent (and hence form a basis) if and only if the columns of C are linearly independent. Such a matrix is called *nonsingular* (see also Definition 2.72). The aforesaid establishes a one-to-one correspondence between the set of all bases of V and the set of nonsingular matrices of order n.

We can extend the law of matrix multiplication to the case where entries in one of the two matrices are vectors (this makes sense because of how operations on a vector space are defined). Then equality (2.6) can be rewritten in matrix form as follows:

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C. \tag{2.7}$$

Let $x \in V$ be a vector. In bases $\{e_1, \dots, e_n\}$ and $\{e'_1, \dots, e'_n\}$, it is expressed as

$$x = x_1 e_1 + \cdots + x_n e_n = x'_1 e'_1 + \cdots + x'_n e'_n.$$

Put

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Then

$$x = (e'_1, \dots, e'_n)X' = (e_1, \dots, e_n)CX'.$$

This gives the formula for the change of coordinates under the transition from the basis $\{e_1, \dots, e_n\}$ to the basis $\{e'_1, \dots, e'_n\}$:

$$X = CX' \tag{2.8}$$

or, in greater detail,

$$x_i = \sum_j c_{ij} x'_j, \qquad i = 1, \dots, n. \tag{2.9}$$

The notions of basis and dimension can be extended to infinite-dimensional vector spaces. For this, we need to define the linear combination of an infinite system of vectors. In a purely algebraic situation there is no other way but to restrict our attention to the case of linear combinations where only finitely many coefficients are nonzero.

Let $\{a_i : i \in I\}$ be a system of vectors enumerated by an infinite set I. A linear combination of vectors a_i, $i \in I$, is an expression of the form $\sum_{i \in I} \lambda_i a_i$ where only finitely many coefficients λ_i are nonzero. Thus, the sum is finite, hence it makes sense. Just as in the case of finite systems of vectors, this definition leads to the notions of linear expression, linear dependence, and basis.

The dimension of a space is the cardinality of its basis. In particular, a vector space with a countable basis is called *countable-dimensional*.

Example 2.43. Consider the set of all sequences (infinite rows) of elements of a field K. Clearly, it is a vector space with respect to operations of addition and multiplication by elements of K that are defined just as they are for rows of finite length. We say that a sequence is *finitary* if only a finite number of its entries are nonzero. Finitary sequences form a subspace in the space of all sequences. Denote it by K^∞. As its basis vectors, we can take sequences of the following form:

$$e_i = (0, \dots, 0, 1, 0, \dots), \qquad i = 1, 2, \dots$$

(the identity is in the ith place). Hence K^∞ is countable-dimensional.

Just as in Proposition 1.66, we can prove that every countable-dimensional space over K is isomorphic to K^∞.

Exercise 2.44. Prove that $\mathbb{R}$ regarded as a vector space over $\mathbb{Q}$ is not countable-dimensional.

Exercise 2.45. Prove that any countable set in K^∞ that spans K^∞ contains a basis.

Exercise 2.46. Prove that every uncountable set of vectors in a countable-dimensional space is linearly dependent (hence, every basis of such a space is countable).

Exercise 2.47. Prove that any (finite or countable) linearly independent system of vectors in a countable-dimensional space can be completed to a basis.

Exercise 2.48. Prove that a subspace of a countable-dimensional vector space is at most countable-dimensional (i.e., either finite-dimensional or countable-dimensional). Give an example of a countable-dimensional subspace of a countable-dimensional vector space that does not coincide with the whole space.

Exercises 2.45–2.48 are analogues of Theorems 2.27, 2.28, 2.38, and 2.40 for countable-dimensional spaces. Similar statements can be proved for spaces of uncountable dimension, but this requires the use of set theory (transfinite induction or Zorn's lemma). On the other hand, this purely algebraic approach has a restricted area of applications. Usually, a space of uncountable dimension is endowed with a topology, which gives meaning to infinite sums of vectors.

The notion of dimension is closely related to those of rank of a matrix and rank of a system of vectors.

Definition 2.49. The *rank of a system of vectors* is the dimension of its linear span. The *rank of a matrix* is the rank of the system of its rows.

The rank of a matrix A is denoted $\operatorname{rk} A$.

Two systems of vectors $\{a_1, a_2, \ldots, a_n\}$ and $\{b_1, b_2, \ldots, b_m\}$ are called *equivalent* if every vector b_j can be expressed as a linear combination of $a_1, a_2, \ldots, a_n$ and, vice versa, every vector a_i can be expressed as a linear combination of $b_1, b_2, \ldots, b_m$. Obviously, this holds if and only if the corresponding spans coincide:

$$\langle a_1, a_2, \ldots, a_n \rangle = \langle b_1, b_2, \ldots, b_m \rangle.$$

Thus, equivalent systems of vectors have the same rank.

The definition of an elementary transformation implies that rows of a matrix A' obtained from another matrix A using an elementary transformation can be expressed as a linear combination of the rows of A. But as A can be obtained from A' using the inverse transformation, its rows can be expressed as a linear combination of the rows of A'. Therefore, the systems of rows of A and A' are equivalent and the ranks of these matrices are equal.

This is useful for calculating the rank of a matrix.

Proposition 2.50. *The rank of a matrix is equal to the number of nonzero rows of the matrix in step form to which it is reduced by elementary transformations.*

Proof. Since the rank of a matrix does not change under elementary transformations, it suffices to prove that the rank of a matrix in step form equals the number of its nonzero rows. This will follow if we can prove that nonzero rows of a matrix in step form are linearly independent.

Consider a matrix in step form (2.2). Assume that a linear combination of its nonzero rows with coefficients $\lambda_1, \lambda_2, \ldots, \lambda_r$ equals 0. The j_1th coordinate of this linear combination is $\lambda_1 a_{j_1}$, thus $\lambda_1 = 0$. This, together with the formula for the j_2th coordinate, implies $\lambda_2 a_{j_2} = 0$ making $\lambda_2 = 0$. Continuing further, we see that all coefficients $\lambda_1, \lambda_2, \ldots, \lambda_r$ are zero, as required. □

In particular, the number of nonzero rows in a matrix in step form, to which a given matrix is reduced, is constant, regardless of the sequence of elementary transformations chosen.

Proposition 2.50 combined with the discussion of systems of linear equations in step form in Section 2.1 imply the following theorem:

Theorem 2.51. *(i)* (Kronecker-Cappelli Theorem.) *A system of linear equations is compatible if and only if the rank of its matrix of coefficients equals the rank of its extended matrix.*

(ii) A compatible system of linear equations is determined if and only if the rank of its matrix of coefficients equals the number of variables.

2.3. Linear Maps

Every algebraic theory considers maps that are more general than isomorphisms. Usually, these maps are called homomorphisms or, in the case of vector spaces, linear maps. While isomorphisms fully preserve inner properties of algebraic structures and their elements, homomorphisms do so only partially.

Definition 2.52. Let U and V be vector spaces over a field K. A map

$$\varphi : U \to V$$

is called *linear* if

(i) $\varphi(a + b) = \varphi(a) + \varphi(b)$ for any $a, b \in U$;

(ii) $\varphi(\lambda a) = \lambda\varphi(a)$ for any $\lambda \in K$, $a \in U$.

This definition is different from that of an isomorphism between two vector spaces only in that it does not require the map to be bijective.

Observe that under a linear map the zero vector is mapped into the zero vector and the opposite of a vector into the opposite of its image. Indeed,

$$\varphi(0) = \varphi(0 \cdot 0) = 0\varphi(0) = 0,$$
$$\varphi(-a) = \varphi((-1)a) = (-1)\varphi(a) = -\varphi(a).$$

It is also easy to show that

$$\varphi(a - b) = \varphi(a) - \varphi(b).$$

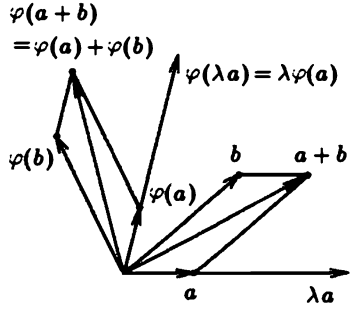


Figure 2.1

Example 2.53. A rotation is a linear map (and even an isomorphism) from E^2 to itself (see Figure 2.1).

Example 2.54. An orthogonal projection onto a plane defines a linear map (but not an isomorphism) from E^3 to the space of geometric vectors on this plane.

Example 2.55. Differentiation is a linear map from the space of all functions continuously differentiable on a given interval of the real line to the space of functions continuous on this interval.

Example 2.56. The map

$$f \mapsto \int_a^b f(x)dx$$

is a linear map of functions continuous on $[a, b]$ to $\mathbb{R}$ regarded as a one-dimensional vector space over $\mathbb{R}$.

A linear map $\varphi : U \to V$ is uniquely determined by the images of the basis vectors of U. Indeed, let $\{e_i : i \in I\}$ be a basis of U. Then for every vector $x = \sum_i x_i e_i$, we have

$$\varphi(x) = \sum_i x_i \varphi(e_i).$$

On the other hand, if $v_i \in V$ $(i \in I)$ are arbitrary vectors, then the map $\varphi : U \to V$ defined as

$$\varphi(x) = \sum_i x_i v_i$$

is easily seen to be linear. Also, $\varphi(e_i) = v_i$.

These considerations lead us towards a more analytical description of linear maps. We shall provide it for the spaces of rows. Let

$$\varphi : K^n \to K^m$$

be a linear map. Apply it to the unit rows $e_1, e_2, \ldots, e_n$ of the space K^n (see Example 1.64). We get rows

$$\varphi(e_j) = (a_{1j}, a_{2j}, \ldots, a_{mj}) \in K^m, \qquad j = 1, 2, \ldots, n.$$

The numbers a_{ij} $(i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n)$ form an $m \times n$ matrix A. It is called the *matrix of the linear map* φ. (Notice that the coordinates of the row $\varphi(e_j)$ form the jth column of A.)

For any row

$$x = (x_1, x_2, \ldots, x_n) = \sum_j x_j e_j \in K^n,$$

we have

$$\varphi(x) = \sum_j x_j \varphi(e_j) = \left(\sum_j a_{1j} x_j, \sum_j a_{2j} x_j, \ldots, \sum_j a_{mj} x_j \right).$$

Therefore, if we put

$$\varphi(x) = y = (y_1, y_2, \ldots, y_m),$$

we can express $y_1, y_2, \ldots, y_m$ in terms of $x_1, x_2, \ldots, x_n$ as

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, 2, \ldots, m. \tag{2.10}$$

Conversely, if $A = (a_{ij})$ is an arbitrary $m \times n$ matrix, the map $\varphi : K^n \to K^m$ defined by (2.10) is linear and has A as its matrix. Thus we established a one-to-one correspondence between linear maps from K^n into K^m and $m \times n$ matrices.

In a similar way, we can determine the matrix of a linear map $\varphi : U \to V$ between two arbitrary finite-dimensional vector spaces. Namely, its jth column contains coordinates of the image of the jth basis vector of U. Of course, this matrix depends on the choices of bases in the spaces U and V.

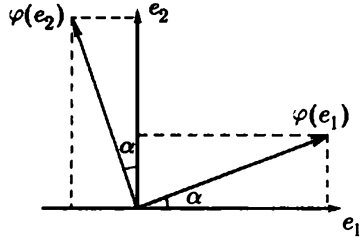


Figure 2.2

Example 2.57. Choose an orthonormal basis $\{e_1, e_2\}$ in E^2. Let φ be a rotation through an angle α (Figure 2.2). Then

$$\begin{aligned} \varphi(e_1) &= e_1 \cos\alpha + e_2 \sin\alpha, \\ \varphi(e_2) &= -e_1 \sin\alpha + e_2 \cos\alpha. \end{aligned}$$

This means that the matrix of φ is

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{2.11}$$

Observe that in this case $U = V$. Also, though not required by definition, we used the same basis $\{e_1, e_2\}$ twice: first as the basis of U and then as the basis of V.

Example 2.58. Here we will determine the matrix of the projection in Example 2.54. Fix a basis $\{e_1, e_2\}$ in the plane of projection. Complete it to a basis of the whole space with a vector e_3 orthogonal to this plane. Since under projection e_1 and e_2 are mapped to themselves and e_3 to 0, the matrix in question has the form

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

(for these choices of bases).

Unlike an isomorphism, a linear map might be neither injective nor surjective. Violations of these two properties provide us with two subspaces associated to any linear map: its kernel and image.

Definition 2.59. The *image* of a linear map $\varphi : U \to V$ is the subset

$$\operatorname{Im}\varphi = \{\varphi(a) : a \in U\} \subset V,$$

and its *kernel*, the subset

$$\operatorname{Ker}\varphi = \{a \in U : \varphi(a) = 0\} \subset U.$$

It is easy to see that $\operatorname{Im}\varphi$ is a subspace of V and $\operatorname{Ker}\varphi$, a subspace of U. For example, let us prove the second claim. If $a, b \in \operatorname{Ker}\varphi$, i.e., $\varphi(a) = \varphi(b) = 0$, then

$$\varphi(a+b) = \varphi(a) + \varphi(b) = 0 + 0 = 0,$$

i.e., $a + b \in \operatorname{Ker}\varphi$. Furthermore, if $a \in \operatorname{Ker}\varphi$, then for any $\lambda \in K$,

$$\varphi(\lambda a) = \lambda\varphi(a) = \lambda 0 = 0,$$

i.e., $\lambda a \in \operatorname{Ker}\varphi$. Finally, $0 \in \operatorname{Ker}\varphi$, as we have already shown that $\varphi(0) = 0$.

Example 2.60. The kernel of the projection map in Example 2.54 is the set of vectors orthogonal to the plane of projection.

Example 2.61. The kernel of the differentiation map in Example 2.55 is the set of constant functions. Its image is the space of all continuous functions. The latter follows from the existence of antiderivative of any continuous function (this is shown in advanced calculus).

Theorem 2.62. *A linear map $\varphi : U \to V$ is injective if and only if* $\operatorname{Ker}\varphi = 0$. *More precisely, for any $b \in \operatorname{Im}\varphi$, the set of solutions of the equation*

$$\varphi(x) = b \tag{2.12}$$

is of the form $a + \operatorname{Ker}\varphi$, where a is one of the solutions of this equation.

(Here $a + \operatorname{Ker}\varphi$ is understood as the collection of sums of the form $a + y$ with $y \in \operatorname{Ker}\varphi$.)

Notice immediately that by definition $\operatorname{Ker}\varphi$ is the set of solutions of the equation

$$\varphi(x) = 0. \tag{2.13}$$

Proof. Injectivity of φ means that for any $b \in \operatorname{Im}\varphi$ equation (2.12) has a unique solution. Thus it suffices to prove only the second claim of the theorem.

Let $\varphi(a) = b$. If $y \in \operatorname{Ker}\varphi$, then

$$\varphi(a+y) = \varphi(a) + \varphi(y) = b + 0 = b.$$

Conversely, if $\varphi(x) = b$, then

$$\varphi(x-a) = \varphi(x) - \varphi(a) = b - b = 0,$$

i.e., $y = x - a \in \operatorname{Ker}\varphi$. Hence,

$$x = a + y \in a + \operatorname{Ker}\varphi.$$

□

Let $\varphi : K^n \to K^m$ be a linear map with matrix A and $b = (b_1, b_2, \ldots, b_m)$. Then equation (2.12) in the coordinate form is simply the system of linear equations (2.1), and equation (2.13) is the system of homogeneous linear equations with the same coefficients:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0. \end{cases} \tag{2.14}$$

In this way we see that the set of solutions of system (2.14) is a subspace of K^n. Also, the set of solutions of system (2.1), if it is nonempty, is the sum of one of its solutions and this subspace. But what is the dimension of the space of solutions of (2.14)? The answer is given by the following theorem.

Theorem 2.63. *Let $\varphi : K^n \to K^m$ be a linear map with matrix A. Then*

$$\dim \operatorname{Ker}\varphi = n - \operatorname{rk} A.$$

Proof. Using elementary transformations, we reduce system (2.14) to step form. In view of Proposition 2.50, the number of nonzero equations in this step form equals $r = \operatorname{rk} A$. Hence, a generic solution of (2.14) has r principal variables and, up to renumeration of variables, is of the following form (*cf.* (2.3)):

$$\begin{cases} x_1 = c_{11}x_{r+1} + c_{12}x_{r+2} + \cdots + c_{1,n-r}x_n, \\ x_2 = c_{21}x_{r+1} + c_{22}x_{r+2} + \cdots + c_{2,n-r}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_r = c_{r1}x_{r+1} + c_{r2}x_{r+2} + \cdots + c_{r,n-r}x_n. \end{cases} \tag{2.15}$$

Assigning the value of 1 to one of the free variables $x_{r+1}, x_{r+2}, \ldots, x_n$ and 0 to the rest of them, we obtain the following solutions of system (2.14):

$$\begin{aligned} u_1 &= (c_{11}, c_{21}, \ldots, c_{r1}, 1, 0, \ldots, 0), \\ u_2 &= (c_{12}, c_{22}, \ldots, c_{r2}, 0, 1, \ldots, 0), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ u_{n-r} &= (c_{1,n-r}, c_{2,n-r}, \ldots, c_{r,n-r}, 0, 0, \ldots, 1). \end{aligned}$$

To prove the theorem, it remains to show that these solutions form a basis of $\operatorname{Ker}\varphi$.

For any $\lambda_1, \lambda_2, \ldots, \lambda_{n-r} \in K$, a linear combination

$$u = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_{n-r} u_{n-r}$$

is a solution of system (2.14) where the free variables assume values $\lambda_1, \lambda_2, \ldots, \lambda_{n-r}$. The values of the principal variables are uniquely determined by the values of the free ones (in accordance with (2.15)). Thus, every solution of system (2.14) is a linear combination of $u_1, u_2, \ldots, u_{n-r}$. On the other hand, if $u = 0$, then $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-r} = 0$; therefore, $u_1, u_2, \ldots, u_{n-r}$ are linearly independent. □

Given a system of homogeneous linear equations, any basis of the space of its solutions is called a *fundamental system of solutions*. The above proof provides a working algorithm for constructing such a system of solutions.

Let $\varphi : U \to V$ be a linear map between finite-dimensional vector spaces and $\{e_1, e_2, \ldots, e_n\}$ a basis of U. Then for any

$$a = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n \in U,$$

we have

$$\varphi(a) = a_1 \varphi(e_1) + a_2 \varphi(e_2) + \cdots + a_n \varphi(e_n).$$

Hence,

$$\operatorname{Im} \varphi = (\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_n)). \tag{2.16}$$

Theorem 2.64. $\dim \operatorname{Im} \varphi + \dim \operatorname{Ker} \varphi = \dim U$.

Proof. Choose a basis of U in a special way: first, choose a basis $\{e_1, \ldots, e_k\}$ of $\operatorname{Ker} \varphi$ and then complete it to a basis of U. Our choice implies $\varphi(e_1) = \cdots = \varphi(e_k) = 0$; thus it follows from (2.16) that

$$\operatorname{Im} \varphi = (\varphi(e_{k+1}), \ldots, \varphi(e_n)).$$

We claim that the vectors $\varphi(e_{k+1}), \ldots, \varphi(e_n)$ are linearly independent; the theorem will follow from this.

To prove this claim, assume

$$\lambda_1 \varphi(e_{k+1}) + \cdots + \lambda_{n-k} \varphi(e_n) = 0.$$

Consider a vector

$$a = \lambda_1 e_{k+1} + \cdots + \lambda_{n-k} e_n.$$

The previous equality implies that $\varphi(a) = 0$, i.e.,

$$a \in \operatorname{Ker} \varphi = \langle e_1, \ldots, e_k \rangle.$$

Since $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ are linearly independent, this is possible only for $\lambda_1 = \cdots = \lambda_{n-k} = 0$ and the claim follows. □

Corollary 2.65. *If $\varphi : K^n \to K^m$ is a linear map with a matrix A, then*

$$\dim \operatorname{Im} \varphi = \operatorname{rk} A.$$

Proof. The proof follows from comparing the statement of Theorems 2.63 and 2.64. □

Corollary 2.66. *The rank of the system of columns of any matrix* (column rank) *equals the rank of its system of rows* (row rank).

Proof. Let $\varphi : K^n \to K^m$ be a linear map with a matrix A and $e_1, e_2, \ldots, e_n$ the unit rows of K^n. It follows from (2.16) that the dimension of $\operatorname{Im}\varphi$ equals the rank of the system of columns of A. Comparing this result with the previous corollary completes the proof. □

Example 2.67. A field K can be viewed as a (one-dimensional) vector space over itself. A linear map $\varphi : V \to K$ is called a *linear function* on V. If φ is a nonzero linear function, then $\operatorname{Im}\varphi = K$. Thus, if $\dim V = n$, Theorem 2.64 implies that

$$\dim \operatorname{Ker}\varphi = n - 1.$$

Example 2.68. Let X be the set of tetrahedron's edges and Y the set of its faces. To any function f on X with values in K, we assign a function g on Y defined as

$$g(y) = \sum_{x \subset y} f(x).$$

That is, the value of g on a face equals the sum of the values of f on the sides of this face. This defines the following linear map:

$$\varphi : F(X, K) \to F(Y, K)$$

(see Example 1.55). It is not difficult to prove that whenever $\operatorname{char} K \neq 2$, φ is surjective. For this, it suffices to show that $\operatorname{Im}\varphi$ contains δ-functions of all faces (see Example 2.33). A function f for which $\varphi(f)$ is a δ-function of the bottom face is shown in Figure 2.3, left (its value on unmarked edges is 0).

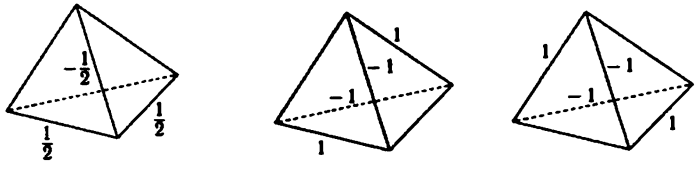


Figure 2.3

We have

$$\dim F(X, K) = 6, \quad \dim F(Y, K) = 4;$$

therefore, according to Theorem 2.64,

$$\dim \operatorname{Ker}\varphi = 6 - 4 = 2.$$

Functions comprising a basis of $\operatorname{Ker}\varphi$ are shown in Figure 2.3, two pictures on the right.

Exercise 2.69. For a linear map φ in the above example, find $\dim \operatorname{Ker}\varphi$ when $\operatorname{char} K = 2$.

Since the columns of a matrix A are the rows of its transposed matrix $A^\top$ (see Section 1.9), Corollary 2.66 implies that

$$\operatorname{rk} A^\top = \operatorname{rk} A.$$

We can define elementary column transformations just as we defined elementary row transformations of a matrix. They correspond to elementary row transformations of the transposed matrix. Thus the rank of a matrix does not change not only under elementary row transformations but also under elementary column transformations.

Remark 2.70. Elementary column transformations are equivalent to multiplying the matrix on the right by elementary matrices.

We turn now to operations on linear maps.

Linear maps $U \to V$ can be added together and multiplied by numbers, just as functions:

$$(\varphi + \psi)(a) = \varphi(a) + \psi(a),$$
$$(\lambda\varphi)(a) = \lambda\varphi(a).$$

They form a vector space with respect to these operations.

Also, if

$$\varphi : V \to W, \qquad \psi : U \to V$$

are linear maps, then their product (composition)

$$\varphi\psi : U \to W$$

is a linear map as well. Indeed,

$$\begin{aligned}(\varphi\psi)(a+b) &= \varphi(\psi(a+b)) = \varphi(\psi(a) + \psi(b)) \\ &= \varphi(\psi(a)) + \varphi(\psi(b)) = (\varphi\psi)(a) + (\varphi\psi)(b), \\ (\varphi\psi)(\lambda a) &= \varphi(\psi(\lambda a)) = \varphi(\lambda\psi(a)) = \lambda\varphi(\psi(a)) = \lambda(\varphi\psi)(a).\end{aligned}$$

Multiplication of linear maps is related to linear operations as follows:

$$\varphi(\psi + \omega) = \varphi\psi + \varphi\omega, \qquad (\varphi + \psi)\omega = \varphi\omega + \psi\omega,$$
$$(\lambda\varphi)\psi = \varphi(\lambda\psi) = \lambda(\varphi\psi) \qquad \forall \lambda \in K.$$

As an example, we prove the first distributive law. Let

$$\varphi : V \to W, \qquad \psi : U \to V, \quad \omega : U \to V$$

be linear maps. For any $a \in U$ we have

$$(\varphi(\psi+\omega))(a) = \varphi((\psi+\omega)(a)) = \varphi(\psi(a)+\omega(a))$$
$$= \varphi(\psi(a)) + \varphi(\omega(a)) = (\varphi\psi)(a) + (\varphi\omega)(a) = (\varphi\psi+\varphi\omega)(a).$$

Multiplication of linear maps is associative, just as multiplication of maps in general. Indeed, let M, N, P, Q be sets and

$$\varphi : P \to Q, \quad \psi : N \to P, \quad \omega : M \to N$$

arbitrary maps. Then for any $a \in M$ we have

$$((\varphi\psi)\omega)(a) = (\varphi\psi)(\omega(a)) = \varphi(\psi(\omega(a))),$$
$$(\varphi(\psi\omega))(a) = \varphi((\psi\omega)(a)) = \varphi(\psi(\omega(a))),$$

thus

$$(\varphi\psi)\omega = \varphi(\psi\omega).$$

Operations on linear maps of spaces of rows correspond to the same operations on matrices. This is clear for linear operations (addition and multiplication by numbers). To prove this for multiplication, let

$$\varphi\colon K^n \to K^m, \quad \psi\colon K^p \to K^n$$

be linear maps with matrices $A = (a_{ij})$ and $B = (b_{jk})$, respectively. Let $e_1, \ldots, e_p$ be unit rows of the space K^p. Then

$$\psi(e_k) = (b_{1k}, b_{2k}, \ldots, b_{nk}),$$

$$(\varphi\psi)(e_k) = \varphi(\psi(e_k)) = \left(\sum_j a_{1j}b_{jk}, \sum_j a_{2j}b_{jk}, \ldots, \sum_j a_{mj}b_{jk}\right).$$

Therefore, the matrix of the map $\varphi\psi$ is $C = (c_{ik})$, where

$$c_{ik} = \sum_j a_{ij}b_{jk}.$$

This means that $C = AB$ as claimed.

Example 2.71. In the language of linear maps, the matrix equality proved in Example 1.78 says that the product of rotations through angles α and β is the rotation through the angle $\alpha+\beta$ (see Example 2.57). As the latter statement is geometrically obvious, we thus proved the formulas for the sine and cosine of the sum of two angles.

The properties of matrix operations obtained in Section 1.9 by direct calculations can be deduced now from corresponding properties of operations on linear maps.

Obviously, the identity map

$$\mathrm{id} : V \to V$$

is linear. The matrix of the identity map $\mathrm{id} : K^n \to K^n$ is the identity matrix E of the nth order. Thus the properties of the identity matrix (1.10) are simply a translation into the matrix language of obvious equalities

$$\varphi \cdot \mathrm{id} = \varphi, \qquad \mathrm{id} \cdot \varphi = \varphi,$$

where $\varphi\colon K^n \to K^m$ is the linear map determined by a matrix A and "id" stands for the identity maps of spaces K^n and K^m in the first and second equalities, respectively.

Recall that a map is invertible if and only if it is bijective. If $\varphi : U \to V$ is a bijective linear map, then the inverse map $\varphi^{-1}V \to U$ is also linear. Indeed, for any $a, b \in V$, let $c, d \in U$ be vectors such that $\varphi(c) = a$, $\varphi(d) = b$. Then $\varphi(c + d) = a + b$, hence

$$\varphi^{-1}(a + b) = c + d = \varphi^{-1}(a) + \varphi^{-1}(b).$$

The second condition for linearity is checked in the same way.

Definition 2.72. An $n \times n$ square matrix A is called *nonsingular* if $\mathrm{rk}\, A = n$.

In other words, A is nonsingular if its rows (or columns) are linearly independent.

Theorem 2.73. *A square matrix is invertible if and only if it is nonsingular.*

Proof. Let $\varphi : K^n \to K^n$ be a linear map determined by a matrix A. According to the discussion above, A is invertible if and only if the map φ is bijective. By Theorem 2.62 this happens if and only if

$$\mathrm{Im}\,\varphi = K^n, \qquad \mathrm{Ker}\,\varphi = 0.$$

In view of Theorem 2.63 and Corollary 2.65 both of these conditions hold if and only if $\mathrm{rk}\, A = n$. □

Finding the inverse matrix of A is the same as solving the matrix equation

$$AX = E$$

(where X is an unknown square matrix). Such an equation can be solved just like equation (2.4) by multiplying it by elementary matrices on the left. This is equivalent to elementary row transformations of the "extended" matrix $(A|E)$. Reducing the left half of this matrix to the identity matrix (which is possible because A is nonsingular), we obtain the inverse matrix on the right.

Example 2.74. Here we find a matrix inverse of

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}.$$

For this, we perform the following elementary transformations:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -3 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & -5 & 2 \\ 0 & 1 & 3 & -1 \end{array}\right).$$

Thus,

$$A^{-1} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}.$$

Exercise 2.75. Using linear maps, prove that the rank of the product of two matrices (not necessarily square) does not exceed the rank of each of them. Also prove that if one of these matrices is nonsingular, then the rank of the product equals the rank of the other matrix.

2.4. Determinants

In the previous section we explained how to find out whether a matrix is nonsingular or, equivalently, if a system of n vectors in an n-dimensional space is linearly independent. In each particular case this question can be answered by reducing the matrix to step form by elementary row transformations. However, it is of interest to formulate a general condition for matrix entries that would tell us when this matrix is nonsingular. We will give an example of such a condition for geometric vectors.

A pair of noncollinear vectors $a_1, a_2 \in E^2$ is said to be *positively oriented* if the turn from a_1 to a_2 (through the angle less than π) is in the positive direction, i.e., counterclockwise. For any vectors a_1, a_2 consider the parallelogram with sides a_1, a_2. Denote by $\operatorname{area}(a_1, a_2)$ the oriented area of this parallelogram, i.e., its area taken with the positive sign if the pair $\{a_1, a_2\}$ is positively oriented, and with the negative sign otherwise. If the vectors a_1 and a_2 are collinear, put $\operatorname{area}(a_1, a_2) = 0$. The value of $|\operatorname{area}(a_1, a_2)|$ measures, in some sense, the degree of linear independence of a_1 and a_2.

The function $\operatorname{area}(a_1, a_2)$ with vector arguments a_1 and a_2 has the following properties:

(i) it is linear in a_1 and a_2 (see Example 2.67);

(ii) $\operatorname{area}(a_2, a_1) = -\operatorname{area}(a_1, a_2)$;

(iii) if $\{e_1, e_2\}$ is a positively oriented orthonormal basis, then $\operatorname{area}(e_1, e_2) = 1$.

The last two properties are obvious. To prove the first, consider the area of a parallelogram as the product of its base and height. Then,

$$\operatorname{area}(a_1, a_2) = |a_1| h_2,$$

where $|a_1|$ is the length of a_1 and h_2 is the signed length of the projection of a_2 onto the line orthogonal to a_1 (Figure 2.4). Since projection is a linear

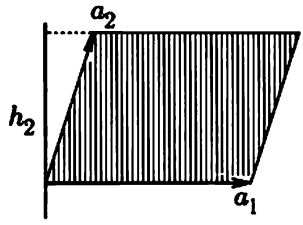


Figure 2.4

map, it follows that area(a_1, a_2) is linear in a_2. Similarly, if we choose a_2 as the base, we can prove that area(a_1, a_2) is linear in a_1.

Properties (i)–(iii) are sufficient to calculate area(a_1, a_2). Express vectors a_1, a_2 in the positively oriented orthonormal basis $\{e_1, e_2\}$:

$$a_1 = a_{11}e_1 + a_{12}e_2,$$
$$a_2 = a_{21}e_1 + a_{22}e_2.$$

Then

$$\begin{aligned}\text{area}(a_1, a_2) &= \text{area}(a_{11}e_1 + a_{12}e_2, a_{21}e_1 + a_{22}e_2)\\ &= a_{11}a_{21}\text{area}(e_1, e_1) + a_{11}a_{22}\text{area}(e_1, e_2) + a_{12}a_{21}\text{area}(e_2, e_1)\\ &\quad + a_{12}a_{22}\text{area}(e_2, e_2) = a_{11}a_{22} - a_{12}a_{21}.\end{aligned}$$

The expression $a_{11}a_{22} - a_{12}a_{21}$ is called the determinant of the matrix $A = (a_{ij})$ of order 2. The discussion above implies that vectors a_1 and a_2 are linearly independent if and only if the matrix composed of their coordinates has a nonzero determinant.

Similarly, it can be shown that the oriented volume vol(a_1, a_2, a_3) of a parallelepiped formed by vectors a_1, a_2, a_3 has the following properties:

(i) it is linear in each of its three arguments a_1, a_2, a_3;

(ii) it changes the sign when any two arguments are interchanged;

(iii) vol$(e_1, e_2, e_3) = 1$ for an arbitrary positively oriented orthonormal basis $\{e_1, e_2, e_3\}$.

(A triple $\{a_1, a_2, a_3\}$ is said to be positively oriented if the turn from a_1 to a_2 is in the positive direction if viewed from a_3's side.)

Using these properties, we can express vol(a_1, a_2, a_3) in terms of the coordinates of a_1, a_2, a_3 in a positively oriented orthonormal basis as follows

(try doing the calculations yourself!):

$$\begin{aligned}\mathrm{vol}(a_1, a_2, a_3) =& a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}.\end{aligned}$$

The expression on the right-hand side of this equality is called the determinant of a matrix $A = (a_{ij})$ of order 3. Thus, vectors a_1, a_2, a_3 are linearly independent if and only if the matrix composed of their coordinates has a nonzero determinant.

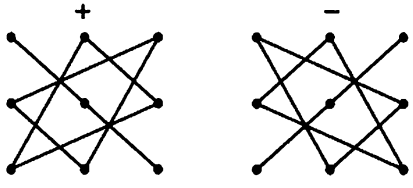


Figure 2.5

The determinant of a matrix $A = (a_{ij})$ of order 3 is an algebraic sum of all possible products of three entries of A, each chosen from a different row and different column. Two schemes in Figure 2.5 show which of these products are taken with the plus and with the minus signs.

The determinant of a matrix A is denoted either by $\det A$ or by the same matrix with parentheses replaced with vertical lines.

Example 2.76.

$$\begin{vmatrix}\cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha\end{vmatrix} = \cos^2\alpha + \sin^2\alpha = 1.$$

Example 2.77.

$$\begin{vmatrix}1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9\end{vmatrix} = \begin{aligned}&1\cdot5\cdot9 + 2\cdot6\cdot7 + 3\cdot4\cdot8 - 3\cdot5\cdot7 - 2\cdot4\cdot9 - 1\cdot6\cdot8 \\ &= 45 + 84 + 96 - 105 - 72 - 48 = 0.\end{aligned}$$

In the case of arbitrary dimension and arbitrary field, we do not have the notions such as area or volume. Hence it is natural to define the determinant as a function with properties similar to (i)–(iii). We begin by introducing necessary definitions.

Let V be a vector space over a field K and $f(a_1, a_2, \ldots, a_m)$ a function of m vectors in the space V that takes values in K.

Definition 2.78. The function $f(a_1, a_2, \ldots, a_m)$ is called *multilinear* (more precisely, *m-linear*) if it is linear in each argument.

For example, linearity in the first argument means that

$$f(a_1' + a_1'', a_2, \dots, a_m) = f(a_1', a_2, \dots, a_m) + f(a_1'', a_2, \dots, a_m),$$
$$f(\lambda a_1, a_2, \dots, a_m) = \lambda f(a_1, a_2, \dots, a_m).$$

Definition 2.79. A multilinear function $f(a_1, a_2, \dots, a_m)$ is said to be *skew-symmetric* if its value is multiplied by -1 when any two of its arguments are interchanged.

A skew-symmetric multilinear function has an important property: it equals zero whenever any two of its arguments take the same value (if $\operatorname{char} K \neq 2$). Indeed, when these two arguments are interchanged, the value of the function does not change, and yet it is multiplied by -1. Hence, it equals zero.

Remark 2.80. If $\operatorname{char} K = 2$, the last property should be taken as the definition of skew-symmetry. In fact, this property implies skew-symmetry as defined above. To prove this, notice that when checking if a function is skew-symmetric in any two of its arguments, the other arguments are fixed. Thus it suffices to consider the case of a bilinear (i.e., 2-linear) function. Let f be a bilinear function that becomes zero whenever the values of its arguments are equal. Then for any $a, b \in V$, we have

$$0 = f(a+b, a+b) = f(a,a) + f(a,b) + f(b,a) + f(b,b) = f(a,b) + f(b,a),$$

implying $f(b,a) = -f(a,b)$.

Now we will introduce notions necessary to define the explicit analytic expression of the determinant of a matrix of order n (similar to those that we obtained for $n = 2, 3$).

A sequence $(k_1, k_2, \dots, k_n)$ of numbers $1, 2, \dots, n$ taken in any order is called an *arrangement* of n elements. Notice that k_1 can assume n possible values; k_2, $n-1$ values if k_1 is fixed; k_3, $n-2$ values if k_1 and k_2 are fixed, etc. Hence, the total number of arrangements is

$$n(n-1)(n-2)\cdots 2 \cdot 1 = n!$$

The arrangement $(1, 2, \dots, n)$ is called *trivial*.

We say that a pair of numbers forms an *inversion* in a given arrangement if the greater of them stands to the left of the lesser. An arrangement is called *even* (respectively, *odd*) if it contains an even (respectively, odd) number of inversions. We also define the *sign* of an arrangement which we set equal to 1 if the arrangement is even and -1 if it is odd. The sign of an arrangement $(k_1, k_2, \dots, k_n)$ is denoted $\operatorname{sign}(k_1, k_2, \dots, k_n)$.

Example 2.81. For $n = 3$ the even arrangements are $(1,2,3)$ (no inversions), $(2,3,1)$ (two inversions), and $(3,1,2)$ (two inversions). The odd

arrangements are $(1,3,2)$ (one inversion), $(3,2,1)$ (three inversions), and $(2,1,3)$ (one inversion).

Example 2.82. The trivial arrangement does not contain any inversions and is thus even. Conversely, in the arrangement $(n, n-1, \dots, 2, 1)$ every pair forms an inversion. Therefore, the number of inversions in this arrangement equals

$$\binom{n}{2} = \frac{n(n-1)}{2} \equiv \left[\frac{n}{2}\right] \pmod 2.$$

Hence,

$$\operatorname{sign}(n, n-1, \dots, 2, 1) = (-1)^{n(n-1)/2} = (-1)^{[n/2]}.$$

The interchange of positions of two elements in an arrangement is called a *transposition* of these elements.

Proposition 2.83. *Any transposition changes the sign of an arrangement.*

Proof. When a transposition is applied to adjacent elements, only their relative position changes and the number of inversions decreases or increases by 1. Hence, in this case the sign changes. A transposition of elements i and j separated by s elements can be achieved by the application of $2s+1$ transpositions of adjacent elements: first interchange i with all intermediate elements and j, then interchange j with all intermediate elements. As we showed above, every time the sign of the arrangement will change; therefore, in the end it will be the opposite of the original. □

Corollary 2.84. *For $n > 1$, the number of even arrangements of n elements is equal to the number of odd arrangements.*

Proof. Write down all even arrangements and transpose the first two elements in each of them. We will obtain all odd arrangements, once each. □

Now we can state and prove the main theorem.

Theorem 2.85. *For any $c \in K$, there exists a unique skew-symmetric n-linear function f on the space K^n that satisfies the following condition:*

$$f(e_1, e_2, \dots, e_n) = c \tag{2.17}$$

(where $e_1, e_2, \dots, e_n$ are the unit rows in K^n). This function has the following form:

$$f(a_1, a_2, \dots, a_n) = c \sum_{(k_1, k_2, \dots, k_n)} \operatorname{sign}(k_1, k_2, \dots, k_n) a_{1k_1} a_{2k_2} \cdots a_{nk_n}, \tag{2.18}$$

where a_{ik} denotes the kth component of the row a_i and the sum is taken over all arrangements of n elements.

Proof. (i) Assume that f is a skew-symmetric n-linear function satisfying condition (2.17). Then

$$f(a_1, a_2, \ldots, a_n) = f\left(\sum_{k_1} a_{1k_1} e_{k_1}, \sum_{k_2} a_{2k_2} e_{k_2}, \ldots, \sum_{k_n} a_{nk_n} e_{k_n}\right)$$
$$= \sum_{k_1, k_2, \ldots, k_n} a_{1k_1} a_{2k_2} \cdots a_{nk_n} f(e_{k_1}, e_{k_2}, \ldots, e_{k_n}).$$

Since f is skew-symmetric, if any of the numbers $k_1, k_2, \ldots, k_n$ are equal, then $f(e_{k_1}, e_{k_2}, \ldots, e_{k_n}) = 0$. If they are all distinct, then

$$f(e_{k_1}, e_{k_2}, \ldots, e_{k_n}) = c\,\mathrm{sign}(k_1, k_2, \ldots, k_n).$$

Indeed, if this equality holds for some arrangement $(k_1, k_2, \ldots, k_n)$, then it holds for any arrangement obtained from $(k_1, k_2, \ldots, k_n)$ using a transposition, since under a transposition both sides of this equality are multiplied by -1. By condition (2.17) this equality holds for the trivial arrangement. But it is obvious that any arrangement can be obtained from the trivial one by a successive application of transpositions. Therefore, this equality holds for any arrangement and we obtain the expression (2.18) for $f(a_1, a_2, \ldots, a_n)$. We conclude that if there is a function f that satisfies all conditions of the theorem, then it has the form (2.18) and is thus unique.

(ii) We have to prove now that a function f determined by (2.18) is a multilinear skew-symmetric function satisfying condition (2.17). Linearity in each of the arguments is clear, since for any i formula (2.18) can be represented as

$$f(a_1, a_2, \ldots, a_n) = \sum_j a_{ij} u_j,$$

where $u_1, \ldots, u_n$ do not depend on a_i. Condition (2.17) is satisfied as well because in the expression for $f(e_1, e_2, \ldots, e_n)$ the summand corresponding to the trivial arrangement equals 1 and other summands are zero. It remains to check that f is skew-symmetric.

Consider what happens when the arguments a_i and a_j are interchanged. We can split the set of all arrangements into the pairs of arrangements obtained from each other by the transposition of k_i and k_j. According to Proposition 2.83, the products $a_{1k_1} a_{2k_2} \cdots a_{nk_n}$ corresponding to arrangements from such a pair appear in (2.18) with opposite signs. When a_i is interchanged with a_j, the products interchange too, hence all of the expression is multiplied by -1. □

Remark 2.86. If $\mathrm{char}\, K = 2$, skew-symmetry should be understood in the sense of Remark 2.80. The proof of the above theorem then says that

whenever $a_i = a_j$, the summands in (2.18) that correspond to arrangements from each of the above pairs cancel.

The function satisfying the conditions of Theorem 2.85 for c = 1 is denoted det.

Definition 2.87. The *determinant* of a square matrix $A = (a_{ij})$ of order n is the number

$$\det A = \det(a_1, a_2, \ldots, a_n),$$

where $a_1, a_2, \ldots, a_n$ are the rows of A.

Therefore,

$$\det A = \sum_{(k_1,k_2,\ldots,k_n)} \operatorname{sign}(k_1, k_2, \ldots, k_n) a_{1k_1} a_{2k_2} \cdots a_{nk_n}. \tag{2.19}$$

When $n = 2$ or 3, we recover the expressions provided in the first part of this section.

Similarly, by identifying each matrix with the collection of its rows, we can consider every function of n elements of K^n as a function of a square matrix of order n and vice versa.

The uniqueness condition from Theorem 2.85 can be now stated as follows:

Corollary 2.88. *If f is a skew-symmetric multilinear function of matrix rows, then*

$$f(A) = f(E) \det A. \tag{2.20}$$

When $n \geq 4$, it is rather difficult to calculate the determinant directly from (2.19). There exist much simpler ways to calculate determinants. They are based on determinants' properties that we will prove below.

Proposition 2.89. *The determinant of a matrix does not change under an elementary transformation of the first type.*

Proof. We add to the first row of A the second row multiplied by c. Denote the new matrix by A'. We have

$$\begin{aligned} \det A' &= \det(a_1 + ca_2, a_2, \ldots, a_n) \\ &= \det(a_1, a_2, \ldots, a_n) + \mathrm{c} \det(a_2, a_2, \ldots, a_n) = \det A. \end{aligned}$$

□

We know that when two rows are interchanged, the determinant is multiplied by -1. Also when a row is multiplied by a number, the determinant is multiplied by this number. Thus, we know how the determinant changes under any elementary row transformation of the given matrix. Since any

matrix can be reduced to step form and every square matrix in step form is triangular (but maybe not strictly triangular), it remains to figure out how to calculate the determinant of a triangular matrix.

Proposition 2.90. *The determinant of a triangular matrix equals the product of its diagonal entries.*

Proof. For any matrix, the product of diagonal entries is contained in (2.19) with the '+' sign because it corresponds to the trivial arrangement. When the matrix is triangular, all other summands in this expression are zero. Indeed, if $a_{1k_1}a_{2k_2}\cdots a_{nk_n} \neq 0$, then

$$k_1 \geq 1, \quad k_2 \geq 2, \quad \ldots, \quad k_n \geq n.$$

However, since

$$k_1 + k_2 + \cdots + k_n = 1 + 2 + \cdots + n,$$

this is possible only if

$$k_1 = 1, \quad k_2 = 2, \quad \ldots, \quad k_n = n.$$

□

Besides providing us with a practical method of calculating determinants, Propositions 2.89 and 2.90 allow us to answer the question which prompted us to introduce the determinant in the first place.

Theorem 2.91. *A square matrix A is nonsingular if and only if* $\det A \neq 0$.

Proof. Reduce A to step form by elementary row transformations. If at some point we used transformations of the second or third type, then the determinant could have changed but its equality or inequality to zero would have been preserved. A is nonsingular if and only if the matrix in step form is strictly triangular, but this is equivalent to its having a nonzero determinant. □

We continue studying properties of the determinant.

Theorem 2.92. $\det A^\top = \det A$.

Proof. Just like the determinant of A, the determinant of $A^\top$ is the algebraic sum of all possible products of n entries of A, one from each row and each column. So, we have only to check that the each product appears in the expressions for $\det A$ and $\det A^\top$ with the same sign.

Lemma 2.93. *Let $a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$ be a product of n entries of the matrix A, one from each row and each column. Then the product*

$$\operatorname{sign}(i_1, i_2, \ldots, i_n)\operatorname{sign}(j_1, j_2, \ldots, j_n)$$

does not change under the permutation of the factors in $a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$.

Proof of Lemma 2.93. Each permutation of factors in $a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n}$ is achieved by subsequently exchanging two neighboring factors. At each such exchange, each $\operatorname{sign}(i_1, i_2, \ldots, i_n)$ and $\operatorname{sign}(j_1, j_2, \ldots, j_n)$ changes to the opposite one, hence their product remains the same. □

We continue with the proof of Theorem 2.92. To find what sign with which a product $a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n}$ enters $\det A$, we must permute the factors ordering them by row numbers, i.e., write $a_{i_1 j_1} \cdots a_{i_n j_n} = a_{1 k_1} \cdots a_{n k_n}$, and then the sign in $\det A$ is $\operatorname{sign}(k_1, k_2, \ldots, k_n)$. To find the sign of the same product in $\det A^\top$, we must order its factors by their column numbers, i.e., write $a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n} = a_{l_1 1} a_{l_2 2} \cdots a_{l_n n}$, and then the sign in $\det A^\top$ is $\operatorname{sign}(l_1, l_2, \ldots, l_n)$. Since $\operatorname{sign}(1, 2, \ldots, n) = 1$, Lemma 2.93 shows that

$$\operatorname{sign}(k_1, k_2, \ldots, k_n) = \operatorname{sign}(l_1, l_2, \ldots, l_n).$$

This means that the product we are considering appears in $\det A$ and $\det A^\top$ with the same sign. □

It follows from Theorem 2.92 that every property of the determinant still holds if we replace rows by columns and columns by rows in its statement. In particular, we have

Corollary 2.94. *The determinant is a skew-symmetric multilinear function of matrix columns.*

Theorem 2.95. *Let matrix A be of the form*

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix},$$

where B and C are square matrices. Then

$$\det A = \det B \cdot \det C.$$

Proof. When B and D are fixed, the determinant of A is a skew-symmetric multilinear function of its lower rows, hence, a skew-symmetric multilinear function of rows of C. By Corollary 2.84, we have

$$\det A = \det \begin{pmatrix} B & D \\ 0 & E \end{pmatrix} \cdot \det C.$$

Moreover, when D is fixed, the first multiplier in the above expression is a skew-symmetric multilinear function of columns of B. Hence,

$$\det \begin{pmatrix} B & D \\ 0 & E \end{pmatrix} = \det \begin{pmatrix} E & D \\ 0 & E \end{pmatrix} \cdot \det B = \det B$$

(because the matrix $\left(\begin{smallmatrix} E & D \\ 0 & E \end{smallmatrix}\right)$ is triangular and has 1's on the diagonal). □

Due to Theorem 2.92, a similar formula holds for matrices with a zero upper right-hand corner.

Example 2.96. Here we will calculate the so-called *Vandermonde determinant*:

$$V(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix}.$$

Subtracting from each column (starting with the last one) the previous column multiplied by x_1 and applying Theorem 2.95, we have

$$\begin{aligned} V(x_1, x_2, \ldots, x_n) &= \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \ldots & x_2^{n-2}(x_2 - x_1) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \ldots & x_n^{n-2}(x_n - x_1) \end{vmatrix} \\ &= (x_2 - x_1) \cdots (x_n - x_1) V(x_2, \ldots, x_n). \end{aligned}$$

Continuing further, we finally obtain

$$V(x_1, x_2, \ldots, x_n) = \prod_{i>j} (x_i - x_j). \tag{2.21}$$

Let A be an arbitrary (not necessarily square) matrix. Any matrix formed by the entries of A whose positions are at the intersections of some selected rows and columns is called a *submatrix* of A. We remark that the selected rows and columns do not have to be adjacent.

The determinant of a square submatrix of order k is called a *minor* of order k of the matrix A. Sometimes we will abuse the language and call the square submatrix itself a minor. In particular, if A is a square matrix of order n, then its minor of order $n-1$ obtained by omitting the ith row and jth column is called the *complementary minor* of the entry a_{ij}. It is denoted M_{ij}. The number

$$A_{ij} = (-1)^{i+j} M_{ij}$$

is called the *cofactor* of the entry a_{ij}. Its meaning will become clear from the following lemma.

Lemma 2.97.

$$\begin{vmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ & & \ldots & & \\ 0 & \ldots & a_{ij} & \ldots & 0 \\ & & \ldots & & \\ a_{n1} & \ldots & a_{nj} & \ldots & a_{nn} \end{vmatrix} = a_{ij} A_{ij}.$$

(The left-hand side is the determinant of the matrix obtained from $A = (a_{ij})$ by replacing all elements of the ith row except for a_{ij} with zeros.)

Proof. Exchange the ith row with all rows above it and the jth column with all columns to its left. In doing this we will exchange rows $i-1$ times and columns $j-1$ times, thus the determinant will be multiplied by

$$(-1)^{i-1+j-1} = (-1)^{i+j}.$$

In the end, we will have the following determinant:

$$\begin{vmatrix} a_{ij} & 0 & \dots & 0 \\ a_{1j} & a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{nj} & a_{n1} & \dots & a_{nn} \end{vmatrix},$$

where the lower right-hand corner is the complementary minor of a_{ij}. By Theorem 2.95, this determinant equals $a_{ij}M_{ij}$. Taking into account the original sign change, we obtain the statement of the lemma. □

Theorem 2.98. *For any square matrix A,*

$$\det A = \sum_j a_{ij}A_{ij} = \sum_i a_{ij}A_{ij}.$$

The first of these formulas is called the *expansion of the determinant along the ith row* and the second the *expansion of the determinant along the jth column*.

Proof. Since every summand in expression (2.19) of $\det A$ contains exactly one element from the ith row, the previous lemma means that the sum of terms that contain a_{ij} equals $a_{ij}A_{ij}$. The formula for the expansion along a row follows. The formula for the expansion along a column is deduced in the same way. □

Remark 2.99. Signs $(-1)^{i+j}$, when put onto the matrix, form a checkerboard pattern and the main diagonal consists of pluses.

Example 2.100. We calculate the determinant Δ from Example 2.77 expanding it along the second row:

$$\Delta = -4\begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 5\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} - 6\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} = -4\cdot(-6) + 5\cdot(-12) - 6\cdot(-6) = 0.$$

Example 2.101. We will calculate the determinant of order n of the type

$$\Delta_n = \begin{vmatrix} 2 & 1 & 0 & \dots & 0 & 0 \\ 1 & 2 & 1 & \dots & 0 & 0 \\ 0 & 1 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 & 1 \\ 0 & 0 & 0 & \dots & 1 & 2 \end{vmatrix}.$$

Expanding along the first row and then expanding the second of the resulting determinants along the first column, we obtain

$$\Delta_n = 2\Delta_{n-1} - \begin{vmatrix} 1 & 1 & \dots & 0 & 0 \\ 0 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 2 & 1 \\ 0 & 0 & \dots & 1 & 2 \end{vmatrix} = 2\Delta_{n-1} - \Delta_{n-2},$$

thus

$$\Delta_n - \Delta_{n-1} = \Delta_{n-1} - \Delta_{n-2}.$$

This means that the sequence $(\Delta_1, \Delta_2, \Delta_3, \dots)$ is an arithmetic progression. Since $\Delta_1 = 2$, $\Delta_2 = 3$, its difference is 1 and

$$\Delta_n = n + 1.$$

Theorem 2.102. *For any two square matrices A, B,*

$$\det AB = \det A \cdot \det B.$$

Proof. It is easy to see that the rows $c_1, \dots, c_n$ of the matrix AB are obtained from the rows $a_1, \dots, a_n$ of the matrix A by multiplication by B:

$$c_i = a_i B, \qquad i = 1, \dots, n.$$

It follows that for a fixed matrix B, $\det AB$ is a skew-symmetric multilinear function of the rows of A. Indeed, let $a_1 = a_1' + a_1''$ for some rows a_1', a_1''. Then

$$\begin{aligned} \det(a_1 B, a_2 B, \dots, a_n B) &= \det((a_1' + a_1'')B, a_2 B, \dots, a_n B) \\ &= \det(a_1' B + a_1'' B, a_2 B, \dots, a_n B) \\ &= \det(a_1' B, a_2 B, \dots, a_n B) + \det(a_1'' B, a_2 B, \dots, a_n B). \end{aligned}$$

Other properties are checked similarly. Now, using Corollary 2.84, we have

$$\det AB = \det EB \cdot \det A = \det A \cdot \det B.$$

□

Example 2.103. Consider the parallelepiped formed by vectors $a_1, a_2, a_3 \in E^3$. We will express its nonoriented volume V in terms of the lengths $|a_1|, |a_2|, |a_3|$ of its edges and the angles

$$\alpha_1 = \widehat{a_2 a_3}, \quad \alpha_2 = \widehat{a_3 a_1}, \quad \alpha_3 = \widehat{a_1 a_2}$$

(see Figure 2.6).

Let $A = (a_{ij})$ be a matrix consisting of the coordinates of the vectors a_i in an orthonormal basis. We know (see the beginning of this section) that

$$V = \pm \det A.$$

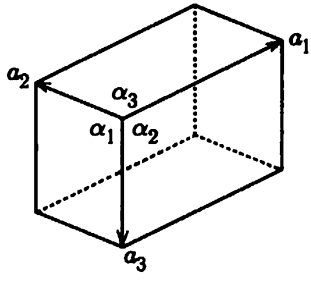


Figure 2.6

Thus

$$V^2 = (\det A)^2 = \det A \cdot \det A^{\mathsf{T}} = \det AA^{\mathsf{T}}.$$

The rules for matrix multiplication imply that the (i, j)th entry of the matrix AA^{T} is the inner product

$$(a_i, a_j) = |a_i||a_j| \cos \widehat{a_i a_j}.$$

Therefore,

$$V^2 = \begin{vmatrix} |a_1|^2 & |a_1||a_2| \cos \alpha_3 & |a_1||a_3| \cos \alpha_2 \\ |a_2||a_1| \cos \alpha_3 & |a_2|^2 & |a_2||a_3| \cos \alpha_1 \\ |a_3||a_1| \cos \alpha_2 & |a_3||a_2| \cos \alpha_1 & |a_3|^2 \end{vmatrix}$$

$$= |a_1|^2 |a_2|^2 |a_3|^2 \begin{vmatrix} 1 & \cos \alpha_3 & \cos \alpha_2 \\ \cos \alpha_3 & 1 & \cos \alpha_1 \\ \cos \alpha_2 & \cos \alpha_1 & 1 \end{vmatrix}$$

and we have

$$V = |a_1||a_2||a_3| \sqrt{1 + 2 \cos \alpha_1 \cos \alpha_2 \cos \alpha_3 - \cos^2 \alpha_1 - \cos^2 \alpha_2 - \cos^2 \alpha_3}.$$

2.5. Several Applications of Determinants

As we saw in the previous section (Theorem 2.91), determinants tell us whether a square matrix is nonsingular (and, hence, invertible); this is why we introduced them. Variations on this theme lead to various applications of determinants in the theory of linear equations and matrix theory. We consider several such applications in this section.

Consider a system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{2.22}$$

Denote by A its coefficient matrix and by A_i $(i = 1, 2, \ldots, n)$ the matrix obtained from A by replacing its ith column with the column of free terms.

Theorem 2.104. *If* $\det A \neq 0$, *then system (2.22) has a unique solution, which can be expressed by the formulas*

$$x_i = \frac{\det A_i}{\det A}, \qquad i = 1, 2, \ldots, n.$$

These formulas are called *Cramer's rules*.

Proof. Under any elementary transformation of system (2.22), the corresponding elementary row transformation acts on the matrices A and A_i $(i = 1, 2, \ldots, n)$. Hence the ratios on the right-hand sides of Cramer's rules do not change. By elementary row transformations, A can be reduced to the identity matrix. Therefore, it suffices to prove the theorem for the case of $A = E$.

If $A = E$, the system is of the form

$$\begin{cases} x_1 & & & = b_1, \\ & x_2 & & = b_2, \\ & & \ldots & \\ & & x_n & = b_n. \end{cases}$$

This system obviously has the unique solution $x_i = b_i$ $(i = 1, 2, \ldots, n)$. On the other hand,

$$\det A = \det E = 1, \qquad \det A_i = \begin{vmatrix} 1 & 0 & \ldots & b_1 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & b_2 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_i & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_{n-1} & \ldots & 1 & 0 \\ 0 & 0 & \ldots & b_n & \ldots & 0 & 1 \end{vmatrix} = b_i,$$

hence Cramer's rules are valid in this case. □

When $\det A = 0$, A is reduced to a matrix in step form which is not strictly triangular. If so, system (2.22) is either incompatible or underdetermined. In this case it is dangerous to interpret Cramer's rules in any way. They are simply inapplicable (and indeed, we obtained them assuming that $\det A \neq 0$), and we should look for other ways to solve the system.

Exercise 2.105. Prove that if $\det A = 0$ but $\det A_i \neq 0$ for some i, system (2.22) is incompatible.

Exercise 2.106. Show that if

$$\det A = \det A_1 = \cdots = \det A_n = 0,$$

then system (2.22) can be either underdetermined or incompatible. (Construct examples to show that either possibility might occur.)

We remark that Cramer's rules are not the best practical method for solving systems of linear equations, except maybe in the case $n = 2$. Their meaning is mostly theoretical. In particular, they allow us to obtain the following explicit formulas for the entries of the inverse matrix.

Theorem 2.107. *Let $A = (a_{ij})$ be a nonsingular square matrix. Then*

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}.$$

(Here A_{ij} stands for the cofactor of a_{ij}; see Section 2.4.)

Proof. The matrix A^{-1} is the solution of the matrix equation

$$AX = E.$$

This equation splits into n equations with respect to columns $X_1, X_2, \dots, X_n$ of the matrix X:

$$AX_j = E_j, \tag{2.23}$$

where E_j is the jth column of E.

In its coordinate form, equation (2.23) is a system of n linear equations with respect to entries $x_{1j}, x_{2j}, \dots, x_{nj}$ of the column X_j. Its coefficient matrix is A and the column of free terms, the column E_j. According to Cramer's rules,

$$x_{ij} = \frac{1}{\det A} \begin{vmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ & & \dots & & \\ a_{j1} & \dots & 1 & \dots & a_{jn} \\ & & \dots & & \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{vmatrix} = \frac{A_{ji}}{\det A}.$$

This completes the proof. □

Example 2.108. For a nonsingular matrix of order 2,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

we have

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

It would be helpful if you remember this simple formula.

Exercise 2.109. Let A be a nonsingular integer (= having integer entries) square matrix. Prove that the matrix A^{-1} is integer if and only if $\det A = \pm 1$.

Finally note that the rank of a matrix can also be found by calculating certain determinants.

Theorem 2.110. *The rank of a matrix equals the largest of the orders of its nonzero minors.*

Proof. Let the rank of a matrix A be r and let $s > r$. Then every s rows of A are linearly dependent and, moreover, so are the rows of any submatrix of A of order s that we obtain from the corresponding rows of A. Therefore, any minor of order s equals zero. Now, consider a submatrix formed by some r linearly independent rows of A. Its rank equals r too; hence it has r linearly independent columns. The minor of order r formed by these columns is nonzero. □

Exercise 2.111. Prove a stronger form of the above theorem: if a matrix A contains a nonzero minor of order r and all minors of order $r+1$ obtained from it by adding to this minor a row and a column are zero, then $\operatorname{rk} A = r$.

Exercise 2.112. Prove that in a matrix of rank r, any minor of rank r that is formed by taking intersection of r linearly independent rows with r linearly independent columns is nonzero.

Exercise 2.113. A *corner minor* of order k of a square matrix A is a determinant of a submatrix of order k in the upper left-hand corner of A. Prove that if all corner minors of A are nonzero, then we can reduce it to a triangular matrix by adding to each row a linear combination of preceding rows. Conclude from this that A can be uniquely presented as $A = UB$, where U is a lower triangular matrix with ones on the main diagonal and B is an upper triangular matrix.

Elements of Polynomial Algebra

3.1. Polynomial Algebra: Construction and Basic Properties

A function of a real variable x is called a *polynomial* if it can be presented as

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

where $a_0, a_1, a_2, \ldots, a_n$ are real numbers (some or even all of which can be equal to zero). It can be shown—we will do this below in a more general setup—that such a presentation is unique up to terms with zero coefficients, i.e., if

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n \quad \forall x \in \mathbb{R},$$

then $a_k = b_k$ for $k = 0, 1, 2, \ldots, n$.

It is obvious that the sum and the product of polynomials as well as the product of a polynomial and a number, are also polynomials. This means that polynomials form a subalgebra in the algebra of all functions of a real variable (see Example 1.72). This subalgebra is called the *polynomial algebra* over $\mathbb{R}$ and is denoted $\mathbb{R}[x]$.

It follows from the above discussion that the polynomials $1, x, x^2, \ldots$ comprise a basis of the algebra $\mathbb{R}[x]$. Its multiplication table looks quite simple:

$$x^k x^l = x^{k+l}.$$

We can try to treat polynomials over any field K in a similar way, but there is a natural difficulty in this approach: formally different polynomials can be equal for all values of the variable. For instance, polynomials x and

x^2 over $\mathbb{Z}_2$ assume the same value 0 for $x = 0$ and 1 for $x = 1$. All the same, we would like to view these polynomials as different. The solution lies in the formal definition which, in fact, identifies a polynomial with the sequence of its coefficients.

Consider the vector space K^∞ of finitary sequences of elements of a field K (for definition, see Example 2.43). We will start enumerating the terms of such sequences from 0. Also, let e_k $(k = 0, 1, 2, \dots)$ denote a sequence whose kth term is 1 and others are 0. Sequences $e_0, e_1, e_2, \dots$ comprise a basis of the space K^∞.

We can turn this space K^∞ into an algebra by providing the following rule for the multiplication of basis vectors:

$$e_k e_l = e_{k+l}.$$

Commutativity and associativity of integer addition imply that multiplication of basis vectors, hence of all elements in the algebra we have just obtained, is commutative and associative. The element e_0 is its identity. This algebra is called the *polynomial algebra* over K and is denoted $K[x]$ (but any letter can be used in place of x).

In order to go back to the usual presentation of polynomials, we agree, first, to identify elements of the type ae_0, $a \in K$, of the algebra $K[x]$ with the corresponding elements of the field K. Second, we denote the element e_1 as x (because the letter x was used in the notation for the algebra). Then, according to the definition of operations in $K[x]$, we see that $e_k = x^k$ and

$$\begin{aligned}(a_0, a_1, a_2, \dots, a_n, 0, \dots) &= a_0 e_0 + a_1 e_1 + a_2 e_2 + \cdots + a_n e_n \\ &= a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.\end{aligned}$$

The numbers $a_0, a_1, a_2, \dots$ are called the *coefficients* of the polynomial. The last nonzero coefficient is called the *leading coefficient* and its index, the *degree* of the polynomial. The degree of a polynomial f is denoted $\deg f$. The degree of the zero polynomial is not defined but sometimes it is useful to assume that it equals $-\infty$.

It is easy to see that

$$\deg(f + g) \le \max\{\deg f, \deg g\}, \tag{3.1}$$

$$\deg fg = \deg f + \deg g. \tag{3.2}$$

As an example, let us prove the latter equality. Let

$$\begin{aligned} f &= a_0 + a_1 x + \cdots + a_n x^n, \qquad a_n \neq 0, \\ g &= b_0 + b_1 x + \cdots + b_m x^m, \qquad b_m \neq 0. \end{aligned}$$

Then, when multiplying f by g, we get only one term of degree $n + m$, namely $a_n b_m x^{n+m}$, and no terms of a higher degree. As there are no zero

divisors in the field, $a_n b_m \neq 0$, hence

$$\deg fg = n + m = \deg f + \deg g.$$

The above discussion shows that there are no zero divisors in the algebra $K[x]$. It also implies that the only invertible elements of $K[x]$ are the polynomials of zero degree, i.e., nonzero elements of the field K.

Remark 3.1. It is customary to denote a polynomial as $f(x)$ or even f if it is clear from the context what letter stands for the variable.

Remark 3.2. It is often useful to order powers of x in a polynomial from the highest one down:

$$f = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n.$$

Remark 3.3. It is possible to consider any commutative associative ring with unity in place of K (see Remark 1.83). In such a case, all of the previous discussion remains valid except for the last part related to equality (3.2). There we must additionally request that K contain no zero divisors.

Remark 3.4. The product of two finitary sequences $(a_0, a_1, a_2, \dots)$ and $(b_0, b_1, b_2, \dots)$ in the ring $K[x]$ is a sequence $(c_0, c_1, c_2, \dots)$ whose terms are determined by the following formulas:

$$c_k = \sum_{l=0}^{k} a_l b_{k-l}.$$

These formulas make sense for any two arbitrary (not necessarily finitary) sequences. In this way, we obtain a commutative associative algebra with unity called *the algebra of formal power series* over K. It is denoted $K[[x]]$. Its elements are usually written as formal infinite sums of the form

$$a_0 + a_1 x + a_2 x^2 + \cdots.$$

Just as $K[x]$, the algebra $K[[x]]$ has no zero divisors; however, the proof is different (try to find it!).

Every polynomial

$$f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \tag{3.3}$$

determines a K-valued function on K, whose value at $c \in K$ is by definition equal to

$$f(c) = a_0 + a_1 c + a_2 c^2 + \cdots + a_n c^n.$$

The sum and product of polynomials (and hence the product of a polynomial and a number) can be rewritten in a canonical form (3.3) by applying transformations that use only the properties of operations on $K[x]$. The same properties are valid in K as well, thus we will obtain the same result

whether we replace $x = c$ before or after these transformations. This means that

$$(f+g)(c) = f(c) + g(c), \quad (fg)(c) = f(c)g(c), \quad (\lambda f)(c) = \lambda f(c),$$

i.e., operations over polynomials and the same operations over the corresponding functions lead to equal answers.

As we showed in the beginning of this section, different polynomials can sometimes determine the same function. However, this happens only if the field K is finite.

Theorem 3.5. *If the field K is infinite, different polynomials over K determine different functions.*

Proof. Choose polynomials $f, g \in K[x]$ that determine the same function. Then their difference $h = f - g$ determines the zero function, i.e., $h(c) = 0$ for all $c \in K$. Assume $h \neq 0$ and let

$$h = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}, \qquad a_{n-1} \neq 0.$$

Consider distinct $x_1, x_2, \ldots, x_n \in K$ (here we use the fact that K is infinite). Regard the following collection of equalities:

$$\begin{cases} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{n-1} x_1^{n-1} = 0, \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{n-1} x_2^{n-1} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_{n-1} x_n^{n-1} = 0, \end{cases}$$

as a (square) system of homogeneous linear equations with respect to $a_0, a_1, a_2, \ldots, a_{n-1}$. The determinant of the coefficient matrix of this system is the Vandermonde determinant $V(x_1, x_2, \ldots, x_n)$ (Example 2.96), hence it is nonzero. Therefore, this system has only the zero solution, and this contradicts our assumption. □

Remark 3.6. Even when the field K is finite, the set of all polynomials over K is infinite (but countable). However, the set of all K-valued functions over K is finite. Thus, there must exist polynomials that determine the same function. Nonetheless, Theorem 3.5 and its proof remain valid for polynomials whose degree is less than the number of elements of K.

Exercise 3.7. The so-called *interpolation problem* consists in finding a polynomial of degree $< n$ that assumes given values $y_1, y_2, \ldots, y_n \in K$ at given (distinct) points $x_1, x_2, \ldots, x_n \in K$. (In particular, when $n = 2$, this is called linear interpolation.) Prove that the interpolation problem has a unique solution for any $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$.

In the standard meaning of the word, it is usually impossible to divide one polynomial by another in the algebra $K[x]$. However, the so-called division with a remainder is possible, just as the similar procedure of division with a remainder in the ring of integers.

Theorem 3.8. *Let $f, g \in K[x]$ and $g \neq 0$. Then there exist polynomials q and r such that $f = qg + r$ and either $r = 0$ or $\deg r < \deg g$. Moreover, polynomials q and r are uniquely determined by the above conditions.*

Assuming that $\deg 0 = -\infty$, we can write $\deg r < \deg q$ in all the cases.

To find such polynomials q and r means exactly *to divide f by g with a remainder*. Here q is called the *incomplete quotient* and r the *remainder* of the division of f by g. A polynomial f is divisible by g in the algebra $K[x]$ if and only if $r = 0$.

Proof. (i) First we prove that it is always possible to divide with a remainder. If $\deg f < \deg g$, then we can take $q = 0$, $r = f$. If $\deg f \geq \deg g$, then q and r can be found via the standard procedure of "long division." Namely, let

$$f = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n,$$
$$g = b_0x^m + b_1x^{m-1} + \cdots + b_{m-1}x + b_m,$$

where $a_0, b_0 \neq 0$. Consider the polynomial

$$f_1 = f - \frac{a_0}{b_0}x^{n-m}g.$$

Its degree is less than that of f. If $\deg f_1 < \deg g$, we can take

$$q = \frac{a_0}{b_0}x^{n-m}, \quad r = f_1.$$

Otherwise, we deal with f_1 as we have just dealt with f. Finally, we obtain a polynomial

$$q = c_0x^{n-m} + c_1x^{n-m-1} + \cdots + c_{n-m}$$

such that $\deg(f - qg) < \deg g$. This is the incomplete quotient of the division of f by g, and the polynomial $r = f - qg$ is the remainder. (ii) Now we have to prove that the polynomials q and r are uniquely determined by the assumptions of the theorem. Let

$$f = q_1g + r_1 = q_2g + r_2,$$

where $\deg r_1 < \deg g$ and $\deg r_2 < \deg g$. Then

$$r_1 - r_2 = (q_2 - q_1)g.$$

Thus, assuming $q_1 \neq q_2$,

$$\deg(r_1 - r_2) = \deg(q_2 - q_1) + \deg g \geq \deg g,$$

which is a contradiction. Hence, $q_1 = q_2$ and $r_1 = r_2$. □

Division with a remainder by a linear binomial $x - c$ has a particular meaning. In this case the remainder has degree < 1, i.e., it lies in the field K. Therefore, when dividing with a remainder a polynomial f by $x - c$, we obtain an expression of the form

$$f(x) = (x - c)q(x) + r, \qquad r \in K.$$

This implies that

$$f(c) = r,$$

i.e., that the remainder equals the value of the polynomial f at the point c. This statement is known as *Bezout's Theorem*.

There exists a simple algorithm for division with a remainder by $x - c$, called *Horner's Scheme*.

Namely, let

$$\begin{aligned} a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \\ = (x - c)(b_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-2}x + b_{n-1}) + r. \end{aligned}$$

By comparing the coefficients at equal powers of x, we obtain the following chain of equalities:

$$\begin{aligned} a_0 &= b_0, \\ a_1 &= b_1 - cb_0, \\ a_2 &= b_2 - cb_1, \\ &\cdots\cdots\cdots\cdots \\ a_{n-1} &= b_{n-1} - cb_{n-2}, \\ a_n &= r - cb_{n-1}. \end{aligned}$$

From this we obtain the following recursive expressions for the coefficients $b_0, b_1, \ldots, b_{n-1}$ and r:

$$\begin{aligned} b_0 &= a_0, \\ b_1 &= a_1 + cb_0, \\ b_2 &= a_2 + cb_1, \\ &\cdots\cdots\cdots\cdots \\ b_{n-1} &= a_{n-1} + cb_{n-2}, \\ r &= a_n + cb_{n-1}. \end{aligned}$$

It is useful to record the original data and the results of all calculations in a table like this:

	a_0	a_1	a_2	$\ldots$	a_{n-1}	a_n
c	b_0	b_1	b_2	$\ldots$	b_{n-1}	r

Starting with b_1, every number in the second line of the table can be

found as the sum of the number above it and the number to the left of it multiplied by c.

Example 3.9. We determine here the value of the polynomial

$$f = 2x^6 - 11x^4 - 19x^3 - 7x^2 + 8x + 5$$

at $x = 3$. The Horner's Scheme gives us

	2	0	−11	−19	−7	8	5
3	2	6	7	2	−1	5	20

Thus, $f(3) = 20$.

3.2. Roots of Polynomials: General Properties

An element c of a field K is called a *root* of a polynomial $f \in K[x]$ (or of the corresponding *algebraic equation* $f(x) = 0$) if $f(c) = 0$. Bezout's theorem (see the previous section) implies

Theorem 3.10. *An element c of a field K is a root of a polynomial $f \in K[x]$ if and only if f is divisible by $x - c$.*

We can use this fact to prove the following theorem.

Theorem 3.11. *The number of roots of a nonzero polynomial does not exceed its degree.*

Proof. Let c_1 be a root of our polynomial f. Then

$$f = (x - c_1)f_1, \qquad f_1 \in K[x].$$

Let c_2 be a root of the polynomial f_1. Then

$$f_1 = (x - c_2)f_2, \qquad f_2 \in K[x],$$

hence,

$$f = (x - c_1)(x - c_2)f_2.$$

Continuing further, we finally obtain the following presentation of f:

$$f = (x - c_1)(x - c_2)\cdots(x - c_m)g, \tag{3.4}$$

where $g \in K[x]$ has no roots. Numbers $c_1, c_2, \ldots, c_m$ are all the roots of the polynomial f. Indeed, for any $c \in K$, we have

$$f(c) = (c - c_1)(c - c_2)\cdots(c - c_m)g(c),$$

and as $g(c) \neq 0$, $f(c) = 0$ only if $c = c_i$ for some i. Therefore, the number of roots of f does not exceed m (it can be less than m since some of the roots $c_1, c_2, \ldots, c_m$ may coincide). However,

$$m = \deg f - \deg g \leq \deg f.$$

□

Remark 3.12. We have actually proved this theorem while proving Theorem 3.5. On the other hand, we can deduce Theorem 3.5 from this one without using the theory of linear equations. Namely, if distinct polynomials f and g over an infinite field K define the same function, then all elements of K are the roots of the nonzero polynomial $h = f - g$. This contradicts Theorem 3.11.

The proof of Theorem 3.11 suggests that some roots should be counted several times. We state this idea more rigorously.

A root c of a polynomial f is called *simple* if f is not divisible by $(x-c)^2$, and *multiple* otherwise. The *multiplicity* of a root c is the maximum k such that f is divisible by $(x-c)^k$. Thus, a simple root is a root of multiplicity 1. Sometimes it is useful to assume that a number which is not a root of a given polynomial has multiplicity 0.

Clearly, c is a root of a polynomial f of multiplicity k if and only if

$$f = (x-c)^k g, \tag{3.5}$$

where $g(c) \neq 0$.

Now we can prove a refined version of Theorem 3.11.

Theorem 3.13. *The number of roots of a polynomial, counted with multiplicities (i.e., a root of multiplicity k is counted k times), does not exceed the degree of this polynomial. Moreover, these numbers are equal if and only if the polynomial is a product of linear factors.*

Proof. We can rewrite (3.4) by grouping together the same factors:

$$f = (x-c_1)^{k_1}(x-c_2)^{k_2}\dots(x-c_s)^{k_s}g \tag{3.6}$$

for distinct $c_1, c_2, \dots, c_s$. Clearly $c_1, c_2, \dots, c_s$ are all the roots of the polynomial f. Then, by distinguishing the factor $(x-c_i)^{k_i}$ in (3.6), we can write

$$f = (x-c_i)^{k_i}h_i, \qquad \text{where} \quad h_i(c_i) \neq 0.$$

Hence, c_i is a root of multiplicity k_i.

Therefore, the number of roots of f counted with multiplicities equals

$$k_1 + k_2 + \dots + k_s = \deg f - \deg g,$$

which implies the theorem. □

Remark 3.14. It is assumed that a polynomial of zero degree factors into a product of zero linear terms.

If a polynomial

$$f = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$$

factors into linear terms, then it can be written as

$$f = a_0(x - c_1)(x - c_2)\cdots(x - c_n),$$

where $c_1, c_2, \ldots, c_n$ are the roots of f. Moreover, the number of occurrences of each root in this expression equals its multiplicity. By comparing the coefficients of the corresponding powers of x in these two presentations of a polynomial f, we obtain the following *Viète's formulas*:

$$\begin{aligned}
&c_1 + c_2 + \cdots + c_n = -\frac{a_1}{a_0},\\
&c_1c_2 + c_1c_3 + \cdots + c_{n-1}c_n = \frac{a_2}{a_0},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\sum_{i_1<i_2<\cdots<i_k} c_{i_1}c_{i_2}\cdots c_{i_k} = (-1)^k\frac{a_k}{a_0},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&c_1c_2\cdots c_n = (-1)^n\frac{a_n}{a_0}.
\end{aligned}$$

The left-hand side of the kth Viète's formula contains the sum of all products of k roots of the polynomial f. Up to multiplication by -1, this is the coefficient of x^{n-k} in the product $(x - c_1)(x - c_2)\cdots(x - c_n)$.

Example 3.15. The complex roots of 1 of degree 5,

$$\varepsilon_k = \cos\frac{2\pi k}{5} + i\sin\frac{2\pi k}{5}, \qquad k = 0, 1, 2, 3, 4$$

(see also Figure 3.1), are the roots of the polynomial $x^5 - 1$.

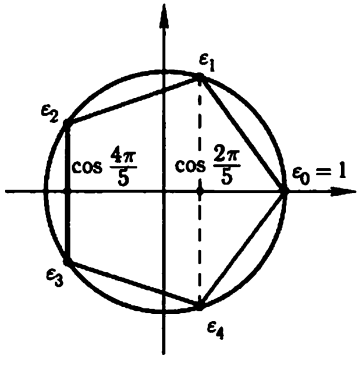


Figure 3.1

According to the first Viète's formula, their sum is 0. Hence the sum of their real parts is also 0:

$$2\cos\frac{4\pi}{5} + 2\cos\frac{2\pi}{5} + 1 = 0.$$

Put $\cos\frac{2\pi}{5} = x$. Then $\cos\frac{4\pi}{5} = 2x^2 - 1$ and we have

$$4x^2 + 2x - 1 = 0.$$

This implies

$$\cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}, \quad \cos\frac{4\pi}{5} = -\frac{\sqrt{5}+1}{4}.$$

Exercise 3.16. Let n be a prime number. Use Exercise 1.51 and the last of Viète's formulas to prove *Wilson's Theorem*:

$$(n-1)! \equiv -1 \pmod{n}.$$

A polynomial f is called *monic* if $a_0 = 1$. Viète's formulas express the coefficients of a monic polynomial in terms of its roots (whenever the number of roots counted with multiplicities equals the degree of the polynomial).

Example 3.17. We can explicitly write out the monic polynomial of degree 4,

$$f = x^4 + a_1x^3 + a_2x^2 + a_3x + a_4,$$

which has the root 1 of multiplicity 2 and the simple roots 2, 3. By Viète's formulas,

$$\begin{aligned} -a_1 &= 1+1+2+3 = 7, \\ a_2 &= 1\cdot1+1\cdot2+1\cdot3+1\cdot2+1\cdot3+2\cdot3 = 17, \\ -a_3 &= 1\cdot1\cdot2+1\cdot1\cdot3+1\cdot2\cdot3+1\cdot2\cdot3 = 17, \\ a_4 &= 1\cdot1\cdot2\cdot3 = 6. \end{aligned}$$

Thus,

$$f = x^4 - 7x^3 + 17x^2 - 17x + 6.$$

There exists another interpretation of the multiplicity of a root, at least in the case $\operatorname{char} K = 0$. To state it, we need to introduce differentiation of polynomials.

The rules for differentiation of functions of a real variable imply that a derivative of a polynomial is also a polynomial. Denote by D a map from $\mathbb{R}[x]$ into itself that assigns to each polynomial its derivative. The map D has the following properties:

(i) it is linear;

(ii) $D(fg) = (Df)g + f(Dg)$;

(iii) $Dx = 1$.

This suggests how me may define differentiation of polynomials over any field K, in particular when the definition from analysis does not make sense.

Proposition 3.18. *There exists a unique map $D\colon K[x] \to K[x]$ satisfying properties (i)–(iii).*

Proof. Let D be such a map. Then

$$D1 = D(1 \cdot 1) = (D1) \cdot 1 + 1 \cdot (D1) = D1 + D1,$$

thus $D1 = 0$. We will prove by induction that $Dx^n = nx^{n-1}$. For $n = 1$ this is true by assumption. The transition from $n-1$ to n consists in the following calculation:

$$Dx^n = D(x^{n-1}x) = (Dx^{n-1})x + x^{n-1}(Dx) = (n-1)x^{n-2} \cdot x + x^{n-1} = nx^{n-1}.$$

This means that the map D is uniquely determined for the basis vectors $1, x, x^2, \ldots$, hence on the whole space $K[x]$ as well.

On the other hand, we can construct a linear map $D : K[x] \to K[x]$ by defining it on the basis vectors by the following formulas:

$$D1 = 0, \quad Dx^n = nx^{n-1}, \qquad n = 1, 2, \ldots.$$

It remains to check that this map satisfies property (ii). By linearity it suffices to check this property for basis vectors only. We have

$$D(x^m x^n) = Dx^{m+n} = (m+n)x^{m+n-1},$$
$$(Dx^m)x^n + x^m(Dx^n) = mx^{m-1}x^n + nx^m x^{n-1} = (m+n)x^{m+n-1}.$$

□

The polynomial Df is called the *derivative* of f and is denoted f' as usual.

We can make a substitution $x = \mathrm{c} + y$ in a polynomial $f \in K[x]$ and then express it as a polynomial in $y = x - c$ (of the same degree):

$$f = b_0 + b_1(x - \mathrm{c}) + b_2(x - \mathrm{c})^2 + \cdots + b_n(x - \mathrm{c})^n. \tag{3.7}$$

Obviously, if c is a root of the polynomial f, its multiplicity equals the exponent of the first nonzero term in this presentation.

Proposition 3.19. *If* $\operatorname{char} K = 0$, *then the coefficients of $f \in K[x]$ regarded as a polynomial in $x - c$ are*

$$b_k = \frac{f^{(k)}(\mathrm{c})}{k!}.$$

(Here as usual $f^{(k)}$ stands for the kth derivative of f.)

Proof. Differentiate equality (3.7) k times and substitute $x = \mathrm{c}$. □

Therefore,

$$f = f(c) + \frac{f'(c)}{1!}(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n.$$

This expression is called *Taylor's formula* for polynomials.

Taylor's formula, together with the previous discussion, implies

Theorem 3.20. *When* $\operatorname{char} K = 0$, *the multiplicity of a root* c *of a polynomial* $f \in K[x]$ *equals the least order of the derivative of* f *that is nonzero at* c.

Corollary 3.21. *Under the same condition, every root of multiplicity* k *of a polynomial* f *is a root of multiplicity* $k-1$ *of its derivative.*

Remark 3.22. When $\operatorname{char} K > 0$, the multiplicity of c can be less than the number in Theorem 3.20. Actually, this number might not exist at all. For instance, if n is a prime number, the first (and thus all other) derivative of the polynomial $x^n \in \mathbb{Z}_n[x]$ is zero, yet this polynomial has root 0 of multiplicity n.

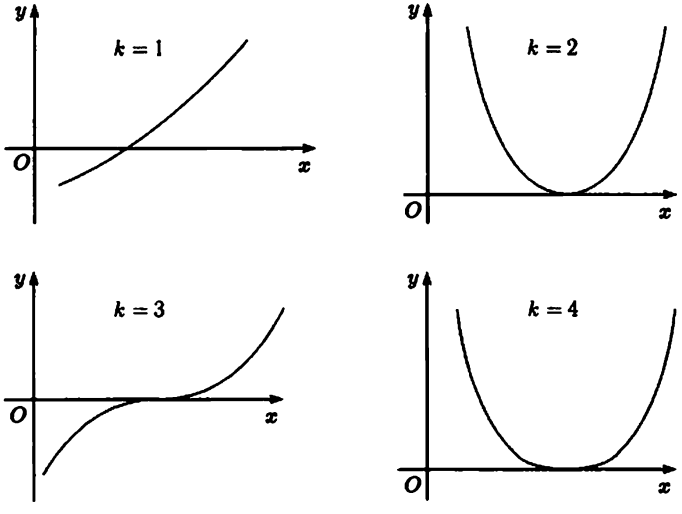


Figure 3.2

In the case $K = \mathbb{R}$, Theorem 3.20 provides a geometric interpretation of multiplicity. Namely, if the multiplicity of a root c of a polynomial $f \in K[x]$ equals k, then f behaves like $b(x-c)^k$, $b \neq 0$, in a neighborhood of c. This means that when $k = 1$, the graph of f simply crosses the x-axis at c, and when $k > 1$, the graph and the axis are tangent of order $k-1$. Moreover,

the sign of $f(x)$ changes when passing c for k odd and does not change for k even (Figure 3.2).

When $\operatorname{char} K = 0$, the coefficients in (3.7), and hence the values of derivatives of f at c, can be found by successive divisions with a remainder of f by $x - c$. In particular, after the first division, we obtain the incomplete quotient b_0 and the remainder

$$f_1 = b_1 + b_2(x - c) + \cdots + b_n(x - c)^{n-1};$$

after dividing f_1 by $x - c$ we obtain the remainder b_1; etc.

Example 3.23. We will express the polynomial

$$f = x^5 - 5x^4 + 7x^3 - 2x^2 + 4x - 8 \in \mathbb{R}[x]$$

as a polynomial in the powers of $x - 2$. For this, we perform consecutive divisions with remainder by $x - 2$ using Horner's Scheme. The result of each division will be used as the top row for the next one:

	1	−5	7	−2	4	−8
2	1	−3	1	0	4	0
	1	−1	−1	−2	0	
	1	1	1	0		
	1	3	7			
	1	5				
	1					

Thus,

$$f = 7(x - 2)^3 + 5(x - 2)^4 + (x - 2)^5.$$

In particular, we see that as a root of f, 2 has multiplicity 3. Moreover,

$$f'''(2) = 3! \cdot 7 = 42, \quad f^{(4)}(2) = 4! \cdot 5 = 120, \quad f^{(5)}(2) = 5! \cdot 1 = 120.$$

3.3. Fundamental Theorem of Algebra of Complex Numbers

In the previous section we obtained the upper bound for the number of the roots of a polynomial. However, the theory we developed does not help us to determine whether this polynomial has any roots at all. Indeed, there exist polynomials of positive degree that have no roots, e.g., the polynomial x^2+1 over the field $\mathbb{R}$ of real numbers. This situation is the reason behind the construction of the field $\mathbb{C}$ of complex numbers. If there were polynomials of positive degree over $\mathbb{C}$ having no roots, we would have needed to construct further extensions, but fortunately, this is not so. This fact is actually a theorem called the *Fundamental Theorem of Algebra of Complex Numbers*.

Theorem 3.24. *Every polynomial of positive degree over the field of complex numbers has a root.*

A field over which every polynomial of positive degree has at least one root is called *algebraically closed*. Thus, Theorem 3.24 says that the field $\mathbb{C}$ of complex numbers is algebraically closed.

There exist several proofs of this theorem. Each of them involves some analysis as it must somehow use the definition of the field of real numbers, which is not purely algebraic. The proof provided below is almost completely analytical.

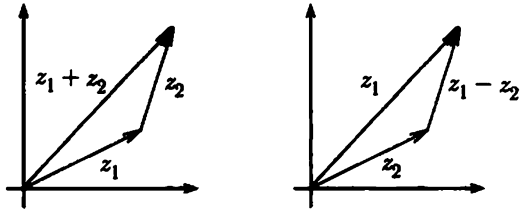


Figure 3.3

We need to introduce the notion of the limit of a sequence of complex numbers. Before we do this, recall that the absolute value $|z|$ of a complex number z is the length of a vector representing this number. It follows that $|z_1 - z_2|$ is the distance between the points representing the numbers z_1 and z_2. Geometry (Figure 3.3) implies that

$$|z_1 + z_2| \le |z_1| + |z_2|, \qquad ||z_1| - |z_2|| \le |z_1 - z_2|.$$

(Equality is attained when the corresponding triangle degenerates into a line segment.)

Definition 3.25. A sequence of complex numbers z_k, $k \in \mathbb{N}$, converges to a complex number z (notation: $z_k \to z$) if $|z_k - z| \to 0$.

Lemma 3.26. *Let* $z_k = x_k + y_k i$, $z = x + yi$ $(x_k, y_k, x, y \in \mathbb{R})$. *Then*

$$z_k \to z \iff x_k \to x \text{ and } y_k \to y.$$

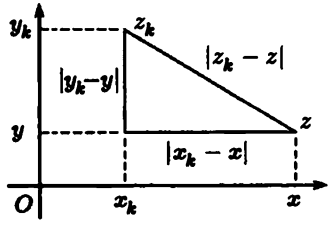


Figure 3.4

Proof. Since (Figure 3.4)

$$|z_k - z| = \sqrt{|x_k - x|^2 + |y_k - y|^2},$$

we have

$$x_k \to x \text{ and } y_k \to y \implies z_k \to z.$$

The converse follows from the inequalities

$$|x_k - x| \le |z_k - z|, \qquad |y_k - y| \le |z_k - z|.$$

□

Lemma 3.27. $z_k \to z \implies |z_k| \to |z|$.

Proof. This is a consequence of

$$||z_k| - |z|| \le |z_k - z|.$$

□

Lemma 3.28. *$z_k \to z$ and $w_k \to w \implies z_k + w_k \to z + w$ and $z_k w_k \to zw$.*

Proof. Same as for real number sequences:

$$|(z_k+w_k)-(z+w)| = |(z_k-z)+(w_k-w)| \le |z_k-z|+|w_k-w| \to 0,$$
$$|z_k w_k - zw| = |(z_k-z)w_k + z(w_k-w)| \le |z_k-z||w_k|+|z||w_k-w| \to 0.$$

□

Corollary 3.29. *Let $z_k \to z$ and let $f \in \mathbb{C}[z]$ be a polynomial. Then $f(z_k) \to f(z)$.*

(Here we follow the standard convention from analysis: the variable and its value are denoted by the same symbol.)

Lemma 3.30. *If $|z_k| \to \infty$ and $f \in \mathbb{C}[z]$ is a polynomial of positive degree, then $|f(z_k)| \to \infty$.*

Proof. Let

$$f = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n, \qquad a_0 \neq 0.$$

Then

$$|f(z_k)| = |z_k|^n \left| a_0 + \frac{a_1}{z_k} + \cdots + \frac{a_{n-1}}{z_k^{n-1}} + \frac{a_n}{z_k^n} \right|$$
$$\ge |z_k|^n \left(|a_0| - \frac{|a_1|}{|z_k|} - \cdots - \frac{|a_{n-1}|}{|z_k|^{n-1}} - \frac{|a_n|}{|z_k|^n} \right).$$

The expression in parentheses tends to $|a_0|$. Therefore, the whole product tends to ∞, and so does $|f(z_k)|$. □

The following lemma is crucial for the proof of the main theorem.

Lemma 3.31 (D'Alembert's Lemma). *Let $f \in \mathbb{C}[z]$ be a polynomial of positive degree and $f(z_0) \neq 0$. Then in any neighborhood of z_0, there exists z such that $|f(z)| < |f(z_0)|$.*

Proof. Express f as a polynomial in $z - z_0$ and divide it by $f(z_0)$. In this expression, some coefficients that immediately follow the free term can be zero, so in general we have

$$\frac{f(z)}{f(z_0)} = 1 + c_p(z-z_0)^p + c_{p+1}(z-z_0)^{p+1} + \cdots + c_n(z-z_0)^n \quad (c_p \neq 0). \tag{3.8}$$

We need to prove that there exists a z such that

$$\left|\frac{f(z)}{f(z_0)}\right| < 1.$$

The idea of the proof is that we choose z very close to z_0; whether the inequality holds or not, will now depend only on the first two terms in (3.8).

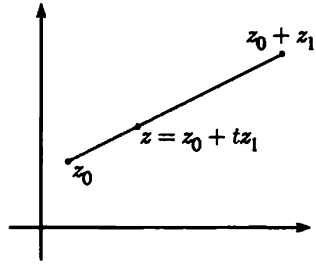


Figure 3.5

Let us look for z of the form

$$z = z_0 + tz_1$$

(Figure 3.5), where $t \in (0,1)$ and z_1 is a complex number satisfying the condition $c_p z_1^p = -1$.

We now have

$$\frac{f(z)}{f(z_0)} = 1 - t^p + t^{p+1}\varphi(t),$$

where φ is a polynomial of degree $n - p - 1$ (with complex coefficients). If C is the maximum absolute value of coefficients of φ, then

$$|\varphi(t)| \leq A = (n-p)C.$$

Hence

$$\left|\frac{f(z)}{f(z_0)}\right| \leq 1 - t^p + At^{p+1} = 1 - t^p(1 - At) < 1$$

for $t < \frac{1}{A}$. □

Proof of Theorem 3.24. Let $f \in \mathbb{C}[z]$ be a polynomial of positive degree. Put

$$M = \inf_z |f(z)|.$$

By the definition of the greatest lower bound, there exists a sequence of complex numbers z_k such that

$$|f(z_k)| \to M. \tag{3.9}$$

If the sequence $|z_k|$ is unbounded, it contains a subsequence that tends to infinity. This contradicts (3.9) by Lemma 3.30.

Therefore, there exists $C > 0$ such that

$$|z_k| \le C \qquad \forall k.$$

Consider the algebraic form of z_k:

$$z_k = x_k + y_k i.$$

Then

$$|x_k| \le |z_k| \le C, \qquad |y_k| \le |z_k| \le C.$$

By the Bolzano–Weierstrass theorem, the sequence x_k contains a converging subsequence. Passing to this subsequence and changing notation, we can assume that

$$x_k \to x_0.$$

Similarly, by passing to another subsequence, we can assume that

$$y_k \to y_0.$$

Then by Lemma 3.26,

$$z_k \to z_0 = x_0 + y_0 i,$$

hence

$$|f(z_k)| \to |f(z_0)| = M.$$

If $M > 0$, this contradicts the definition of M by D'Alembert's lemma. Therefore, $M = 0$, i.e., $f(z_0) = 0$. □

Corollary 3.32. *In the algebra $\mathbb{C}[x]$ every polynomial splits into a product of linear factors.*

Indeed, according to the main theorem the polynomial g from (3.4) must have zero degree, i.e., it must be just a number.

In view of Theorem 3.13, we also have

Corollary 3.33. *Every polynomial of degree n over $\mathbb{C}$ has n roots (counted with multiplicities).*

3.4. Roots of Polynomials with Real Coefficients

A polynomial of degree n with real coefficients can have $\leq n$ real roots (and might have none at all). But as any polynomial with complex coefficients, it always has exactly n complex roots (counted with multiplicities). Imaginary roots of a polynomial with real coefficients have a special property.

Theorem 3.34. *If c is an imaginary root of a polynomial $f \in \mathbb{R}[x]$, then $\bar{c}$ is also a root of this polynomial. Moreover, it has the same multiplicity as c.*

Proof. Let

$$f = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n, \qquad a_0, a_1, \ldots, a_n \in \mathbb{R}.$$

Complex conjugation is an automorphism of the field $\mathbb{C}$ (Section 1.5), hence if $f(c) = 0$, then

$$\begin{aligned} f(\bar{c}) &= a_0\bar{c}^n + a_1\bar{c}^{n-1} + \cdots + a_{n-1}\bar{c} + a_n \\ &= \bar{a}_0\bar{c}^n + \bar{a}_1\bar{c}^{n-1} + \cdots + \bar{a}_{n-1}\bar{c} + \bar{a}_n = \overline{f(c)} = \bar{0} = 0, \end{aligned}$$

i.e., $\bar{c}$ is also a root of f. Similarly, we can prove that

$$f^{(k)}(c) = 0 \iff f^{(k)}(\bar{c}) = 0.$$

Thus, the multiplicities of c and $\bar{c}$ are equal. □

Corollary 3.35. *In the algebra $\mathbb{R}[x]$ every nonzero polynomial factors into a product of linear terms and quadratic terms with negative discriminants.*

Proof. Observe that if c is an imaginary number, then the quadratic polynomial

$$(x - c)(x - \bar{c}) = x^2 - (c + \bar{c})x + c\bar{c}$$

has real coefficients. Its discriminant is obviously negative.

Now, let

$$c_1, \ldots, c_s, c_{s+1}, \ldots, c_{s+t}, \bar{c}_{s+1}, \ldots, \bar{c}_{s+t}$$

be all (distinct) complex roots of the polynomial $f \in \mathbb{R}[x]$ with

$$c_1, \ldots, c_s \in \mathbb{R}, \qquad c_{s+1}, \ldots, c_{s+t} \notin \mathbb{R}.$$

If the multiplicity of the root c_i is k_i, then

$$\begin{aligned} f = a_0&(x - c_1)^{k_1} \cdots (x - c_s)^{k_s} \\ &\times [(x - c_{s+1})(x - \bar{c}_{s+1})]^{k_{s+1}} \cdots [(x - c_{s+t})(x - \bar{c}_{s+t})]^{k_{s+t}} \end{aligned}$$

(where a_0 is the leading coefficient of f). After multiplying linear factors in square brackets, the corollary follows. □

Example 3.36.

$$\begin{aligned} x^5 - 1 &= (x-1)\left(x - \left(\cos\frac{2\pi}{5} + \imath \sin\frac{2\pi}{5}\right)\right)\left(x - \left(\cos\frac{2\pi}{5} - \imath \sin\frac{2\pi}{5}\right)\right) \\ &\quad \times \left(x - \left(\cos\frac{4\pi}{5} + \imath \sin\frac{4\pi}{5}\right)\right)\left(x - \left(\cos\frac{4\pi}{5} - \imath \sin\frac{4\pi}{5}\right)\right) \\ &= (x-1)\left(x^2 - 2x\cos\frac{2\pi}{5} + 1\right)\left(x^2 - 2x\cos\frac{4\pi}{5} + 1\right) \\ &= (x-1)\left(x^2 - \frac{\sqrt{5}-1}{2}x + 1\right)\left(x^2 + \frac{\sqrt{5}+1}{2}x + 1\right) \end{aligned}$$

(see Example 3.15).

Example 3.37. The polynomial f in Example 3.23 factors as

$$f = (x-2)^3(x^2 + x + 1).$$

Theorem 3.34 also implies that every polynomial $f \in \mathbb{R}[x]$ of odd degree has at least one real root. Still, there is another easy proof of this. Namely, if the leading coefficient of a polynomial f is positive, then

$$\lim_{x \to +\infty} f(x) = +\infty, \qquad \lim_{x \to -\infty} f(x) = -\infty.$$

Hence, f assumes both positive and negative values. By the Intermediate Value Theorem it follows that at some point f equals 0.

Clearly, it is useful to have a way of determining precisely the number of real roots. By calculating the values of a polynomial at some points, we can find that they have different signs at a and b. This means that there exists at least one root in the interval (a, b). To be more precise, there exist an odd number of roots there (counted with multiplicities). This reasoning allows us to bound the number of real roots from below.

Example 3.38. Given the polynomial

$$f = x^4 + x^2 - 4x + 1,$$

we find that

$$f(0) = 1 > 0, \qquad f(1) = -1 < 0, \qquad f(2) = 13 > 0.$$

Therefore, f has roots in both intervals $(0, 1)$ and $(1, 2)$. It is not hard to show that $f(x) > 0$ for $x \leq 0$ and also for $x \geq 2$. Thus, all real roots of the polynomial f lie in the interval $(0, 2)$. However, their exact number remains unknown because in either of the intervals $(0, 1)$ and $(0, 2)$ there might be as many as three roots.

There exist methods that, in principle, provide the number of real roots of any polynomial as well as the number of its roots on any interval of the real line. However, in practice they require a great deal of calculations. Below we will state a theorem that—though it does not give the precise number—requires none. Its subject is not just the number of all real roots but also the number of positive (or negative) roots. This theorem generalizes the following obvious statement: if all coefficients of a polynomial are nonnegative, then it has no positive roots.

To formulate this theorem, we need a technical notion.

Consider a finite sequence of real numbers

$$a_0, a_1, a_2, \dots, a_n.$$

We say that there is a *change of sign* in the kth position of this sequence if $a_k \neq 0$ and the sign of a_k is the opposite of the sign of the last nonzero number that precedes it. (If a_k is the first nonzero term in the sequence, there is no change of sign in the kth position.)

Theorem 3.39 (Descartes Theorem). *The number of positive roots (counted with multiplicities) of a polynomial $f \in \mathbb{R}[x]$ does not exceed the number of changes of sign in the sequence of its coefficients and is comparable to it modulo 2. If all complex roots of f are real, then these numbers are equal.*

Denote by $N(f)$ the number of positive roots of f and by $L(f)$ the number of changes of sign in the sequence of its coefficients. Clearly, these numbers do not change when f is multiplied by -1; therefore, we can assume that the leading coefficient of f is positive. Moreover, if 0 is a root of f of multiplicity k, then after dividing f by x^k these numbers do not change either. Thus we can assume that the free term of f is nonzero.

Lemma 3.40. $N(f) \equiv L(f) \pmod 2$.

Proof. Let

$$f = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n \qquad (a_0 > 0,\ a_n \neq 0).$$

Then $f(0) = a_n$ and $f(x) > 0$ for large enough x. When we move along the real line to the right, $f(x)$ changes its sign when we pass a simple root. When we pass a root of multiplicity k, the sign of $f(x)$ changes by $(-1)^k$, as if it changed k times. Thus $N(f)$ is even if $a_n > 0$ and odd if $a_n < 0$. The same can be said about $L(f)$. □

Lemma 3.41. $N(f) \leq N(f') + 1$.

Proof. By Rolle's theorem, the derivative of f has a root between any two roots of f. Moreover, every root of f of multiplicity k is a root of

the derivative of f of multiplicity $k-1$ (Corollary 3.21). Thus, $N(f') \geq N(f) - 1$. □

Lemma 3.42. $L(f') \leq L(f)$.

Proof. Clear. □

The number of negative roots of a polynomial f equals the number of positive roots of the polynomial

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

Lemma 3.43. $L(f) + L(\bar{f}) \leq n = \deg f$.

Proof. We obtain coefficients of the polynomial $\bar{f}$ from those of f by multiplying every other one by -1. First assume that all coefficients $a_0, a_1, \ldots, a_n$ of the polynomial f are nonzero. Then if a change of sign occurs in the sequence $a_0, a_1, \ldots, a_n$ in the kth position, it does not occur in the sequence of coefficients of $\bar{f}$ and vice versa. Hence, in this case $L(f) + L(\bar{f}) = n$.

In the general case, some of the coefficients $a_0, a_1, \ldots, a_n$ might be zero. When we replace them with some arbitrary nonzero numbers, $L(f)$ and $L(\bar{f})$ can only increase. We proved above that their sum will then become n, thus $L(f) + L(\bar{f}) \leq n$. □

Proof of Theorem 3.39. We will prove the inequality $N(f) \leq L(f)$ by induction on $\deg f$. If $\deg f = 0$, then $N(f) = L(f) = 0$. Assume now that $\deg f = n > 0$. Then $\deg f' = n - 1$. Lemmas 3.41 and 3.42 and the induction assumption imply that

$$N(f) \leq N(f') + 1 \leq L(f') + 1 \leq L(f) + 1.$$

But by Lemma 3.40, the equality $N(f) = L(f) + 1$ is impossible. Thus, $N(f) \leq L(f)$.

Assume now that all roots of f are real. We can assume that 0 is not a root. Then, by the above inequality and Lemma 3.43,

$$n = N(f) + N(\bar{f}) \leq L(f) + L(\bar{f}) \leq n,$$

hence

$$N(f) = L(f), \qquad N(\bar{f}) = L(\bar{f}).$$

□

Example 3.44. For the polynomial f of Example 3.38, we have $L(f) = 2$. Thus $N(f) \leq 2$. But we have established already that $N(f) \geq 2$. Therefore, $N(f) = 2$.

Example 3.45. The polynomial $f = x^2 - x + 1$ has no positive (and even real) roots, yet $L(f) = 2$. So, in this case $N(f) < L(f)$.

By applying Descartes' theorem to the polynomial

$$g(x) = f(c+x) = f(c) + \frac{f'(c)}{1!}x + \frac{f''(c)}{2!}x^2 + \cdots + \frac{f^{(n)}(c)}{n!}x^n,$$

we can deduce certain facts about the number of roots of the polynomial f on the interval $(c, +\infty)$. In particular, if all coefficients of g are nonnegative, then it has no positive roots (this is the trivial case of Descartes' theorem), hence no real root of f exceeds c.

Example 3.46. Here we will find the bounds for the real roots of the polynomial

$$f = x^5 - 5x^3 - 10x^2 + 2.$$

Using Horner's Scheme, we can calculate $f(3)$:

	1	0	−5	−10	0	2
3	1	3	4	2	6	20

We see that $f(3) = 20 > 0$. Moreover, all coefficients of the incomplete quotient are positive. Hence, all derivatives of f are positive at $x = 3$ (see Example 3.23), and so all its real roots are less than 3. Consider now the polynomial

$$\bar{f}(x) = -f(-x) = x^5 - 5x^3 + 10x^2 - 2.$$

Again, we use Horner's Scheme to calculate the values of $\bar{f}$ and its derivatives at $x = 1$:

	1	0	−5	10	0	−2
1	1	1	−4	6	6	4
	1	2	−2	4	10	
	1	3	1	5		

We see that

$$\bar{f}(1) = 4 > 0, \qquad \bar{f}'(1) = 10 > 0, \qquad \bar{f}''(1) = 2 \cdot 5 > 0.$$

The values of other derivatives at $x = 1$ are also positive because the last row of the table contains only positive numbers. Therefore, all real roots of $\bar{f}$ are less than 1. This means that all real roots of f are greater that -1. Therefore, all real roots of f belong to the interval $(-1, 3)$.

Exercise 3.47. By studying the derivative of the polynomial $\bar{f}$, prove that the polynomial f of the previous example has only one negative root.

Now we turn to the approximate calculation of roots.

If a polynomial $f \in \mathbb{R}[x]$ is known to have only one root on a given interval, this root can be calculated with any precision by calculating the values of the polynomial at particular points. We explain this in the following example.

Example 3.48. As we demonstrated (see Example 3.44), the polynomial f in Example 3.38 has exactly one root in the interval $(1, 2)$. We will find this root within 0.01. We know that $f(1) < 0$. By calculating $f(x)$ for $x = 1.1, 1.2, 1.3$, we see that

$$f(1.2) < 0, \qquad f(1.3) > 0.$$

Hence, this root lies in the interval $(1.2, 1.3)$. By calculating $f(x)$ for $x =$ 1.21, 1.22, 1.23, 1.24, 1.25, we see that

$$f(1.24) < 0, \qquad f(1.25) > 0.$$

Thus the root lies in the interval $(1.24, 1.25)$.

Certainly, there exist better methods for approximate calculations of roots. They can be applied to algebraic equations of any degree and some of them, even to transcendental equations. However, they lie beyond the scope of this course: they belong to computational mathematics rather than algebra.

Remark 3.49. If a polynomial has a multiple root but its coefficients are known only approximately (but with any degree of precision), then we cannot prove that this multiple root exists because under any perturbation of the coefficients, however small, it may separate into simple roots or simply cease to exist.

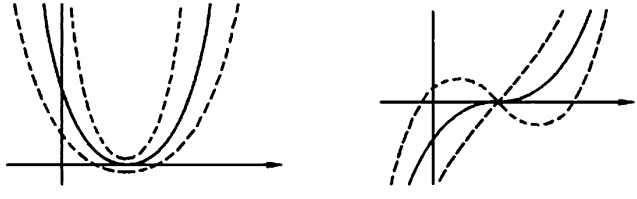


Figure 3.6

For instance in the case of a double root, we can never distinguish between situations shown in Figure 3.6, left, and in the case of a triple root, in Figure 3.6, right.

3.5. Factorization in Euclidean Domains

Factorization of polynomials over $\mathbb{C}$ into linear factors, as well as factorization of polynomials over $\mathbb{R}$ into products of linear and quadratic factors, is similar to factorization of integers into primes. Such a factorization exists for polynomials over any field but there factors can be of any degree. The problem of finding such a factorization can be viewed as a generalization of

the problem of finding all roots of a polynomial (they are equivalent over $\mathbb{C}$). This problem does not have a general solution which is valid over any field. In this section we will prove that such a factorization is unique. Simultaneously we will prove that the factorization of an integer into primes is unique; you probably learned the latter fact in high school but did not see the proof.

To make an argument encompassing both cases, let us introduce several additional notions.

Definition 3.50. A commutative and associative ring with unity and without zero divisors is called an *integral domain*.

For instance, the rings $\mathbb{Z}$ of integers and $K[x]$ of polynomials over a field K are integral domains. Moreover, the polynomial ring over an integral domain is itself an integral domain (see Remark 3.3).

Let A be an integral domain. We say that an element $a \in A$ is *divisible* by $b \in A$ (notation $a \vdots b$) or, equivalently, that b *divides* a (notation $b|a$) if there exists an element $q \in A$ such that $a = qb$. Elements a and b are called *associated* (notation $a \sim b$) if either of the following equivalent conditions holds:

(i) $b|a$ and $a|b$;

(ii) $a = cb$, where c is invertible.

The next definition describes axiomatically the common property of the ring of integers and the ring of polynomials over any field—division with a remainder.

Definition 3.51. Let A be an integral domain which is not a field. We call A *Euclidean* if there exists a function

$$N\colon A \setminus \{0\} \to \mathbb{Z}_+$$

(called a *norm*) that satisfies the following conditions:

(i) $N(ab) \geq N(a)$ and the equality holds if and only if b is invertible;

(ii) for any $a, b \in A$, where $b \neq 0$, there exist $q, r \in A$ such that $a = qb + r$ and either $r = 0$ or $N(r) < N(b)$.

Remark 3.52. Condition (ii) means that we can "divide with a remainder." Uniqueness (i.e., the existence of only one such pair (q, r)) is not required here.

Remark 3.53. The second part of condition (i) can be deduced from other conditions. Indeed, assume that b is not invertible. Then a is not divisible by ab. Divide a by ab with a remainder:

$$a = q(ab) + r.$$

Since $r = a(1 - qb)$, we have

$$N(a) \leq N(r) < N(ab).$$

Our main examples of Euclidean domains are the ring $\mathbb{Z}$ of integers and the ring $K[x]$ of polynomials over a field K. In the first case, we can take as a norm the absolute value of an integer and in the second, the degree of a polynomial.

Other Euclidean domains exist as well.

Example 3.54. Complex numbers of the form $c = a + bi$, where $a, b \in \mathbb{Z}$, are called *Gaussian integers*. They form a subring of $\mathbb{C}$ denoted $\mathbb{Z}[i]$. The domain $\mathbb{Z}[i]$ is Euclidean with respect to the norm

$$N(c) = |c|^2 = a^2 + b^2.$$

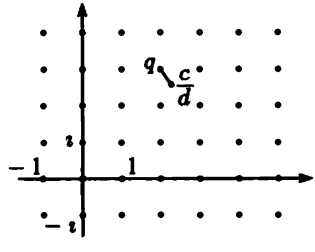


Figure 3.7

Indeed, it is clear that

$$N(cd) = N(c)N(d)$$

and, since $N(1) = 1$, the invertible elements of the ring $\mathbb{Z}[i]$ are the elements with norm 1, i.e., ± 1 and $\pm i$. It follows that condition (i) of Definition 3.51 holds. Now we have to prove that it is possible to divide with a remainder in $\mathbb{Z}[i]$. Let $c, d \in \mathbb{Z}[i]$, $d \neq 0$. Consider a Gaussian integer q nearest to $\frac{c}{d}$. It is easy to see that $\left|\frac{c}{d} - q\right| \leq 1/\sqrt{2}$ (see Figure 3.7). Put $r = c - qd$. Then $c = qd + r$ and

$$N(r) = |c - qd|^2 = \left|\frac{c}{d} - q\right|^2 |d|^2 \leq \frac{1}{2} N(d) < N(d).$$

Exercise 3.55. Prove that the ring of rational numbers of the form $2^{-n}m$, $m \in \mathbb{Z}, n \in \mathbb{Z}_+$, is a Euclidean domain.

Definition 3.56. The *greatest common divisor* of elements a and b of an integral domain is a common divisor of a and b divisible by all their common divisors. It is denoted (a, b) or $\mathrm{GCD}\{a, b\}$.

The greatest common divisor, if it exists, is defined up to association relation (see above). However, it may not exist at all. For instance, elements x^5 and x^6 in the ring of polynomials without a linear term do not possess the greatest common divisor.

Theorem 3.57. *For any two elements a, b of a Euclidean domain, there exists the greatest common divisor d. It can be presented in the form $d = au + bv$, where u, v are some elements of the ring.*

Proof. If $b = 0$, then $d = a = a\cdot 1 + b\cdot 0$. If b divides a, then $d = b = a\cdot 0 + b\cdot 1$. Otherwise, divide a by b with a remainder, then b by this remainder, then the first remainder by the second, etc. Since the norms of the remainders decrease, at some point the remainder will be zero. We obtain the following chain of equalities:

$$\begin{aligned} a &= q_1 b + r_1, \\ b &= q_2 r_1 + r_2, \\ r_1 &= q_3 r_2 + r_3, \\ &\cdots\cdots\cdots\cdots \\ r_{n-2} &= q_n r_{n-1} + r_n, \\ r_{n-1} &= q_{n+1} r_n. \end{aligned}$$

We will prove that the last nonzero remainder r_n is in fact the greatest common divisor of a and b.

By moving up this chain, we consecutively obtain the following:

$$r_n | r_{n-1}, \quad r_n | r_{n-2}, \quad \ldots, \quad r_n | r_1, \quad r_n | b, \quad r_n | a.$$

Thus, r_n is a common divisor of a and b.

Again, by moving up the chain, we consecutively obtain

$$\begin{aligned} r_1 &= au_1 + bv_1, \\ r_2 &= au_2 + bv_2, \\ r_3 &= au_3 + bv_3, \\ &\cdots\cdots\cdots\cdots \\ r_n &= au_n + bv_n, \end{aligned}$$

where u_i, v_i, $i = 1, \ldots, n$, are some elements of our ring (for example, $u_1 = 1$, $v_1 = -q_1$). Hence, r_n can be presented in the form $au + bv$. This implies that any common divisor of a and b divides r_n. □

We now turn to factorization into prime factors.

Definition 3.58. A noninvertible element p of an integral domain is called *prime* if it cannot be presented as $p = ab$, where a and b are noninvertible elements.

In other words, an element p is prime if its every divisor is associated with either 1 or p. In this sense, prime elements of the ring $\mathbb{Z}$ are numbers of the type $\pm p$, where p is a prime number.

Prime elements of the ring $K[x]$ over a field K are traditionally called *irreducible polynomials*. Thus, an irreducible polynomial is a polynomial of positive degree that does not factor into two polynomials of positive degree.

Clearly, every polynomial of the first degree is irreducible. The fundamental theorem of algebra of complex numbers implies that these are the only irreducible polynomials over $\mathbb{C}$. In turn, Corollary 3.35 implies that all irreducible polynomials over $\mathbb{R}$ are polynomials of the first degree and polynomials of the second degree with negative discriminants. In the next section we will discuss the question of irreducibility of polynomials over $\mathbb{Q}$ and, in particular, we will see that they can be of any degree.

Now let A be any Euclidean domain.

Lemma 3.59. *If a prime element p of A divides the product $a_1 a_2 \cdots a_n$, then it divides at least one of the factors $a_1, a_2, \ldots, a_n$.*

Proof. We will prove this statement by induction on n. For $n = 2$, assume that p does not divide a_1. Then $(p, a_1) = 1$, hence there exist $u, v \in A$ such that $pu + a_1 v = 1$. Multiplying this equality by a_2, we obtain

$$pua_2 + a_1 a_2 v = a_2.$$

This implies that p divides a_2.

For $n > 2$, consider the product $a_1 a_2 \cdots a_n$ in the form $a_1(a_2 \cdots a_n)$. According to the above, $p|a_1$ or $p|a_2 \cdots a_n$. In the second case, the induction statement implies that $p|a_i$, where i is one of the indices $2, \ldots, n$. □

Theorem 3.60. *In a Euclidean domain, every noninvertible nonzero element factors into prime factors. This factorization is unique up to an arrangement of factors and multiplication by invertible elements.*

Remark 3.61. When speaking of factorization into prime factors, we do not exclude the case of factorization into just one factor.

Proof. Call a noninvertible nonzero element $a \in A$ good if it can be factored into primes. Assume that there exist bad elements. Pick one with the least norm and denote it by a. It cannot be prime. Hence $a = bc$, where b and c are good elements. But then it is obvious that a is good, which contradicts our assumption. Therefore, every noninvertible nonzero element of A can be factored into prime factors.

We will now prove by induction that if

$$a = p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m, \tag{3.10}$$

where p_i, q_j are prime elements, then $m = n$ and that, after renumbering the factors, we get $p_i \sim q_i$ for $i = 1, 2, \ldots, n$.

For $n = 1$, this statement is obvious. When $n > 1$, we have $p_1 | q_1 q_2 \cdots q_m$ and by Lemma 3.59 there exists i such that $p_1 | q_i$. We can assume that $i = 1$ and $p_1 = q_1$. After cancelling p_1 in (3.10), we obtain

$$p_2 \cdots p_n = q_2 \cdots q_m.$$

The induction assumption thus implies that $m = n$ and, after a change of indices, $p_i \sim q_i$ for $i = 2, \ldots, n$. This completes the proof. □

Corollary 3.62. *Let $a = p_1^{k_1} \cdots p_s^{k_s}$ be a factorization of $a \in A$ into prime factors such that $p_i \not\sim p_j$ for $i \neq j$. Then every divisor d of a has the following form:*

$$d = c p_1^{l_1} \cdots p_s^{l_s},$$

where $0 \leq l_i \leq k_i$ $(i = 1, \ldots, s)$ and c is an invertible element.

Proof. Let $a = qd$. Factor q and d into prime factors. After multiplying these factorizations, we will obtain a factorization of a. Comparing it with the one in this corollary's statement completes the proof. □

Exercise 3.63. Prove that in a Euclidean domain

a) $b|a$, $c|a$ and $(b, c) = 1 \implies bc|a$;

b) $c|ab$ and $(b, c) = 1 \implies c|a$.

Exercise 3.64. The *least common multiple* of elements a and b of an integral domain is their common multiple (i.e., an element divisible by both a and b) that divides all their common multiples. It is denoted $[a, b]$ or $\mathrm{LCM}\{a, b\}$. Prove that in a Euclidean domain every two elements a and b have the least common multiple $[a, b]$ and also that

$$(a, b)[a, b] \sim ab.$$

Exercise 3.65. Factor the elements 2, 3, and 5 of the ring $\mathbb{Z}[i]$ (see Example 3.54) into primes. What is the principal difference between these three cases?

It is well known that there exist infinitely many prime numbers. Recall the argument that proves this. (Assume that $p_1, p_2, \ldots, p_n$ are the only prime numbers. Then the number $p_1 p_2 \cdots p_n + 1$ is not divisible by any of them, which is clearly impossible.) The same argument shows that there exist infinitely many monic irreducible polynomials over any field K. If K is infinite, then this result if of little interest: clearly in this case there exist infinitely many monic polynomials of the first degree. However, if K is finite, this result implies that there exist irreducible polynomials of arbitrarily high

degree. Actually, in this case there exist irreducible polynomials of *any* degree.

Exercise 3.66. List irreducible polynomials of degree ≤ 4 over the field $\mathbb{Z}_2$ and prove that there exist exactly 6 irreducible polynomials of degree 5.

3.6. Polynomials with Rational Coefficients

Every integer factors into primes uniquely, and this implies

Theorem 3.67. *If a polynomial*

$$f = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \in \mathbb{Z}[x]$$

has a rational root $\frac{u}{v}$, *where* $u, v \in \mathbb{Z}$, $(u, v) = 1$, *then* $u|a_n$, $v|a_0$.

Proof. By assumption,

$$0 = v^n f\left(\frac{u}{v}\right) = a_0u^n + a_1u^{n-1}v + \cdots + a_{n-1}uv^{n-1} + a_nv^n.$$

All terms on the right-hand side, except for the last one, are divisible by u. Therefore, the last term must be divisible by u as well. But as u and v are relatively prime, a_n is divisible by u (see Exercise 3.63(b)). The proof that a_0 is divisible by v is similar. □

Corollary 3.68. *If a monic polynomial with integer coefficients has a rational root, then this root is an integer.*

Obviously, every polynomial with rational coefficients is proportional to a polynomial with integer coefficients. Thus Theorem 3.67 suggests how one can find all rational roots of a given polynomial with rational coefficients in a finite number of steps. Of course, usually there exist no such roots. The specially chosen example below is one of the exceptions that only prove the rule.

Example 3.69. According to Theorem 3.67, the only candidates for the rational roots of the polynomial

$$f = 2x^4 - 7x^3 + 4x^2 - 2x - 3$$

are

$$\pm\frac{1}{2}, \qquad \pm 1, \qquad \pm\frac{3}{2}, \qquad \pm 3.$$

Tryouts provide two roots:

$$x_1 = 3, \qquad x_2 = -\frac{1}{2}.$$

The next theorem can be considered a generalization of Theorem 3.67.

Theorem 3.70 (Gauss Lemma). *If a polynomial with integer coefficients factors into a product of two polynomials with rational coefficients, then it factors into a product of polynomials with integer coefficients proportional to the ones from the first factorization.*

In other words, if $f \in \mathbb{Z}[x]$ and $f = gh$, where $g, h \in \mathbb{Q}[x]$, then there exists $\lambda \in \mathbb{Q}^*$ such that $\lambda g, \lambda^{-1}h \in \mathbb{Z}[x]$.

Before we prove this theorem, we need to introduce a few useful notions.

A polynomial $f \in \mathbb{Z}[x]$ is called *primitive* if its coefficients taken together are relatively prime, i.e., if they do not have a common divisor greater than 1. If such a divisor exists, it can be carried out. Thus, every polynomial with integer coefficients is proportional to a primitive polynomial (determined uniquely up to multiplication by ± 1). Hence so is every polynomial with rational coefficients.

Let p be a prime number. We define *reduction modulo* p of a polynomial

$$f = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \in \mathbb{Z}[x]$$

as a polynomial

$$[f]_p = [a_0]_p x^n + [a_1]_p x^{n-1} + \cdots + [a_{n-1}]_p x + [a_n]_p \in \mathbb{Z}_p[x]$$

whose coefficients are residue classes modulo p of coefficients of f. From the definition of operations on residue classes, it follows that

$$[f+g]_p = [f]_p + [g]_p,$$
$$[fg]_p = [f]_p[g]_p$$

for any $f, g \in \mathbb{Z}[x]$.

Proof of Theorem 3.70. Let $f \in \mathbb{Z}[x]$ and $f = gh$, where $g, h \in \mathbb{Q}[x]$. The preceding discussion implies that g and h are proportional to some primitive polynomials g_1 and h_1. We have

$$f = \mu g_1 h_1, \qquad \mu \in \mathbb{Q}.$$

Let $\mu = \frac{u}{v}$, where $u, v \in \mathbb{Z}$, $(u, v) = 1$. If we can show that $v = \pm 1$, the theorem will follow. Assume this is not so and let p be a prime divisor of v. Reduce the equality

$$vf = ug_1h_1$$

modulo p. We have

$$0 = [u]_p[g_1]_p[h_1]_p.$$

However, $[u]_p \neq 0$ because u and v are relatively prime by assumption. On the other hand, $[g_1]_p \neq 0$ and $[h_1]_p \neq 0$ because g_1 and h_1 are primitive polynomials, hence their coefficients cannot be all divisible by p. This contradicts the absence of zero divisors in $\mathbb{Z}_p[x]$. □

Corollary 3.71. *If a polynomial $f \in \mathbb{Z}[x]$ factors into a product of two polynomials of positive degree in $\mathbb{Q}[x]$, then it factors into such a product in $\mathbb{Z}[x]$.*

This makes it much easier to determine whether a polynomial is irreducible over $\mathbb{Q}[x]$.

Example 3.72. Let p be a prime number. We will prove that the "*cyclotomic polynomial*,"

$$f = x^{p-1} + x^{p-2} + \cdots + x + 1 = \frac{x^p - 1}{x - 1},$$

is irreducible over $\mathbb{Q}$. (All complex roots of this polynomial are nontrivial pth roots of 1. Together with 1 they cut the circle $|z| = 1$ into p equal parts.) By the binomial formula (see, e.g., Section 1.6), the following equality holds in $\mathbb{Z}_p[x]$:

$$x^p - 1 = (x - 1)^p.$$

Thus,

$$[f]_p = (x - 1)^{p-1}.$$

If $f = gh$, where $g, h \in \mathbb{Z}[x]$ are polynomials of positive degree, then $[f]_p = [g]_p[h]_p$. Therefore,

$$[g]_p = (x - 1)^k, \qquad [h]_p = (x - 1)^l, \qquad k, l > 0,\ k + l = p - 1.$$

Hence,

$$[g(1)]_p = [g]_p(1) = 0, \qquad [h(1)]_p = [h]_p(1) = 0,$$

i.e., $g(1)$ and $h(1)$ are divisible by p. But then $f(1) = g(1)h(1)$ is divisible by p^2, which is false as $f(1) = p$.

An algorithm due to Kronecker allows us to determine whether any given polynomial with integer coefficients is irreducible over $\mathbb{Q}$. It is based on the following considerations.

Let $f \in \mathbb{Z}[x]$ be a polynomial of degree n without integer roots. Assume that it factors into a product of two polynomials of positive degree with integer coefficients:

$$f = gh.$$

Then the degree of one of them, say g, does not exceed $m = \left[\frac{n}{2}\right]$.

Let us assign in turn distinct integer values $x_0, x_1, \ldots, x_m$ to the variable x. Equalities

$$f(x_i) = g(x_i)h(x_i)$$

imply that $g(x_i) | f(x_i)$ for $i = 0, 1, \ldots, m$. The polynomial g is uniquely determined by its values at points $x_0, x_1, \ldots, x_m$. By considering all collections of divisors $d_0, d_1, \ldots, d_m$ of integers $f(x_0), f(x_1), \ldots, f(x_m)$ and determining for each the interpolation polynomial of degree $\leq m$ (a polynomial that

assumes values $d_0, d_1, \ldots, d_m$ at points $x_0, x_1, \ldots, x_m$), we will create the list of all candidates for g (there will be a finite number of them). Those that have fractional coefficients can be removed from the list immediately. Checking others, we can determine if any of them divide f. This will solve the question of f's irreducibility.

3.7. Polynomials in Several Variables

A function of real variables $x_1, x_2, \ldots, x_n$ is called a *polynomial* if it can be presented as

$$f(x_1, x_2, \ldots, x_n) = \sum_{k_1, k_2, \ldots, k_n} a_{k_1 k_2 \ldots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}, \tag{3.11}$$

where the summation is taken over a finite set of collections $(k_1, k_2, \ldots, k_n)$ of nonnegative integers. (Formally, it can be said that the summation is taken over all such collections but only finitely many coefficients $a_{k_1 k_2 \ldots k_n}$ are nonzero.) Polynomials form a subalgebra of the algebra of all functions of $x_1, x_2, \ldots, x_n$. It is called the algebra of polynomials in $x_1, x_2, \ldots, x_n$ over $\mathbb{R}$ and is denoted $\mathbb{R}[x_1, x_2, \ldots, x_n]$.

We will show (see Theorem 3.76 below) that the presentation of a polynomial in the form (3.11) is unique, i.e., that the coefficients of a polynomial are determined by its values.

When we attempt to define polynomials in n variables over an arbitrary field K, we hit the same problem as in the case of one variable. This makes it necessary to give a formal definition such as the following one.

Consider an infinite-dimensional algebra over K with a basis

$$\{e_{k_1 k_2 \ldots k_n} : k_1, k_2, \ldots, k_n \in \mathbb{Z}_+\}$$

and the multiplication table

$$e_{k_1 k_2 \ldots k_n} e_{l_1 l_2 \ldots l_n} = e_{k_1 + l_1, k_2 + l_2, \ldots, k_n + l_n}.$$

Obviously, this algebra is commutative and associative. Its identity is the element $e_{00\ldots0}$. This algebra is called the *polynomial algebra* over K and is denoted $K[x_1, x_2, \ldots, x_n]$.

We identify the elements of the form $ae_{00\ldots0}$, $a \in K$, with the corresponding elements of K. Introduce the following notation:

$$\begin{aligned} e_{10\ldots0} &= x_1, \\ e_{01\ldots0} &= x_2, \\ &\ldots\ldots\ldots \\ e_{00\ldots1} &= x_n. \end{aligned}$$

Then

$$e_{k_1 k_2 \ldots k_n} = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

and every element

$$\sum_{k_1,k_2,\dots,k_n} a_{k_1k_2\dots k_n} e_{k_1k_2\dots k_n} \in K[x_1, x_2, \dots, x_n]$$

is presented in the usual form (3.11).

The polynomial (3.11) is called *homogeneous* of degree d if

$$a_{k_1k_2\dots k_n} = 0 \qquad \text{for} \qquad k_1 + k_2 + \cdots + k_n \neq d.$$

Homogeneous polynomials of a fixed degree d form a finite-dimensional subspace. Indeed, there exist only a finite number of collections $(k_1, k_2, \dots, k_n)$ of nonnegative integers such that

$$k_1 + k_2 + \cdots + k_n = d.$$

Exercise 3.73. Prove that the dimension of the space of homogeneous polynomials of degree d in n variables equals

$$\frac{n(n+1)\cdots(n+d-1)}{d!}$$

(*the number of d choices out of n with repetitions*).

Every polynomial can be uniquely presented as a sum of homogeneous polynomials of degrees $0, 1, 2, \dots$, called its *homogeneous components*. (Only finitely many of them are nonzero.)

The (total) *degree* of a nonzero polynomial is the maximum degree of its nonzero terms—which is the same as the maximum degree of its nonzero homogeneous components. The degree of a polynomial f is denoted $\deg f$. It satisfies the following properties:

$$\deg(f+g) \le \deg f + \deg g, \tag{3.12}$$

$$\deg(fg) = \deg f + \deg g. \tag{3.13}$$

The first is obvious; as for the second, we will prove it later.

On the other hand, every polynomial $f \in K[x_1, x_2, \dots, x_n]$ can be uniquely presented as

$$f(x_1, x_2, \dots, x_n) = \sum_{k=0}^{\infty} f_k(x_2, \dots, x_n) x_1^k, \tag{3.14}$$

where $f_0, f_1, f_2, \dots$ are some polynomials in $x_2, \dots, x_n$ and only finitely many of them are nonzero. The greatest index among the nonzero polynomials f_k is called the degree of f in x_1 and is denoted $\deg_{x_1} f$.

Using presentation (3.14), we can regard the ring $K[x_1, x_2, \dots, x_n]$ as a ring of polynomials in x_1 with coefficients from $K[x_2, \dots, x_n]$:

$$K[x_1, x_2, \dots, x_n] = K[x_2, \dots, x_n][x_1]. \tag{3.15}$$

Remark 3.74. We speak here about rings but not about algebras because $K[x_1, x_2, \dots, x_n]$ is by definition an algebra over K, while $K[x_2, \dots, x_n][x_1]$ is an algebra over $K[x_2, \dots, x_n]$. However, if we regard $K[x_2, \dots, x_n][x_1]$ as an algebra over K (this is possible because $K[x_2, \dots, x_n] \supset K$), we can speak about equality between the algebras.

Proposition 3.75. *The algebra $K[x_1, x_2, \dots, x_n]$ has no zero divisors.*

Proof. We actually proved in Section 3.1 that the ring of polynomials in one variable over an integral domain is also an integral domain (and, in particular, has no zero divisors), see Remark 3.3. Therefore, we can prove the statement by induction starting with the field K and using equality (3.15). □

Now we can prove property (3.13). Write polynomials f and g as sums of their homogeneous components:

$$f = f_0 + f_1 + \cdots + f_d, \qquad \deg f_k = k,\ f_d \neq 0,$$
$$g = g_0 + g_1 + \cdots + g_e, \qquad \deg g_k = k,\ g_e \neq 0.$$

Clearly, when crossmultiplying, we will not get terms of degrees $> d + e$, and the sum of all terms of degree $d + e$ will be equal to $f_d g_e$. By the above proposition, $f_d g_e \neq 0$. Hence,

$$\deg fg = d + e = \deg f + \deg g.$$

As in the case $n = 1$, every polynomial in n variables over a field K determines a K-valued function on K^n.

Theorem 3.76. *If the field K is infinite, then different polynomials in n variables over K determine different functions.*

Proof. As in the case of polynomials in one variable (see the proof of Theorem 3.5), it suffices to prove that a nonzero polynomial determines a nonzero function. We will prove this by induction on n.

For $n = 1$, our claim is Theorem 3.5. Assume now that a polynomial $f \in K[x_1, x_2, \dots, x_n]$, $n > 1$, determines the zero function. Present it in the form (3.14) and assign some values to variables $x_2, \dots, x_n$. We obtain a polynomial in one variable x_1 with coefficients in K that vanishes for any value of x_1. By Theorem 3.5, all its coefficients are zero. Thus, every polynomial $f_k \in K[x_2, \dots, x_n]$ vanishes for any values of $x_2, \dots, x_n$, i.e., it determines the zero function. By the induction assumption, it follows that $f_k = 0$ for all k. But then $f = 0$ too. □

Remark 3.77. If the field K is finite, the theorem and its proof remain valid in the case of polynomials whose degree in every variable is less than the number of elements of K (cf. Remark 3.6).

Exercise 3.78. Prove that if a field K contains q elements, the functions determined by monomials $x_1^{k_1} \cdots x_n^{k_n}$ for $k_1, \ldots, k_n < q$ form a basis in the space of all K-valued functions on K^n.

If $n > 1$, the terms of a polynomial in n variables cannot be ordered by degrees: there may be several terms of the same degree. However, it is sometimes useful to have an order of some sort. In such a case, the *lexicographic* (i.e., dictionary-style) order is used. For this, first the exponents of x_1 are compared, then, if they turn out to be equal, the exponents of x_2, etc. If a monomial u is lexicographically greater than a monomial v, we denote this as $u \succ v$. According to the definition, this means that the first variable that has different exponents in u and v, has a greater exponent in u than in v.

Proposition 3.79. *The relation of lexicographic order of monomials has the following properties:*

(i) if $u \succ v$ and $v \succ w$, then $u \succ w$ (transitivity);

(ii) if $u \succ v$, then $uw \succ vw$ for any monomial w;

(iii) if $u_1 \succ v_1$ and $u_2 \succ v_2$, then $u_1 u_2 \succ v_1 v_2$.

The first property is the one that gives us the right to call the relation $\succ$ an order.

Proof. (i) For the first variable that has different exponents in the monomials u, v, w, denote these exponents by k, l, m, respectively. Then

$$k \geq l \geq m,$$

and at least one of these inequalities is strict, thus $k > m$.

(ii) When we multiply by w, we add the same number to the exponent of each variable in both u and v. Therefore, the inequality (or equality) relations between these exponents do not change, and in comparing monomials we use these relations only.

(iii) The previous property implies

$$u_1 u_2 \succ v_1 u_2 \succ v_1 v_2.$$

□

Example 3.80. The terms of the following polynomial are ordered lexicographically:

$$x_1^2 x_2 + x_1 x_2^2 x_3 + 2 x_1 x_3^2 + x_2 x_3^3 - x_2 x_3^2 + 3.$$

Notice that the term $x_1 x_2^2 x_3$ is lexicographically less than $x_1^2 x_2$, even though the degree of the former is greater.

Among nonzero terms of a nonzero polynomial $f \in K[x_1, x_2, \ldots, x_n]$, there exists one that is lexicographically greater than the others. It is called the *leading monomial* of the polynomial f.

Proposition 3.81. *The leading monomial of the product of nonzero polynomials equals the product of their leading monomials.*

Proof. It suffices to prove the statement for any pair of polynomials. Let f_1, f_2 be nonzero polynomials. Let u_1, u_2 be the leading monomials of f_1, f_2, and v_1, v_2 arbitrary terms of f_1, f_2. If $v_1 \neq u_1$ or $v_2 \neq u_2$, then by Proposition 3.79,

$$u_1 u_2 \succ v_1 v_2.$$

Therefore, after gathering terms in the product $f_1 f_2$, we see that the product $u_1 u_2$ still remains a nonzero term that is greater than all others. □

3.8. Symmetric Polynomials

Definition 3.82. A polynomial $f \in K[x_1, x_2, \ldots, x_n]$ is called *symmetric* if it remains the same for any arrangement of the variables.

Since any arrangement can be achieved as a sequence of permutations of two elements, a polynomial is symmetric if it does not change when any two variables are interchanged.

Clearly, a homogeneous component of a symmetric polynomial is also a symmetric polynomial.

Example 3.83. *Power sums*

$$s_k = x_1^k + x_2^k + \cdots + x_n^k, \qquad k = 1, 2, \ldots,$$

are obviously symmetric polynomials.

Example 3.84. The following symmetric polynomials are called *elementary symmetric polynomials*:

$$\begin{aligned}
\sigma_1 &= x_1 + x_2 + \cdots + x_n, \\
\sigma_2 &= x_1 x_2 + x_1 x_3 + \cdots + x_{n-1} x_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\sigma_k &= \sum_{i_1 < i_2 < \cdots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k}, \\
&\cdots\cdots\cdots\cdots \\
\sigma_n &= x_1 x_2 \cdots x_n.
\end{aligned}$$

Example 3.85. The Vandermonde determinant

$$V(x_1, x_2, \ldots, x_n) = \prod_{i>j}(x_i - x_j)$$

(see Example 2.96) is a product of differences of two variables. Under a permutation of variables, it can only change by ± 1 and this happens when the minuend and subtrahend change places. The number of such occurrences equals the number of inversions in the arrangements. Therefore,

$$V(x_{k_1}, x_{k_2}, \ldots, x_{k_n}) = \operatorname{sign}(k_1, k_2, \ldots, k_n) V(x_1, x_2, \ldots, x_n).$$

Hence, the Vandermonde determinant is not a symmetric polynomial itself but its square is:

$$V(x_1, x_2, \ldots, x_n)^2 = \prod_{i>j}(x_i - x_j)^2.$$

Example 3.86. Under any permutation of the variables x_1, x_2, x_3, x_4, the polynomials

$$h_1 = x_1x_2 + x_3x_4, \qquad h_2 = x_1x_3 + x_2x_4, \qquad h_3 = x_1x_4 + x_2x_3$$

are themselves permuted. Therefore, any symmetric polynomial in h_1, h_2, h_3 will be also symmetric in x_1, x_2, x_3, x_4. In particular, their product is such:

$$h_1h_2h_3 = (x_1x_2 + x_3x_4)(x_1x_3 + x_2x_4)(x_1x_4 + x_2x_3).$$

Exercise 3.87. Prove that the polynomial

$$(x_1 + x_2 - x_3 - x_4)(x_1 - x_2 + x_3 - x_4)(x_1 - x_2 - x_3 + x_4)$$

is symmetric.

Symmetric polynomials are useful in the study of algebraic equations in one variable. The key is Viète's formulas (see Section 3.2) that express elementary symmetric polynomials of the roots of an algebraic equation via its coefficients (if the number of the roots of the equation in the given field equals its degree). Clearly, only symmetric polynomials in the roots of an equation are well defined: the value of any other polynomial depends on the order of roots. On the other hand, we will show that any symmetric polynomial in the roots of an algebraic equation can be expressed via the coefficients of this equation.

Example 3.88. The polynomial $s_2 = x_1^2 + x_2^2 + \cdots + x_n^2$ is symmetric. It is clear that

$$s_2 = \sigma_1^2 - 2\sigma_2. \tag{3.16}$$

Thus, the sum of squares of roots of an algebraic equation

$$x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n = 0$$

equals $a_1^2 - 2a_2$.

Obviously, sums and products of symmetric polynomials as well as the products of symmetric polynomials and numbers are symmetric polynomials. In other words, symmetric polynomials form a subalgebra in the algebra of all polynomials.

Therefore, if $F \in K[X_1, X_2, \ldots, X_m]$ is a polynomial in m variables and $f_1, f_2, \ldots, f_m \in K[x_1, x_2, \ldots, x_n]$ are symmetric polynomials, then $F(f_1, f_2, \ldots, f_m)$ is a symmetric polynomial in $x_1, x_2, \ldots, x_n$. It is natural to ask if there exist symmetric polynomials $f_1, f_2, \ldots, f_m$ such that any symmetric polynomial can be expressed through them as above. It happens that they do exist, and the elementary symmetric polynomials $\sigma_1, \sigma_2, \ldots, \sigma_n$ are exactly such.

Theorem 3.89. *Any symmetric polynomial can be uniquely presented as a polynomial in elementary symmetric polynomials.*

Two lemmas are required for the proof.

Lemma 3.90. *Let $u = ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ be the leading monomial of a symmetric polynomial f. Then*

$$k_1 \geq k_2 \geq \cdots \geq k_n. \tag{3.17}$$

Proof. Assume that $k_i < k_{i+1}$ for some i. Apart from u, f must contain the monomial

$$u' = ax_1^{k_1}\cdots x_i^{k_{i+1}}x_{i+1}^{k_i}\cdots x_n^{k_n}$$

obtained from u by exchanging x_i and x_{i+1}. It is easy to see that $u' \succ u$. This contradicts the assumption that u is the leading monomial of f. □

Lemma 3.91. *For any monomial $u = x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ satisfying (3.17), there exist nonnegative integers $l_1, l_2, \ldots, l_n$ such that the leading monomial of the product of symmetric polynomials $\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}$ equals u. Moreover, this condition determines the numbers $l_1, l_2, \ldots, l_n$ uniquely.*

Proof. The leading monomial of σ_k equals $x_1x_2\cdots x_k$. By Proposition 3.81, the leading monomial of $\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}$ equals

$$x_1^{l_1}(x_1x_2)^{l_2}\cdots(x_1x_2\cdots x_n)^{l_n} = x_1^{l_1+l_2+\cdots+l_n}x_2^{l_2+\cdots+l_n}\cdots x_n^{l_n}.$$

Setting it equal to the monomial u, we obtain the following system of linear equations:

$$\begin{cases} l_1 + l_2 + \cdots + l_n = k_1, \\ \qquad l_2 + \cdots + l_n = k_2, \\ \qquad\qquad \cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad l_n = k_n, \end{cases}$$

which obviously has the unique solution

$$l_i = k_i - k_{i+1}, \quad i = 1, 2, \ldots, n-1, \qquad l_n = k_n. \tag{3.18}$$

The assumptions of this lemma imply that the numbers $l_1, l_2, \ldots, l_n$ defined by 3.18 are nonnegative. □

Remark 3.92. The equation $l_1 + l_2 + \cdots + l_n = k_1$ shows that the total degree of the monomial $X_1^{l_1} X_2^{l_2} \cdots X_n^{l_n}$ equals the degree of u in x_1.

Proof of Theorem 3.89. Let $f \in K[x_1, x_2, \ldots, x_n]$ be a symmetric polynomial. We need to find $F \in K[X_1, X_2, \ldots, X_n]$ such that

$$F(\sigma_1, \sigma_2, \ldots, \sigma_n) = f.$$

If $f = 0$, then we can take $F = 0$. Otherwise, let $u_1 = ax_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$ be the leading monomial of f. By Lemma 3.90 the inequalities (3.17) hold. By Lemma 3.91, there exists a monomial $F_1 \in K[X_1, X_2, \ldots, X_n]$ such that the leading monomial of the polynomial $F_1(\sigma_1, \sigma_2, \ldots, \sigma_n)$ equals u_1. Consider the symmetric polynomial

$$f_1 = f - F_1(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

If $f_1 = 0$, we can take $F = F_1$. Otherwise, let u_2 be the leading monomial of f_1. Clearly it is less than u_1. There exists a monomial $F_2 \in K[X_1, X_2, \ldots, X_n]$ such that the leading monomial of $F_2(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is equal to u_2. Consider the symmetric polynomial

$$f_2 = f_1 - F_2(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

If $f_2 = 0$, we can take $F = F_1 + F_2$. Otherwise, continuing as above, we obtain a sequence of symmetric polynomials $f, f_1, f_2, \ldots$ whose leading monomials satisfy the following inequalities:

$$u_1 \succ u_2 \succ \cdots.$$

By Lemma 3.90, the exponent of any variable in any monomial u_m does not exceed the exponent of x_1 in this monomial, which, in turn, does not exceed k_1. Therefore, there exist only finitely many possible collections of exponents of variables in u_m, and thus the algorithm we have just described will stop at some point. This means that $f_M = 0$ for some M. Then we can take $F = F_1 + F_2 + \cdots + F_M$.

It remains to prove that the polynomial F is uniquely determined. Assume that F and G are polynomials such that

$$F(\sigma_1, \sigma_2, \ldots, \sigma_n) = G(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

Consider their difference $H = F - G$. Then

$$H(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0.$$

We have to prove that $H = 0$. Assume that this is not the case and let $H_1, H_2, \ldots, H_s$ be all nonzero terms of H. Denote by w_i, $i = 1, 2, \ldots, s$, the

leading monomial of the polynomial

$$H_i(\sigma_1, \sigma_2, \ldots, \sigma_n) \in K[x_1, x_2, \ldots, x_n].$$

By Lemma 3.91 neither monomial $w_1, w_2, \ldots, w_s$ is proportional to any other. Consider the greatest of them. Assume it is w_1. By construction, w_1 is greater than all other monomials in $H_1(\sigma_1, \sigma_2, \ldots, \sigma_n)$ and all monomials in the polynomials $H_i(\sigma_1, \sigma_2, \ldots, \sigma_n)$, $i = 2, \ldots, s$. Thus, after we gather terms in

$$H_1(\sigma_1, \sigma_2, \ldots, \sigma_n) + H_2(\sigma_1, \sigma_2, \ldots, \sigma_n) + \cdots + H_s(\sigma_1, \sigma_2, \ldots, \sigma_n) = H(\sigma_1, \sigma_2, \ldots, \sigma_n),$$

the monomial w_1 will not disappear. Hence, this sum is nonzero, which contradicts our assumption. □

Remark 3.93. According to Remark 3.92, for any m we have

$$\deg F_m = \deg_{x_1} u_m \leq \deg_{x_1} u_1 = \deg_{x_1} f (= k_1).$$

Thus,

$$\deg F = \deg_{x_1} f. \tag{3.19}$$

The proof of the above theorem provides an algorithm for expressing a given symmetric polynomial as a polynomial in $\sigma_1, \sigma_2, \ldots, \sigma_n$.

Example 3.94. We express here the polynomial

$$f = s_3 = x_1^3 + x_2^3 + \cdots + x_n^3$$

as a polynomial in $\sigma_1, \sigma_2, \ldots, \sigma_n$. Our calculations are collected in the following table:

m	u_m	$F_m(\sigma_1, \sigma_2, \ldots, \sigma_n)$	f_m
1	x_1^3	$\sigma_1^3 = \sum_i x_i^3 + 3\sum_{i \neq j} x_i^2 x_j + 6\sum_{i<j<k} x_i x_j x_k$	$-3\sum_{i \neq j} x_i^2 x_j - 6\sum_{i<j<k} x_i x_j x_k$
2	$-3x_1^2 x_2$	$-3\sigma_1\sigma_2 = -3\sum_{i \neq j} x_i^2 x_j - 9\sum_{i<j<k} x_i x_j x_k$	$3\sum_{i<j<k} x_i x_j x_k$
3	$3x_1 x_2 x_3$	$3\sigma_3 = 3\sum_{i<j<k} x_i x_j x_k$	0

Therefore,

$$s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3. \tag{3.20}$$

There exists a more practical approach to homogeneous symmetric polynomials, which we will explain in the next example.

Example 3.95. Here we express the polynomial

$$f = (x_1x_2 + x_3x_4)(x_1x_3 + x_2x_4)(x_1x_4 + x_2x_3)$$

from Example 3.86 as a polynomial in $\sigma_1, \sigma_2, \sigma_3, \sigma_4$. In the notation of the proof of Theorem 3.89, we have $u_1 = x_1^3x_2x_3x_4$. Without any calculation we can find the candidates for $u_2, u_3, \ldots$ up to coefficients. First of all, their exponents should satisfy the inequalities in Lemma 3.90. Second, since f is homogeneous of degree 6, the sum of the exponents must equal 6. Third, they must be less than u_1. We present the possible exponents of monomials that satisfy these conditions in the table. They are ordered lexicographically, starting with the exponents in u_1. In the right column we put the corresponding products of elementary functions in accordance with (3.18):

3	1	1	1	$\sigma_1^2\sigma_4$
2	2	2	0	σ_3^2
2	2	1	1	$\sigma_2\sigma_4$

Thus, we can claim that

$$f = \sigma_1^2\sigma_4 + a\sigma_3^2 + b\sigma_2\sigma_4.$$

In order to find the coefficients a and b, let us assign some values to the variables x_1, x_2, x_3, x_4. The calculations are collected in the following table with resulting equations in the rightmost column:

x_1	x_2	x_3	x_4	σ_1	σ_2	σ_3	σ_4	f	
1	1	1	0	3	3	1	0	1	$a = 1$
1	1	−1	−1	0	−2	0	1	8	$-2b = 8$

Thus, $a = 1$ and $b = -4$. We conclude that

$$f = \sigma_1^2\sigma_4 + \sigma_3^2 - 4\sigma_2\sigma_4.$$

In the case of a nonhomogeneous symmetric polynomial, the above method can be applied to each homogeneous component; then we add up the expressions that we get.

Remark 3.96. Without any changes, this theory can be carried over to the more general case of a commutative associative ring K with unity. For instance, in the case $K = \mathbb{Z}$, we obtain the following result: every symmetric polynomial with integer coefficients can be presented as a polynomial with integer coefficients in elementary symmetric polynomials.

Theorem 3.89 together with Viète's formulas allows us to find any symmetric polynomial in the roots of a given algebraic equation. Let $f \in$

$K[x_1, x_2, \ldots, x_n]$ be a symmetric polynomial and $F \in K[X_1, X_2, \ldots, X_n]$ a polynomial such that

$$f = F(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

Now, let $c_1, c_2, \ldots, c_n$ be the roots of the following algebraic equation:

$$a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = 0, \qquad a_0 \neq 0.$$

Then

$$f(c_1, c_2, \ldots, c_n) = F\left(-\frac{a_1}{a_0}, \frac{a_2}{a_0}, \ldots, (-1)^n \frac{a_n}{a_0}\right). \tag{3.21}$$

Remark 3.97. Let $\deg_{x_1} f = k$. Then $\deg F = k$ (see Remark 3.93). By multiplying equality (3.21) by a_0^k, we obtain on the right a homogeneous polynomial in $a_0, a_1, a_2, \ldots, a_n$ of degree k.

Example 3.98. Let c_1, c_2, c_3, c_4 be the roots of the equation

$$x^4 + px^2 + qx + r = 0. \tag{3.22}$$

We will find the equation of the third degree whose roots are the numbers

$$d_1 = c_1c_2 + c_3c_4, \qquad d_2 = c_1c_3 + c_2c_4, \qquad d_3 = c_1c_4 + c_2c_3.$$

It has the form

$$y^3 + a_1y^2 + a_2y + a_3 = 0.$$

According to Viète's formulas

$$a_1 = -(d_1 + d_2 + d_3), \quad a_2 = d_1d_2 + d_1d_3 + d_2d_3, \quad a_3 = -d_1d_2d_3.$$

We have $d_i = h_i(c_1, c_2, c_3, c_4)$, where h_1, h_2, h_3 are the polynomials of Example 3.86. Then

$$h_1 + h_2 + h_3 = \sigma_2,$$

$$h_1h_2 + h_1h_3 + h_2h_3 = \sum_{i \neq j,k, j<k} x_i^2 x_j x_k = \sigma_1\sigma_3 - 4\sigma_4,$$

$$h_1h_2h_3 = \sigma_1^2\sigma_4 + \sigma_3^2 - 4\sigma_2\sigma_4.$$

(The last equality is the result of Example 3.95.) Viète's formulas give

$$\sigma_1(c_1, c_2, c_3, c_4) = 0,$$

$$\sigma_2(c_1, c_2, c_3, c_4) = p,$$

$$\sigma_3(c_1, c_2, c_3, c_4) = -q,$$

$$\sigma_4(c_1, c_2, c_3, c_4) = r.$$

Therefore,

$$a_1 = -p, \qquad a_2 = -4r, \qquad a_3 = 4pr - q^2,$$

i.e., our equation has the form

$$y^3 - py^2 - 4ry + (4pr - q^2) = 0. \tag{3.23}$$

Exercise 3.99. Prove that in the notation of Example 3.98,

$$(c_1 + c_2 - c_3 - c_4)^2 = 4(d_1 - p),$$
$$(c_1 - c_2 + c_3 - c_4)^2 = 4(d_2 - p),$$
$$(c_1 - c_2 - c_3 + c_4)^2 = 4(d_3 - p)$$

and

$$(c_1 + c_2 - c_3 - c_4)(c_1 - c_2 + c_3 - c_4)(c_1 - c_2 - c_3 + c_4) = -8q \tag{3.24}$$

(see Exercise 3.87).

Using the results of this exercise, we can reduce solving equation (3.22) to solving equation (3.23) (assuming that $\operatorname{char} K \neq 2$). Namely, by adding up the following equalities taken with appropriate signs:

$$c_1 + c_2 + c_3 + c_4 = 0,$$
$$c_1 + c_2 - c_3 - c_4 = 2\sqrt{d_1 - p},$$
$$c_1 - c_2 + c_3 - c_4 = 2\sqrt{d_2 - p},$$
$$c_1 - c_2 - c_3 + c_4 = 2\sqrt{d_3 - p},$$

we obtain that

$$c_{1,2,3,4} = \frac{1}{2}\left(\pm\sqrt{d_1 - p} \pm \sqrt{d_2 - p} \pm \sqrt{d_3 - p}\right).$$

Here the number of minuses must be even. The values of square roots should be chosen so that their product is equal to $-q$ (see equation (3.24)).

Equation (3.23) is called the *cubic resolution* of equation (3.22).

3.9. Cubic Equations

In solving a quadratic equation, a major role is played by the discriminant. Its turning into zero indicates a multiple root, and its sign (in the case of the field of real numbers), the number of real roots.

Let us explain the meaning of the discriminant $D(\varphi)$ of a quadratic polynomial

$$\varphi = a_0x^2 + a_1x + a_2 \in \mathbb{C}[x].$$

Let c_1, c_2 be its roots. Then

$$D(\varphi) = a_1^2 - 4a_0a_2 = a_0^2\left[\left(\frac{a_1}{a_0}\right)^2 - \frac{4a_2}{a_0}\right] = a_0^2[(c_1+c_2)^2 - 4c_1c_2] = a_0^2(c_1-c_2)^2.$$

When $a_0, a_1, a_2 \in \mathbb{R}$, this formula explains very well the connection between the discriminant and the properties of the roots that we mentioned above. Namely, three possibilities can occur:

(i) $c_1, c_2 \in \mathbb{R}$, $c_1 \neq c_2$; then $c_1 - c_2$ is a nonzero real number and $D(\varphi) > 0$;

(ii) $c_1 = c_2 \in \mathbb{R}$; then $c_1 - c_2 = 0$ and $D(\varphi) = 0$;

(iii) $c_1 = \bar{c}_2 \notin \mathbb{R}$; then $c_1 - c_2$ is a nonzero imaginary number and $D(\varphi) < 0$.

More importantly, this formula shows how to determine the discriminant of an arbitrary polynomial

$$\varphi = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n \in K[x], \qquad a_0 \neq 0.$$

Assume first that the polynomial φ has n roots $c_1, c_2, \ldots, c_n \in K$. We define its *discriminant* $D(\varphi)$ by the formula

$$D(\varphi) = a_0^{2n-2} \prod_{i>j} (c_i - c_j)^2. \tag{3.25}$$

(The exponent of a_0 does not matter much; it will become clear later why we chose this one.)

In other words, $D(\varphi)$ is a product of a_0^{2n-2} and the value of the symmetric polynomial

$$f = \prod_{i>j} (x_i - x_j)^2$$

(see Example 3.85) computed at the roots of φ. The procedure described in Section 3.8 allows us to express $D(\varphi)$ via the coefficients of φ. Since

$$\deg_{x_1} f = 2n - 2,$$

by Remark 3.97, this expression will be a homogeneous polynomial Δ of degree $2n - 2$ in $a_0, a_1, \ldots, a_n$:

$$D(\varphi) = \Delta(a_0, a_1, \ldots, a_n). \tag{3.26}$$

In order to find Δ, we do not need to know that the polynomial φ has exactly n roots in K. This makes it possible to determine the discriminant of any polynomial φ by formula (3.26).

Remark 3.100. Since f has integer coefficients (see Remark 3.96), Δ has integer coefficients too.

Remark 3.101. It can be shown (see Theorem 9.114) that for any polynomial $\varphi \in K[x]$ of degree n, there exists a field extension L of K where φ has n roots. (For instance, when $K = \mathbb{R}$, one can take $L = \mathbb{C}$.) The above-described procedure for computing the discriminant does not depend on the field over which the polynomial φ is considered (provided its coefficients lie in this field). Thus, formula (3.25) is valid for $D(\varphi)$ if we take the roots of φ in L as $c_1, c_2, \ldots, c_n$.

Definition (3.25) of the discriminant clearly implies that the polynomial $\varphi \in \mathbb{C}[x]$ has multiple roots if and only if $D(\varphi) = 0$. This shows that having multiple roots is a special case: if one is to select coefficients of a polynomial randomly, the probability of it having multiple roots is zero.

Now, let φ be a cubic polynomial with real coefficients. Let c_1, c_2, c_3 be its complex roots. Then

$$D(\varphi) = a_0^4(c_1 - c_2)^2(c_1 - c_3)^2(c_2 - c_3)^2.$$

Three possibilities can occur (up to reordering the roots):

(i) c_1, c_2, c_3 are distinct real numbers; then $D(\varphi) > 0$;

(ii) $c_1, c_2, c_3 \in \mathbb{R}$, $c_2 = c_3$; then $D(\varphi) = 0$;

(iii) $c_1 \in \mathbb{R}$, $c_2 = \bar{c}_3 \notin \mathbb{R}$; then

$$\begin{aligned} D(\varphi) &= a_0^4[(c_1 - c_2)(c_1 - \bar{c}_2)]^2(c_2 - \bar{c}_2)^2 \\ &= a_0^4|c_1 - c_2|^4(c_2 - \bar{c}_2)^2 < 0. \end{aligned}$$

Thus, we come to the same conclusion as in the case of the quadratic equation: all roots of the polynomial φ are real if and only if $D(\varphi) \geq 0$.

Exercise 3.102. Let φ be a polynomial of any degree with real coefficients, and suppose it has no multiple complex roots. Prove that

$$\operatorname{sign} D(\varphi) = (-1)^t,$$

where t is the number of pairs of conjugate imaginary roots of φ.

We will explicitly express the discriminant of a cubic equation through its coefficients. First, let us make several general observations that will simplify our calculations.

Any polynomial can be made monic by dividing it by its leading coefficient. This does not change its roots. Furthermore, any monic polynomial

$$\varphi = x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n$$

over a field of zero characteristic (or, more generally, over a field whose characteristic does not divide n) can be transformed into a polynomial of the form

$$\psi = y^n + b_2y^{n-2} + \cdots + b_{n-1}y + b_n$$

by substituting

$$x = y - \frac{a_1}{n}.$$

Here the coefficient of y^{n-1} is zero. A polynomial of this type is called *depressed*. When $n = 2$, this approach leads to the formula for the roots of the quadratic equation. When $n > 2$, this is not so but at least the equation (hence, the problem) looks simpler.

Now we will find the discriminant of the depressed cubic polynomial

$$\varphi = x^3 + px + q. \tag{3.27}$$

Using the method of Example 3.95, we express the symmetric polynomial

$$f = (x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2$$

via elementary symmetric functions $\sigma_1, \sigma_2, \sigma_3$. The polynomial f is homogeneous of degree 6. Its leading monomial is $x_1^4 x_2^2$. The following table contains all possible leading monomials of symmetric polynomials that can occur in the process described in the proof of Theorem 3.89. The right column contains the corresponding products of elementary symmetric polynomials:

4	2	0	$\sigma_1^2\sigma_2^2$
4	1	1	$\sigma_1^3\sigma_3$
3	3	0	σ_2^3
3	2	1	$\sigma_1\sigma_2\sigma_3$
2	2	2	σ_3^2

We see that

$$f = \sigma_1^2\sigma_2^2 + a\sigma_1^3\sigma_3 + b\sigma_2^3 + c\sigma_1\sigma_2\sigma_3 + d\sigma_3^2. \tag{3.28}$$

To compute $D(\varphi)$, we need to make the following substitution in (3.28):

$$\sigma_1 = 0, \qquad \sigma_2 = p, \qquad \sigma_3 = -q.$$

Hence, coefficients a and c will not affect the final result and we do not have to find them.

To find b and d, let us assign values from the left column of the following table to variables x_1, x_2, x_3. The rightmost column contains the resulting equations:

x_1	x_2	x_3	σ_1	σ_2	σ_3	f	
1	−1	0	0	−1	0	4	$-b = 4$
2	−1	−1	0	−3	2	0	$-27b + 4d = 0$

Thus, $b = -4$, $d = -27$, and

$$D(\varphi) = -4p^3 - 27q^2. \tag{3.29}$$

Example 3.103. How many real roots does the polynomial

$$\varphi = x^3 - 0.3x^2 - 4.3x + 3.9$$

have? By substituting

$$y = x - 0.1,$$

we obtain the depressed polynomial

$$\psi = y^3 - 4.33y + 3.468$$

(its coefficients can be found from Horner's Scheme as in Example 3.23). Now,

$$D(\varphi) = D(\psi) = 4 \cdot 4.33^3 - 27 \cdot 3.468^2 = 0.0013 > 0.$$

Therefore, the polynomial φ has 3 distinct real roots.

Remark 3.104. The discriminant of a generic cubic equation

$$\varphi = a_0x^3 + a_1x^2 + a_2x + a_3$$

equals

$$D(\varphi) = a_1^2a_2^2 - 4a_1^3a_3 - 4a_0a_2^3 + 18a_0a_1a_2a_3 - 27a_0^2a_3^2.$$

We will explain now how to solve a cubic equation.

Assume that the base field K contains a nontrivial (i.e., different from 1) cubic root of unity; denote it ω. Then 1, ω, and ω^{-1} are all cubic roots of unity and by Viète's formula,

$$\omega + \omega^{-1} = -1. \tag{3.30}$$

Consider linear polynomials

$$h_1 = x_1 + \omega x_2 + \omega^{-1}x_3, \qquad h_2 = x_1 + \omega^{-1}x_2 + \omega x_3.$$

They interchange when x_2 and x_3 are interchanged. When x_1 and x_2 are interchanged, h_1 becomes ωh_2 and h_2 becomes $\omega^{-1}h_1$. It follows that the polynomials

$$f = h_1^3 + h_2^3, \qquad g = h_1h_2$$

are symmetric. We can express them via elementary symmetric polynomials as

$$f = 2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3, \qquad g = \sigma_1^2 - 3\sigma_2.$$

Now let c_1, c_2, c_3 be the roots of polynomial (3.27). Put

$$d_1 = c_1 + \omega c_2 + \omega^{-1}c_3, \qquad d_2 = c_1 + \omega^{-1}c_2 + \omega c_3.$$

The above implies

$$d_1^3 + d_2^3 = -27q, \qquad d_1d_2 = -3p,$$

hence

$$d_1^3d_2^3 = -27p^3.$$

Therefore, d_1^3 and d_2^3 are the roots of the quadratic equation

$$x^2 + 27qx - 27p^3 = 0.$$

By solving it, we find

$$d_1^3 = 27\left(-\frac{q}{2} + \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}\right), \tag{3.31}$$

$$d_2^3 = 27\left(-\frac{q}{2} - \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}\right). \tag{3.32}$$

Observe that the expression under the root sign differs from the discriminant of (3.27) only by the factor of $-\frac{1}{108}$.

Adding up the equalities

$$\begin{aligned} c_1 + \quad c_2 + \quad c_3 &= 0, \\ c_1 + \omega c_2 + \omega^{-1} c_3 &= d_1, \\ c_1 + \omega^{-1} c_2 + \omega c_3 &= d_2 \end{aligned}$$

and taking (3.30) into account, we obtain

$$c_1 = \frac{1}{3}(d_1 + d_2).$$

Since the order of roots is chosen arbitrarily, this formula in fact produces all three roots. We obtain them all when we assign to d_1 and d_2 different values of cubic roots from expressions (3.31) and (3.32) such that the following relation holds (we deduced it above):

$$d_1 d_2 = -3p. \tag{3.33}$$

Therefore, we obtain the following formula:

$$c_{1,2,3} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}},$$

called *Cardano's Formula*.

Remark 3.105. Cardano's Formula makes sense if it is possible to extract all of its square and cubic roots. In particular, if we use this formula to solve a cubic equation with real coefficients, in general we have to work with complex numbers even if we are interested only in real roots. This happens, for instance, in the case of a positive discriminant when all roots are real: here the number under the square root is negative.

Example 3.106. We will find here the roots of the polynomial ψ of Example 3.103. We have

$$\frac{p^3}{27} + \frac{q^2}{4} = -\frac{1}{108} D(\psi) \approx -0.0000120,$$

so under one of the cubic roots in Cardano's formula we have the following number:

$$-\frac{q}{2}+\sqrt{\frac{p^3}{27}+\frac{q^2}{4}} \approx -1.734 + 0.00347$$
$$\approx 1.73400[\cos(\pi - 0.00200) + i\sin(\pi - 0.00200)].$$

Under the other cubic root we have this number's complex conjugate. Condition (3.33) means here that when we extract cubic roots, we need to combine their conjugate values. Recall that when we add complex conjugates, we get double their real part. Therefore,

$$c_1 \approx 2\sqrt[3]{1.73400}\cos\frac{\pi - 0.00200}{3} \approx 1.20278,$$
$$c_2 \approx 2\sqrt[3]{1.73400}\cos\frac{\pi + 0.00200}{3} \approx 1.20001,$$
$$c_3 \approx -2\sqrt[3]{1.73400}\cos\frac{0.00200}{3} \approx -2.40277.$$

3.10. Field of Rational Fractions

Just as the ring of integers can be extended to the field of rational numbers, any integral domain can be extended to a field.

Let A be an integral domain. Consider the set of pairs (a, b), where $a, b \in A$, $b \neq 0$. Define an equivalence relation on it by the following rule:

$$(a_1, b_1) \sim (a_2, b_2) \iff a_1 b_2 = a_2 b_1.$$

Clearly, this relation is symmetric and reflexive; let us prove that it is transitive. If $(a_1, b_1) \sim (a_2, b_2)$ and $(a_2, b_2) \sim (a_3, b_3)$, then

$$a_1 b_2 b_3 = a_2 b_1 b_3 = a_3 b_1 b_2.$$

Cancelling b_2, we have

$$a_1 b_3 = a_3 b_1,$$

i.e., $(a_1, b_1) \sim (a_3, b_3)$.

The above definition implies that

$$(a, b) \sim (ac, bc) \tag{3.34}$$

for any $c \neq 0$. On the other hand, any equivalence $(a_1, b_1) \sim (a_2, b_2)$ is a corollary of equivalences of the form (3.34), as the following chain of equivalences demonstrates:

$$(a_1, b_1) \sim (a_1 b_2, b_1 b_2) = (a_2 b_1, b_1 b_2) \sim (a_2, b_2).$$

(We first multiplied both entries in (a_1, b_1) by b_2 and then cancelled b_1 in both entries of the resulting pair.)

Define now addition and multiplication of pairs by the following rules:

$$(a_1, b_1) + (a_2, b_2) = (a_1 b_2 + a_2 b_1, b_1 b_2),$$
$$(a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2).$$

We will prove that the equivalence relation defined above agrees with these operations. By the preceding discussion, it suffices to show that when we multiply both entries in one of the pairs (a_1, b_1) or (a_2, b_2) by the same element c, their sum and product get replaced by equivalent pairs. But it is clear that when we do this, both entries in the sum and the product are multiplied by c.

We write the equivalence class containing the pair (a, b) as a "fraction" $\frac{a}{b}$ or a/b (these are just symbols for now; they do not imply the actual operation of division). By the above, operations of addition and multiplication of fractions are performed in accordance with the following rules:

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1 b_2 + a_2 b_1}{b_1 b_2}, \qquad \frac{a_1}{b_1}\frac{a_2}{b_2} = \frac{a_1 a_2}{b_1 b_2}.$$

We will prove now that fractions form a field with respect to these operations.

Any finite set of fractions has a common denominator, and the addition of fractions with the same denominator comes down to the addition of their numerators. Therefore, addition of fractions is commutative and associative. The fraction $\frac{0}{1}$ $(= \frac{0}{b}$ for any $b \neq 0)$ serves as the zero for addition of fractions, and the fraction $-\frac{a}{b}$ is the opposite element of the fraction $\frac{a}{b}$. Thus, fractions form an abelian group with respect to their addition.

Commutativity and associativity of multiplication of fractions is obvious. The following chain of equalities proves the distributive law for the fractions:

$$\left(\frac{a_1}{b} + \frac{a_2}{b}\right)\frac{a_3}{b_3} = \frac{(a_1 + a_2)a_3}{bb_3} = \frac{a_1 a_3 + a_2 a_3}{bb_3} = \frac{a_1}{b}\frac{a_3}{b_3} + \frac{a_2}{b}\frac{a_3}{b_3}.$$

The fraction $\frac{1}{1}$ is the identity for multiplication of fractions. For $a \neq 0$, the fraction $\frac{b}{a}$ is the inverse of the fraction $\frac{a}{b}$.

The field we have just constructed is called the *quotient field* (or the *field of fractions*) of the ring A and is denoted $Q(A)$.

Addition and multiplication of fractions of the form $\frac{a}{1}$ come down to the corresponding operations on their numerators. Besides, $\frac{a}{1} = \frac{b}{1}$ only if $a = b$. Therefore, fractions of this form comprise a subring isomorphic to A. We identify a fraction of the form $\frac{a}{1}$ with the element a of A and thus obtain the embedding of the ring A into the field $Q(A)$. Moreover, since

$$\frac{a}{b}\frac{b}{1} = \frac{a}{1},$$

the fraction $\frac{a}{b}$ equals the ratio of elements a and b of the ring A in the field $Q(A)$. Thus, the notation $\frac{a}{b}$ acquires a real meaning.

In view of (3.34), a fraction does not change when its numerator and denominator are multiplied or divided (if possible) by the same element of the ring A. If A is a Euclidean domain, then by cancelling their greatest common divisor in the numerator and denominator, each fraction can be presented as $\frac{a}{b}$ with $(a,b)=1$. This form of a fraction is called *reduced*. (By abuse of the language, the fraction itself is usually called reduced.)

Proposition 3.107. *Any form of fraction over a Euclidean domain can be obtained from its reduced form by multiplying its numerator and denominator by the same element.*

Proof. Let $\frac{a}{b}=\frac{a_0}{b_0}$ with $(a_0,b_0)=1$. The equality $ab_0=a_0b$ implies that $b_0|a_0b$, hence $b_0|b$. Put $b=cb_0$; then clearly $a=ca_0$. □

Corollary 3.108. *The reduced form of a fraction over a Euclidean domain is determined uniquely up to multiplication of the numerator and denominator by the same invertible element.*

The quotient field of the ring $\mathbb{Z}$ of integers is the field $\mathbb{Q}$ of rational numbers. The quotient field of the ring $K[x]$ of polynomials over a field K is called the *field of rational fractions* (or *rational functions*) over the field K and is denoted $K(x)$.

Every rational fraction determines a K-valued function on K defined whenever its denominator (from its reduced form) is nonzero. Namely, the value of the fraction $\frac{f}{g}$, $f,g\in K[x]$, at $c\in K$ is the number $\frac{f(c)}{g(c)}$. It is easy to see that the operations of addition and multiplication of fractions correspond to similar operations on the functions that they determine (in their common domain).

Exercise 3.109. Prove that if rational fractions $\frac{f_1}{g_1}$ and $\frac{f_2}{g_2}$ over an infinite field K determine functions that coincide on their common domain, then $\frac{f_1}{g_1}=\frac{f_2}{g_2}$.

A rational fraction $\frac{f}{g}$ is called *proper* if $\deg f<\deg g$. Obviously, the sum and product of proper rational fractions are themselves proper.

Proposition 3.110. *Every rational fraction can be uniquely presented as a sum of a polynomial and a proper rational fraction.*

Proof. Let $f,g\in K[x]$, $g\neq 0$. Divide f by g with a remainder in the ring $K[x]$:

$$f=qg+r,\qquad q,r\in K[x],\quad \deg r<\deg g. \tag{3.35}$$

Then

$$\frac{f}{g}=q+\frac{r}{g}, \tag{3.36}$$

where $\frac{r}{g}$ is a proper rational fraction.

Now let

$$\frac{f}{g} = q_1 + \frac{r_1}{g_1}$$

be some other presentation of $\frac{f}{g}$ as a sum of a polynomial and a proper rational fraction. Then

$$q - q_1 = \frac{r_1}{g_1} - \frac{r}{g}$$

and we obtain a contradiction because a nonzero polynomial cannot be equal to a proper rational fraction. □

The polynomial q from equality (3.36) is called the *regular part* of the rational fraction $\frac{f}{g}$.

Proposition 3.111. *Every proper rational fraction of the form*

$$\frac{f}{g_1 g_2 \cdots g_s},$$

where $g_1, g_2, \ldots, g_s$ are pairwise relatively prime, can be presented as a sum of proper rational fractions with denominators $g_1, g_2, \ldots, g_s$.

Proof. We will prove this proposition by induction on s. For $s = 2$, by Theorem 3.60 there exist polynomials u_1 and u_2 such that $g_1 u_1 + g_2 u_2 = f$. By dividing this equality by g, we obtain

$$\frac{f}{g} = \frac{u_2}{g_1} + \frac{u_1}{g_2}.$$

Since the fraction $\frac{f}{g}$ is proper, the sum of regular parts of the fractions $\frac{u_2}{g_1}$ and $\frac{u_1}{g_2}$ must be zero. We can delete them from our expression and thus obtain the presentation of $\frac{f}{g}$ as a sum of proper rational fractions with denominators g_1 and g_2.

For $s > 2$, observe that the polynomials g_1 and $g_2 \cdots g_s$ are relatively prime; hence the first part of the proof implies that $\frac{f}{g}$ can be presented as a sum of proper rational fractions with denominators g_1 and $g_2 \cdots g_s$. By induction, the latter fraction, in turn, can be presented as a sum of proper rational fractions with denominators $g_2, \ldots, g_s$. □

Exercise 3.112. Prove that the presentation from Proposition 3.111 is unique.

We will now describe a theory used in calculus for integrating rational functions.

Definition 3.113. A rational fraction $\frac{f}{g}$ over a field K is called *primitive* if $g = p^k$, where $p \in K[x]$ is irreducible, and $\deg f < \deg p$.

In particular any fraction of the form

$$\frac{a}{(x-c)^k}, \qquad a, c \in K,$$

is primitive. In the case of $K = \mathbb{C}$, all primitive fractions are of this form. In the case of $K = \mathbb{R}$, there also exist primitive fractions of the form

$$\frac{ax+b}{(x^2+px+q)^k}, \qquad a, b, p, q \in \mathbb{R},$$

where $p^2 - 4q < 0$.

Theorem 3.114. *Every proper rational fraction $\frac{f}{g}$ can be presented as a sum of primitive fractions. More precisely, if $g = p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s}$ is the factorization of the polynomial g into irreducible factors, the fraction $\frac{f}{g}$ can be presented as a sum of primitive fractions with denominators*

$$p_1, p_1^2, \ldots, p_1^{k_1}, p_2, p_2^2, \ldots, p_2^{k_2}, \ldots, p_s, p_s^2, \ldots, p_s^{k_s}.$$

Proof. By Proposition 3.111, the fraction $\frac{f}{g}$ can be presented as a sum of proper fractions with denominators $p_1^{k_1}, p_2^{k_2}, \ldots, p_s^{k_s}$. Thus, we only have to prove the theorem for $g = p^k$, where p is an irreducible polynomial. In this case, after dividing f by p with a remainder, we obtain

$$\frac{f}{p^k} = \frac{f_1}{p^{k-1}} + \frac{r}{p^k}, \qquad \deg r < \deg p.$$

The second summand is a primitive fraction and the first is a proper rational fraction, since it is a difference of such. Continuing further, we finally present the fraction $\frac{f}{p^k}$ as a sum of primitive fractions with denominators $p, p^2, \ldots, p^k$. □

Remark 3.115. By Exercise 3.112, the presentation from the above theorem is unique.

Example 3.116. Let

$$g = (x - c_1)(x - c_2) \cdots (x - c_n),$$

where $c_1, c_2, \ldots, c_n$ are distinct. Then

$$\frac{f}{g} = \frac{a_1}{x - c_1} + \frac{a_2}{x - c_2} + \cdots + \frac{a_n}{x - c_n},$$

where $a_1, a_2, \ldots, a_n \in K$. Fix i, $1 \leq i \leq n$. To find a_i, multiply both sides of the above equation by g and let $x = c_i$. Then we obtain

$$f(c_i) = a(c_i - c_1) \cdots (c_i - c_{i-1})(c_i - c_{i+1}) \cdots (c_i - c_n) = a_i g'(c_i),$$

thus

$$a_i = \frac{f(c_i)}{g'(c_i)}.$$

Therefore,

$$\frac{f}{g} = \sum_{i=1}^{n} \frac{f(c_i)}{g'(c_i)(x - c_i)} \tag{3.37}$$

(assuming that $\deg f < \deg g$). It is interesting to notice that by multiplying both sides of this equality by g, we obtain the *Lagrange Interpolation Formula*

$$f = \sum_{i=1}^{n} b_i \frac{(x - c_1) \cdots (x - c_{i-1})(x - c_{i+1}) \cdots (x - c_n)}{(c_i - c_1) \cdots (c_i - c_{i-1})(c_i - c_{i+1}) \cdots (c_1 - c_n)}.$$

This formula defines the polynomial f of degree $< n$ that assumes the values $b_1, b_2, \ldots, b_n$ at the points $c_1, c_2, \ldots, c_n$.

Exercise 3.117. Prove the equality:

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

where $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1}$ are complex roots of unity of degree n.

Exercise 3.118. Present the fraction $\frac{1}{x^p - x}$ as a sum of primitive fractions over the field $\mathbb{Z}_p$, p prime.

Example 3.119. In the previous example, we used the method of undetermined coefficients. It can be used in a more general situation. For example, let us present the rational fraction

$$\frac{x}{(x+1)(x^2+1)^2}$$

as a sum of primitive fractions over $\mathbb{R}$. By Theorem 3.114,

$$\frac{x}{(x+1)(x^2+1)^2} = \frac{a}{x+1} + \frac{bx+c}{x^2+1} + \frac{dx+e}{(x^2+1)^2},$$

where a, b, c, d, e are some real numbers. To find them, we multiply the above equality by $(x+1)(x^2+1)^2$:

$$x = a(x^2+1)^2 + (bx+c)(x+1)(x^2+1) + (dx+e)(x+1).$$

By subsequently assigning $x = -1$ and $x = i$ in this equality, we obtain $-1 = 4a$, $i = (di+e)(i+1) = (e-d) + (d+e)i$. Thus,

$$a = -\frac{1}{4}, \qquad d = e = \frac{1}{2}.$$

Also, by comparing free terms and coefficients of x^4, we obtain $0 = a + c + e$ and $0 = a + b$, implying

$$b = \frac{1}{4}, \qquad c = -\frac{1}{4}.$$

Finally, we have

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

Chapter 4

Elements of Group Theory

4.1. Definitions and Examples

In the first chapter we introduced the concept of an abelian group. In particular, the additive group of a ring, the multiplicative group of a field, and the additive group of a vector space are all abelian groups. Most important examples of nonabelian groups come up as transformation groups.

Call a map of a set X into itself a *transformation*.

Definition 4.1. A *transformation group* of a set X is a collection G of bijective transformations of X satisfying the following conditions:

(i) if $\varphi, \psi \in G$, then $\varphi\psi \in G$;

(ii) if $\varphi \in G$, then $\varphi^{-1} \in G$;

(iii) $\mathrm{id} \in G$.

Example 4.2. The collection $S(X)$ of all bijective transformations of a set X is a transformation group. If X is infinite, this group is too big to be interesting. If X is finite, we can assume that $X = \{1, 2, \ldots, n\}$. In this case, $S(X)$ is called the *group of permutations* or the *symmetric group* on n elements and is denoted S_n. A permutation $\sigma \in S_n$ can be written as a table

$$\sigma = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix},$$

where the top row contains numbers $1, 2, \ldots, n$ in some order and the bottom row, their images, i.e., $j_k = \sigma(i_k)$. Fix the order of numbers in the

top row (e.g., order them increasingly). Then we see that the number of arrangements is $n!$ (see Section 2.4). Every permutation can be written in $n!$ ways. Here is an example of multiplication of permutations:

$$\begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix}\begin{pmatrix}1&2&3&4\\4&1&3&2\end{pmatrix} = \begin{pmatrix}4&1&3&2\\2&3&1&4\end{pmatrix}\begin{pmatrix}1&2&3&4\\4&1&3&2\end{pmatrix} = \begin{pmatrix}1&2&3&4\\2&3&1&4\end{pmatrix}.$$

(We first rewrote the second permutation so that its top row coincided with the bottom row of the first permutation.)

Example 4.3. The motions of the Euclidean plane E^2 (respectively, the Euclidean space E^3) form a group denoted Isom E^2 (respectively, Isom E^3). This fact is an axiom in the version of axioms of Euclidean geometry that takes the concept of a motion as the basic one. In the other version, where the basic notion is the distance between points, a motion is defined as a distance-preserving transformation. That the motions form a group is then an easy theorem.

Remark 4.4. In the previous chapters, we denoted by E^2 (respectively, E^3) the set of vectors on the Euclidean plane (respectively, in the Euclidean space). Here the symbol E^2 (respectively, E^3) denotes the Euclidean plane (respectively, space) itself. However, if we fix a point o on the plane (respectively, in the space) which we will call the *origin*, then we can identify points with their position vectors with respect to o. We will freely use this identification in what follows.

Remark 4.5. In the version of axioms of Euclidean geometry that takes the concept of motion as its basic concept, the statement that every bijective distance-preserving transformation is a motion is a (not so difficult) theorem.

Example 4.6. The properties of linear maps proven in Section 2.3 imply that the bijective linear maps of a vector space V form a transformation group. It is called the *general linear group* of V and is denoted $\mathrm{GL}(V)$.

Example 4.7. Call the following transformation of a vector space V,

$$t_a\colon x \mapsto x + a,$$

a *parallel translation* of V along the vector $a \in V$. It is easy to see that

$$t_a t_b = t_{a+b}, \qquad t_a^{-1} = t_{-a}, \qquad \mathrm{id} = t_0. \tag{4.1}$$

These formulas show that the collection $\mathrm{Tran}\, V$ of all parallel translations of V is a transformation group.

Exercise 4.8. Consider the collection of all strictly monotone continuous functions f on the interval $[0, 1]$ such that $f(0) = 0$, $f(1) = 1$. Prove that it is a transformation group of the interval $[0, 1]$.

Studying properties of the operation of multiplication in transformation groups, we come to the following notion of a group. The difference from the notion of an abelian group is that there is no requirement of commutativity.

Definition 4.9. A *group* is a set G with an operation of multiplication that satisfies the following properties:

(i) $(ab)c = a(bc)$ for any $a, b, c \in G$ (*associativity*);

(ii) there exists an element $e \in G$ (the *identity*) such that $ae = ea = a$ for any $a \in G$;

(iii) for any $a \in G$, there exists an element $a^{-1} \in G$ (the *inverse*) such that $aa^{-1} = a^{-1}a = e$.

A group G is called *abelian* or *commutative* if

$$ab = ba \qquad \forall a, b \in G.$$

In the above definition of a group the operation is called a multiplication. Addition is commonly used for abelian groups only (though, in principle, it does not matter how we denote and call the group operation).

Just as in the case of abelian groups, we can prove that the identity and inverse elements are unique in every group. As for division, in a non-abelian group, one should distinguish between the left and the right division. Namely, for any $a, b \in G$, the equation $ax = b$ has a unique solution $a^{-1}b$ and the equation $xa = b$ has a unique solution ba^{-1}.

In any group,

$$(ab)^{-1} = b^{-1}a^{-1}.$$

Indeed,

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aa^{-1} = e.$$

Every transformation group is a group with respect to the operation of taking composition of transformations. Indeed, associativity of this operation is well known (see Section 2.3), the identity transformation serves as the identity of the group, and the inverse element is the inverse transformation.

Example 4.10. Nonsingular square matrices of order n over a field K form a multiplicative group denoted $\mathrm{GL}_n(K)$. Since there exists a one-to-one correspondence between the square matrices of order n and linear maps of the space K^n (where nonsingular matrices correspond to invertible linear maps and matrix multiplication corresponds to composition of transformations),

the group $\mathrm{GL}_n(K)$ is isomorphic to the group $\mathrm{GL}(K^n)$ and also to the group $\mathrm{GL}(V)$ for any n-dimensional vector space V over K.

The group $\mathrm{GL}_n(K)$ is the group of invertible elements of the ring $\mathrm{L}_n(K)$ of all matrices. If A is an associative ring with unity, the set of its invertible elements is a multiplicative group as well. We denote this group as A^*. A particular example of such a group is the multiplicative group K^* of a field K (which consists of all nonzero elements of this field). Note that $K^* = \mathrm{GL}_1(K)$.

Example 4.11. According to formulas (4.1), the group $\mathrm{Tran}\, V$ is isomorphic to the additive group of the space V.

Example 4.12. A finite group can be defined by its multiplication table. For instance, the set $G = \{e, a, b, \mathrm{c}\}$ with the following multiplication table:

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

is an abelian group. Indeed, e is the identity and every element is its own inverse. Furthermore, it is easy to see that every permutation of elements a, b, c is an automorphism of G with respect to the operation defined as above. Thus, if we exclude the trivial cases involving the identity and take commutativity into account, associativity is implied by the following equalities:

$$a^2b = a(ab) = b, \qquad (ab)\mathrm{c} = a(b\mathrm{c}) = e.$$

Exercise 4.13. Prove that the set $G = \{A, B, C, D, E, F\}$ with the following multiplication table:

	A	B	C	D	E	F
A	F	E	D	C	B	A
B	C	D	E	F	A	B
C	B	A	F	E	D	C
D	E	F	A	B	C	D
E	D	C	B	A	F	E
F	A	B	C	D	E	F

is a group isomorphic to the group S_3.

Exercise 4.14. Let G be a set with associative multiplication that contains an element e (the *right identity*) such that $ae = a$ for every $a \in G$. Also

assume that for each $a \in G$, there exists an element a^{-1} (the *right inverse*) such that $aa^{-1} = e$. Prove that G is a group.

Definition 4.15. A *subgroup* of a group G is a subset $H \subset G$ that satisfies the following conditions:

(i) if $a, b \in H$, then $ab \in H$;

(ii) if $a \in H$, then $a^{-1} \in H$;

(iii) $e \in H$.

Remark 4.16. Since $aa^{-1} = e$, we can request H to be nonempty instead of using condition (iii).

Clearly, a subgroup is a group with respect to the same operation.

Comparing Definitions 4.1 and 4.15, we see that a transformation group of a set X is just a subgroup of the group $S(X)$.

Example 4.17. Let f be a polynomial in n variables. Then

$$\operatorname{Sym} f = \{\sigma \in S_n \colon f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) = f(x_1, x_2, \dots, x_n)\}$$

is a subgroup of the group S_n. Indeed, let $\sigma, \tau \in \operatorname{Sym} f$. Set $x_{\sigma(i)} = y_i$. Then

$$\begin{aligned} f(x_{\sigma\tau(1)}, x_{\sigma\tau(2)}, \dots, x_{\sigma\tau(n)}) &= f(y_{\tau(1)}, y_{\tau(2)}, \dots, y_{\tau(n)}) = f(y_1, y_2, \dots, y_n) \\ &= f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) = f(x_1, x_2, \dots, x_n). \end{aligned}$$

Other subgroup axioms are clearly satisfied. In particular, the polynomial f is symmetric if and only if $\operatorname{Sym} f = S_n$. As an example of a polynomial with a smaller but nontrivial symmetry, consider the polynomial $f = x_1x_2 + x_3x_4$ (in four variables). It is easy to see that $\operatorname{Sym} f$ consists of eight permutations that preserve the partition of the set $\{1, 2, 3, 4\}$ into two subsets $\{1, 2\}$ and $\{3, 4\}$. (The subsets can be permuted as well as elements within each of them; see also Example 4.81.)

Example 4.18. Similarly, linear transformations of the space K^n that preserve a given polynomial in n variables form a subgroup of $\mathrm{GL}_n(K)$. Linear transformations of the space $\mathbb{R}^n$ that preserve the polynomial $x_1^2 + x_2^2 + \dots + x_n^2$ are called *orthogonal*. They form a subgroup of $\mathrm{GL}_n(\mathbb{R})$ called the *orthogonal group* and denoted O_n. Since in Cartesian coordinates of E^2 (respectively, E^3), the polynomial $x^2 + y^2$ (respectively, $x^2 + y^2 + z^2$) expresses the square of the length of a vector, the orthogonal transformations of E^2 (respectively, E^3) are length-preserving linear transformations of E^2. The condition

$$\varphi = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{O}_2$$

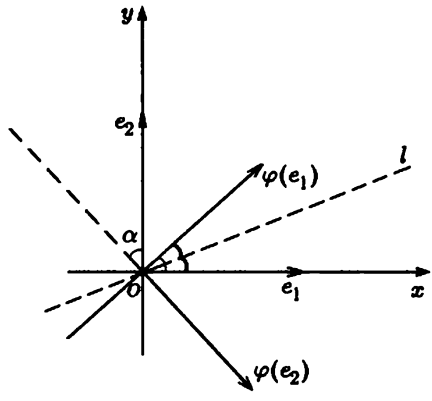


Figure 4.1

means that

$$(ax + by)^2 + (cx + dy)^2 = x^2 + y^2,$$

i.e.,

$$a^2 + c^2 = b^2 + d^2 = 1, \qquad ab + cd = 0.$$

The equation $a^2 + c^2 = 1$ implies that there exists an angle α such that

$$a = \cos\alpha, \qquad c = \sin\alpha.$$

The remaining two equations show that

$$b = \pm\sin\alpha, \qquad d = \mp\cos\alpha.$$

Therefore,

$$\varphi = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

or

$$\varphi = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

In the first case, we already know that φ is a rotation through the angle α (see Example 2.57). In the second case, φ is a reflection through a line l forming the angle $\frac{\alpha}{2}$ with the x-axis (see Figure 4.1). These two cases differ in that φ preserves the orientation of the plane in the first case and changes it in the second. We will prove in Chapter 6 that every orientation-preserving orthogonal transformation of the space E^3 is a rotation about some line.

Example 4.19. Motions of the Euclidean plane that preserve the origin o form a subgroup of the group Isom E^2. Denote it H. Since addition of vectors and their multiplication by numbers can be defined in geometric terms,

an origin-preserving motion is a linear map. Moreover, since it is distance-preserving, it is orthogonal. Conversely, since the distance between points a and b is the length of the vector $a - b$, every orthogonal transformation preserves distances between points, hence is a motion. Thus, $H = \mathrm{O}_2$. Similarly, the group of origin-preserving motions of the Euclidean space is O_3.

Example 4.20. Let F be a figure on the Euclidean plane. Then

$$\operatorname{Sym} F = \{\varphi \in \operatorname{Isom} E^2 \colon \varphi(F) = F\}$$

is a subgroup of $\operatorname{Isom} E^2$. It is called the *symmetry group* of the figure F. For instance, the symmetry group of a circle whose center is the origin o, is the group O_2. The symmetry group of a regular polygon with center at o is a subgroup of O_2. It contains rotations about o through angles that are multiples of $\frac{2\pi}{n}$ and reflections through lines passing through o and a vertex or a midpoint of an edge. Thus, this group contains $2n$ elements (n rotations and n reflections). It is called the *dihedral group* and is denoted D_n.

Example 4.21. The expression for the determinant of a product (see Theorem 2.102) implies that matrices with determinant 1 form a subgroup of the group $\mathrm{GL}_n(K)$. This subgroup is called the *unimodular group* (or *the special linear group*) and is denoted $\mathrm{SL}_n(K)$.

Example 4.22. Integer matrices with determinant 1 form a subgroup of the group $\mathrm{SL}_n(\mathbb{R})$ denoted $\mathrm{SL}_n(\mathbb{Z})$ (see Exercise 2.109).

Example 4.23. The set of nonsingular diagonal matrices of order n is an abelian subgroup of the group $\mathrm{GL}_n(K)$.

Exercise 4.24. Prove that the set of strictly upper triangular matrices of order n is a subgroup of the group $\mathrm{GL}_n(K)$.

4.2. Groups in Geometry and Physics

The goal of this section is to acquaint you with the role of groups in geometry and physics.

In the nineteenth century mathematicians realized that Euclidean geometry is not the only possible one. Even if we accept that "the space we live in" satisfies the laws of Euclidean geometry (this is true only as an approximation), it makes sense to study geometry of other spaces that appear in mathematics. Thus, the question arises: what do we mean by a geometry? The answers are different, depending on which property of Euclidean geometry we want to generalize.

In particular, one can generalize the concept of the group of motions. In his 1872 lecture that became known as the *Erlangen Program*, the German mathematician Felix Klein defined geometry as a science that studies

properties of figures invariant under the action of a given transformation group.

More precisely, let X be a set and G a group of transformations of X. Call a figure $F_1 \subset X$ *equivalent* (or, in the language of elementary geometry, *congruent*) to a figure $F_2 \subset X$ with respect to the group G if there exists $\varphi \in G$ such that $F_2 = \varphi(F_1)$. The notation is $F_1 \overset{G}{\sim} F_2$. Let us check that this is indeed an equivalence relation:

(i) $F \overset{G}{\sim} F$ because $F = \mathrm{id}(F)$ and $\mathrm{id} \in G$;

(ii) if $F_1 \overset{G}{\sim} F_2$, i.e., $F_2 = \varphi(F_1)$ for some $\varphi \in G$, then $F_2 \overset{G}{\sim} F_1$ because $F_1 = \varphi^{-1}(F_2)$ and $\varphi^{-1} \in G$;

(iii) if $F_1 \overset{G}{\sim} F_2$ and $F_2 \overset{G}{\sim} F_3$, i.e., $F_2 = \varphi(F_1)$ and $F_3 = \psi(F_2)$ for $\varphi, \psi \in G$, then $F_1 \overset{G}{\sim} F_3$ because $F_3 = \psi\varphi(F_1)$ and $\psi\varphi \in G$.

So we see that the three axioms of an equivalence relation correspond precisely to the three axioms of a transformation group. One of the principal problems of geometry is to find the necessary and sufficient conditions for equivalence of figures (recall triangle congruence theorems in Euclidean geometry). For this we have to consider quantities that are invariant under the action of the group G (such as distances between points or the measure of an angle in Euclidean geometry). Relations between these quantities are provided by geometric theorems (such as, for instance, the Pythagorean theorem or the theorem stating that the medians of a triangle intersect at one point).

Of course, not every transformation group leads to a geometry which is interesting and also important for some applications. All such geometries are connected to quite rich transformation groups, and there are not many of them. The minimal condition here is transitivity.

Definition 4.25. A transformation group G of a set X is called *transitive* if for any $x, y \in X$ there exists a transformation $\varphi \in G$ such that $y = \varphi(x)$.

(This means that in the corresponding geometry all points are equivalent in the sense of the definition of equivalent figures given above.)

Example 4.26. The group $\mathrm{Tran}\, V$ of parallel translations of a vector space V (see Example 4.7) is transitive. Indeed, for any $x, y \in V$, we have

$$y = t_{y-x}x.$$

However, the group $\mathrm{Tran}\, V$ is still too small to define an interesting geometry. An example of an interesting geometry that is different from Euclidean one is affine geometry.

Let V be a vector space, $\varphi \in \mathrm{GL}(V)$, and $a \in V$. Then

$$\varphi t_a \varphi^{-1} = t_{\varphi(a)}. \tag{4.2}$$

Indeed, for every $x \in V$, we have

$$(\varphi t_a \varphi^{-1})(x) = \varphi(\varphi^{-1}(x) + a) = x + \varphi(a) = t_{\varphi(a)} x.$$

Proposition 4.27. *For each subgroup $G \subset \mathrm{GL}(V)$, the set*

$$\mathrm{Tran}\, V \cdot G = \{t_a \varphi \colon a \in V, \varphi \in G\}$$

is a transitive transformation group of the space V.

Proof. For $a, b \in V$, $\varphi, \psi \in \mathrm{GL}(V)$, formulas (4.1) and (4.2) imply

$$(t_a \varphi)(t_b \psi) = t_a (\varphi t_b \varphi^{-1}) \varphi\psi = t_{a+\varphi(b)} \varphi\psi \in \mathrm{Tran}\, V \cdot G.$$

It follows that

$$(t_a \varphi)^{-1} = t_{-\varphi^{-1}(a)} \varphi^{-1} \in \mathrm{Tran}\, V \cdot G.$$

Therefore, $\mathrm{Tran}\, V \cdot G$ is a transformation group. It is transitive because its subgroup $\mathrm{Tran}\, V$ is transitive. □

In particular, we can take $G = \mathrm{GL}(V)$. The resulting group

$$\mathrm{GA}(V) = \mathrm{Tran}\, V \cdot \mathrm{GL}(V) \tag{4.3}$$

is called the *full affine group* of V, and its elements, (bijective) *affine transformations*. The corresponding geometry is called *affine geometry*.

In the case of $V = E^2$, we obtain affine geometry of the Euclidean plane.

Proposition 4.28. *The group of motions of the Euclidean plane is a subgroup of the group $\mathrm{GA}(E^2)$ equal to $\mathrm{Tran}\, E^2 \cdot \mathrm{O}_2$.*

Proof. First of all, observe that all parallel translations and orthogonal transformations are motions. Now pick a motion f. Let $a = f(o)$. Then the motion $\varphi = t_a^{-1} f$ does not move the point o and thus belongs to the group O_2 (see Example 4.19). Therefore,

$$f = t_a \varphi \in \mathrm{Tran}\, E^2 \cdot \mathrm{O}_2.$$

□

The group of motions of the Euclidean space is described similarly.

Corollary 4.29. *If figures $F_1, F_2 \subset E^2$ are congruent in Euclidean geometry, they are congruent in affine geometry.*

The group $\mathrm{GA}(E^2)$ is larger than the group of motions. An example of an affine transformation which is not a motion is a homothety (with a coefficient $\neq \pm 1$) or a contraction along an axis. Thus, the group $\mathrm{GA}(E^2)$ is richer than the group of motions, and figures that are not congruent in Euclidean geometry may become congruent in affine geometry. For instance, all circles are congruent in affine geometry.

Exercise 4.30. Prove that in affine geometry all triangles are congruent.

Affine geometry lacks the notion of distance between two points. However, as the following exercise demonstrates, in affine geometry, there exists an invariant of three points lying on the same line.

Exercise 4.31. Prove that an affine transformation preserves the ratio at which a point divides an interval.

Affine geometry of the Euclidean space is defined similarly.

Within the framework of the transformation group approach to geometry, it is possible to construct projective, conformal, Lobachevsky, and other geometries used in mathematics and its applications.

In physics, transformation groups describe the symmetry of physical laws, in particular, the spacetime symmetry.

A point in spacetime is given by its three spatial coordinates x, y, z and the temporal coordinate t, so that the spacetime with a fixed frame of reference can be identified with $\mathbb{R}^4$. A transition to another frame of reference gives a transformation of the space $\mathbb{R}^4$. In classical, as well as relativistic, mechanics (more precisely, in special relativity), there exists the notion of inertial frames of reference. In them, all mechanical laws have the same form. Transitions from one inertial frame to others form a group of transformations of $\mathbb{R}^4$. This group determines the laws of physics uniquely. The difference between classical and relativistic mechanics is that they consider different transformation groups.

The symmetry group of spacetime in classical mechanics is the *Galileo group* defined as follows:

$$G = \mathrm{Tran}\,\mathbb{R}^4 \cdot H \cdot \mathrm{O}_3,$$

where O_3 is the group of orthogonal transformations of spatial coordinates and H, the group of transformations of the form

$$(x, y, z, t) \mapsto (x + at, y + bt, z + ct, t)$$

corresponding to transitions to a new frame of reference that moves at a steady speed and in a straight line with respect to the old one. This description of the Galileo group shows that in classical mechanics time is absolute

in the sense that the difference of temporal coordinates of two events is the same in every inertial frame of reference.

According to the principles of relativistic mechanics, the symmetry group of spacetime is the *Poincaré group*

$$G = \operatorname{Tran} \mathbb{R}^4 \cdot \mathrm{O}_{3,1},$$

where $\mathrm{O}_{3,1}$ is the group of linear transformations preserving the polynomial

$$x^2 + y^2 + z^2 - t^2$$

(in the unit system where the speed of light is 1). The group $\mathrm{O}_{3,1}$ contains the group O_3 that does not act on the temporal coordinate. Nontrivial examples of transformations from $\mathrm{O}_{3,1}$ are the *Lorentz transformations*

$$(x, y, z, t) \mapsto (x \cosh a + t \sinh a, y, z, x \sinh a + t \cosh a),$$

which mix the temporal and the spacial coordinates. Their form shows that time is not absolute in relativistic mechanics.

The Poincaré group first appeared in the works of Lorentz and Poincaré as the symmetry group of the laws of electrodynamics (Maxwell's laws). Einstein's role was in stating boldly that mechanical laws should have the same symmetry group.

Transformation groups also lie at the foundation of crystallography and the theory of elementary particles. For instance, in crystallography they describe the symmetries of crystal structures, thus, of the physical properties of crystals (for crystal structures of table salt, diamond, and graphite, see Figure 4.2).

4.3. Cyclic Groups

Just as in the group $\mathbb{R}^*$, the integer *powers* of an element $g \in G$ can be defined in any group G as

$$g^k = \begin{cases} \underbrace{gg\cdots g}_{k} & \text{if } k > 0, \\ e & \text{if } k = 0, \\ \underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{-k} & \text{if } k < 0. \end{cases}$$

The following property holds:

$$g^k g^l = g^{k+l}. \tag{4.4}$$

This is clear for $k, l > 0$. Consider the case of $k > 0$, $l < 0$, $k + l > 0$. Then

$$g^k g^l = \underbrace{gg\cdots g}_{k}\underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{-l} = \underbrace{gg\cdots g}_{k+l} = g^{k+l}.$$

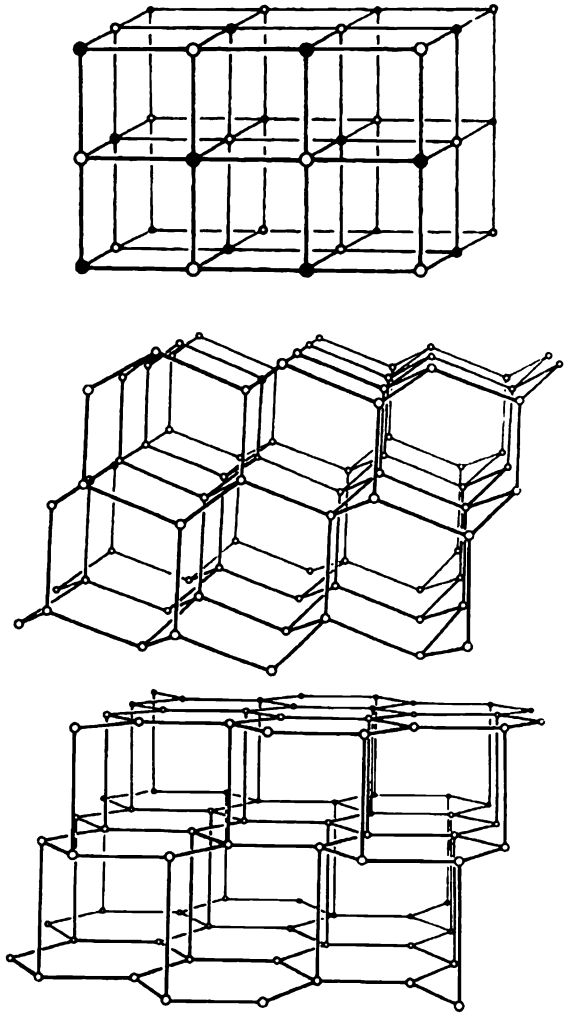


Figure 4.2

Other cases are proved similarly.

Equality (4.4) implies that

$$(g^k)^{-1} = g^{-k}.$$

Furthermore, by definition $e = g^0$. Thus, the powers of an element g form a subgroup of the group G. It is called the *cyclic subgroup generated by* g and is denoted $\langle g \rangle$.

Two distinct cases are possible: either all powers of g are different or not. In the first case, the subgroup $\langle g \rangle$ is infinite. Let us concentrate on the second case for now.

Let $g^k = g^l$, $k > l$. Then $g^{k-l} = e$. The least natural number n such that $g^n = e$ is called the *order* of the element g and is denoted $\operatorname{ord} g$.

Proposition 4.32. *If* $\operatorname{ord} g = n$, *then*

(i) $g^m = e \iff n|m$;

(ii) $g^k = g^l \iff k \equiv l \pmod{n}$.

Proof. (i) Divide m by n with a remainder:

$$m = qn + r, \qquad 0 \le r < n.$$

Then by the definition of the order,

$$g^m = (g^n)^q \cdot g^r = g^r = e \quad \iff \quad r = 0.$$

(ii) By the above,

$$g^k = g^l \quad \iff \quad g^{k-l} = e \quad \iff \quad n|(k-l) \quad \iff \quad k \equiv l \pmod{n}.$$

□

Corollary 4.33. *If* $\operatorname{ord} g = n$, *then the subgroup* $\langle g \rangle$ *contains* n *elements.*

Proof. Indeed,

$$\langle g \rangle = \{e, g, g^2, \ldots, g^{n-1}\} \tag{4.5}$$

and all these elements are distinct. □

When there does not exist a natural n such that $g^n = e$ (i.e., when we are in the first of the two cases), we put $\operatorname{ord} g = \infty$. Observe that $\operatorname{ord} e = 1$ and the orders of other group elements are greater than 1.

In an additive group, we do not speak about powers of an element g but rather about its *multiples* and denote them kg. Thus, the order of an element g of an additive group G is the least natural n such that

$$ng := \underbrace{g + g + \cdots + g}_{n} = 0$$

(if such n exists).

Example 4.34. The characteristic of a field (see Section 1.6) is the order of any nonzero element in its additive group.

Example 4.35. Obviously, in a finite group the order of any element is finite. Here we will show how to calculate the orders of elements of the group S_n. A permutation $\tau \in S_n$ is called a *cycle* of length p if it cyclically permutes the numbers $i_1, i_2, \ldots, i_p$, i.e., if $\tau(i_1) = i_2$, $\tau(i_2) = i_3, \ldots, \tau(i_p) = i_1$, and does not permute the other numbers. The notation is $(i_1 i_2 \ldots i_p)$.

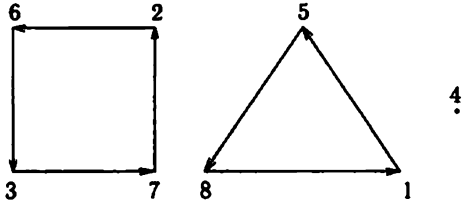


Figure 4.3

Clearly, the order of a cycle of length p equals p. Cycles τ_1 and τ_2 are called *disjoint* if the sets of numbers that they actually permute have an empty intersection. In this case, $\tau_1\tau_2 = \tau_2\tau_1$. Any permutation decomposes into a product of disjoint cycles uniquely. For example,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 4 & 8 & 3 & 2 & 1 \end{pmatrix} = (2637)(158),$$

as shown in Figure 4.3, where the arrows show the action of σ. If a permutation σ decomposes into a product of disjoint cycles of lengths $p_1, p_2, \ldots, p_s$, then

$$\operatorname{ord}\sigma = \operatorname{GCD}\{p_1, p_2, \ldots, p_s\}.$$

For example, for the permutation σ in Figure 4.3, $\operatorname{ord}\sigma = 12$.

Exercise 4.36. Prove that the order of any element of the group S_n does not exceed

$$e^{n/e} \approx 1.44^n.$$

Example 4.37. The order of a complex number c in the group $\mathbb{C}^*$ is finite if and only if this number is a root of unity. This happens if and only if $|c| = 1$ and $\arg c$ is commensurable with π, i.e., $\frac{\arg c}{\pi} \in \mathbb{Q}$.

Exercise 4.38. Prove that $\tan^{-1}\frac{3}{4}$ is not commensurable with π.

Example 4.39. Let us find the elements of finite order in the group $\operatorname{Isom} E^2$ of planar motions. Let $\varphi \in \operatorname{Isom} E^2$, $\varphi^n = \mathrm{id}$. For any point $p \in E^2$, the points

$$p, \varphi p, \varphi^2 p, \ldots, \varphi^{n-1} p$$

are cyclically permuted by φ, so their center of mass o is fixed under φ. Therefore, φ is either a rotation about o through an angle $\frac{2\pi k}{n}$ for some k or a reflection through a line passing through o.

Example 4.40. We will find here the order of the matrix

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

in $\mathrm{GL}_2(\mathbb{R})$. We have

$$A^2 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \qquad A^3 = -E,$$

thus

$$A^4 = -A, \qquad A^5 = -A^2, \qquad A^6 = -A^3 = E,$$

and $\operatorname{ord} A = 6$. Of course, this example was not chosen randomly: the probability that a randomly chosen matrix A in $\mathrm{GL}_2(\mathbb{R})$ has a finite order, is zero.

Proposition 4.41. *If* $\operatorname{ord} g = n$, *then*

$$\operatorname{ord} g^k = \frac{n}{(n,k)}. \tag{4.6}$$

Proof. Let

$$(n,k) = d, \qquad n = n_1 d, \qquad k = k_1 d,$$

so that $(n_1, k_1) = 1$. We have

$$(g^k)^m = e \iff n | km \iff n_1 | k_1 m \iff n_1 | m.$$

Therefore, $\operatorname{ord} g^k = n_1$. □

Definition 4.42. A group G is called *cyclic* if there exists an element $g \in G$ such that $G = \langle g \rangle$. Every such element is called a *generator* of the group G.

Example 4.43. The additive group $\mathbb{Z}$ of integers is cyclic since it is generated by the element 1.

Example 4.44. The additive group $\mathbb{Z}_n$ of residue classes modulo n is cyclic since it is generated by the element $[1]$.

Example 4.45. The multiplicative group C_n of complex roots of unity of order n is cyclic. Indeed, these roots are the numbers

$$\varepsilon_k = \cos\frac{2\pi k}{n} + \imath \sin\frac{2\pi k}{n}, \qquad k = 0, 1, \ldots, n-1.$$

Clearly, $\varepsilon_k = \varepsilon_1^k$. Therefore, the group C_n is generated by the element ε_1.

Exercise 4.46. Prove that the group $\mathbb{Z}_n^*$ of invertible elements of the ring $\mathbb{Z}_n$ (see Exercise 1.50) is cyclic for $n \leq 7$ and is not cyclic for $n = 8, 9$.

It is easy to see that in an infinite cyclic group $G = \langle g \rangle$, there are only two generators: g and g^{-1}. For instance, the only generators of the group $\mathbb{Z}$ are 1 and -1.

We call the number of elements of a finite group G its *order* and denote it $|G|$. The order of a finite cyclic group equals the order of its generator. Proposition 4.41 thus implies

Proposition 4.47. *The element g^k of a cyclic group $G = \langle g \rangle$ of order n is a generator if and only if $(n, k) = 1$.*

Example 4.48. The generators of the group C_n (see Example 4.45) are called the *nth primitive roots* of unity. These are the roots of the form ε_k for $(n, k) = 1$. For instance, the 12th primitive roots of unity are ε_1, ε_5, ε_7, ε_{11}.

Cyclic groups are the simplest groups imaginable. (In particular, they are abelian.) The following theorem describes them completely.

Theorem 4.49. *Every infinite cyclic group is isomorphic to the group $\mathbb{Z}$. Every finite cyclic group of order n is isomorphic to the group $\mathbb{Z}_n$.*

Proof. If $G = \langle g \rangle$ is an infinite cyclic group, then by (4.4), the map $f\colon \mathbb{Z} \to G$, $k \mapsto g^k$, is an isomorphism.

Now, let $G = \langle g \rangle$ be a finite cyclic group of order n. Consider the map

$$f\colon \mathbb{Z}_n \to G, \qquad [k] \mapsto g^k \qquad (k \in \mathbb{Z}).$$

Since

$$[k] = [l] \iff k \equiv l \,(\mathrm{mod}\, n) \iff g^k = g^l,$$

the map is well defined and bijective. The equality

$$f(k + l) = f(k) f(l)$$

follows from (4.4) as well. Thus, f is an isomorphism. □

To understand the structure of a group, it is important to know its subgroups. All subgroups of a cyclic group can be easily described.

Theorem 4.50. *(i) Every subgroup of a cyclic group is cyclic.*

(ii) In a cyclic subgroup of order n, the order of any subgroup divides n. For any divisor q of n, there exists one and only one subgroup of order q.

Proof. (i) Let $G = \langle g \rangle$ be a cyclic group and H its subgroup, different from $\{e\}$. (The trivial subgroup is obviously cyclic.) Observe that if $g^{-m} \in H$ for some $m \in \mathbb{N}$, then $g^m \in H$ as well. Let m be the least natural number

such that $g^m \in H$. We will prove that $H = \langle g^m \rangle$. Let $g^k \in H$. Divide k by m with a remainder:

$$k = qm + r, \qquad 0 \le r < m.$$

Then

$$g^r = g^k (g^m)^{-q} \in H,$$

hence $r = 0$ by definition of m. Thus, $g^k = (g^m)^q$.

(ii) If $|G| = n$, the above deduction applied to $k = n$ (in this case $g^k = e \in H$) shows that $n = qm$. Also,

$$H = \{e, g^m, g^{2m}, \dots, g^{(q-1)m}\} \tag{4.7}$$

and H is the only subgroup of order q in G. Conversely, if q is a divisor of n and $n = qm$, then the subset H defined by (4.7) is a subgroup of order q. □

Corollary 4.51. *In a cyclic subgroup of a prime order, every nontrivial subgroup coincides with the whole group.*

Example 4.52. In the group $\mathbb{Z}$, every subgroup is of the form $m\mathbb{Z}$, where $m \ge 0$.

Example 4.53. In the group C_n of roots of unity of order n, every subgroup is a group C_q of roots of unity of order q for some q such that $q|n$.

4.4. Generating Sets

Let S be a subset of a group G. Denote by $\langle S \rangle$ the collection of all products of the form

$$g_1^{\varepsilon_1} g_2^{\varepsilon_2} \cdots g_k^{\varepsilon_k}, \qquad g_1, g_2, \dots, g_k \in S; \quad \varepsilon_1, \varepsilon_2, \dots, \varepsilon_k = \pm 1. \tag{4.8}$$

This is the smallest subgroup of G that contains S. Indeed, if a subgroup contains S, then it should contain all products of the form (4.8). Conversely, $\langle S \rangle$ itself is a subgroup as the following equalities demonstrate:

$$(g_1^{\varepsilon_1} g_2^{\varepsilon_2} \cdots g_k^{\varepsilon_k})(g_{k+1}^{\varepsilon_{k+1}} g_{k+2}^{\varepsilon_{k+2}} \cdots g_{k+l}^{\varepsilon_{k+l}}) = g_1^{\varepsilon_1} g_2^{\varepsilon_2} \cdots g_{k+l}^{\varepsilon_{k+l}},$$
$$(g_1^{\varepsilon_1} g_2^{\varepsilon_2} \cdots g_k^{\varepsilon_k})^{-1} = g_k^{-\varepsilon_k} \cdots g_2^{-\varepsilon_2} g_1^{-\varepsilon_1}.$$

We say that $\langle S \rangle$ is a *subgroup generated by* S. In particular, if S consists of just one element g, then $\langle S \rangle = \langle g \rangle$ is a cyclic subgroup generated by the element g (as defined in the previous section).

Remark 4.54. It is convenient to assume that the products (4.8) contain the empty product ($k = 0$) that is equal to e by definition.

Definition 4.55. A group G is *generated* by a set S or, equivalently, S is a *generating set* of G if $G = \langle S \rangle$.

Of course, any group G is generated by the subset $S = G$. However, we should look for smaller generating sets.

Example 4.56. The dihedral group D_n (see Example 4.20) is generated by the rotation φ through the angle $\frac{2\pi}{n}$ and any reflection $\psi \in D_n$. Indeed, φ generates the cyclic subgroup C_n of all rotations in D_n. Multiplying elements of this subgroup by ψ, we obtain all reflections in D_n.

The following two theorems contain important examples of generating sets.

A permutation which is a cycle of length 2 (see Example 4.35) is called a *transposition*.

Theorem 4.57. *The group S_n is generated by transpositions.*

Proof. Notice that every transposition is the inverse of itself. Thus, the theorem says that every permutation decomposes into a product of transpositions.

When multiplying a permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ k_1 & k_2 & \dots & k_n \end{pmatrix} \tag{4.9}$$

by the transposition (ij) from the left, we exchange i and j in the bottom row. Recall that such an operation itself is called a transposition. Obviously, by applying successive transpositions, we can reduce any permutation $(k_1, k_2, \dots, k_n)$ to the trivial one: first, if $k_1 \neq 1$, exchange k_1 and 1 and put 1 into the first place, then similarly put 2 into the second place, etc. Therefore, there exist transpositions $\tau_1, \tau_2, \dots, \tau_s$ such that

$$\tau_s \cdots \tau_2 \tau_1 \sigma = \mathrm{id}\,.$$

Hence,

$$\sigma = \tau_1 \tau_2 \cdots \tau_s.$$

□

Exercise 4.58. Prove that the group S_n is generated by *adjacent transpositions* $(12), (23), \dots, (n-1\,n)$ and that the minimum number of adjacent transpositions required to produce a permutation $\sigma \in S_n$ as their product equals the number of inversions in the bottom row of the standard form (4.9) of σ.

Theorem 4.59. *The group $\mathrm{GL}_n(K)$ is generated by elementary matrices.*

(For the definition of elementary matrices, see Section 2.1.)

Proof. Recall that the inverse of an elementary matrix is also an elementary matrix (see Section 2.1). Thus, the theorem says that every nonsingular matrix decomposes into a product of elementary matrices.

By multiplying a matrix $A \in \mathrm{GL}_n(K)$ by an elementary matrix from the left, we perform an elementary row transformation. We know that any nonsingular matrix can be reduced to the identity matrix by elementary row transformations. I.e., there exist elementary matrices $U_1, U_2, \ldots, U_s$ such that

$$U_s \cdots U_2 U_1 A = E.$$

Hence,

$$A = U_1^{-1} U_2^{-1} \cdots U_s^{-1}.$$

□

Exercise 4.60. Prove that the group $\mathrm{SL}_2(\mathbb{Z})$ (see Example 4.22) is generated by the matrices

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Exercise 4.61. Prove that the group of planar motions is generated by reflections. (*Hint*: prove first that any rotation or parallel translation is a product of two reflections.)

4.5. Cosets

Let G be a group and H its subgroup. We say that elements $g_1, g_2 \in G$ are *congruent modulo* H if

$$g_1^{-1} g_2 \in H, \tag{4.10}$$

i.e., if $g_2 = g_1 h$ for some $h \in H$. The notation is

$$g_1 \equiv g_2 \pmod{H}.$$

This definition generalizes that of congruence modulo n in the case of $G = \mathbb{Z}$, $H = n\mathbb{Z}$.

Let us prove that the relation of congruence modulo H is an equivalence relation:

(i) $g \equiv g \pmod{H}$ as $g^{-1} g = e \in H$;

(ii) if $g_1 \equiv g_2 \pmod{H}$, i.e., $g_1^{-1} g_2 \in H$, then $g_2 \equiv g_1 \pmod{H}$ because

$$g_2^{-1} g_1 = (g_1^{-1} g_2)^{-1} \in H;$$

(iii) if $g_1 \equiv g_2 \pmod{H}$ and $g_2 \equiv g_3 \pmod{H}$, i.e., $g_1^{-1} g_2$, $g_2^{-1} g_3 \in H$, then $g_1 \equiv g_3 \pmod{H}$ because

$$g_1^{-1} g_3 = (g_1^{-1} g_2)(g_2^{-1} g_3) \in H.$$

The classes of this equivalence are called the (*left*) *cosets* of the subgroup H in the group G. Clearly, a coset that contains an element g is of the form

$$gH = \{gh\colon h \in H\}.$$

The subgroup H itself is also one of the cosets.

Since multiplication in a group is not necessarily commutative, we obtain, in general, another equivalence relation by considering, instead of (4.10), a similar condition

$$g_2 g_1^{-1} \in H. \tag{4.11}$$

The classes of this equivalence are called the *right cosets* of the subgroup H in the group G. They are of the form

$$Hg = \{hg\colon h \in H\}.$$

Observe that the inversion $g \mapsto g^{-1}$ establishes a one-to-one correspondence between the sets of left and right cosets. Namely,

$$(gH)^{-1} = Hg^{-1}.$$

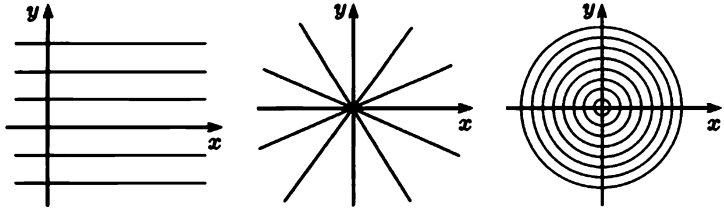


Figure 4.4

Example 4.62. On the complex plane, cosets of the subgroup $\mathbb{R}$ in the additive subgroup $\mathbb{C}$ are depicted as lines parallel to the real axis (Figure 4.4, left).

Example 4.63. Cosets of the subgroup $\mathbb{R}^*$ of positive numbers in the multiplicative group $\mathbb{C}^*$ are rays pointing from the origin (Figure 4.4, center).

Example 4.64. Cosets of the subgroup

$$\mathbb{T} = \{z \in \mathbb{C}^*\colon |z| = 1\}$$

in the group $\mathbb{C}^*$ are concentric circles with the center at the origin (Figure 4.4, right).

Example 4.65. In the case of $G = \mathrm{GL}_n(K)$, $H = \mathrm{SL}_n(K)$ (see Example 4.21), condition (4.10), as well as (4.11), translates into $\det g_1 = \det g_2$. Thus, the left and the right cosets coincide here (although $\mathrm{GL}_n(K)$ is not abelian). Each of them is a collection of matrices with a fixed value of determinant.

Example 4.66. In the group $G = S_n$, consider the subgroup H of permutations that do not move the number n. Permutations $\sigma_1, \sigma_2 \in S_n$ belong to the same left coset of H if $\sigma_1^{-1}\sigma_2(n) = n$, i.e., if

$$\sigma_1(n) = \sigma_2(n).$$

Therefore, there exist n left cosets $P_1, P_2, \ldots, P_n$, where

$$P_k = \{\sigma \in S_n \colon \sigma(n) = k\}.$$

On the other hand, permutations $\sigma_1, \sigma_2 \in S_n$ belong to the same right coset of H if $\sigma_2\sigma_1^{-1}(n) = n$, i.e., if

$$\sigma_1^{-1}(n) = \sigma_2^{-1}(n).$$

Therefore, there exist n right cosets $Q_1, Q_2, \ldots, Q_n$, where

$$Q_k = \{\sigma \in S_n \colon \sigma(k) = n\}.$$

Observe that the right and left cosets are different (except for $Q_n = P_n = H$).

The set of left cosets of H in G is denoted G/H. If finite, the number of cosets of H in G (it does not matter, left or right) is called the *index* of the subgroup H and is denoted $|G : H|$.

Theorem 4.67 (Lagrange's Theorem). *Let G be a finite group and H its subgroup. Then*

$$|G| = |G : H||H|.$$

Proof. All cosets gH contain the same number of elements, which is equal to $|H|$. Since they are equivalence classes, they do not intersect, and the order of G is equal to the product of the number of classes and $|H|$. □

Corollary 4.68. *The order of a subgroup of a finite group divides the order of the group.*

We have already proved this in the case of cyclic groups (in Theorem 4.50).

Corollary 4.69. *The order of any element of a finite group divides the order of the group.*

Proof. The order of an element equals the order of the cyclic subgroup that it generates. Thus, this corollary follows from Corollary 4.68. □

Corollary 4.70. *Every finite group of a prime order is cyclic.*

Proof. In view of Corollary 4.68, such a group coincides with the cyclic group generated by any element which is not the identity. □

Corollary 4.71. *If $|G| = n$, then $g^n = e$ for every $g \in G$.*

Proof. Let $\operatorname{ord} g = m$. By Corollary 4.69, $m|n$. Hence, $g^n = e$. □

Example 4.72. If p is a prime number, then the multiplicative group $\mathbb{Z}_p^*$ of the field $\mathbb{Z}_p$ is an abelian group of order $p-1$. Therefore, $g^{p-1} = 1$ for every $g \in \mathbb{Z}_p^*$. Hence,

$$a^{p-1} \equiv 1 \pmod{p}$$

for any integer a that is not divisible by p. The latter statement is known as *Fermat's Little Theorem*. (For a different proof, see Exercise 1.43.)

For any n, the order of the group $\mathbb{Z}_n^*$ of invertible elements of the ring $\mathbb{Z}_n$ equals the number of elements in the sequence $1, 2, \ldots, n$ that are relatively prime to n. Denote $|\mathbb{Z}_n^*|$ by $\varphi(n)$. This defines a function φ on the set of natural numbers. It is called *Euler's function*. Applying Corollary 4.71 to the group $\mathbb{Z}_n^*$, we see that

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

for any integer a which is relatively prime to n. This generalization of Fermat's little theorem is known as *Euler's Theorem*.

It is easy to see that $\varphi(125) = 125 - 25 = 100$. Thus, $2^{100} \equiv 1 \pmod{125}$. We have already obtained this result in Example 1.48 by direct computation.

Cosets arise naturally in the study of transformation groups.

Let G be a transformation group of a set X. We say that points $x, y \in X$ are equivalent with respect to G (notation: $x \overset{G}{\sim} y$) if there exists an element $g \in G$ such that $y = gx$. This is a particular case of the equivalence of figures defined in Section 4.2, thus it is an equivalence relation. The equivalence class of a point $x \in X$ is called its *orbit*. In other words, the orbit of x is the set

$$Gx = \{gx \colon g \in G\}.$$

In particular, transitive transformation groups (Definition 4.25) are groups of transformations with only one orbit.

The subgroup

$$G_x = \{g \in G \colon gx = x\}$$

is called the *stabilizer* of x.

Example 4.73. The group of motions of the Euclidean plane is transitive. The stabilizer of the origin is the orthogonal group O_2 (see Example 4.19).

Example 4.74. Orbits of the group O_2 are circles centered at the origin o, as well as o itself. The stabilizer of a point $p \neq o$ consists of the identity transformation and the reflection through the line op. The stabilizer of o is all of O_2.

Example 4.75. The group S_n is transitive on the set $\{1, 2, \ldots, n\}$. The stabilizer of n is the subgroup $H \simeq S_{n-1}$ considered in Example 4.66.

The following theorem generalizes (the first part of) Example 4.66.

Theorem 4.76. *There exists a one-to-one correspondence between an orbit Gx and the set G/G_x of cosets, which maps a point $y = gx \in Gx$ to the coset gG_x.*

Proof. For $g_1, g_2 \in G$, we have

$$g_1 \equiv g_2 \,(\mathrm{mod}\, G_x) \iff g_1^{-1} g_2 \in G_x \iff g_1^{-1} g_2 x = x \iff g_1 x = g_2 x.$$

Thus, all elements of the same coset of G_x in G map the point x to the same point. More precisely, all elements of the coset G_x map x to $y = gx$ and they are the only such elements. Thus we obtain the correspondence in the theorem's statement. □

If finite, the number of elements of an orbit Gx is called its *length* and is denoted $|Gx|$.

Corollary 4.77. *If G is a finite group, then*

$$|G| = |Gx||G_x|. \tag{4.12}$$

The above formula implies that the orders of stabilizers of all points of an orbit are equal. Actually, there exists a precise relation between the stabilizers of points of the same orbit, whether G is finite or not. We state it as an exercise.

Exercise 4.78. Prove that

$$G_{gx} = gG_x g^{-1}.$$

Example 4.79. Let $K \subset E^3$ be a cube. Consider its symmetry group

$$G = \mathrm{Sym}\, K = \{\varphi \in \mathrm{Isom}\, E^3 \colon \varphi(K) = K\}.$$

Clearly, this is a finite group. Moreover, a symmetry of a cube is fully determined by how it permutes the vertices. Thus we can view G as a transformation group of the set V of K's vertices. Since a cube is a regular polyhedron, a vertex of a cube can be mapped into any other vertex by some transformation in G, i.e., the group G is transitive on V. Therefore,

$$|G| = 8|G_v|$$

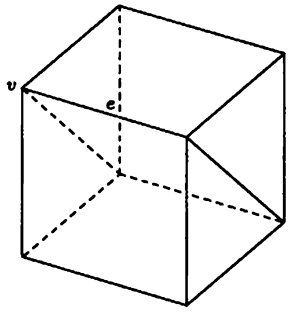


Figure 4.5

for a vertex v. Similarly, if we regard the group G_v as a transformation group of the set of edges adjacent to v, we can show that

$$|G_v| = 3|G_{v,e}|,$$

where $G_{v,e}$ is a subgroup of G_v that stabilizes the edge e. The group $G_{v,e}$ consists of the identity transformation and the reflection through the plane passing through e and the center of the cube (see Figure 4.5). Thus,

$$|\operatorname{Sym} K| = 8 \cdot 3 \cdot 2 = 48.$$

Exercise 4.80. Prove the result from Example 4.79 by treating the group $\operatorname{Sym} K$ first as a transformation group of the set of faces of the cube, and then as a transformation group of the set of edges of the cube.

Similarly, one can determine the orders of the symmetry groups of other regular polyhedra (see Figure 4.6). (For the definition of regular polyhedra, see Section 7.3).

Example 4.81. Let G be a transformation group of the polynomial algebra $K[x_1, x_2, x_3, x_4]$ consisting of all permutations of variables x_1, x_2, x_3, x_4. It is isomorphic to S_4, hence $|G| = 4! = 24$. Consider the polynomial $f = x_1x_2 + x_3x_4$. By permuting the variables, we can obtain from f the following three polynomials:

$$x_1x_2 + x_3x_4, \qquad x_1x_3 + x_2x_4, \qquad x_1x_4 + x_2x_3.$$

This implies that $|Gf| = 3$. Using (4.12), we have

$$|G_f| = \frac{|G|}{|Gf|} = \frac{24}{3} = 8.$$

Note that if we identify G with the group S_4, G_f becomes exactly the subgroup that we denoted $\operatorname{Sym} f$ in Example 4.17.

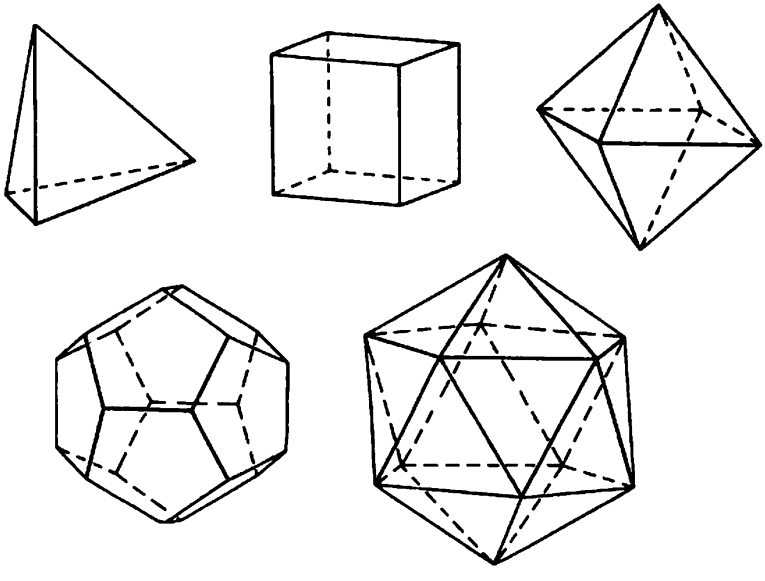


Figure 4.6

The relation of congruence modulo n in the additive group of integers agrees with the operation of addition. This allows us to define addition on the quotient set. Similarly, we can define an operation on the set of cosets of a subgroup in some other cases; although, this is not always possible.

Definition 4.82. A subgroup H of a group G is called *normal* if

$$gH = Hg \qquad \forall g \in G \tag{4.13}$$

or, equivalently,

$$gHg^{-1} = H \qquad \forall g \in G. \tag{4.14}$$

The notation is $H \lhd G$ (or $G \rhd H$).

In order for a subgroup H to be normal, it is sufficient (but not necessary) that every element of the group G commute with every element of H. In particular, every subgroup of an abelian group is normal.

Theorem 4.83. *The relation of congruence modulo a subgroup H agrees with the operation of multiplication on the group G if and only if H is normal.*

Proof. Agreement of the operation of multiplication with the relation of congruence modulo a subgroup H means the following:

$$g_1 \equiv g_1' \,(\mathrm{mod}\, H),\ g_2 \equiv g_2' \,(\mathrm{mod}\, H) \implies g_1 g_2 \equiv g_1' g_2' \,(\mathrm{mod}\, H)$$

or, equivalently,

$$(g_1 h_1)(g_2 h_2) \equiv g_1 g_2 \,(\mathrm{mod}\, H)$$

for any $g_1, g_2 \in G$ and any $h_1, h_2 \in H$. By definition, the last condition is equivalent to

$$g_2^{-1} h_1 g_2 \in H.$$

Since g_2 can be taken to be any element of G and h_1, any element of H, this is equivalent to condition (4.14) of normality. □

Exercise 4.84. Prove that every equivalence relation in a group that agrees with the group operation is the relation of congruence modulo some (normal) subgroup.

Therefore, if $H \lhd G$, the operation of multiplication on the group G defines the operation of multiplication on the set G/H as follows:

$$(g_1 H)(g_2 H) = g_1 g_2 H.$$

Associativity of the operation on group G survives the passage to the quotient. It also possesses the identity: the coset eH. Every coset gH has an inverse, namely, the coset $g^{-1}H$. Thus, G/H is a group. This group is called the *quotient group* of G by H.

Clearly, every quotient group of an abelian group is also abelian.

Example 4.85. The quotient group $\mathbb{Z}/n\mathbb{Z}$ is the group $\mathbb{Z}_n$ of residue classes.

Example 4.86. Cosets of the subgroup $\mathbb{R}$ in $\mathbb{C}$ (Example 4.62) are lines $L_a = \{z\colon \Im z = a\}$ $(a \in \mathbb{R})$. The operation of addition in $\mathbb{C}/\mathbb{R}$ is given by $L_a + L_b = L_{a+b}$, hence the quotient group $\mathbb{C}/\mathbb{R}$ is isomorphic to the group $\mathbb{R}$.

Example 4.87. Cosets of the subgroup $\mathbb{T}$ in the group $\mathbb{C}^*$ (Example 4.64) are circles $C_r = \{z \in \mathbb{C}^*\colon |z| = r\}$, $r > 0$. The operation of multiplication on $\mathbb{C}^*/\mathbb{T}$ is given by $C_r C_s = C_{rs}$, hence the quotient group $\mathbb{C}^*/\mathbb{T}$ is isomorphic to the group $\mathbb{R}_+^*$.

Example 4.88. As we have seen earlier (Example 4.65), the left cosets of the subgroup $\mathrm{SL}_n(K)$ in the group $\mathrm{GL}_n(K)$ coincide with the right cosets. They are

$$M_a = \{A \in \mathrm{GL}_n(K)\colon \det A = a\}, \qquad a \in K^*.$$

Thus, $\mathrm{SL}_n(K)$ is a normal subgroup. The operation of multiplication in the quotient group is given by $M_a M_b = M_{ab}$, hence the quotient group $\mathrm{GL}_n(K)/\mathrm{SL}_n(K)$ is isomorphic to the group K^*.

Example 4.89. In the group S_n, the subgroup H considered in Example 4.65 (it is isomorphic to S_{n-1}) is not normal whenever $n \geq 3$.

Exercise 4.90. Prove that every quotient group of a cyclic group is cyclic.

Exercise 4.91. Prove that the group of diagonal matrices is not a normal subgroup of $\mathrm{GL}_n(K)$ whenever $n \geq 2$ and $|K| \geq 3$.

4.6. Homomorphisms

Relations between different algebraic structures of the same type are established via homomorphisms. The notion of homomorphism is different from that of isomorphism as it does not require maps to be bijective. We have already encountered homomorphisms once. Namely, homomorphisms of vector spaces are precisely their linear maps.

Let us define a group homomorphism rigorously.

Definition 4.92. A *homomorphism* of a group G into a group H is a map $f\colon G \to H$ such that

$$f(ab) = f(a)f(b) \qquad \forall a, b \in G.$$

Here are several general properties of group homomorphisms:

(i) $f(e) = e$. Indeed, let $f(e) = h \in H$. Then

$$h^2 = f(e)^2 = f(e^2) = f(e) = h,$$

implying $h = e$.

(ii) $f(a^{-1}) = f(a)^{-1}$ because

$$f(a)f(a^{-1}) = f(aa^{-1}) = f(e) = e.$$

(iii) $\operatorname{Im} f = \{f(a)\colon a \in G\}$ is a subgroup of H (called the *image* of the homomorphism f). This follows from the definition of a homomorphism and the above-mentioned properties.

(iv) $\operatorname{Ker} f = \{a \in G\colon f(a) = e\}$ is a normal subgroup of G (called the *kernel* of the homomorphism f). Indeed,

$$a, b \in \operatorname{Ker} f \implies f(ab) = f(a)f(b) = e^2 = e \implies ab \in \operatorname{Ker} f,$$

$$a \in \operatorname{Ker} f \implies f(a^{-1}) = f(a)^{-1} = e^{-1} = e \implies a^{-1} \in \operatorname{Ker} f,$$

$$e \in \operatorname{Ker} f,$$

$$a \in \operatorname{Ker} f,\ g \in G \implies f(gag^{-1}) = f(g)f(a)f(g)^{-1} = f(g)ef(g)^{-1}$$
$$= f(g)f(g)^{-1} = e \implies gag^{-1} \in \operatorname{Ker} f.$$

(v) $f(g_1) = f(g_2) \iff g_1 \equiv g_2 \pmod{\operatorname{Ker} f}$; in particular, the homomorphism f is injective if and only if $\operatorname{Ker} f = \{e\}$. Indeed,

$$f(g_1) = f(g_2) \iff f(g_1^{-1}g_2) = e \iff g_1^{-1}g_2 \in \operatorname{Ker} f$$
$$\iff g_1 \equiv g_2 \pmod{\operatorname{Ker} f}.$$

Therefore, a homomorphism $f: G \to H$ is an isomorphism (i.e., a bijection) if and only if $\operatorname{Im} f = H$ and $\operatorname{Ker} f = \{e\}$. This is sometimes written as $f: G \xrightarrow{\sim} H$. If groups G and H are isomorphic (i.e., if there exists an isomorphism $f: G \xrightarrow{\sim} H$), we write $G \simeq H$.

A homomorphism of a group into itself is called an *endomorphism*.

An isomorphism of a group into itself is called an *automorphism*.

Example 4.93. Let K be a ring. The distributive law $a(b + c) = ab + ac$ implies that the map $x \mapsto ax$ (left multiplication by a) is an endomorphism of the additive group of the ring K (a similar statement holds for the right multiplication).

Example 4.94. Let G be an additive (respectively, multiplicative) abelian group. For any $n \in \mathbb{Z}$, the map $x \mapsto nx$ (respectively $x \mapsto x^n$) is an endomorphism of the group G. (In general, this is not true for a nonabelian group.) In the case $G = \mathbb{C}^*$, the kernel of this homomorphism is the group C_n of nth roots of unity.

Example 4.95. The basic identity for the exponential function implies that the map $x \mapsto e^x$ is a homomorphism of the additive group $\mathbb{R}$ to the multiplicative group $\mathbb{R}^*$. Its image is the subgroup $\mathbb{R}_+^*$ of positive numbers and its kernel is trivial.

Example 4.96. The map $x \mapsto \cos x + i \sin x$ is a homomorphism of the group $\mathbb{R}$ to the group $\mathbb{C}^*$. Its image is $\mathbb{T}$ and the kernel, $2\pi\mathbb{Z}$.

Example 4.97. The formula for multiplication of determinants implies that the map

$$\det: \mathrm{GL}_n(K) \to K^*, \qquad A \mapsto \det A$$

is a homomorphism. Its kernel is the unimodular group $\mathrm{SL}_n(K)$.

Example 4.98. For an element σ of the group S_n,

$$\sigma = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix},$$

its sign, denoted $\operatorname{sign}\sigma$, is the product of the signs of permutations in the top and bottom rows:

$$\operatorname{sign}\sigma = \operatorname{sign}(i_1, i_2, \dots, i_n) \cdot \operatorname{sign}(j_1, j_2, \dots, j_n).$$

It does not depend on the way we present σ since we can pass from any presentation to another by interchanging columns, and any such interchange reverses the signs of both the top and the bottom permutations. Hence, their product is preserved. Note that this statement is basically the same as Lemma 2.93.

The map

$$\operatorname{sign}\colon S_n \to C_2 = \{\pm 1\}, \qquad \sigma \mapsto \operatorname{sign}\sigma$$

is a homomorphism. Indeed, when multiplying a permutation σ by τ, we can assume that the top row of σ coincides with the bottom row of τ:

$$\sigma = \begin{pmatrix} j_1 & j_2 & \dots & j_n \\ k_1 & k_2 & \dots & k_n \end{pmatrix}, \qquad \tau = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}.$$

Hence,

$$\sigma\tau = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ k_1 & k_2 & \dots & k_n \end{pmatrix}$$

and

$$\begin{aligned} \operatorname{sign}\sigma\tau &= \operatorname{sign}(i_1, i_2, \dots, i_n) \cdot \operatorname{sign}(k_1, k_2, \dots, k_n) \\ &= [\operatorname{sign}(i_1, i_2, \dots, i_n) \operatorname{sign}(j_1, j_2, \dots, j_n)] \\ &\quad \times [\operatorname{sign}(j_1, j_2, \dots, j_n) \operatorname{sign}(k_1, k_2, \dots, k_n)] \\ &= \operatorname{sign}\tau \cdot \operatorname{sign}\sigma = \operatorname{sign}\sigma \cdot \operatorname{sign}\tau. \end{aligned}$$

The kernel of the homomorphism sign is called the *alternating group* and is denoted A_n. The following terminology is also used: a permutation σ such that $\operatorname{sign}\sigma = 1$ (respectively, $\operatorname{sign}\sigma = -1$) is called *even* (respectively, *odd*). Thus A_n is the group of even permutations.

Exercise 4.99. Deduce the following formula for the sign of a cyclic permutation:

$$\operatorname{sign}(i_1 i_2 \dots i_p) = (-1)^{p-1}.$$

Using this, prove that the sign of any permutation σ equals $(-1)^{m-s}$, where s is the number of disjoint cycles into which σ factors out and m, the number of symbols that it actually permutes (i.e., does not leave in place).

Theorem 4.100 (Homomorphism Theorem). *Let $f\colon G \to H$ be a group homomorphism. Then*

$$\operatorname{Im} f \simeq G/\operatorname{Ker} f.$$

More precisely, there exists an isomorphism

$$\varphi\colon \operatorname{Im} f \xrightarrow{\sim} G/\operatorname{Ker} f$$

that maps each element $h = f(g) \in \operatorname{Im} f$ to the coset $g \operatorname{Ker} f$.

Proof. The proof of this theorem is similar to that of Theorem 4.76. The property (v) implies that the homomorphism f maps all elements of the coset $g \operatorname{Ker} f$, and only them, to the element $h = f(g) \in \operatorname{Im} f$. This shows that the map φ in the statement of the theorem is well defined and bijective. It remains to show that φ is a homomorphism.

Let $g_1, g_2 \in G$, $f(g_1) = h_1$, $f(g_2) = h_2$. Then $f(g_1 g_2) = h_1 h_2$ and

$$\varphi(h_1 h_2) = g_1 g_2 \operatorname{Ker} f = (g_1 \operatorname{Ker} f)(g_2 \operatorname{Ker} f) = \varphi(h_1)\varphi(h_2).$$

□

Corollary 4.101. *For a finite group G,*

$$|G| = |\operatorname{Im} f||\operatorname{Ker} f|.$$

(Compare this formula with (4.12).)

Example 4.102. Consider the homomorphism

$$f: \mathbb{C} \to \mathbb{R}, \qquad z \mapsto \Im z.$$

We have $\operatorname{Im} f = \mathbb{R}$, $\operatorname{Ker} f = \mathbb{R}$, thus

$$\mathbb{C}/\mathbb{R} \simeq \mathbb{R}.$$

This was already shown in Example 4.86.

Example 4.103. Consider the homomorphism

$$f: \mathbb{C}^* \to \mathbb{R}_+^*, \qquad z \mapsto |z|.$$

We have $\operatorname{Im} f = \mathbb{R}_+^*$, $\operatorname{Ker} f = \mathbb{T} = \{z \in \mathbb{C}^* : |z| = 1\}$, thus

$$\mathbb{C}^*/\mathbb{T} \simeq \mathbb{R}_+^*.$$

This was already shown in Example 4.87.

Example 4.104. The map

$$f: \mathbb{C}^* \to \mathbb{T}, \qquad z \mapsto \frac{z}{|z|}$$

is also a homomorphism. Here $\operatorname{Im} f = \mathbb{T}$, $\operatorname{Ker} f = \mathbb{R}_+^*$. Thus,

$$\mathbb{C}^*/\mathbb{R}_+^* \simeq \mathbb{T}.$$

(The corresponding partition into cosets was described in Example 4.63.)

Example 4.105. Consider the homomorphism

$$f: \mathbb{R} \to \mathbb{T}, \qquad x \mapsto \cos 2\pi x + \imath \sin 2\pi x$$

(see Example 4.96). Since $\operatorname{Ker} f = \mathbb{Z}$, we obtain

$$\mathbb{R}/\mathbb{Z} \simeq \mathbb{T}.$$

Example 4.106. Just as above, by considering the homomorphism det from Example 4.97, we see that

$$\mathrm{GL}_n(K)/\mathrm{SL}_n(K) \simeq K^*.$$

This was already shown in Example 4.88.

Example 4.107. By considering the homomorphism sign from Example 4.98, we see that

$$S_n/A_n \simeq C_2.$$

In particular, this implies that

$$|A_n| = \frac{1}{2}n!.$$

Example 4.108. By definition, any affine transformation f is a composition of a parallel translation and a linear map φ (see Section 4.2). The latter is called the *linear part* or the *differential* of f and is denoted df. The formula

$$(t_a\varphi)(t_b\psi) = t_{a+\varphi(b)}\varphi\psi$$

from the proof of Proposition 4.27 implies that the map

$$d\colon \mathrm{GA}(V) \to \mathrm{GL}(V), \qquad f \mapsto df$$

is a homomorphism. Clearly,

$$\operatorname{Im} d = \mathrm{GL}(V), \qquad \operatorname{Ker} d = \operatorname{Tran} V,$$

hence,

$$\mathrm{GA}(V)/\operatorname{Tran} V \simeq \mathrm{GL}(V).$$

Example 4.109. Let $\triangle = A_1A_2A_3$ be an equilateral triangle. Establish a correspondence between $\operatorname{Sym}\triangle$ and S_3 by the following rule: $\varphi \in \operatorname{Sym}\triangle$ corresponds to $\sigma \in S_3$ if

$$\varphi(A_i) = A_{\sigma(i)}.$$

We thus obtain a homomorphism

$$f\colon \operatorname{Sym}\triangle \to S_3.$$

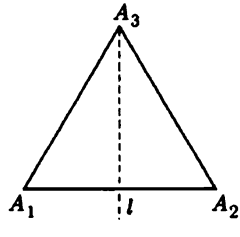


Figure 4.7

If a planar motion preserves three points which do not lie on the same line, then it is trivial; thus $\operatorname{Ker} f = \{\mathrm{id}\}$. We will prove now that $\operatorname{Im} f = S_3$. Since $\operatorname{Im} f$ is a subgroup of S_3 and S_3 is generated by transpositions, it suffices to check that every transposition belongs to $\operatorname{Im} f$, i.e., it is achieved by a motion $\varphi \in \operatorname{Sym} \Delta$. This is indeed so: for instance, the transposition (12) is achieved by a reflection through l in Figure 4.7. Therefore,

$$\operatorname{Sym} \Delta \simeq S_3.$$

Similarly, one can prove that the symmetry group of a regular tetrahedron is isomorphic to S_4 (do this!).

Example 4.110. When we permute the variables x_1, x_2, x_3, x_4, we automatically permute the polynomials

$$x_1x_2 + x_3x_4, \qquad x_1x_3 + x_2x_4, \qquad x_1x_4 + x_2x_3. \tag{4.15}$$

By enumerating them in some way, we establish a homomorphism

$$f\colon S_4 \to S_3.$$

Let us prove that $\operatorname{Im} f = S_3$. It suffices to check that every transposition of polynomials (4.15) is performed by a permutation of the variables x_1, x_2, x_3, x_4. This is indeed so: for instance, the transposition of the first two polynomials (4.15) is performed by the interchange of the two variables x_2 and x_3.

The subgroup $\operatorname{Ker} f$ is the so-called *Klein 4-group*

$$V_4 = \{e, (12)(34), (13)(24), (14)(23)\}.$$

By the homomorphism theorem $V_4 \lhd S_4$ and $S_4/V_4 \simeq S_3$. It is easy to see that the group V_4 is isomorphic to the group in Example 4.12.

Exercise 4.111. Prove that for any $n \in \mathbb{N}$, there exists the following "unusual" isomorphism:

$$\mathbb{C}^*/C_n \simeq \mathbb{C}^*.$$

Exercise 4.112. Let p be a prime number. Determine orders of the groups $\mathrm{GL}_2(\mathbb{Z}_p)$ and $\mathrm{SL}_2(\mathbb{Z}_p)$.

It is obvious that a composition of homomorphisms $F \to G$ and $G \to H$ is a homomorphism $F \to H$.

Example 4.113. Consider the composition of homomorphisms

$$\det\colon \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^* \qquad \text{and} \qquad \operatorname{sign}\colon \mathbb{R}^* \to C_2 = \{\pm 1\},$$

where sign denotes the sign of a real number. In this way we obtain the homomorphism

$$\varepsilon\colon \mathrm{GL}_n(\mathbb{R}) \to C_2.$$

For $n = 2$, it has the following geometric meaning. If $\varepsilon(A) = 1$ (respectively, $\varepsilon(A) = -1$), then the linear transformation of E^2 determined by the matrix A preserves (respectively, reverses) the orientation in the sense that it maps every positively oriented basis to a positively (respectively, negatively) oriented basis. A similar interpretation is possible for the case $n = 3$ as well.

Example 4.114. The composition of homomorphisms

$$d\colon \mathrm{GA}(\mathbb{R}^n) \to \mathrm{GL}(\mathbb{R}^n) = \mathrm{GL}_n(\mathbb{R}) \qquad \text{and} \qquad \varepsilon\colon \mathrm{GL}_n(\mathbb{R}) \to C_2$$

is a homomorphism

$$\mathrm{GA}(\mathbb{R}^n) \to C_2. \tag{4.16}$$

For $n = 2$ and 3, this allows us to extend the notion of orientation-preserving linear transformations to affine transformations of the Euclidean plane and space. Namely, an affine transformation preserves (respectively, reverses) orientation if its differential preserves (respectively, reverses) orientation. In particular, we can speak about orientation-preserving or orientation-reversing motions. (We did this before but without defining these notions rigorously.)

Example 4.115. Let $G \subset \mathrm{Isom}\, E^n$ ($n = 2$ or 3) be a subgroup that contains orientation-reversing motions. By restricting homomorphism (4.16) to G, we see that the subset of orientation-preserving motions in G is a subgroup of index 2. We denote it by G_+.

Example 4.116. In particular, we call the subgroup $\mathrm{Sym}_+ K \subset \mathrm{Sym}\, K$ the *group of rotations of the cube* K. Since $|\mathrm{Sym}\, K| = 48$ (see Example 4.79) and $\mathrm{Sym}_+ K$ is a subgroup of index 2,

$$|\mathrm{Sym}_+ K| = 24.$$

We will now prove that

$$\mathrm{Sym}_+ K \simeq S_4.$$

Enumerate the four diagonals of the cube K in some way. Then one can associate to a motion $\varphi \in \mathrm{Sym}_+ K$ the permutation that it performs on the set of diagonals. We obtain a homomorphism

$$f\colon \mathrm{Sym}_+ K \to S_4.$$

Let us prove that $\mathrm{Im}\, f = S_4$; this will imply that f is an isomorphism, since $|\mathrm{Sym}_+ K| = |S_4|$. For this, it suffices to check that every transposition belongs to $\mathrm{Im}\, f$. This is indeed so: for instance, transposition (12) is achieved by rotation through π about the line l in Figure 4.8.

Exercise 4.117. Prove that the group D_4 (the symmetry group of a square) is isomorphic to the group $\mathrm{Sym}(x_1x_2 + x_3x_4)$ (see Examples 4.17 and 4.81).

Exercise 4.118. Prove that $\mathrm{SL}_2(\mathbb{Z}_2) \simeq S_3$.

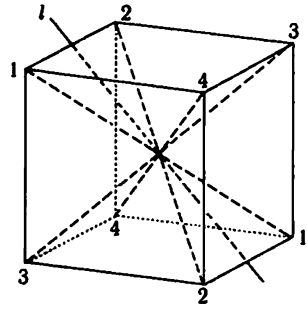


Figure 4.8

By the definition of the operation on the quotient group G/N, the map

$$\pi\colon G \to G/N, \qquad g \mapsto gN$$

is a homomorphism. It is called the *canonical homomorphism* of the group G onto the quotient group G/N. Obviously, its kernel is the subgroup N.

Let $f\colon G \to H$ be a surjective homomorphism. Put $\operatorname{Ker} f = N$. By Theorem 4.100, $H \simeq G/N$. Also, if we identify H with G/N via the isomorphism in that theorem, the homomorphism f coincides with the canonical homomorphism from G onto G/N. Thus, Theorem 4.100 can be interpreted as a statement that the only surjective homomorphisms are the canonical homomorphisms onto quotient groups.

Vector Spaces

This and the next two chapters are devoted to linear algebra and related geometric theory. We already began their study in Chapter 2. Linear algebra is the most applied branch of algebra. Every mathematician needs its machinery, just like the machinery of calculus.

We should warn you, though, against regarding linear algebra simply as matrix manipulation. This approach ignores its ideology, in particular, the geometric ideas that are hidden behind its concepts. A reader who decides to choose this easy path will forfeit a lot. Such a reader will cover scores of pages with formulas or overload the computer in situations that look obvious to those who truly know linear algebra.

Except for the main definitions, several examples, and situations with statements to the contrary, all vector spaces in the linear algebra chapters are assumed to be finite-dimensional. Unless the base field is specified, it is denoted by K.

5.1. Relative Position of Subspaces

Obviously the *intersection* $U \cap W$ of two subspaces U and W of a vector space V is also a subspace. This is the largest space contained in both U and W.

Definition 5.1. The *sum* of two subspaces U and W is the collection of vectors of the form $u + w$, where $u \in U$ and $w \in W$.

This is the smallest subspace containing both U and W.

Definition 5.2. A basis of a space V *agrees* with a subspace U if U is a linear span of some basis vectors (i.e., if it is one of the "coordinate subspaces" with respect to this basis).

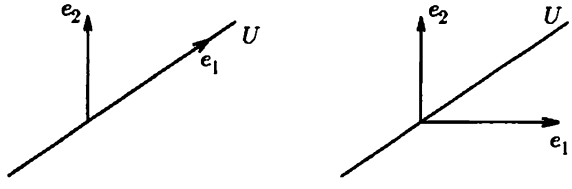


Figure 5.1

For instance, the basis $\{e_1, e_2\}$ agrees with the subspace U in Figure 5.1, left, but not in Figure 5.1, right.

It is easy to see that for any subspace, there exists a basis that agrees with it. The following amazing generalization is less obvious but still true.

Theorem 5.3. *For any two subspaces $U, W \subset V$, there exists a basis of the space V that agrees with both U and W.*

Proof. Let $\{e_1, \dots, e_p\}$ be a basis of $U \cap W$. We can complete it to a basis of U with vectors $e_{p+1}, \dots, e_k$ and to a basis of W with vectors $e_{k+1}, \dots, e_{k+l-p}$ (here $p = \dim U \cap W$, $k = \dim U$, $l = \dim W$). We will prove that the vectors $e_1, \dots, e_{k+l-p}$ are linearly independent. Then we can complete them to a basis of V and thus obtain a basis that agrees with both U and W.

Assume that

$$\sum_{i=1}^{k+l-p} \lambda_i e_i = 0.$$

Consider the vector

$$x = \sum_{i=1}^{k} \lambda_i e_i = -\sum_{i=k+1}^{k+l-p} \lambda_i e_i.$$

The first expression for x implies that it belongs to U, and the second means that it belongs to W. Thus $x \in U \cap W$ and

$$x = \sum_{i=1}^{p} \mu_i e_i = -\sum_{i=k+1}^{k+l-p} \lambda_i e_i.$$

Since the vectors $e_1, \dots, e_p, e_{k+1}, \dots, e_{k+l-p}$ are linearly independent, it follows that $x = 0$ and $\lambda_i = 0$ for $i = k+1, \dots, k+l-p$. Furthermore,

since the vectors $e_1, \ldots, e_k$ are linearly independent, the equality

$$\sum_{i=1}^{k} \lambda_i e_i = 0$$

implies that $\lambda_i = 0$ for $i = 1, \ldots, k$. □

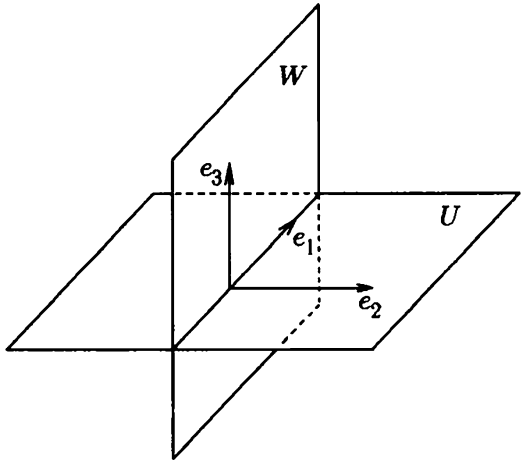


Figure 5.2

Figure 5.2 illustrates the proof in the case of $p = 1$, $k = l = 2$.

Corollary 5.4. $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

Proof. In the notation of Theorem 5.3, the vectors $e_1, \ldots, e_{k+l-p}$ form a basis of the subspace $U + W$. Hence,

$$\dim(U + W) = k + l - p.$$

□

In the case of three subspaces, a similar theorem does not hold.

Exercise 5.5. Give an example demonstrating the preceding statement.

To describe the relative position of an arbitrary (finite) number of subspaces is in general difficult (and, in some sense, impossible). However, we are mostly interested in one particular case where this is easily done.

Definition 5.6. Subspaces $U_1, \ldots, U_k$ are called *linearly independent* if the equality $u_1 + \cdots + u_k = 0$, $u_i \in U_i$, implies that $u_1 = \cdots = u_k = 0$; otherwise, the subspaces are called *linearly dependent*.

For two subspaces U and W, the statement of linear independence is equivalent to the statement $U \cap W = 0$. An expected generalization of this for a larger number of subspaces is not true.

Exercise 5.7. Give an example of three linearly dependent subspaces such that the intersection of any two of them is zero.

Definition 5.8. The *sum* $U_1 + \cdots + U_k$ of subspaces $U_1, \ldots, U_k \subset V$ is the collection of vectors of the form $u_1 + \cdots + u_k$, where $u_i \in U_i$.

This is the smallest subspace containing all subspaces $U_1, \ldots, U_k$.

Proposition 5.9. *The following properties of a system of subspaces* $U_1, \ldots, U_k \subset V$ *are equivalent:*

(i) $U_1, \ldots, U_k$ *are linearly independent;*

(ii) the union of bases of subspaces $U_1, \ldots, U_k$ *is linearly independent;*

(iii) $\dim(U_1 + \cdots + U_k) = \dim U_1 + \cdots + \dim U_k$.

Proof. (i)$\Longleftrightarrow$(ii). Let $\{e_{i1}, \ldots, e_{in_i}\}$ be a basis of the space U_i, $i = 1, \ldots, k$. Assume that there exists a nontrivial linear dependence of vectors e_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$, e.g.,

$$\sum_{i,j} \lambda_{ij} e_{ij} = 0.$$

Then the sum of vectors

$$x_i = \sum_j \lambda_{ij} e_{ij} \in U_i, \qquad i = 1, \ldots, k,$$

is zero and some of them are nonzero. Thus, the subspaces $U_1, \ldots, U_k$ are linearly dependent.

Conversely, if the subspaces $U_1, \ldots, U_k$ are linearly dependent, there exist vectors $x_i \in U_i$, $i = 1, \ldots, k$, whose sum is zero but some of them are nonzero. By expressing each of them in the basis of their respective subspaces, we obtain a nontrivial linear dependence of vectors e_{ij}.

(ii)$\Longleftrightarrow$(iii). Since the union of bases of subspaces $U_1, \ldots, U_k$ spans the sum $U_1 + \cdots + U_k$, each of the properties (ii) and (iii) is equivalent to the fact that this union is the basis of the space $U_1 + \cdots + U_k$. Thus these properties are equivalent. □

The sum of linearly independent subspaces $U_1, \ldots, U_k$ is called their *direct sum* and is denoted $U_1 \oplus \cdots \oplus U_k$. Every vector u of the direct sum can be *uniquely* presented in the form $u = u_1 + \cdots + u_k$, where $u_i \in U_i$. Here the vector u_i is called the *projection* of the vector u onto the subspace U_i.

Observe that the projection of a vector onto a subspace U_i depends not only on this subspace but also on other summands in the direct sum $U_1 \oplus \cdots \oplus U_k$.

Example 5.10. A square matrix A is called *symmetric* if $A^\top = A$ and *skew-symmetric* if $A^\top = -A$. Symmetric (respectively, skew-symmetric) matrices form a subspace $\mathrm{L}_n^+(K)$ (respectively, $\mathrm{L}_n^-(K)$) of the space $\mathrm{L}_n(K)$ of all matrices. If $\operatorname{char} K \neq 2$, every matrix A can be presented as a sum of a symmetric and a skew-symmetric matrix:

$$A = \frac{1}{2}(A + A^\top) + \frac{1}{2}(A - A^\top).$$

On the other hand, under the same restriction it is obvious that a matrix which is simultaneously symmetric and skew-symmetric, must be zero. This implies that

$$\mathrm{L}_n(K) = \mathrm{L}_n^+(K) \oplus \mathrm{L}_n^-(K).$$

Example 5.11. One can similarly prove that the space of all functions on the real line is a direct sum of the subspaces of odd and even functions. (In this example both subspaces and the vector space itself are infinite-dimensional.)

Example 5.12. Let $\{e_1, \ldots, e_n\}$ be a basis of a vector space V. Then

$$V = \langle e_1 \rangle \oplus \cdots \oplus \langle e_n \rangle.$$

The projection of a vector $x \in V$ onto $\langle e_i \rangle$ equals $x_i e_i$, where x_i is the ith coordinate of the vector x in the basis $\{e_1, \ldots, e_n\}$. It depends not only on e_i but also on other basis vectors.

Definitions 5.6 and 5.8 can be generalized to the case of infinitely many subspaces, but there we should consider sums of vectors with only a finite number of nonzero summands.

Example 5.13. Consider the algebra $A = K[x_1, \ldots, x_n]$ of polynomials in n variables. Denote by A_d the subspace of homogeneous polynomials of degree d. Since any polynomial can be uniquely presented as a sum of homogeneous polynomials of nonequal degrees, we have

$$A = A_0 \oplus A_1 \oplus A_2 \oplus \cdots = \bigoplus_{d=0}^{\infty} A_d.$$

Moreover, here

$$A_d A_e \subset A_{d+e}. \tag{5.1}$$

A presentation of an algebra A as a direct sum of subspaces A_d, $d \in \mathbb{Z}$, that satisfy condition (5.1) is called a *grading*. An algebra with a grading

is called a *graded algebra*. (Some of the subspaces A_d might be zero. For instance, in the above example $A_d = 0$ for $d < 0$.)

Exercise 5.14. Consider the algebra $A = \mathrm{L}_n(K)$ of matrices. Denote by A_d the linear span of matrix units E_{ij} such that $j - i = d$. Prove that the subspaces A_d determine a grading of the algebra A. (Here $A_d = 0$ for $|d| \geq n$.)

5.2. Linear Functions

Vectors spaces and their subspaces are the world where the characters of linear algebra dwell. Apart from vectors, the simplest of them are linear functions which, as we will see, are in some sense dual to vectors.

Definition 5.15. A *linear function* (or a *linear form*) on a space V is a function $\alpha : V \to K$ that satisfies the following properties:

(i) $\alpha(x + y) = \alpha(x) + \alpha(y)$;

(ii) $\alpha(\lambda x) = \lambda\alpha(x)$.

In other words, a linear function is a linear map from the space V to the field K regarded as a (one-dimensional) vector space.

Example 5.16. It is shown in elementary geometry that the function $\alpha(x) = (a, x)$, $a \in E^3$, is a linear function on the space E^3.

Example 5.17. The function $\alpha(f) = f(x_0)$, $x_0 \in X$, is a linear function on the space $F(X, K)$ of K-valued functions on the set X (see Example 1.55).

Example 5.18. The function $\alpha(f) = f'(x_0)$, $x_0 \in \mathbb{R}$, is a linear function on the space $C^1(\mathbb{R})$ of differentiable functions on the real line.

Example 5.19. The function $\alpha(f) = \int_a^b f(x)\,dx$ is a linear function on the space $C[a, b]$ of continuous functions on the interval $[a, b]$.

Example 5.20. The *trace* of a square matrix is the sum of entries on its main diagonal. We denote the trace of a matrix X by $\operatorname{tr} X$. The function $\alpha(X) = \operatorname{tr} X$ is a linear function on the space $\mathrm{L}_n(K)$ of square matrices.

If $x_1, \ldots, x_n$ are the coordinates of a vector x in the basis $\{e_1, \ldots, e_n\}$, then

$$\alpha(x) = a_1x_1 + \cdots + a_nx_n, \tag{5.2}$$

where $a_i = \alpha(e_i)$. Thus, a linear function is completely determined by its values on the basis vectors. These are called the *coefficients* of α in this particular basis. Coefficients may be arbitrary: for any collection $a_1, \ldots, a_n \in K$, the function α defined by (5.2) is linear.

Linear functions form a subspace in the space $F(V, K)$ of all K-valued functions on V.

Definition 5.21. The space of linear functions on V is called the *dual space* of V and is denoted V^*.

Let $\{e_1, \dots, e_n\}$ be a basis of the space V. Linear functions $\varepsilon_1, \dots, \varepsilon_n \in V^*$ defined as

$$\varepsilon_i(x) = x_i$$

are called the *coordinate functions* with respect to the basis $\{e_1, \dots, e_n\}$. They make up a basis of the space V^* and we say that it is *dual* to the basis $\{e_1, \dots, e_n\}$. Its definition implies that for any vector $x \in V$,

$$x = \sum_i \varepsilon_i(x) e_i. \tag{5.3}$$

The following condition also defines the dual basis:

$$\varepsilon_i(e_j) = \delta_{ij} := \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j \end{cases} \qquad (\textit{Kronecker symbol}).$$

The above discussion implies that $\dim V^* = \dim V$, hence the spaces V and V^* are isomorphic, although there exists no natural (particular) isomorphism between them. However, the second dual space $V^{**} = (V^*)^*$ is naturally isomorphic to the space V.

From the definition of the operations in the space V^*, it follows that for any vector $x \in V$, the function defined as

$$f_x(\alpha) = \alpha(x)$$

is linear.

Theorem 5.22. *The map $x \mapsto f_x$ is an isomorphism from the space V to the space V^{**}.*

Proof. The definition of a linear function implies that $f_{x+y} = f_x + f_y$ and $f_{\lambda x} = \lambda f_x$. It remains to check that the map $x \mapsto f_x$ is bijective. Let $\{e_1, \dots, e_n\}$ be a basis of V and $\{\varepsilon_1, \dots, \varepsilon_n\}$ the dual basis of V^*. Then

$$f_{e_i}(\varepsilon_j) = \varepsilon_j(e_i) = \delta_{ij},$$

hence $\{f_{e_1}, \dots, f_{e_n}\}$ is the basis of V^{**} dual to $\{\varepsilon_1, \dots, \varepsilon_n\}$. The map $x \mapsto f_x$ sends a vector with coordinates $x_1, \dots, x_n$ in the basis $\{e_1, \dots, e_n\}$ of V to the vector with the same coordinates in the basis $\{e_1, \dots, e_n\}$ of the space V^{**}. Therefore, this map is indeed bijective. □

In the sequel, we will identify the spaces V and V^{**} via the above isomorphism, i.e., regard every vector $x \in V$ also as a linear function on V^* (and write $x(\alpha)$ instead of $f_x(\alpha)$). With this convention, the roles of spaces V and V^* are absolutely symmetric.

Corollary 5.23. *Every basis of the space V^* is dual to some basis of the space V.*

Exercise 5.24. Prove that linear functions $\varepsilon_1, \dots, \varepsilon_n$ (here $n = \dim V$) form a basis of the space V^* if and only if there exists no nonzero vector $x \in V$ such that $\varepsilon_1(x) = \dots = \varepsilon_n(x) = 0$.

Exercise 5.25. Let V be the space of polynomials of degree $\leq n$ over a field K. Prove that linear functions $\varepsilon_0, \varepsilon_1 \dots, \varepsilon_n$, defined as

$$\varepsilon_i(f) = f(x_i),$$

where $x_0, x_1, \dots, x_n$ are some elements of K, form a basis of the space V^*. Determine the dual basis of V and show that in this case formula (5.3) becomes Lagrange's interpolation formula.

Exercise 5.26. Here V is the same as in Exercise 5.25. Assume that $\operatorname{char} K = 0$. Prove that linear functions defined as

$$\varepsilon_i(f) = f^{(i)}(x_0),$$

where $x_0 \in K$, form a basis of the space V^*. Determine the dual basis of the space V and show that in this case formula (5.3) becomes Taylor's formula.

Remark 5.27. Theorem 5.22 does not hold for infinite-dimensional spaces. If the space V is infinite-dimensional, then the space V^* and, furthermore, the space V^{**} are of larger dimensions. For instance, let $V = K^\infty$, the space of finitary sequences (see Example 2.43). This space is of countable dimension. Linear functions on V have the form

$$\alpha(x_1, x_2, \dots) = a_1 x_1 + a_2 x_2 + \cdots$$

(since the sequence $(x_1, x_2, \dots)$ is finitary, the sum is actually finite). Here $a_1, a_2, \dots$ are arbitrary elements of the field K. Therefore, the space V^* is isomorphic to the space of *all* sequences, which can be shown (try to do it!) to be of uncountable dimension.

There exists a one-to-one correspondence between the subspaces of the spaces V and V^*: a k-dimensional subspace of V corresponds to an $(n-k)$-dimensional subspace of V^* (here $n = \dim V$).

Definition 5.28. The *annihilator* of a subspace $U \subset V$ is the subspace

$$U^\circ = \{\alpha \in V^* \colon \alpha(x) = 0 \ \forall x \in U\}.$$

Theorem 5.29. $\dim U^\circ = \dim V - \dim U$.

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis of V such that $U = \langle e_1, \ldots, e_k \rangle$. Let $\{\varepsilon_1, \ldots, \varepsilon_n\}$ be the dual basis of the space V^*. Then $U^\circ = \langle \varepsilon_{k+1}, \ldots, \varepsilon_n \rangle$. □

Since we identified the spaces V and V^{**}, we can assume that the annihilator of a subspace $W \subset V^*$ lies in V. By definition,

$$W^\circ = \{x \in V : \alpha(x) = 0 \ \forall \alpha \in W\}.$$

Theorem 5.30. *$(U^\circ)^\circ = U$ for every subspace $U \subset V$.*

Proof. We continue using the notation of Theorem 5.29. It is clear that $(U^\circ)^\circ = \langle e_1, \ldots, e_k \rangle = U$. □

Corollary 5.31. *Every subspace of V is the annihilator of some subspace in V^*.*

Consider the following system of homogeneous linear equations:

$$\sum_{j=1}^{n} a_{ij} x_j = 0, \qquad i = 1, \ldots, m. \tag{5.4}$$

We interpret $x_1, \ldots, x_n$ as coordinates of a vector x of an n-dimensional vector space V in a basis $\{e_1, \ldots, e_n\}$. Then the system (5.4) can be rewritten as

$$\alpha_i(x) = 0, \qquad i = 1, \ldots, m,$$

where $\alpha_1, \ldots, \alpha_m \in V^*$ are the linear functions that appear on the left-hand side of the equations in system (5.4). The set of solutions of this system is the annihilator of the subspace $\langle \alpha_1, \ldots, \alpha_m \rangle \subset V^*$. Observe that the dimension of this subspace equals the rank of the coefficient matrix of system (5.4). Thus Theorem 2.63 on the dimension of the space of solutions of a homogeneous linear system is a direct corollary of Theorem 5.29.

In this context, Corollary 5.31 can be reformulated as follows:

Theorem 5.32. *Every subspace is the set of solutions of some system of homogeneous linear equations.*

5.3. Bilinear and Quadratic Functions

Vector space axioms do not incorporate all elementary geometry of vectors in the Euclidean space since they lack notions such as the length of a vector and the angle between two vectors. The length and the angle can be expressed via the inner product of vectors. One of the basic properties of the inner product of geometric vectors is its linearity in every factor. In this section, we will consider functions of two vector arguments which generalize the inner product.

Definition 5.33. A *bilinear function* (or a *bilinear form*) on a vector space V is a function $\alpha\colon V \times V \to K$ that is linear in every argument.

Example 5.34. As shown in elementary geometry, the inner product of geometric vectors is a bilinear function on the space E^3.

Example 5.35. The function

$$\alpha(f,g) = \int_a^b f(x)g(x)\,dx$$

is a bilinear function on the space $C[a,b]$.

Example 5.36. The function

$$\alpha(X,Y) = \operatorname{tr} XY$$

is a bilinear function on the space $\mathrm{L}_n(K)$.

Example 5.37. The determinant of matrices of order 2 regarded as a function of the rows of the matrix is a bilinear function on the space K^2.

Let $\{e_1,\dots,e_n\}$ be a basis of the space V. For vectors $x = \sum_i x_i e_i$ and $y = \sum_j y_j e_j$, we have

$$\alpha(x,y) = \sum_{i,j} a_{ij}x_i y_j, \qquad \text{where } a_{ij} = \alpha(e_i,e_j). \tag{5.5}$$

The matrix $A = (a_{ij})$ is called the *matrix of the bilinear function* α in the basis $\{e_1,\dots,e_n\}$. The above formula implies that every bilinear function is uniquely determined by its matrix.

Formula (5.5) can be restated in matrix notation:

$$\alpha(x,y) = X^\top A Y, \tag{5.6}$$

where X and Y are the columns of coordinates of vectors x and y, respectively.

When the basis changes as follows:

$$(e'_1,\dots,e'_n) = (e_1,\dots,e_n)C,$$

coordinates of vectors change too:

$$X = CX', \qquad Y = CY'.$$

Substituting this into (5.6), we obtain

$$\alpha(x,y) = (X')^\top C^\top A C Y'.$$

This implies that in the new basis $\{e'_1,\dots,e'_n\}$, the matrix of α becomes

$$A' = C^\top A C. \tag{5.7}$$

The main goal of the theory of bilinear functions is to reduce the matrix of a bilinear function to the simplest possible form by choosing the right basis. So, it is important to know the properties of the matrix of a bilinear function that do not depend on the choice of a basis.

Definition 5.38. The *kernel* of a bilinear function α is the subspace

$$\operatorname{Ker}\alpha = \{y \in V\colon \alpha(x,y) = 0\ \forall x \in V\}.$$

The function α is called *nondegenerate* if $\operatorname{Ker}\alpha = 0$.

All bilinear functions in Examples 5.34–5.37 are nondegenerate. For instance, the inner product is nondegenerate because $(y,y) > 0$ whenever $y \neq 0$. Nondegeneracy of the bilinear function in Example 5.35 is proved similarly.

Exercise 5.39. Prove that the bilinear functions in Examples 5.36 and 5.37 are nondegenerate.

Clearly, if $\{e_1,\dots,e_n\}$ is a basis of the space V, then

$$\operatorname{Ker}\alpha = \{y \in V\colon \alpha(e_i,y) = 0,\ i = 1,\dots,n\}.$$

Writing these conditions in the coordinate form, we obtain the system of homogeneous linear equations whose coefficient matrix is the matrix A of the function α. Therefore,

$$\dim\operatorname{Ker}\alpha = n - \operatorname{rk}A. \tag{5.8}$$

In particular, $\operatorname{Ker}\alpha = 0$ if and only if $\operatorname{rk}A = n$, i.e., when A is nonsingular.

Formula (5.8) implies that the rank of the matrix of a bilinear function α does not depend on the choice of a basis. It is called the *rank of the bilinear function* α and is denoted $\operatorname{rk}\alpha$.

Definition 5.40. A bilinear function α is called *symmetric* (respectively, *skew-symmetric*) if $\alpha(x,y) = \alpha(y,x)$ (respectively, $\alpha(x,y) = -\alpha(y,x)$) for $x,y \in V$.

For instance, bilinear functions in Examples 5.34 and 5.35 are symmetric.

The bilinear function in Example 5.36 is also symmetric. Indeed, if $X = (x_{ij})$, $Y = (y_{ij})$, then

$$\operatorname{tr}XY = \sum_{i,j} x_{ij}y_{ji} = \sum_{i,j} y_{ji}x_{ij} = \sum_{i,j} y_{ij}x_{ji} = \operatorname{tr}YX.$$

The bilinear function in Example 5.37 is skew-symmetric. Of course, there also exist bilinear functions that are neither symmetric nor skew-symmetric.

A bilinear function is symmetric (respectively, skew-symmetric) if and only if its matrix A is symmetric (respectively, skew-symmetric), i.e., $A^\top = A$ (respectively, $A^\top = -A$).

Definition 5.41. Let α be a symmetric bilinear function over a field K of characteristic $\neq 2$. The function $q\colon V \to K$ defined as

$$q(x) = \alpha(x, x)$$

is called the *quadratic function* (or the *quadratic form*) associated to α.

In coordinate notation, a quadratic function is written as

$$q(x) = \sum_{i,j} a_{ij} x_i x_j, \tag{5.9}$$

i.e., it is a homogeneous polynomial of the second degree.

A symmetric bilinear function α can be reconstructed from the corresponding quadratic function q as follows:

$$\alpha(x, y) = \frac{1}{2}[q(x + y) - q(x) - q(y)]. \tag{5.10}$$

This bilinear function α is called the *polarization* of the quadratic function q.

Thus, there exists a one-to-one correspondence between symmetric bilinear and quadratic functions. Using this correspondence, we can transfer all notions introduced for symmetric bilinear functions (matrix, rank, nondegeneracy, etc.) to quadratic functions. In the future, speaking of a quadratic function, we will have in mind the corresponding symmetric bilinear function and vice versa.

The geometric picture associated with the inner product of vectors can be useful in the study of arbitrary bilinear functions. This is the origin of the terminology below.

Let α be a symmetric or a skew-symmetric bilinear function over a field K of characteristic $\neq 2$. Vectors $x, y \in V$ are called *orthogonal* (with respect to α) if $\alpha(x, y) = 0$; the notation is $x \perp y$. This relation is clearly symmetric: if $x \perp y$, then $y \perp x$. Observe also that when α is skew-symmetric, every vector is orthogonal to itself.

Definition 5.42. The *orthogonal complement* of a subspace U (with respect to α) is the subspace

$$U^\perp = \{y \in V\colon \alpha(x, y) = 0 \ \forall x \in U\}.$$

In particular, $V^\perp = \operatorname{Ker} \alpha$.

Proposition 5.43. *If the function α is nondegenerate, then*

$$\dim U^\perp = \dim V - \dim U \qquad \textit{and} \qquad (U^\perp)^\perp = U.$$

Proof. Fix a basis $\{e_1, \ldots, e_k\}$ of U. Then

$$U^\perp = \{y \in V : \alpha(e_i, y) = 0,\ i = 1, \ldots, k\}. \tag{5.11}$$

Rewriting these conditions in the coordinate form, we obtain a system of homogeneous linear equations. They are linearly independent because, for any $\lambda_1, \ldots, \lambda_k$ some of which are nonzero, the linear function

$$\sum_{i=1}^{k} \lambda_i \alpha(e_i, y) = \alpha\left(\sum_{i=1}^{k} \lambda_i e_i, y\right)$$

is nonzero (due to the nondegeneracy of α). Hence,

$$\dim U^\perp = n - k,$$

where $n = \dim V$.

This formula also implies

$$\dim(U^\perp)^\perp = n - (n - k) = k = \dim U.$$

Clearly, $(U^\perp)^\perp \supset U$. Thus, $(U^\perp)^\perp = U$. □

Definition 5.44. A subspace U is *nondegenerate* with respect to the function α if the restriction of α to U is nondegenerate.

Proposition 5.45. *$V = U \oplus U^\perp$ if and only if the subspace U is nondegenerate.*

Proof. It follows from (5.11) that for any U,

$$\dim U^\perp \geq \dim V - \dim U.$$

On the other hand,

$$U \cap U^\perp = \operatorname{Ker} \alpha|_U.$$

Therefore, if $U \cap U^\perp = 0$, the subspace U is nondegenerate. Conversely, if U is nondegenerate, then $U \cap U^\perp = 0$. Thus,

$$\dim(U + U^\perp) = \dim U + \dim U^\perp \geq \dim V,$$

implying $U + U^\perp = V$. □

Consider now a symmetric bilinear function α.

A basis $\{e_1, \ldots, e_n\}$ of a space V is called *orthogonal* (with respect to α) if all its vectors are pairwise orthogonal. The matrix of α is diagonal in an orthogonal basis and the function α itself and the corresponding quadratic function q have the form

$$\alpha(x, y) = a_1 x_1 y_1 + \cdots + a_n x_n y_n, \tag{5.12}$$

$$q(x) = a_1 x_1^2 + \cdots + a_n x_n^2. \tag{5.13}$$

Theorem 5.46. *For every symmetric bilinear function, there exists an orthogonal basis.*

Proof. We will prove this theorem by induction on $n = \dim V$. When $n = 1$, there is nothing to prove. Let $n > 1$. If $\alpha \equiv 0$, again, there is nothing to prove. If $\alpha \not\equiv 0$, then $q \not\equiv 0$ by (5.10). That is, there exists a vector e_1 such that

$$\alpha(e_1, e_1) = q(e_1) \neq 0.$$

By Proposition 5.45,

$$V = (e_1) \oplus (e_1)^{\perp}.$$

By the induction hypothesis, there exists an orthogonal basis $\{e_2, \dots, e_n\}$ of the space $(e_1)^{\perp}$. By adding the vector e_1 to it, we obtain an orthogonal basis $\{e_1, e_2, \dots, e_n\}$ of V. □

The following theorem describes a more explicit way of constructing an orthogonal basis (under some extra conditions).

First, let $\{e_1, \dots, e_n\}$ be a basis of the space V and A the matrix of the function α in this basis. For each k, we can restrict α to the subspace $V_k = (e_1, \dots, e_k)$. Denote by A_k the matrix of this restriction in the basis $\{e_1, \dots, e_k\}$ of V_k. Notice that A_k is the upper left corner of the matrix A. Call the number $\delta_k = \det A_k$ the *corner minor* of A of order k. Also, let $V_0 = 0$, $\delta_0 = 1$.

Theorem 5.47. *If all corner minors $\delta_1, \dots, \delta_n$ of the matrix A are nonzero, then there exists a unique orthogonal basis $\{f_1, \dots, f_n\}$ of the space V such that*

$$f_k \in e_k + V_{k-1}, \qquad k = 1, \dots, n. \tag{5.14}$$

Also,

$$q(f_k) = \alpha(f_k, f_k) = \frac{\delta_k}{\delta_{k-1}}, \qquad k = 1, \dots, n. \tag{5.15}$$

Proof. We will prove this theorem by induction on n. For $n = 1$, we have

$$f_1 = e_1, \qquad q(f_1) = \delta_1 \left(= \frac{\delta_1}{\delta_0} \right).$$

For $n > 1$, apply first the induction hypothesis to the basis $\{e_1, \dots, e_{n-1}\}$ of the space V_{n-1}. Let $\{f_1, \dots, f_{n-1}\}$ be the basis of V_{n-1} that satisfies the conditions of the theorem. Now we have to construct the vector f_n such that

$$f_n = e_n + \sum_{i=1}^{n-1} \lambda_i f_i \in e_n + V_{n-1}.$$

Observe that

$$q(f_i) = \frac{\delta_i}{\delta_{i-1}} \neq 0 \qquad \text{for} \quad i = 1, \ldots, n-1.$$

Hence, the following orthogonality condition:

$$0 = (f_n, f_i) = (e_n, f_i) + \lambda_i q(f_i), \qquad i = 1, \ldots, n-1,$$

must be satisfied for the right choice of $\lambda_1, \ldots, \lambda_{n-1}$ and this choice is unique. Since $f_n \notin V_{n-1}$, we see that $\{f_1, \ldots, f_n\}$ is a basis of V.

It remains to check that equality (5.15) holds when $k = n$. The transition matrix from the basis $\{e_1, \ldots, e_n\}$ to the basis $\{f_1, \ldots, f_n\}$ is (upper) unitriangular, i.e., triangular with 1's on the diagonal. Hence, its determinant equals 1 and formula (5.7) implies that the determinant of the matrix of α does not change when the basis is changed. However, in the new basis $\{f_1, \ldots, f_n\}$, the matrix of α is diagonal and its diagonal entries are equal to

$$q(f_1), \ \ldots, \ q(f_{n-1}), \ q(f_n).$$

Therefore,

$$\delta_n = q(f_1) \cdots q(f_{n-1}) q(f_n).$$

The same argument applied to the restriction of α to the subspace V_{n-1} (or, for that matter, the induction hypothesis) implies that

$$q(f_n) = \frac{\delta_n}{\delta_{n-1}}.$$

□

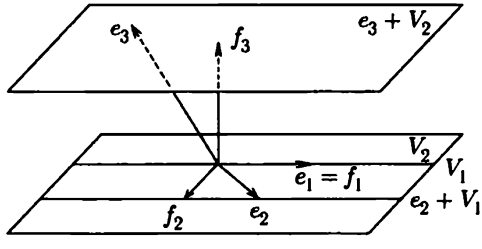


Figure 5.3

The algorithm for constructing the orthogonal basis that we described in the above theorem is called the *Gram-Schmidt orthogonalization procedure*. Figure 5.3 shows its application for the case of the inner product in E^3.

Let $\{e_1, \ldots, e_n\}$ be an orthogonal basis of a space V with respect to a function α. By scaling vectors e_i, we can multiply the numbers $a_i = q(e_i)$ by squares of nonzero elements of the field K. Moreover, when permuting

the basis vectors, we permute these numbers as well. However, as the proof of Theorem 5.46 shows, there is much more freedom in the choice of an orthogonal basis. How do the a_i's change with all this freedom? The answer to this question depends on the field K.

Let $K = \mathbb{C}$. Then by scaling the basis vectors, we can make the a_i's equal to either 0 or 1. Then, after a suitable permutation, the function $q(x)$ assumes the so-called *normal form* :

$$q(x) = x_1^2 + \cdots + x_r^2.$$

The number r is an invariant here, as $r = \operatorname{rk} q$.

Now let $K = \mathbb{R}$. Then by scaling the basis vectors, we can make the a_i's equal to either 0 or ± 1. Again, after a suitable permutation, we obtain the function in the *normal form*:

$$q(x) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{k+l}^2. \tag{5.16}$$

The sum $k + l = \operatorname{rk} q$ is an invariant but are k and l invariants as well? To answer this question, we need to introduce the following important notion.

Definition 5.48. A real quadratic function q is *positive definite* if $q(x) > 0$ for $x > 0$. A real symmetric bilinear function is *positive definite* if the corresponding quadratic function is positive definite.

For example, the inner product of geometric vectors is a positive definite symmetric bilinear function.

One similarly defines *negative definite* quadratic and symmetric bilinear functions.

Obviously, a positive definite quadratic function has the normal form $x_1^2 + \cdots + x_n^2$.

Theorem 5.49. *The index k in the normal form (5.16) of a real quadratic function q is the maximum dimension of a subspace on which q is positive definite.*

Proof. It is clear that q is positive definite on the k-dimensional subspace $\langle e_1, \ldots, e_k \rangle$. Now let U be a subspace on which q is positive definite. Let $W = \langle e_{k+1}, \ldots, e_n \rangle$. Since $q(x) \leq 0$ for every $x \in W$, we have $U \cap W = 0$. It follows that $\dim U \leq k$. □

Similarly, l is the maximum dimension of a subspace on which q is negative definite.

Corollary 5.50 (The Law of Inertia). *The numbers k and l in the normal form (5.16) of a real quadratic function q do not depend on the choice of a basis in which this function has the normal form.*

These numbers are called, respectively, the *positive* and the *negative indices of inertia* of the quadratic function q (and the corresponding symmetric bilinear function α). The pair (k, l) is called the *signature* of q (or α).

Example 5.51. Consider the quadratic function $q(x) = x_1x_2$. With the following (nonsingular) change of coordinates:

$$x_1 = x_1' + x_2', \qquad x_2 = x_1' - x_2',$$

it becomes $q(x) = {x_1'}^2 - {x_2'}^2$. Thus, its signature is $(1, 1)$.

Exercise 5.52. Find the signature of the symmetric bilinear function from Example 5.36 (when $K = \mathbb{R}$).

If its assumptions are satisfied, Theorem 5.47 allows us to find indices of inertia of a real quadratic function from the corner minors $\delta_1, \ldots, \delta_n$ of its matrix in some basis.

Theorem 5.53 (Jacobi Method). *If all corner minors δ_k of a matrix of a real quadratic form q are nonzero, then the negative index of inertia of q equals the number of changes of sign in the sequence*

$$1,\ \delta_1,\ \delta_2,\ \ldots,\ \delta_n. \tag{5.17}$$

(For the definition of the number of changes of sign in a sequence of real numbers, see Section 3.4.)

Proof. The assertion follows directly from Theorem 5.47. □

Observe that the function in the statement of this theorem is necessarily nondegenerate, so the sum of its indices of inertia equals n.

Corollary 5.54. *A real quadratic function is positive definite if and only if all corner minors of its matrix are positive.*

Proof. If all corner minors are positive, then, in particular, they are nonzero and the Jacobi method shows that the function is positive definite. Conversely, if the function is positive definite, then its restriction to any of the spaces V_k (in the notation of Theorem 5.47) is also positive definite and thus nondegenerate. This implies that all of its corner minors are nonzero. Applying the Jacobi method, we see that they all are positive. □

Remark 5.55. We can modify the process of orthogonalization (try it) so that the Jacobi method still work even when some of the corner minors are zero. The only requirement is that no two consecutive numbers in the sequence $\delta_1, \delta_2, \ldots, \delta_n$ be both zero (in particular, $\delta_n = 0$ is allowed but if so, δ_{n-1} must be nonzero).

As we saw in the cases $K = \mathbb{C}$ and $K = \mathbb{R}$, the only possible changes in the diagonal form of the matrix of a quadratic form come from scaling and permutations of the basis vectors. However, this is not always the case.

Let $K = \mathbb{Z}_p$ be the field of residue classes modulo a prime number $p \neq 2$. It is known (Theorem 9.34) that $\mathbb{Z}_p^*$ is a cyclic group. Therefore, $(\mathbb{Z}_p^*)^2 = \{a^2 : a \in \mathbb{Z}_p^*\}$ is a subgroup of index 2. Its elements are called *quadratic residues* and the elements of the other coset, *quadratic nonresidues*. Fix a quadratic nonresidue $\varepsilon \in \mathbb{Z}_p^*$.

Theorem 5.56. *Every nondegenerate quadratic function q over the field $\mathbb{Z}_p$ ($p \neq 2$) can be reduced to one of the following two forms:*

$$x_1^2 + \cdots + x_{n-1}^2 + x_n^2,$$
$$x_1^2 + \cdots + x_{n-1}^2 + \varepsilon x_n^2.$$

The particular form depends on whether the determinant of the matrix of q is a quadratic residue.

Lemma 5.57. *For every nondegenerate quadratic function q in a vector space of dimension $n \geq 2$ over the field $\mathbb{Z}_p$, the equation $q(x) = 1$ has a solution.*

Proof. It suffices to consider the case $n = 2$. We can assume that

$$q(x) = a_1 x_1^2 + a_2 x_2^2,$$

where $a_1, a_2 \neq 0$. The equation $q(x) = 1$ can be presented as

$$a_1 x_1^2 = 1 - a_2 x_2^2.$$

For all possible values of x_1, the left-hand side of the above equation assumes the total of $\frac{p+1}{2}$ distinct values. Similarly, for all possible values of x_2, the right-hand side assumes the total of $\frac{p+1}{2}$ distinct values. Since

$$\frac{p+1}{2} + \frac{p+1}{2} > p,$$

there exist x_1 and x_2 such that the left- and the right-hand side assume the same value. □

Proof of Theorem 5.56. Following the proof of Theorem 5.46, we will choose the vector e_1 so that $q(e_1) = 1$. For $n > 1$ this is possible by the above lemma. When we change the basis, the determinant of the matrix of q gets multiplied by the square of the determinant of the transition matrix. Hence $q(e_n)$ will be a quadratic residue or nonresidue depending on whether the determinant of the matrix of q is such in any basis. □

Exercise 5.58. Prove that a (not necessarily nondegenerate) quadratic function q over the field $\mathbb{Z}_p$ can be reduced to one of the following two forms:

$$x_1^2 + \cdots + x_{r-1}^2 + x_r^2,$$
$$x_1^2 + \cdots + x_{r-1}^2 + \varepsilon x_r^2,$$

where $r = \operatorname{rk} q$.

We now turn to skew-symmetric functions. Quite surprisingly, their structure does not depend on the field K.

Consider a skew-symmetric bilinear function α on an n-dimensional vector space V.

A basis $\{e_1, \ldots, e_n\}$ of V is called *symplectic* (with respect to α) if

$$\alpha(e_{2k-1}, e_{2k}) = -\alpha(e_{2k}, e_{2k-1}) = 1 \qquad \text{for} \quad k = 1, \ldots, m,$$
$$\alpha(e_i, e_j) = 0 \quad \text{in all other cases.}$$

In other words, in this basis the matrix of α is

$$\begin{pmatrix} \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & & & & \\ & \ddots & & \mathbf{0} & \\ & & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & & \\ & \mathbf{0} & & \ddots & \\ & & & & 0 \end{pmatrix},$$

where the number of blocks on the diagonal is m. Clearly, here $\operatorname{rk} \alpha = 2m$.

Theorem 5.59. *For every skew-symmetric bilinear form, there exists a symplectic basis.*

Proof. We will prove this claim by induction on n. For $n = 1$, there is nothing to prove. Let $n > 1$. If $\alpha \equiv 0$, again, there is nothing to prove. If $\alpha \not\equiv 0$, there exist vectors e_1 and e_2 such that $\alpha(e_1, e_2) \neq 0$. After multiplying one of these vectors by a suitable number, we have

$$\alpha(e_1, e_2) = -\alpha(e_2, e_1) = 1.$$

The matrix of the restriction of α to $\langle e_1, e_2 \rangle$ has the form $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ in the basis $\{e_1, e_2\}$. In particular, it is nonsingular. By Proposition 5.45,

$$V = \langle e_1, e_2 \rangle \oplus \langle e_1, e_2 \rangle^{\perp}.$$

By the induction hypothesis, there exists a symplectic basis $\{e_3, e_4, \dots, e_n\}$ in $\langle e_1, e_2\rangle^\perp$. By adding to it vectors e_1 and e_2, we obtain a symplectic basis $\{e_1, e_2, e_3, e_4, \dots, e_n\}$ of the space V. □

Corollary 5.60. *The rank of a skew-symmetric bilinear form is always even.*

5.4. Euclidean Spaces

The properties of operations on geometric vectors, including those of the inner product, are most fully reflected in the concept of a Euclidean vector space.

Definition 5.61. A *Euclidean vector space* is a real vector space with a fixed positive definite symmetric bilinear function.

Usually this function is called the *inner product* and is denoted (,).

Example 5.62. The space of geometric vectors with the standard inner product.

Example 5.63. The space $\mathbb{R}^n$ with the inner product

$$(x, y) = x_1y_1 + \cdots + x_ny_n,$$

where $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$.

Example 5.64. The space $C_2[0, 1]$ of continuous functions on the interval $[0, 1]$ with the following inner product:

$$(f, g) = \int_0^1 f(x)g(x)dx. \tag{5.18}$$

It is possible to define the *length* of a vector and the *angle* between two vectors in a Euclidean space. This is done so that in the case of geometric vectors, these notions coincide with the standard length and angle. Namely, the length $|x|$ of a vector x is defined by the formula

$$|x| = \sqrt{(x, x)}.$$

To define the angle, we first have to prove the following

Proposition 5.65. *For any two vectors x, y in a Euclidean space,*

$$|(x, y)| \le |x||y|, \tag{5.19}$$

and equality is attained if and only if x and y are proportional.

Inequality (5.19) is called the *Cauchy–Schwarz inequality*.

Proof. If $y = \lambda x$, then

$$|(x,y)| = |\lambda||(x,x)| = |\lambda||x|^2 = |x||y|.$$

If the vectors x and y are not proportional, then they form a basis of a two-dimensional space. The matrix of the inner product on this space in the basis $\{x,y\}$ is

$$\begin{pmatrix} (x,x) & (x,y) \\ (x,y) & (y,y) \end{pmatrix}.$$

Since the inner product is positive definite, the determinant of this matrix is positive, implying

$$|(x,y)| < |x||y|.$$

□

The angle $\widehat{xy}$ between two nonzero vectors x and y in a Euclidean space is defined as

$$\cos \widehat{xy} = \frac{(x,y)}{|x||y|}.$$

In particular, the angle $\widehat{xy}$ equals 0 or π if and only if the vectors x and y are proportional; $\widehat{xy} = \frac{\pi}{2}$ if and only if x and y are orthogonal.

The Cauchy–Schwarz inequality is a particular case of a more general inequality that concerns an arbitrary finite system of vectors $\{a_1, \ldots, a_k\}$ in a Euclidean space.

Definition 5.66. The matrix

$$G(a_1, \ldots, a_k) = \begin{pmatrix} (a_1,a_1) & (a_1,a_2) & \ldots & (a_1,a_k) \\ (a_2,a_1) & (a_2,a_2) & \ldots & (a_2,a_k) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (a_k,a_1) & (a_k,a_2) & \ldots & (a_k,a_k) \end{pmatrix}$$

is called the *Gram matrix* of the system $\{a_1, \ldots, a_k\}$.

Theorem 5.67. *For any vectors $a_1, \ldots, a_k$ in a Euclidean space,*

$$\det G(a_1, \ldots, a_k) \geq 0,$$

and equality is attained if and only if the vectors $a_1, \ldots, a_k$ are linearly dependent.

Proof. If $\sum_i \lambda_i a_i = 0$, then $\sum_i \lambda_i (a_i, a_j) = 0$ for all j, implying that the linear combination of rows of $G(a_1, \ldots, a_k)$ with coefficients $\lambda_1, \ldots, \lambda_k$ is zero. Thus, if vectors $a_1, \ldots, a_k$ are linearly dependent, then $\det G(a_1, \ldots, a_k) = 0$. If they are linearly independent, then just as in the case $k = 2$, we can prove that $\det G(a_1, \ldots, a_k) > 0$. □

Exercise 5.68. Find the relation between the dihedral angles of a tetrahedron by considering the Gram matrix corresponding to the system of unit vectors orthogonal to the faces. Using this relation, determine the angle between two faces of a regular tetrahedron.

Definition 5.69. A basis of a Euclidean space where the inner product has the normal form (see Section 5.3) is called *orthonormal*.

A basis $\{e_1, \ldots, e_n\}$ is orthonormal if either of the following equivalent conditions holds:

(i) the inner product in this basis is of the form

$$(x, y) = x_1 y_1 + \cdots + x_n y_n;$$

(ii) the inner square in this basis is of the form

$$(x, x) = x_1^2 + \cdots + x_n^2;$$

(iii) the matrix of the inner product in this basis (i.e., the Gram matrix $G(e_1, \ldots, e_n)$) is the identity matrix;

(iv) $(e_i, e_j) = \delta_{ij}$;

(v) the basis vectors are pairwise orthogonal and have length 1.

The general theory of Section 5.3 implies that every Euclidean space has an orthonormal basis. Of course, such a basis is not unique. Let us describe all orthonormal bases given a choice of one orthonormal basis $\{e_1, \ldots, e_n\}$.

Let

$$(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)C.$$

Then the matrix of the inner product in the basis $\{e'_1, \ldots, e'_n\}$ is

$$C^\top E C = C^\top C$$

(cf. formula (5.7)). Therefore, the basis $\{e'_1, \ldots, e'_n\}$ is orthonormal if and only if

$$C^\top C = E. \tag{5.20}$$

Clearly, the following properties of the matrix C are equivalent:

(i) $C^\top C = E$;

(ii) $\sum_k c_{ki} c_{kj} = \delta_{ij}$ for all i, j;

(iii) $C^\top = C^{-1}$;

(iv) $CC^\top = E$;

(v) $\sum_k c_{ik} c_{jk} = \delta_{ij}$ for all i, j.

Definition 5.70. Matrices satisfying these equivalent properties are called *orthogonal*.

Observe that equality (5.20) implies that $\det C = \pm 1$ (but not vice versa).

The restriction of the inner product to any subspace U of a Euclidean space V is also positive definite, hence it is nondegenerate symmetric bilinear function. Proposition 5.45 implies that

$$V = U \oplus U^{\perp}.$$

It follows that every vector $x \in V$ can be uniquely written as

$$x = y + z, \qquad y \in U, \quad z \in U^{\perp}. \tag{5.21}$$

The vector y is called the *orthogonal projection* of x onto U and is denoted $\mathrm{pr}_U x$. The vector z is called the *orthogonal component* of x with respect to U and is denoted $\mathrm{ort}_U x$.

If $\{e_1, \ldots, e_k\}$ is an orthonormal basis of the subspace U, then the projection $\mathrm{pr}_U x$ can be found as follows:

$$\mathrm{pr}_U x = \sum_{i=1}^{k} (x, e_i) e_i. \tag{5.22}$$

More generally, if $\{e_1, \ldots, e_k\}$ is an orthogonal (but not necessarily orthonormal) basis of the subspace U, then

$$\mathrm{pr}_U x = \sum_{i=1}^{k} \frac{(x, e_i)}{(e_i, e_i)} e_i. \tag{5.23}$$

To construct an orthogonal basis of a Euclidean space V, one can use the *method of orthogonalization* described in Theorem 5.47. In the previous notation, if $\{e_1, \ldots, e_n\}$ is a basis of the space V, then the basis $\{f_1, \ldots, f_n\}$ obtained in the process of orthogonalization is as follows:

$$f_k = \mathrm{ort}_{V_{k-1}} e_k, \qquad k = 1, \ldots, n. \tag{5.24}$$

Since $\{f_1, \ldots, f_{k-1}\}$ is an orthogonal basis of the subspace V_{k-1}, we can find the projection $\mathrm{pr}_{V_{k-1}} e_k$—and thus the vector f_k—from formula (5.23).

Example 5.71. Let V be the space of polynomials of degree ≤ 3 with the inner product defined as in (5.18). We will apply orthogonalization to the basis

$$e_1 = 1, \qquad e_2 = x, \qquad e_3 = x^2, \qquad e_4 = x^3.$$

Notice that $(e_i, e_j) = \frac{1}{i+j-1}$. We have

$$f_1 = e_1 = 1, \qquad (f_1, f_1) = 1,$$

$$f_2 = e_2 - \frac{(e_2, f_1)}{(f_1, f_1)} f_1 = x - \frac{1}{2}, \qquad (f_2, f_2) = (f_2, e_2) = \frac{1}{12},$$

$$f_3 = e_3 - \frac{(e_3, f_2)}{(f_2, f_2)} f_2 - \frac{(e_3, f_1)}{(f_1, f_1)} f_1 = x^2 - x + \frac{1}{6}, \quad (f_3, f_3) = (f_3, e_3) = \frac{1}{180},$$

$$f_4 = e_4 - \frac{(e_4, f_3)}{(f_3, f_3)} f_3 - \frac{(e_4, f_2)}{(f_2, f_2)} f_2 - \frac{(e_4, f_1)}{(f_1, f_1)} f_1 = x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20},$$

$$(f_4, f_4) = (f_4, e_4) = \frac{1}{2800}.$$

Exercise 5.72. By applying orthogonalization to rows of a matrix prove that every matrix $A \in \mathrm{GL}_n(\mathbb{R})$ can be uniquely presented as $A = OB$, where O is an orthogonal matrix and B is a triangular matrix with positive elements on the diagonal.

We define the *distance* ρ between two vectors of a Euclidean space as

$$\rho(x, y) = |x - y|.$$

The distance satisfies the axioms of a metric space. In particular, the triangle axiom holds:

$$\rho(x, z) \le \rho(x, y) + \rho(y, z). \tag{5.25}$$

This inequality follows from

$$|x + y| \le |x| + |y|, \tag{5.26}$$

and, in turn, the latter inequality is easily deduced (do it!) from the Cauchy-Schwarz inequality.

The distance between two subsets X and Y of a metric space is defined as

$$\rho(X, Y) = \inf_{x \in X, y \in Y} \rho(x, y).$$

Theorem 5.73. *The distance from a vector x in a Euclidean space V to a subspace $U \subset V$ equals $|\operatorname{ort}_U x|$, and $\operatorname{pr}_U x$ is the vector of U nearest to x.*

Proof. In Figure 5.4, $y = \operatorname{pr}_U x$, $z = \operatorname{ort}_U x$. For any $y' \in U$, $y' \ne y$, we have

$$\rho(x, y') = |z'| = \sqrt{|z|^2 + |u|^2} > |z| = \rho(x, y).$$

□

Example 5.74. Computations in Example 5.71 imply that the polynomial $\frac{3}{2}x^2 - \frac{3}{5}x + \frac{1}{20}$ is the nearest one to x^3 in the metric of the space $C_2[0, 1]$. The distance between them is $\frac{1}{20\sqrt{7}}$.

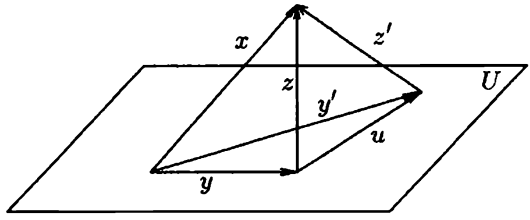


Figure 5.4

The following theorem contains the explicit expression for the distance between a vector x and a subspace U determined by its basis $\{e_1, \ldots, e_k\}$.

Theorem 5.75. $(\rho(x, U))^2 = \dfrac{\det G(e_1, \ldots, e_k, x)}{\det G(e_1, \ldots, e_k)}$.

Proof. If $x \in U$, then $\rho(x, U) = 0$ and $\det G(e_1, \ldots, e_k, x) = 0$. Thus, in this case, the theorem is true.

Assume $x \notin U$ and put $z = \operatorname{ort}_U x$. Applying Theorem 5.47 to the space $U \oplus \langle x \rangle$, we have

$$|z|^2 = (z, z) = \frac{\delta_{k+1}}{\delta_k} = \frac{\det G(e_1, \ldots, e_k, x)}{\det G(e_1, \ldots, e_k)}.$$

□

This formula can be used to calculate the volume of a parallelepiped in a Euclidean space.

A *parallelepiped* on vectors $a_1, \ldots, a_n$ in a Euclidean space is the set

$$P(a_1, \ldots, a_n) = \left\{ \sum_i x_i a_i \colon 0 \le x_i \le 1 \right\}.$$

The *base* of this n-dimensional parallelepiped is the $(n-1)$-dimensional parallelepiped $P(a_1, \ldots, a_{n-1})$ and its *height* is the length of the vector $\operatorname{ort}_{\langle a_1, \ldots, a_{n-1} \rangle} a_n$. For $n = 2, 3$, this definition agrees with the standard one from elementary geometry. Keeping in mind the well-known formulas for the area of a parallelogram and the volume of a three-dimensional parallelepiped, we give the following inductive definition:

Definition 5.76. The *volume* of an n-dimensional parallelepiped $(n > 1)$ is the product of the volume of its base and its height. The volume of a one-dimensional parallelepiped $P(a)$ is the length of the vector a.

The volume of a parallelepiped P is denoted $\operatorname{vol} P$.

Theorem 5.77. $(\operatorname{vol} P(a_1, \ldots, a_n))^2 = \det G(a_1, \ldots, a_n)$.

Proof. We will prove this theorem by induction on n. For $n = 1$, it holds by definition. For $n > 1$, the definition says that

$$\operatorname{vol} P(a_1, \ldots, a_n) = \operatorname{vol} P(a_1, \ldots, a_{n-1}) \cdot h,$$

where h is the length of the vector $\operatorname{ort}_{\langle a_1, \ldots, a_{n-1} \rangle} a_n$, i.e., the distance from the vector a_n to the subspace $\langle a_1, \ldots, a_{n-1} \rangle$. The induction hypothesis together with Theorem 5.75 implies

$$\begin{aligned}(\operatorname{vol} P(a_1, \ldots, a_n))^2 &= \det G(a_1, \ldots, a_{n-1}) \cdot \frac{\det G(a_1, \ldots, a_{n-1}, a_n)}{\det G(a_1, \ldots, a_{n-1})} \\ &= \det G(a_1, \ldots, a_n).\end{aligned}$$

□

In particular, we see that while the base of a parallelepiped depends on what vector we count "last," the volume of the parallelepiped as defined above depends only on the parallelepiped itself. Together with the expressions for the area of a parallelogram and the volume of a three-dimensional parallelepiped, this seems like a reasonable motivation for the above definition. However, the really convincing motivation comes from measure theory, which explains what quantity in general should be called the volume of a set.

Assume that vectors $a_1, \ldots, a_n$ are expressed via vectors of an orthonormal basis $\{e_1, \ldots, e_n\}$ with the help of a matrix A:

$$(a_1, \ldots, a_n) = (e_1, \ldots, e_n)A.$$

Theorem 5.78. $\operatorname{vol} P(a_1, \ldots, a_n) = |\det A|$.

Proof. This follows from the equality

$$G(a_1, \ldots, a_n) = A^{\mathsf{T}} E A = A^{\mathsf{T}} A,$$

which implies

$$\det G(a_1, \ldots, a_n) = (\det A)^2.$$

□

This equality can be regarded as the "geometric meaning" of $|\det A|$. As for the sign of $\det A$, it can be interpreted as the orientation of the system $\{a_1, \ldots, a_n\}$ (with respect to the basis $\{e_1, \ldots, e_n\}$). Recall that when introducing determinants of order n in Section 2.4, we relied on the fact that determinants of order 2 and 3 give the oriented area of a parallelogram and the oriented volume of a parallelepiped, respectively.

We showed in Section 2.2 that the structure of a vector space (over a given field) depends only on its dimension. Is this also true for Euclidean

spaces? To answer this question, we have to understand, first, which Euclidean spaces should be considered "similar" or more precisely, isomorphic. It is natural to accept the following definition:

Definition 5.79. Euclidean vector spaces V and U are *isomorphic* if there exists a bijective map $f\colon V \to U$ which is an isomorphism of vector spaces and satisfies the following condition:

$$(f(a), f(b)) = (a, b) \qquad \forall a, b \in V.$$

The map f is then called an *isomorphism* of the spaces V and U.

Clearly, if Euclidean spaces are isomorphic, their dimensions are the same. The converse also turns out to be true.

Theorem 5.80. *Two Euclidean vector spaces of the same* (finite) *dimension are isomorphic.*

Proof. Let V and U be n-dimensional Euclidean spaces. In each, we choose an orthonormal basis $\{v_1, \ldots, v_n\}$ and $\{u_1, \ldots, u_n\}$, respectively. Let $f\colon V \to U$ be an isomorphism of vector spaces that maps v_i to u_i $(i = 1, \ldots, n)$. Then

$$(f(v_i), f(v_j)) = (u_i, u_j) = \delta_{ij} = (v_i, v_j),$$

implying that

$$(f(a), f(b)) = (a, b)$$

for any $a, b \in V$. □

In particular, any two-dimensional (respectively, three-dimensional) Euclidean space is exactly like E^2 (respectively, E^3). Thus, when we are considering vectors that lie in a two- or three-dimensional subspace, we can evoke theorems from elementary geometry. For instance, in this way we can prove the Cauchy–Schwarz inequality (5.19), the triangle inequality (5.25), and Theorem 5.73.

5.5. Hermitian Spaces

When we try to introduce metric in a complex vector space exactly as we did it for real spaces, we encounter a difficulty: there are no positive definite quadratic functions on a complex space. This difficulty can be circumvented by the introduction of so-called sesquilinear functions (not a very good name but nobody has come up with a better one).

Definition 5.81. Let V be a complex vector space. A function $\alpha\colon V \times V \to \mathbb{C}$ is called *sesquilinear* if it is linear in the second argument and anti-linear in the first. The latter means that

$$\alpha(x_1 + x_2, y) = \alpha(x_1, y) + \alpha(x_2, y),$$
$$\alpha(\lambda x, y) = \overline{\lambda}\alpha(x, y).$$

Remark 5.82. Sometimes a sesquilinear function is defined as linear in the first and anti-linear in the second argument.

The theory of sesquilinear functions is similar to that of bilinear functions. Thus, we are going to present it briefly and will fully discuss only points of significant difference.

Let $\{e_1, \dots, e_n\}$ be a basis of the space V. A sesquilinear function α is determined by the numbers $a_{ij} = \alpha(e_i, e_j)$. Namely,

$$\alpha(x, y) = \sum_{i,j} a_{ij}\overline{x}_i y_j. \tag{5.27}$$

The matrix $A = (a_{ij})$ is called the *matrix of the function* α in the basis $\{e_1, \dots, e_n\}$. A change of basis

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C$$

induces the following change of A:

$$A' = C^*AC, \tag{5.28}$$

where $C^* = \overline{C}^{\mathsf{T}}$. (The bar stands for complex conjugation applied to every entry of the matrix C.) The function α is called nondegenerate if

$$\operatorname{Ker}\alpha := \{y \in V\colon \alpha(x, y) = 0 \ \forall x \in V\} = 0.$$

This condition is equivalent to A being nonsingular.

A sesquilinear function α is called *Hermitian* (respectively, *skew-Hermitian*) if $\alpha(y, x) = \overline{\alpha(x, y)}$ (respectively, $\alpha(y, x) = -\overline{\alpha(x, y)}$). Multiplication by i makes a Hermitian function skew-Hermitian and vice versa.

A function α is Hermitian (respectively, skew-Hermitian) if and only if its matrix A satisfies the condition $A^* = A$ (respectively, $A^* = -A$). Such matrices are called *Hermitian* (respectively, *skew-Hermitian*). Observe that diagonal elements of a Hermitian matrix are real and of a skew-Hermitian matrix, purely imaginary.

To each Hermitian sesquilinear function α corresponds the *Hermitian quadratic function*

$$q(x) = \alpha(x, x).$$

It is easy to see that its values are real. Formulas

$$q(x+y) = q(x) + q(y) + \alpha(x,y) + \alpha(y,x),$$
$$q(x+\imath y) = q(x) + q(y) + \imath\alpha(x,y) - \imath\alpha(y,x)$$

allow us to recover α from q. In particular, if $q \equiv 0$, then $\alpha \equiv 0$.

Let α be a Hermitian sesquilinear function. Just as in the case of symmetric bilinear functions, we define *orthogonal vectors* and the *orthogonal complement* of a subspace with respect to α. The analogue of Proposition 5.45 holds. It implies that every Hermitian sesquilinear function together with the corresponding quadratic function can be presented in the *normal form*:

$$\alpha(x,y) = \overline{x}_1 y_1 + \cdots + \overline{x}_k y_k - \overline{x}_{k+1} y_{k+1} - \cdots - \overline{x}_{k+l} y_{k+l}, \tag{5.29}$$

$$q(x) = |x_1|^2 + \cdots + |x_k|^2 - |x_{k+1}|^2 - \cdots - |x_{k+l}|^2. \tag{5.30}$$

An Hermitian quadratic function q (respectively, the corresponding Hermitian sesquilinear function) is said to be *positive definite* if $q(x) > 0$ for $x \neq 0$. This happens if and only if in the normal form (5.29), we have $k = n$, $l = 0$.

In general, the *Law of Inertia* holds. It says that numbers k and l are determined uniquely. They are called the *positive* and the *negative indices of inertia* of q.

Since for any complex matrix,

$$\det A^* = \overline{\det A},$$

the determinant of a Hermitian matrix is always real. If all corner minors of the matrix of a Hermitian sesquilinear function are nonzero, then just as in the case of a bilinear function, one can perform the orthogonalization and deduce the Jacobi method for determining indices of inertia from corner minors.

A complex analogue of a Euclidean space is a Hermitian space. This is a complex vector space with a fixed positive definite Hermitian sesquilinear function called *inner product* and denoted $(\ ,\ )$.

Example 5.83. The space $\mathbb{C}^n$ with the inner product

$$(x,y) = \overline{x}_1 y_1 + \cdots + \overline{x}_n y_n.$$

Example 5.84. The space of continuous complex-valued functions on the interval $[0,1]$ with the inner product

$$(f,g) = \int_0^1 \overline{f(x)} g(x) dx.$$

In a Hermitian space, the *length* of a vector is defined as

$$|x| = \sqrt{(x,x)}.$$

The Cauchy–Schwarz inequality

$$|(x,y)| \le |x||y|$$

and the triangle inequality

$$|x+y| \le |x| + |y|$$

hold (prove this).

A basis $\{e_1, \dots, e_n\}$ of a Hermitian space is called *orthonormal* if the inner product has the normal form in this basis, i.e.,

$$(e_i, e_j) = \delta_{ij}.$$

The matrix of transition from one orthonormal basis to another satisfies the condition $C^* = C^{-1}$. Such complex matrices are called *unitary*.

Exercise 5.85. Express the condition of unitarity for matrices using matrix entries (in two different ways).

Observe that the absolute value of the determinant of a unitary matrix C equals 1. Indeed, by taking the determinants of both sides in $C^*C = E$, we have

$$\overline{\det C} \cdot \det C = 1,$$

implying $|\det C| = 1$.

Just as in Euclidean spaces, for any subspace U of a Hermitian space V, we have the following decomposition:

$$V = U \oplus U^{\perp}.$$

If $\{e_1, \dots, e_k\}$ is an orthogonal basis of the subspace U, then the orthogonal projection of a vector $x \in V$ onto U is expressed as

$$\mathrm{pr}_U x = \sum_{i=1}^{k} \frac{(e_i, x)}{(e_i, e_i)} e_i.$$

(Notice the difference between this formula and formula (5.23).)

Analogues of Theorems 5.73 and 5.75 also hold for Hermitian spaces.

From a mathematical standpoint, Hermitian spaces are as useful as complex numbers. This will become clear in the next chapter. From a physical standpoint, Hermitian spaces are necessary for constructing an adequate quantum-mechanical view of the world.

Chapter 6

Linear Operators

The theory of linear operators is the crux of linear algebra and the main source of its numerous applications. Just like bilinear functions, linear operators on a finite-dimensional vector space are described by square matrices. So, in some sense, these objects are of equal difficulty (but a symmetric or a skew-symmetric bilinear function is a simpler object than a generic linear operator).

In this chapter, we continue using conventions introduced at the beginning of Chapter 5.

6.1. Matrix of a Linear Operator

Definition 6.1. A *linear operator* (or a *linear transformation*) on a vector space V is a linear map from V to itself.

Explicitly, a linear operator is a map $\mathcal{A} : V \to V$ such that

(i) $\mathcal{A}(x + y) = \mathcal{A}x + \mathcal{A}y$ for any $x, y \in V$;

(ii) $\mathcal{A}(\lambda x) = \lambda \mathcal{A}x$ for any $x \in V$, $\lambda \in K$.

(Usually, we will denote linear operators by script letters.)

Given a basis $\{e_1, \ldots, e_n\}$ in the space V, a linear operator can be described by a matrix.

Definition 6.2. The *matrix of a linear operator* $\mathcal{A}$ in a basis $\{e_1, \ldots, e_n\}$ is the matrix $A = (a_{ij})$ determined by the following equalities:

$$\mathcal{A}e_j = \sum_i a_{ij} e_i. \tag{6.1}$$

In other words, the jth column of A consists of the coordinates of the vector $\mathcal{A}e_j$ in the basis $\{e_1, \ldots, e_n\}$. (Note that unlike the definition of the matrix of a linear map, this one features only one basis!)

Equalities (6.1) can be rewritten as

$$(\mathcal{A}e_1, \ldots, \mathcal{A}e_n) = (e_1, \ldots, e_n)A \tag{6.2}$$

(cf. definition (2.7) of the transition matrix in Chapter 2).

Obviously, for vectors $f_1, \ldots, f_n \in V$, there exists a unique linear operator $\mathcal{A}$ that maps the basis vectors $e_1, \ldots, e_n$ to $f_1, \ldots, f_n$, respectively. This operator maps a vector $x = \sum_i x_i e_i$ to the vector $\sum_i x_i f_i$. Therefore, a linear operator is uniquely determined by its matrix and every square matrix of order n is the matrix of a linear operator (in the given basis).

Let us find the explicit expression for coordinates of the image $y = \mathcal{A}x$ of a vector x. For $x = \sum_j x_j e_j$, we have

$$y = \sum_j x_j \mathcal{A}e_j = \sum_{i,j} a_{ij} x_j e_i = \sum_i y_i e_i,$$

where

$$y_i = \sum_j a_{ij} x_j. \tag{6.3}$$

Denote by X and Y the columns of coordinates of vectors x and y, respectively; then (6.3) can be rewritten in the matrix form as follows:

$$Y = AX \tag{6.4}$$

(cf. expression (2.8) for the change of coordinates in Chapter 2).

Now we will describe how the matrix of a linear operator changes under a transition to another basis. Let

$$(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)C.$$

Since the operator $\mathcal{A}$ is linear, we have

$$\begin{aligned}(\mathcal{A}e'_1, \ldots, \mathcal{A}e'_n) &= (\mathcal{A}e_1, \ldots, \mathcal{A}e_n)C \\ &= (e_1, \ldots, e_n)AC = (e'_1, \ldots, e'_n)C^{-1}AC.\end{aligned}$$

Thus, if we denote by A' the matrix of $\mathcal{A}$ in the basis $\{e'_1, \ldots, e'_n\}$,

$$A' = C^{-1}AC. \tag{6.5}$$

With a change of basis, the matrix of a linear operator on V can be reduced to a simpler form. In particular, this is possible if we know an invariant subspace of V.

Definition 6.3. A subspace $U \subset V$ is *invariant* with respect to an operator $\mathcal{A}$ if

$$\mathcal{A}U \subset U$$

(i.e., $\mathcal{A}u \in U$ for any $u \in U$).

The restriction $\mathcal{A}|_U$ of a linear operator $\mathcal{A}$ to an invariant subspace U is a linear operator on U.

If a basis $\{e_1, \ldots, e_n\}$ of a space V is chosen so that $U = \langle e_1, \ldots, e_k \rangle$ (which is always possible), then the matrix of the operator $\mathcal{A}$ is of the form

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix}, \tag{6.6}$$

where B is the matrix of the operator $\mathcal{A}|_U$ in the basis $\{e_1, \ldots, e_k\}$, C is a square matrix of order $n - k$, and D is a $k \times (n - k)$ matrix. Conversely, if the matrix of $\mathcal{A}$ in a basis $\{e_1, \ldots, e_n\}$ has the form (6.6), where B is a square matrix of order k, then $U = \langle e_1, \ldots, e_k \rangle$ is an invariant subspace.

Things look even better when V splits into a direct sum of two invariant subspaces U and W:

$$V = U \oplus W.$$

If $\{e_1, \ldots, e_k\}$ is a basis of U and $\{e_{k+1}, \ldots, e_n\}$ is a basis of W, then $\{e_1, \ldots, e_n\}$ is a basis of V and in this basis, the matrix of the operator $\mathcal{A}$ has the form

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}, \tag{6.7}$$

where B is the matrix of $\mathcal{A}|_U$ in the basis $\{e_1, \ldots, e_k\}$ and C is the matrix of $\mathcal{A}|_W$ in the basis $\{e_{k+1}, \ldots, e_n\}$.

More generally, let the space V split into a direct sum of k invariant subspaces $V_1, V_2, \ldots, V_k$. Then in the basis of V comprised of bases of these subspaces, the matrix of $\mathcal{A}$ is of the form

$$\begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_n \end{pmatrix}, \tag{6.8}$$

where A_i is the matrix of the operator $\mathcal{A}|_{V_i}$.

Example 6.4. The rotation through an angle α is a linear operator on E^2 (see Example 2.53). We proved in Example 2.56 that in the orthonormal basis $\{e_1, e_2\}$, its matrix is

$$\Pi(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{6.9}$$

In particular, the matrix of the rotation through $\frac{\pi}{2}$ in such a basis is

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

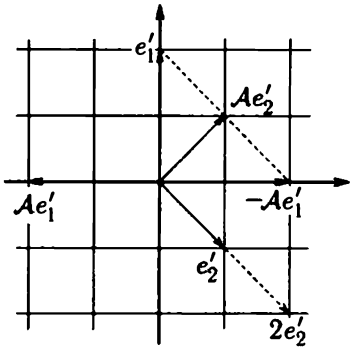


Figure 6.1

Let us find the matrix of this operator in the basis

$$e_1' = 2e_2, \qquad e_2' = e_1 - e_2. \tag{6.10}$$

Figure 6.1 shows that

$$\mathcal{A}e_1' = -e_1' - 2e_2', \qquad \mathcal{A}e_2' = e_1' + e_2'.$$

This implies that

$$A' = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}.$$

Of course, the matrix A' can be found using formula (6.5). It follows from formula (6.10) that

$$C = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}, \qquad C^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix}.$$

Therefore,

$$A' = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}.$$

Example 6.5. Similarly, the rotation about an axis through an angle α is a linear operator on E^3. In an orthonormal basis $\{e_1, e_2, e_3\}$ such that e_3 is collinear with the axis of rotation, this operator has the following matrix:

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \Pi(\alpha) & 0 \\ 0 & 1 \end{pmatrix}.$$

This agrees with the way E^3 splits into the direct sum of two invariant subspaces:

$$E^3 = \langle e_1, e_2 \rangle \oplus \langle e_3 \rangle. \tag{6.11}$$

Example 6.6. In Example 2.54 we considered the orthogonal projection onto a plane as a linear map from the space E^3 to the space of vectors on this plane. However, we may view this map as a linear operator on E^3. In an orthonormal basis chosen so that its first two vectors lie on this plane, the matrix of this linear operator is of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The splitting (6.11) is a splitting into a direct sum of invariant subspaces in this case as well.

Example 6.7. Differentiation is a linear operator on the space of polynomials. Although this space is infinite-dimensional, it is a union of finite-dimensional invariant subspaces that consist of polynomials whose degree is no greater than some fixed bound. In the basis $\{1, x, x^2, \ldots, x^n\}$ of the space of polynomials of degree no greater than n, the operator of differentiation has the following matrix:

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 2 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & n \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

In the basis $\{1, \frac{x}{1!}, \frac{x^2}{2!}, \ldots, \frac{x^n}{n!}\}$, this operator has an even simpler matrix:

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}. \tag{6.12}$$

Example 6.8. Let φ be a bijective transformation of a set X. Then the map φ_* defined as

$$(\varphi_* f)(x) = f(\varphi^{-1}(x)) \tag{6.13}$$

is a linear operator on the space $F(X, K)$ of K-valued functions on X. (One could apply φ rather than its inverse to the argument; the reason for our

choice will be explained in Chapter 10.) For instance, let $X = \mathbb{R}$, $K = \mathbb{R}$, and $\varphi(x) = x + a$, $a \in \mathbb{R}$. Then

$$(\varphi_* f)(x) = f(x - a).$$

(The graph of $\varphi_* f$ is obtained from that of f by the shift to the right by a.) Since

$$\cos(x - a) = \cos a \cdot \cos x + \sin a \cdot \sin x,$$
$$\sin(x - a) = -\sin a \cdot \cos x + \cos a \cdot \sin x,$$

the subspace $\langle \cos x, \sin x \rangle$ is invariant with respect to φ_*. In the basis $\{\cos x, \sin x\}$, the matrix of the restriction of φ_* to this subspace is

$$\begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix} = \Pi(a).$$

Example 6.9. For any algebra A, the transformation

$$L_a \colon x \mapsto ax, \qquad a \in A,$$

called the *left multiplication by* a, is a linear operator. For example, regard the field $\mathbb{C}$ as an algebra over $\mathbb{R}$. Equalities

$$(a + bi) \cdot 1 = a + bi, \qquad (a + bi) \cdot i = -b + ai$$

show that the matrix of the operator L_{a+bi} in the basis $\{1, i\}$ is

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Exercise 6.10. Determine the matrix of the left multiplication by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in the algebra $\mathrm{L}_2(K)$ in the basis that consists of matrix units. Prove that the subspaces $\langle E_{11}, E_{21} \rangle$ and $\langle E_{12}, E_{22} \rangle$ are invariant.

Linear operators on the same vector space can be added and multiplied by one another and by numbers. These operations are defined just as they are defined for general linear maps (see Section 2.3). They correspond to the same operations on matrices, e.g., the matrix of the product of two linear operators in some basis equals the product of their matrices in this basis.

The properties of operations on linear maps that we proved in Section 2.3 imply that the set of all linear operators on a vector space V is an associative algebra. We denote it $\mathrm{L}(V)$. Observe that if $\dim V = n$, then $\dim \mathrm{L}(V) = \dim \mathrm{L}_n(K) = n^2$.

The algebra $\mathrm{L}(V)$ has a unity. This is the *identity operator*, which we denote $\mathcal{E}$. In each basis, the matrix of $\mathcal{E}$ is the identity matrix E.

A linear operator $\mathcal{A} \in \mathrm{L}(V)$ is invertible if and only if $\operatorname{Ker} \mathcal{A} = 0$ and $\operatorname{Im} \mathcal{A} = V$. In the finite-dimensional case, it follows from Theorem 2.64 that if $\operatorname{Ker} \mathcal{A} = 0$, then automatically $\operatorname{Im} \mathcal{A} = V$ and vice versa. On the other

hand, it is clear that a linear operator is invertible if and only if its matrix is invertible, i.e., nonsingular.

In the general case, the dimension of the space $\operatorname{Im}\mathcal{A}$ is called the *rank of the linear operator* $\mathcal{A}$ and is denoted $\operatorname{rk}\mathcal{A}$. By Corollary 2.65, it is equal to the rank of the matrix A (in any basis).

Formula (6.5) implies that the determinant of the matrix of $\mathcal{A}$ does not depend on the choice of a basis. It is called the *determinant of the linear operator* $\mathcal{A}$ and is denoted $\det\mathcal{A}$.

6.2. Eigenvectors

The main goal of the theory of linear operators is to reduce the matrix of a linear operator to the simplest possible form through the choice of a particular basis.

As we have remarked, in order to achieve this goal, it is useful to know invariant subspaces. Considering them, we come to the notion of an eigenvector.

Definition 6.11. A nonzero vector $e \in V$ is an *eigenvector* of an operator $\mathcal{A}$ if $\mathcal{A}e = \lambda e$ for some $\lambda \in K$. The number $\lambda \in K$ is called the *eigenvalue* of the operator $\mathcal{A}$ corresponding to the vector e.

Obviously, a nonzero vector e is an eigenvector if and only if the one-dimensional subspace $\langle e\rangle$ is invariant. In a basis of eigenvectors (if such exists), the matrix of the operator is diagonal, a dream come true.

Example 6.12. For the operator of differentiation on the space of polynomials, the only (up to multiplication by a number) eigenvector is the polynomial 1 with the eigenvalue 0. So in this case, there is no basis of eigenvectors.

Example 6.13. Eigenvectors of a rotation through an angle $\alpha \neq k\pi$ in the three-dimensional space are the vectors on the axis of rotation (and their eigenvalue is 1). When $\alpha = k\pi$, the vectors orthogonal to the axis of rotation are also eigenvectors (with eigenvalues $(-1)^k$). Thus, in this example the basis of eigenvectors exists only when $\alpha = 0$ or π (if we are to consider only $0 \leq \alpha < 2\pi$).

An eigenvector with an eigenvalue λ exists if and only if the operator $\mathcal{A} - \lambda\mathcal{E}$ is singular, i.e., if $\det(\mathcal{A} - \lambda\mathcal{E}) = 0$. If $A = (a_{ij})$ is a matrix of $\mathcal{A}$ in some basis, then

$$\det(\mathcal{A} - t\mathcal{E}) = \begin{vmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{vmatrix},$$

implying that $\det(\mathcal{A} - t\mathcal{E})$ is a polynomial of degree n in t.

Definition 6.14. The polynomial

$$f_{\mathcal{A}}(t) = (-1)^n \det(\mathcal{A} - t\mathcal{E}) = \det(t\mathcal{E} - \mathcal{A})$$

is called the *characteristic polynomial* of $\mathcal{A}$.

It is easy to see that the coefficient of t^n in the polynomial $f_{\mathcal{A}}(t)$ equals 1 and the coefficient of t^{n-1} equals $-\operatorname{tr} A$, where $\operatorname{tr} A$ is the trace of A (the sum of diagonal entries of A). The free term of $f_{\mathcal{A}}(t)$ equals $f_{\mathcal{A}}(0) = (-1)^n \det A$.

Exercise 6.15. Prove that the coefficient of t^{n-k} in the polynomial $f_{\mathcal{A}}(t)$ equals $(-1)^k \times$(sum of principal minors of A of order k). (A *principal minor* of a square matrix is the determinant of a submatrix which is symmetric with respect to the main diagonal.)

Observe that by definition, the characteristic polynomial of a linear operator does not depend on the choice of a basis. In particular, it follows that the trace of a linear operator does not depend on the choice of a basis.

Actually, we have proved the following

Theorem 6.16. *The eigenvalues of a linear operator are exactly the roots of its characteristic polynomial.*

Corollary 6.17. *Every linear operator on a complex vector space has an eigenvector.*

A linear operator on a real vector space may have no eigenvectors as the example of a planar rotation through an angle $\alpha \neq 0, \pi$ shows. However, the use of complex numbers allows us to obtain some useful information concerning linear operators over reals as well. This is achieved by the so-called *complexification*.

Let V be a real vector space. We will construct from it a complex vector space $V(\mathbb{C})$ just as we constructed the field $\mathbb{C}$ from $\mathbb{R}$. That is, for elements of $V(\mathbb{C})$, we take the pairs (x, y), where $x, y \in V$, and define the addition of such pairs and their multiplication by complex numbers as follows:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2),$$
$$(\lambda + i\mu)(x, y) = (\lambda x - \mu y, \mu x + \lambda y).$$

It is easy to check that in this way we obtain a vector space over $\mathbb{C}$. By definition, addition of pairs of the type $(x, 0)$ and their multiplication by real numbers comes down to the corresponding operations over their first components. Identify every such pair $(x, 0)$ with the vector $x \in V$; then V becomes embedded into $V(\mathbb{C})$ as a real subspace. Moreover,

$$(x, y) = x + iy.$$

Every basis of a space V (over $\mathbb{R}$) is at the same time a basis of $V(\mathbb{C})$ (over $\mathbb{C}$). However, $V(\mathbb{C})$ possesses other bases.

Every linear operator $\mathcal{A}$ in the space V can be uniquely extended to the operator $\mathcal{A}_{\mathbb{C}}$ on the space $V(\mathbb{C})$. In a basis consisting of real vectors, the operator $\mathcal{A}_{\mathbb{C}}$ has the same matrix as $\mathcal{A}$.

The operator $\mathcal{A}_{\mathbb{C}}$ can have complex eigenvalues and corresponding complex eigenvectors. What is their meaning in real terms?

Proposition 6.18. *A vector $x + \imath y$, $x, y \in V$, is an eigenvector of an operator $\mathcal{A}_{\mathbb{C}}$ with an imaginary eigenvalue $\lambda + \imath\mu$ ($\lambda, \mu \in \mathbb{R}, \mu \neq 0$) if and only if $U = \langle x, y\rangle \subset V$ is a two-dimensional invariant subspace for the operator $\mathcal{A}$ and*

$$\begin{aligned} \mathcal{A}x &= \lambda x - \mu y, \\ \mathcal{A}y &= \mu x + \lambda y. \end{aligned} \tag{6.14}$$

The proof of this proposition is obtained via direct calculation. Equalities (6.14) mean that in the basis $\{x, y\}$ of the space U, the operator $\mathcal{A}|_U$ has the following matrix:

$$\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix}. \tag{6.15}$$

They also imply that the vector $x - \imath y$ is an eigenvector of $\mathcal{A}_{\mathbb{C}}$ with the eigenvalue $\lambda - \imath\mu$.

Example 6.19. The operator $\mathcal{A}$ of rotation of the Euclidean plane through an angle α has the matrix $\Pi(\alpha)$ in an orthonormal basis $\{e_1, e_2\}$ (see (6.9)). Therefore, the vector $e_1 + \imath e_2$ is an eigenvector of the operator $\mathcal{A}_{\mathbb{C}}$ with the eigenvalue $\cos\alpha - \imath\sin\alpha$, and $e_1 - \imath e_2$ is an eigenvector with the eigenvalue $\cos\alpha + \imath\sin\alpha$. Thus, the matrix of a rotation can be diagonalized in the complex space.

As a corollary of Proposition 6.18, we obtain the following important

Theorem 6.20. *Every linear operator over the field of real numbers has a one- or two-dimensional invariant subspace.*

For a given eigenvalue λ, the corresponding eigenvectors can be found from the system of homogeneous equations

$$(A - \lambda E)X = 0,$$

where X denotes the column of coordinates of the unknown vector. Together with the zero vector, they comprise the subspace

$$V_\lambda(\mathcal{A}) = \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})$$

called the *eigenspace* of the operator $\mathcal{A}$ corresponding to the eigenvalue λ. Its dimension is $n - \operatorname{rk}(\mathcal{A} - \lambda\mathcal{E})$, where $n = \dim V$.

Theorem 6.21. *Eigenspaces corresponding to distinct eigenvalues* $\lambda_1, \ldots, \lambda_k$ *of an operator* $\mathcal{A}$ *are linearly independent.*

Proof. We will prove this theorem by induction on k. For $k = 1$ there is nothing to prove. Assume $k > 1$ and let

$$e_1 + \cdots + e_{k-1} + e_k = 0, \qquad e_i \in V_{\lambda_i}(\mathcal{A}).$$

Applying $\mathcal{A}$, we obtain

$$\lambda_1 e_1 + \cdots + \lambda_{k-1} e_{k-1} + \lambda_k e_k = 0.$$

By subtracting the original equality multiplied by λ_k, we obtain

$$(\lambda_1 - \lambda_k)e_1 + \cdots + (\lambda_{k-1} - \lambda_k)e_{k-1} = 0.$$

It follows by induction that $e_1 = \cdots = e_{k-1} = 0$. But then $e_k = 0$. □

Corollary 6.22. *If the characteristic polynomial* $f_{\mathcal{A}}(t)$ *has* n *distinct roots, then there exists a basis of eigenvectors of* $\mathcal{A}$.

This condition is not necessary for the existence of a basis of eigenvectors. For instance, all nonzero vectors are eigenvectors of the identity operator $\mathcal{E}$, hence any basis consists of its eigenvectors. Yet, its characteristic polynomial $f_{\mathcal{E}}(t) = (t-1)^n$ has only one root 1 (of multiplicity n).

Consider now two interesting (and important) examples.

Example 6.23. Let $V = U \oplus W$. The linear operator $\mathcal{P}$ defined as

$$\mathcal{P}(y + z) = y, \qquad y \in U,\ z \in W,$$

is called the *projection* onto U along W. Obviously,

$$U = V_1(\mathcal{P}), \qquad W = V_0(\mathcal{P}).$$

In a basis of V obtained from bases of U and W, the operator $\mathcal{P}$ has a diagonal matrix with 0's and 1's on the diagonal.

Exercise 6.24. Prove that a linear operator $\mathcal{P}$ is a projection (for some U and W) if and only if $\mathcal{P}^2 = \mathcal{P}$.

Example 6.25. In the same notation, the linear operator $\mathcal{R}$ defined as

$$\mathcal{R}(y + z) = y - z, \qquad y \in U,\ z \in W,$$

is called the *reflection* through U along W. Obviously,

$$U = V_1(\mathcal{R}), \qquad W = V_{-1}(\mathcal{R}).$$

In a basis of V obtained from bases of U and W, the operator $\mathcal{R}$ has a diagonal matrix with 1's and -1's on the diagonal.

Exercise 6.26. Prove that a linear operator $\mathcal{R}$ is a reflection (for some U and W) if and only if $\mathcal{R}^2 = \mathcal{E}$.

To obtain the necessary and sufficient conditions for the existence of a basis consisting of eigenvectors, we need to prove the following proposition first.

Proposition 6.27. *The characteristic polynomial of the restriction of a linear operator to an invariant subspace divides the characteristic polynomial of this operator.*

Proof. Let $\mathcal{B}$ be the restriction of $\mathcal{A}$ to an invariant subspace $U \subset V$. In a basis of V whose first vectors form a basis of U, the matrix A of the operator $\mathcal{A}$ is of the form (6.6), where B is the matrix of $\mathcal{B}$. Therefore,

$$f_{\mathcal{A}}(t) = f_{\mathcal{B}}(t) \cdot \det(tE - C). \tag{6.16}$$

□

Corollary 6.28. *The dimension of an eigenspace of a linear operator does not exceed the multiplicity of the corresponding root of the characteristic polynomial.*

Proof. Let $\dim V_\lambda(\mathcal{A}) = k$. Then the characteristic polynomial of the restriction of the operator $\mathcal{A}$ to $V_\lambda(\mathcal{A})$ equals $(t - \lambda)^k$. Applying Proposition 6.27 to the subspace $U = V_\lambda(\mathcal{A})$, we complete the proof. □

Example 6.29. Consider the operator of differentiation on the space of polynomials of degree not greater than n. We calculated its matrix in Example 6.7 and can now conclude that its characteristic polynomial is t^{n+1}. This polynomial has root 0 of multiplicity $n + 1$ but the dimension of the corresponding eigenspace is 1 (see Example 6.12). This example shows that the dimension of an eigenspace can be strictly less than the multiplicity of the corresponding root of the characteristic polynomial.

Theorem 6.30. *In order for a basis of eigenvectors of a linear operator $\mathcal{A}$ to exist, the following conditions are necessary and sufficient:*

(i) the characteristic polynomial $f_{\mathcal{A}}(t)$ splits into linear factors;

(ii) the dimension of every eigenspace equals the multiplicity of the corresponding root of the polynomial $f_{\mathcal{A}}(t)$.

Proof. Let $\lambda_1, \ldots, \lambda_s$ be all roots of $f_{\mathcal{A}}(t)$ with multiplicities $k_1, \ldots, k_s$, respectively. Denote the eigenspace corresponding to λ_i by V_i. By Corollary 6.28, $\dim V_i \leq k_i$, thus

$$\sum_i \dim V_i \leq \sum_i k_i \leq n. \tag{6.17}$$

However, the only way to obtain a basis of eigenvectors is to take the union of bases of eigenspaces. In order for this procedure to actually give us a

basis of V, it is necessary and sufficient that

$$\sum_i \dim V_i = n.$$

By (6.17), it is equivalent to require that $\sum_i k_i = n$ and $\dim V_i = k_i$ for all i. The first condition means that $f_{\mathcal{A}}(t)$ splits into linear factors and the second is exactly condition (ii) of this theorem. □

6.3. Linear Operators and Bilinear Functions on Euclidean Space

Let V be a Euclidean space with an orthonormal basis $\{e_1, \dots, e_n\}$.

To each vector $a \in V$, there corresponds the linear function

$$\varphi_a(x) = (x, a). \tag{6.18}$$

Moreover, the coefficients $\varphi_a(e_i) = (e_i, a)$ of the linear function φ_a in the basis $\{e_1, \dots, e_n\}$ equal the coordinates of a in this basis. This implies that the map $a \mapsto \varphi_a$ is an isomorphism between the spaces V and V^*. Notice that this isomorphism does not depend on the choice of a basis. Thus, one can say that for a finite-dimensional Euclidean space, the difference between the space and its dual space disappears. Usually this is stated as "the canonical isomorphism identifies V with its dual" (the isomorphism is the one above).

Similarly, to each linear operator $\mathcal{A}$ on the space V, there corresponds a bilinear function

$$\varphi_{\mathcal{A}}(x, y) = (x, \mathcal{A}y). \tag{6.19}$$

Also, the matrix of the bilinear function $\varphi_{\mathcal{A}}(x, y)$ in the basis $\{e_1, \dots, e_n\}$ coincides with the matrix of the operator $\mathcal{A}$ in this basis. Indeed, $\varphi_{\mathcal{A}}(e_i, e_j) = (e_i, \mathcal{A}e_j)$ is nothing but the ith coordinate of the vector $\mathcal{A}e_j$. It follows that the map $\mathcal{A} \mapsto \varphi_{\mathcal{A}}$ is an isomorphism from the space of linear operators to the space of bilinear functions on V. This isomorphism does not depend on the choice of a basis. However, in a nonorthonormal basis, the matrix of $\varphi_{\mathcal{A}}$ may differ from that of the operator $\mathcal{A}$.

For any bilinear function φ, one can define the "transposed" function

$$\varphi^{\top}(x, y) = \varphi(y, x),$$

whose matrix in every basis is the transposed matrix of φ. The linear operator $\mathcal{A}^*$ corresponding to the function $\varphi_{\mathcal{A}}^{\top}$ is called the *adjoint operator* of $\mathcal{A}$. In other words, the adjoint operator is defined by the following formula:

$$(\mathcal{A}^* x, y) = (x, \mathcal{A}y). \tag{6.20}$$

The matrix of the operator $\mathcal{A}^*$ in an orthonormal basis is the transposed matrix of $\mathcal{A}$.

Symmetric (respectively, skew-symmetric) bilinear functions correspond to the so-called *symmetric* (respectively, *skew-symmetric*) linear operators. They are determined by the property $\mathcal{A}^* = \mathcal{A}$ (respectively, $\mathcal{A}^* = -\mathcal{A}$) or, in matrix terms, by the property that their matrices in an orthonormal basis are symmetric (respectively, skew-symmetric). Symmetric operators are also called *selfadjoint*.

Example 6.31. An *orthogonal projection* onto a subspace is a symmetric operator (prove this).

Linear operators such that $\mathcal{A}^* = \mathcal{A}^{-1}$ are called *orthogonal*. In other words, an operator $\mathcal{A}$ is orthogonal if

$$(\mathcal{A}x, \mathcal{A}y) = (x, y), \tag{6.21}$$

i.e., if $\mathcal{A}$ preserves the inner product of vectors. The identity

$$(x, y) = \frac{1}{2}(|x + y|^2 - |x|^2 - |y|^2)$$

implies that an operator $\mathcal{A}$ is orthogonal if and only if it preserves the vector length.

Example 6.32. A linear operator defined by a motion on the space of geometric vectors is orthogonal.

Example 6.33. An *orthogonal reflection* with respect to a subspace (i.e., a reflection along the orthogonal subspace) is an orthogonal operator.

In matrix terms, the orthogonal operators are characterized by the property that their matrix in an orthonormal basis is orthogonal (see Definition 5.69).

Proposition 6.34. *A linear operator of each of the three types described above (i.e., symmetric, skew-symmetric, or orthogonal) has the following property: if a subspace U is invariant with respect to this operator, then so is its orthogonal complement $U^\perp$.*

Proof. Consider the most difficult case: that of an orthogonal operator $\mathcal{A}$. Notice, first, that the operator $\mathcal{A}|_U$ is also orthogonal, hence nonsingular. Therefore, for any vector $x \in U$, there exists a vector $z \in U$ such that $x = \mathcal{A}z$. Now, consider a vector $y \in U^\perp$. In the above notation, for any $x \in U$, we have

$$(x, \mathcal{A}y) = (\mathcal{A}z, \mathcal{A}y) = (z, y) = 0,$$

implying $\mathcal{A}y \in U^\perp$. □

Using this proposition together with Theorem 6.20, we can obtain by induction the canonical form of a matrix of a linear operator of either of these three types.

Theorem 6.35. *For any symmetric operator $\mathcal{A}$, there exists an orthonormal basis of eigenvectors.*

Proof. It suffices to prove that at least one eigenvector exists. In view of Theorem 6.20, it suffices to do this for a two-dimensional space. In this case, the matrix of a symmetric operator in an orthonormal basis is $\left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$. Its characteristic polynomial is

$$f_{\mathcal{A}}(t) = t^2 - (a + c)t + (ac - b^2),$$

and its discriminant

$$D = (a + c)^2 - 4(ac - b^2) = (a - c)^2 + 4b^2$$

is always nonnegative. Hence, $f_{\mathcal{A}}(t)$ has real roots and thus $\mathcal{A}$ has eigenvectors. □

Corollary 6.36. *The characteristic polynomial of a symmetric polynomial splits into linear factors over $\mathbb{R}$. The dimension of each eigenspace equals the multiplicity of the corresponding root. Eigenspaces corresponding to distinct roots are orthogonal.*

Proof. To prove the last claim of the corollary, it suffices to notice that if $\{e_1, \ldots, e_n\}$ is a basis of eigenvectors of $\mathcal{A}$ and $\mathcal{A}e_i = \lambda_i e_i$, then $V_\lambda(\mathcal{A})$ is the linear span of the e_i's such that $\lambda_i = \lambda$.

This can also be shown directly. Indeed, let $x \in V_\lambda(\mathcal{A})$, $y \in V_\mu(\mathcal{A})$, $\lambda \neq \mu$. Then

$$\lambda(x, y) = (\mathcal{A}x, y) = (x, \mathcal{A}y) = \mu(x, y),$$

implying $(x, y) = 0$. □

Using the correspondence between symmetric operators and symmetric bilinear functions, we obtain the following

Corollary 6.37. *For any quadratic function q on a Euclidean space, there exists an orthonormal basis where its matrix is diagonal, i.e.,*

$$q(x) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2. \tag{6.22}$$

Notice that in the statement of the corollary, "orthonormal" is understood in the sense of the inner product and not in the sense of the bilinear function φ corresponding to q. However, since the matrix of φ in this basis is diagonal, this basis is orthogonal (but not orthonormal) in the sense of φ.

Observe that the numbers $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the corresponding symmetric operator, hence they are defined up to a permutation.

Expression (6.22) is called the *canonical form* of the quadratic function q, and the determination of the orthonormal basis where q has such a form is sometimes called the *reduction to principal axes*.

Using the correspondence between symmetric operators and quadratic functions on a Euclidean space, we can obtain another proof of the existence of an eigenvector of a symmetric operator.

Namely, let q be a quadratic function corresponding to a given symmetric operator $\mathcal{A}$, i.e.,

$$q(x) = (\mathcal{A}x, x).$$

Note that since q is a continuous function, it must have a maximum on the unit sphere S in the space V defined as

$$(x, x) = 1.$$

Proposition 6.38. *Every point where the function q reaches a maximum on the sphere S is an eigenvector of the operator $\mathcal{A}$. The value of this maximum is equal to the corresponding eigenvalue.*

Proof. The tangent space to the sphere S at a point x is determined as

$$(x, dx) = 0,$$

i.e., it is the orthogonal complement of the space $\langle x \rangle$. On the other hand, the differential of q equals

$$dq(x) = (\mathcal{A}dx, x) + (\mathcal{A}x, dx) = 2(\mathcal{A}x, dx).$$

If the function q attains its maximum at a point $e \in S$, then its differential is zero on the tangent space to S at this point. By the above, this implies that the vector $\mathcal{A}e$ is orthogonal to all vectors that are orthogonal to e, thus $\mathcal{A} = \lambda e$. Moreover,

$$q(e) = (\mathcal{A}e, e) = \lambda(e, e) = \lambda.$$

□

In this proof we used only the necessary condition of maximality. It also holds at any critical point of the function q on S, in particular, at any point of minimum. Clearly the maximum is attained at the eigenvector $e \in S$ only if λ is the maximum eigenvalue of $\mathcal{A}$.

A symmetric operator is called *positive definite* if its corresponding quadratic function is positive definite or, equivalently, if all its eigenvalues are positive.

We now turn to discussing linear operators of other types.

Theorem 6.39. *For any skew-symmetric linear operator $\mathcal{A}$, there exists an orthonormal basis where its matrix is of the form*

$$A = \begin{pmatrix} H(a_1) & & & & & & 0 \\ & \ddots & & & & & \\ & & H(a_k) & & & & \\ & & & 0 & & & \\ 0 & & & & \ddots & & \\ & & & & & & 0 \end{pmatrix},$$

where $H(a) = \left(\begin{smallmatrix} 0 & -a \\ a & 0 \end{smallmatrix}\right)$.

Proof. This is obvious since $H(a)$ is the generic matrix form of a skew-symmetric operator in an orthonormal basis of a two-dimensional Euclidean space. □

Theorem 6.40. *For any orthogonal operator $\mathcal{A}$, there exists an orthonormal basis where its matrix is of the form*

$$A = \begin{pmatrix} \Pi(\alpha_1) & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \Pi(\alpha_k) & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 1 & & \\ 0 & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix},$$

where $\Pi(\alpha) = \left(\begin{smallmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{smallmatrix}\right)$.

Observe that if we use matrices $\Pi(\pi) = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\Pi(0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, we can have at most one free diagonal entry equal to -1 and at most one equal to 1.

Proof. It suffices to consider orthogonal operators on one-dimensional and two-dimensional spaces. On a one-dimensional space, an orthogonal operator is a multiplication by ± 1.

On a two-dimensional space, every orthogonal operator α is either a rotation through an angle α or a reflection through a line; this was shown in Example 4.18. In the first case, the matrix of $\mathcal{A}$ in any orthonormal basis is $\Pi(\alpha)$. In the second case, there exists an orthonormal basis where the matrix of the operator $\mathcal{A}$ is of the form $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. □

In particular, on a three-dimensional Euclidean space, the matrix of any orthogonal operator $\mathcal{A}$ in an appropriate basis has one of the following forms:

$$\begin{pmatrix} \Pi(\alpha) & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} \Pi(\alpha) & 0 \\ 0 & -1 \end{pmatrix}.$$

In the first case, the operator α is a rotation through α about an axis; in the second, it is a *mirror rotation*, i.e., a rotation composed with a reflection through the plane that is orthogonal to the axis of rotation.

It is clear that a mirror rotation cannot arise from a continuous motion as it changes the space orientation. Hence, the end result of any real motion of a convex solid with a fixed point, even a very complicated one, is still just like the one after a simple rotation about an appropriate axis through an appropriate angle. This quite nontrivial statement is known as *Euler's Theorem*.

Orthogonal operators on a Euclidean space V form a subgroup of $\mathrm{GL}(V)$ called the *orthogonal group* and denoted $\mathrm{O}(V)$. Accordingly, orthogonal matrices form a subgroup of the group $\mathrm{GL}_n(\mathbb{R})$ denoted O_n (this agrees with the notation introduced in Example 4.18).

As we remarked in Section 5.4, the determinant of an orthogonal matrix equals ± 1. Orthogonal matrices with determinant 1 form a subgroup of index 2 in O. It is denoted SO_n. So, orthogonal operators with determinant 1 form a subgroup of index 2 in $\mathrm{O}(V)$ called the *special orthogonal group* and denoted $\mathrm{SO}(V)$. Geometrically, operators in $\mathrm{SO}(V)$ are interpreted as orientation-preserving orthogonal operators (see Example 4.113).

Example 6.41. The group $\mathrm{O}_2 = \mathrm{O}(E^2)$ consists of rotations (they form the subgroup $\mathrm{SO}_2 = \mathrm{SO}(E^2)$) and reflections through lines.

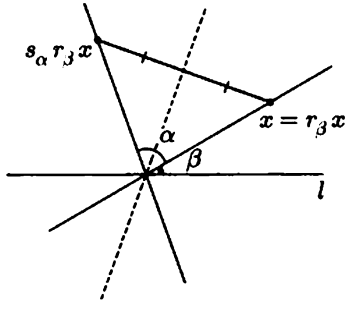


Figure 6.2

Denote by s_α the rotation through α and by r_α, the reflection through a line making the angle α with some fixed line l. Clearly, $s_\alpha s_\beta = s_{\alpha+\beta}$. Next,

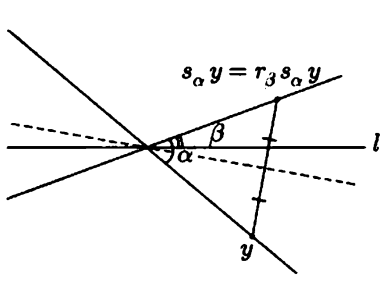


Figure 6.3

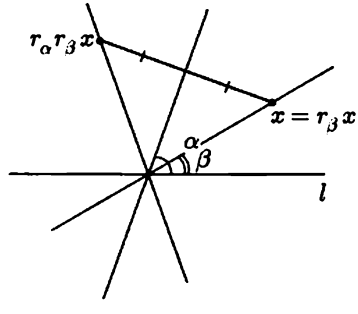


Figure 6.4

the product of a rotation and a reflection changes the orientation, hence it is a reflection. By tracking a single point (see Figures 6.2, 6.3), it is easy to see that

$$s_\alpha r_\beta = r_{\beta+\frac{\alpha}{2}}, \qquad r_\beta s_\alpha = r_{\beta-\frac{\alpha}{2}}.$$

Finally, the product of two reflections preserves the orientation, hence it is a rotation. Again, by tracking a single point (Figure 6.4), it is easy to see that

$$r_\alpha r_\beta = s_{2(\alpha-\beta)},$$

i.e., that the product of two reflections is a rotation through the double angle between their axes.

In particular, it follows that reflections generate the group O_2.

Exercise 6.42. Prove that the group $O(V)$ is generated by reflections through $(n-1)$-dimensional subspaces (here $n = \dim V$).

Every linear operator on a Euclidean space uniquely decomposes into a sum of a symmetric and a skew-symmetric operator (see Example 5.10).

There exists a multiplicative analog of this decomposition where an orthogonal operator replaces the skew-symmetric one (why this happens will become clear in Chapter 12).

Theorem 6.43. *Every nonsingular linear operator on a Euclidean space decomposes uniquely into a product of a positive definite symmetric operator and an orthogonal operator.*

This decomposition of a linear operator is called its *polar decomposition*.

Before proving this theorem, we will prove the following

Proposition 6.44. *Every positive definite symmetric operator $\mathcal{B}$ can be presented uniquely as $\mathcal{B} = \mathcal{C}^2$ for a positive definite symmetric operator $\mathcal{C}$.*

Proof. Let $\lambda_1, \ldots, \lambda_s$ be (distinct) eigenvalues of the operator $\mathcal{B}$ and $V_1, \ldots, V_s$, the respective eigenspaces. By assumption, the λ_i's are positive. Put $\mu_i = \sqrt{\lambda_i}$ (the positive value of the square root). Then the linear operator $\mathcal{C}$ acting on V_i as the multiplication by μ_i satisfies all requirements of this proposition. (In particular, it is symmetric because its matrix is diagonal in an orthonormal basis of eigenvectors of $\mathcal{B}$.)

Conversely, let the operator $\mathcal{C}$ satisfy all the requirements above. Let $\mu_1, \ldots, \mu_s$ be its (distinct) eigenvalues and $W_1, \ldots, W_s$ its eigenspaces. Then the operator $\mathcal{C}^2 = \mathcal{B}$ acts on W_i as the multiplication by μ_i^2. Therefore, with an appropriate reordering, $\mu_i^2 = \lambda_i$ and $W_i = V_i$. This shows that the operator $\mathcal{C}$ is uniquely determined. □

Proof of Theorem 6.43. Let $\mathcal{A}$ be a nonsingular linear operator. Assume that $\mathcal{A} = \mathcal{C}\mathcal{O}$, where $\mathcal{C}$ is a positive definite symmetric operator and $\mathcal{O}$, an orthogonal operator. Then

$$\mathcal{A}\mathcal{A}^* = \mathcal{C}\mathcal{O}\mathcal{O}^*\mathcal{C}^* = \mathcal{C}^2.$$

By Proposition 6.44, this uniquely determines the operator $\mathcal{C}$, hence $\mathcal{O}$.

Conversely, the equality

$$(x, \mathcal{A}\mathcal{A}^* y) = (\mathcal{A}^* x, \mathcal{A}^* y)$$

and the nonsingularity of the operator $\mathcal{A}$ (hence of $\mathcal{A}^*$ as well) implies that $\mathcal{A}\mathcal{A}^*$ is a positive definite symmetric operator. By Proposition 6.44, we can find a positive definite symmetric operator $\mathcal{C}$ such that $\mathcal{A}\mathcal{A}^* = \mathcal{C}^2$. Put $\mathcal{O} = \mathcal{C}^{-1}\mathcal{A}$. Then $\mathcal{A} = \mathcal{C}\mathcal{O}$ and

$$\mathcal{A}\mathcal{A}^* = \mathcal{C}\mathcal{O}\mathcal{O}^*\mathcal{C} = \mathcal{C}^2.$$

After cancelling $\mathcal{C}$ in the above expression we obtain that $\mathcal{O}\mathcal{O}^* = \mathcal{E}$, i.e., that $\mathcal{O}$ is an orthogonal operator. □

Example 6.45. A deformation of a solid with a fixed point can be approximately taken to be a nonsingular linear operator. Let $\mathcal{A} = \mathcal{C}\mathcal{O}$ be the polar decomposition of this operator. Then $\mathcal{O}$ is a rotation about an axis. It is not a true deformation in the sense that it does not create tensions within the solid. On the other hand, by Theorem 6.35 the operator $\mathcal{C}$ is a combination of expansions (or contractions) in three mutually perpendicular directions, i.e., a "pure deformation." It is this operator, called the tensor of deformation, that is used in the statement of Hooke's law.

Exercise 6.46. Prove that every matrix $A \in \mathrm{GL}_n(\mathbb{R})$ can be presented in the form $O_1 D O_2$, where O_1, O_2 are orthogonal matrices and D is a diagonal matrix with positive entries. Is this presentation unique?

A similar theory can be developed for linear operators on a Hermitian space. It is even simpler since in a Hermitian space every linear operator has an eigenvector. We will outline this theory briefly, omitting the proofs which are similar to those given in the Euclidean case.

For an operator $\mathcal{A}$ on a Hermitian space, one defines the *adjoint operator* $\mathcal{A}^*$ by formula (6.20). If an operator $\mathcal{A}$ has the matrix A in some orthonormal basis, then in the same basis the operator $\mathcal{A}^*$ has the matrix A^* (recall that $A^* = \bar{A}^\mathsf{T}$).

A linear operator $\mathcal{A}$ is called *Hermitian* (respectively, *skew-Hermitian*, *unitary*) if $\mathcal{A}^* = \mathcal{A}$ (respectively, $\mathcal{A}^* = -\mathcal{A}$, $\mathcal{A}^* = \mathcal{A}^{-1}$). This is equivalent to its matrix being Hermitian (respectively, skew-Hermitian, unitary) in an orthonormal basis. Hermitian operators are also called *selfadjoint*.

For any of these three types of operators, one can prove the existence of an orthonormal basis of eigenvectors. The eigenvalues of a Hermitian operator are real, of a skew-Hermitian operator, purely imaginary, and the eigenvalues of a unitary operator have absolute value 1.

For instance, let us prove that the eigenvalues of a Hermitian operator $\mathcal{A}$ are real. Let e be an eigenvector of $\mathcal{A}$ with the eigenvalue λ. Then

$$\bar{\lambda}(e,e) = (\mathcal{A}e,e) = (e,\mathcal{A}e) = \lambda(e,e),$$

implying $\bar{\lambda} = \lambda$.

Formula (6.19) defines a bijection between the sets of Hermitian operators and Hermitian sesquilinear functions. In every orthonormal basis the matrices of a Hermitian operator and of the corresponding Hermitian function coincide.

Since for every Hermitian operator there exists an orthonormal basis of eigenvectors, we obtain that for every Hermitian quadratic function q on a Hermitian space, there exists an orthonormal basis in which the matrix of q

is diagonal, i.e.,

$$q(x) = \lambda_1|x_1|^2 + \cdots + \lambda_n|x_n|^2. \tag{6.23}$$

The values of $\lambda_1, \ldots, \lambda_n$ are determined uniquely up to a permutation (as the eigenvalues of the corresponding Hermitian operator). Expression (6.23) is called the *canonical form* of the Hermitian quadratic function q.

A Hermitian operator is called *positive definite* if the corresponding Hermitian quadratic function is positive definite or, equivalently, if all its eigenvalues are positive.

Unitary operators on a Hermitian space V form a subgroup of the group $\mathrm{GL}(V)$ called the *unitary group* and denoted $\mathrm{U}(V)$. Accordingly, unitary matrices form a subgroup of the group $\mathrm{GL}_n(\mathbb{C})$ denoted U_n.

Unitary operators (respectively, matrices) with determinant 1 form a subgroup of $\mathrm{U}(V)$ (respectively, U_n) called the *special unitary group* and denoted $\mathrm{SU}(V)$ (respectively, SU_n).

Every nonsingular linear operator on a Hermitian space decomposes uniquely into a product of a unitary and a positive definite Hermitian operator. Such decomposition of a linear operator is called the *polar decomposition*. In the one-dimensional case, a linear operator is just a complex number and its polar decomposition is the trigonometric form of this number. The trigonometric form of a complex number is related to polar coordinates on the plane; this is why the decomposition is called "polar" in the general case.

The complexification $V(\mathbb{C})$ of a Euclidean space V becomes canonically a Hermitian space if one defines the inner product as

$$(x_1 + iy_1, x_2 + iy_2) = [(x_1, x_2) + (y_1, y_2)] + i\,[(x_1, y_2) - (y_1, x_2)]\,.$$

Here the complex extension $\mathcal{A}_{\mathbb{C}}$ of a symmetric (respectively, skew-symmetric, orthogonal) operator $\mathcal{A}$ becomes a Hermitian (respectively, skew-Hermitian, unitary) operator.

With this in mind, we can prove in yet another way that for a symmetric operator $\mathcal{A}$ on a Euclidean space V, there exists an eigenvector. Namely, let $x + iy$, $x, y \in V$, be an eigenvector of the operator $\mathcal{A}_{\mathbb{C}}$. Since $\mathcal{A}_{\mathbb{C}}$ is Hermitian, the corresponding eigenvalue λ is real, hence

$$\mathcal{A}x = \lambda x, \qquad \mathcal{A}y = \lambda y.$$

At least one of the vectors x, y is nonzero, so it is an eigenvector of $\mathcal{A}$.

6.4. Jordan Canonical Form

It is possible to prove that matrices of special types of linear operators—such as symmetric, Hermitian, and unitary ones that we considered in the

previous section—can be reduced to the diagonal form. In general, there exist obstructions to this reduction described in Theorem 6.30.

The first is that the characteristic polynomial may not split into linear factors, i.e., have less than n roots. This does not happen with linear operators over the field of complex numbers. In the case of an operator over the reals, we can consider its complexification. This eliminates the problem somewhat: a good choice of a basis of complex vectors allows us to understand the action of the original operator on the real space. For instance, we saw in Section 6.2 that to every imaginary eigenvector, there corresponds a two-dimensional invariant subspace of the real space. In Section 9.5 we will show that a similar extension is possible in the general case as well.

The second obstruction is that the dimension of an eigenspace can be less than the multiplicity of the corresponding root of the characteristic polynomial. Then we have to abandon the dream of completely diagonalizing the matrix; however, if the characteristic polynomial splits into linear factors, we can reduce to the so-called Jordan canonical form, which differs little from the diagonal one. This is the theme of this section.

As long as the eigenvectors are not sufficient, we naturally have to consider some more general vectors.

Definition 6.47. A *root vector* of a linear operator $\mathcal{A}$ corresponding to a number $\lambda \in K$ is a vector $e \in V$ such that

$$(\mathcal{A} - \lambda\mathcal{E})^m e = 0$$

for some $m \in \mathbb{Z}_+$. The least such m is called the *height* of the root vector e.

In particular, eigenvectors are root vectors of height 1. It is useful to regard the zero vector as a root vector of height 0 (corresponding to any λ).

Example 6.48. Consider the operator of differentiation on the space $C^\infty(\mathbb{R})$ of infinitely differentiable functions. The eigenvectors corresponding to λ are the functions proportional to $e^{\lambda x}$, and the root vectors are the functions of the form $p(x)e^{\lambda x}$, where $p(x)$ is a polynomial. The height of such a root vector equals $\deg p + 1$. In particular, polynomials are the root vectors corresponding to 0.

If e is a root vector of height $m > 0$, then the vector

$$f = (\mathcal{A} - \lambda\mathcal{E})^{m-1} e$$

is an eigenvector with the eigenvalue λ. Hence, λ is a root of the characteristic polynomial.

It is easy to see that the root vectors corresponding to a root λ form a subspace. It is called the *root subspace* and is denoted $V^\lambda(\mathcal{A})$. Clearly,

$$V^\lambda(\mathcal{A}) \supset V_\lambda(\mathcal{A}).$$

If e is a root vector of height $m > 0$, then $(\mathcal{A} - \lambda\mathcal{E})e$ is a root vector of height $m - 1$. It follows that the root subspace $V^\lambda(\mathcal{A})$ is invariant under $\mathcal{A} - \lambda\mathcal{E}$, hence, under $\mathcal{A}$.

The set of root vectors of height $\leq m$ is nothing but the kernel of the operator $(\mathcal{A} - \lambda\mathcal{E})^m$. Thus, the root subspace $V^\lambda(\mathcal{A})$ is the union of the ascending chain of subspaces

$$\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E}) \subset \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})^2 \subset \cdots .$$

In the finite-dimensional case, this chain stabilizes at some place, hence, $V^\lambda(\mathcal{A}) = \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E})^m$ for some m. In the basis of the space $V^\lambda(\mathcal{A})$ that agrees with this chain, the operator $\mathcal{A} - \lambda\mathcal{E}$ has a niltriangular matrix (i.e., a triangular matrix with zeros on the diagonal), and the operator $\mathcal{A}$, a triangular matrix with λ's on the diagonal. We come to the following two conclusions:

(i) the characteristic polynomial of the restriction of the operator $\mathcal{A}$ to $V^\lambda(\mathcal{A})$ is $(t - \lambda)^k$, where $k = \dim V^\lambda(\mathcal{A})$;

(ii) for $\mu \neq \lambda$, the operator $\mathcal{A} - \mu\mathcal{E}$ is nonsingular on $V^\lambda(\mathcal{A})$.

Exercise 6.49. Prove that the height of a root vector corresponding to the root λ does not exceed $\dim V^\lambda(\mathcal{A})$.

Now we explain why the notion of a root vector is useful. The key is the following

Proposition 6.50. *The dimension of a root subspace equals the multiplicity of the corresponding root of the characteristic polynomial.*

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis of the space V whose first k vectors form a basis of the subspace $V^\lambda(\mathcal{A})$. In this basis, the matrix A of the operator $\mathcal{A}$ has the form (6.6), where B is the matrix of the operator $\mathcal{B} = \mathcal{A}|_{V^\lambda(\mathcal{A})}$. Therefore,

$$f_{\mathcal{A}}(t) = f_{\mathcal{B}}(t) \cdot \det(tE - C) = (t - \lambda)^k \det(tE - C).$$

Let $\mathcal{C}$ be the linear operator on the space $W = \langle e_{k+1}, \ldots, e_n \rangle$ determined by the matrix C. We have to show that λ is not a root of the polynomial $\det(tE - C)$, i.e., not an eigenvalue of the operator $\mathcal{C}$.

Assume the contrary. Then there exists a nonzero vector $e \in W$ such that $\mathcal{C}e = \lambda e$. This means that

$$\mathcal{A}e = \lambda e + u, \qquad u \in V^\lambda(\mathcal{A}),$$

hence $(\mathcal{A} - \lambda\mathcal{E})e = u$ is a root vector. But then so is e, which contradicts the definition of $V^\lambda(\mathcal{A})$. □

Proposition 6.51. *Root subspaces corresponding to different roots* $\lambda_1, \ldots, \lambda_k$ *are linearly independent.*

Proof. The proof is similar to that of Theorem 6.21 (there we were concerned with linear independence of eigenspaces). Let

$$e_1 + \cdots + e_{k-1} + e_k = 0, \qquad e_i \in V^{\lambda_i}(\mathcal{A}).$$

Apply the operator $(\mathcal{A} - \lambda_k \mathcal{E})^m$, where m is the height of e_k, to this equality. We obtain

$$(\mathcal{A} - \lambda_k \mathcal{E})^m e_1 + \cdots + (\mathcal{A} - \lambda_k \mathcal{E})^m e_{k-1} = 0.$$

If we are to use induction on k, the induction hypothesis implies

$$(\mathcal{A} - \lambda_k \mathcal{E})^m e_1 = \cdots = (\mathcal{A} - \lambda_k \mathcal{E})^m e_{k-1} = 0.$$

Since the operator $\mathcal{A} - \lambda_k \mathcal{E}$ is nonsingular on each of the subspaces $V^{\lambda_1}(\mathcal{A})$, $\ldots, V^{\lambda_{k-1}}(\mathcal{A})$, it follows that

$$e_1 = \cdots = e_{k-1} = 0.$$

Thus, also $e_k = 0$. □

Together, Propositions 6.50 and 6.51 imply the following

Theorem 6.52. *If the characteristic polynomial* $f_{\mathcal{A}}(t)$ *splits into linear factors, then*

$$V = \bigoplus_{i=1}^{s} V^{\lambda_i}(\mathcal{A}),$$

where $\lambda_1, \ldots, \lambda_s$ *are (distinct) roots of the polynomial* $f_{\mathcal{A}}(t)$.

Let us study now the action of the operator $\mathcal{A}$ on each root subspace in greater detail.

Definition 6.53. A linear operator $\mathcal{N}$ is *nilpotent* if there exists $m \in \mathbb{Z}_+$ such that $\mathcal{N}^m = 0$. The least such m is called the *height* of the nilpotent operator $\mathcal{N}$.

Example 6.54. The operator of differentiation on the space of polynomials of degree no greater than n is a nilpotent operator of height $n + 1$.

Since $V^{\lambda}(\mathcal{A}) = \operatorname{Ker}(\mathcal{A} - \lambda \mathcal{E})^m$ for some m, the operator

$$\mathcal{N} = (\mathcal{A} - \lambda \mathcal{E})|_{V^{\lambda}(\mathcal{A})}$$

is nilpotent. Thus, our discussion comes down to the study of nilpotent operators.

Let $\mathcal{N}$ be a nilpotent operator on a vector space V.

The *height* of a vector $e \in V$ with respect to $\mathcal{N}$ is the least m such that $\mathcal{N}^m e = 0$, i.e., the height of the vector e viewed as a root vector of

the operator $\mathcal{N}$ (corresponding to the root 0). Obviously, the height of any vector is not greater than the height of $\mathcal{N}$ itself and there exist vectors whose height is exactly the height of $\mathcal{N}$. We denote the height of a vector e by $\operatorname{ht} e$.

Lemma 6.55. *If $e \in V$ is a vector of height m, then the vectors*

$$e,\ \mathcal{N}e,\ \mathcal{N}^2e,\ \ldots,\ \mathcal{N}^{m-1}e$$

are linearly independent.

Proof. Assume that there is a linear dependence

$$\lambda_0 e + \lambda_1 \mathcal{N}e + \lambda_2 \mathcal{N}^2 e + \cdots + \lambda_{m-1}\mathcal{N}^{m-1}e = 0.$$

Let λ_k be the first nonzero coefficient. Then, applying the operator $\mathcal{N}^{m-k-1}$, we obtain

$$\lambda_k \mathcal{N}^{m-1}e = 0,$$

which is not true. □

Definition 6.56. The subspace $\langle e, \mathcal{N}e, \mathcal{N}^2e, \ldots, \mathcal{N}^{m-1}e\rangle$ ($m = \operatorname{ht} e$) is called the *cyclic subspace* of the nilpotent operator $\mathcal{N}$ generated by the vector e.

Clearly, a cyclic subspace is invariant under the action of $\mathcal{N}$. The restriction of the operator $\mathcal{N}$ to the cyclic subspace $\langle e, \mathcal{N}e, \mathcal{N}^2e, \ldots, \mathcal{N}^{m-1}e\rangle$ is of height m. In the basis $\{\mathcal{N}^{m-1}e, \mathcal{N}^{m-2}e, \ldots, \mathcal{N}e, e\}$ it has the matrix

$$J(0) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

called the *nilpotent Jordan block* (of order m) (cf. Example 6.7).

Any vector of a cyclic subspace $\langle e, \mathcal{N}e, \mathcal{N}^2e, \ldots, \mathcal{N}^{m-1}e\rangle$ that does not belong to the subspace $\mathcal{N}U = \langle \mathcal{N}e, \mathcal{N}^2e, \ldots, \mathcal{N}^{m-1}e\rangle$ has height m, hence generates the same cyclic subspace.

Theorem 6.57. *The space V decomposes into a direct sum of cyclic subspaces of the operator $\mathcal{N}$. The number of summands in this decomposition equals* $\dim \operatorname{Ker} \mathcal{N}$.

Proof. We will prove this theorem by induction on $n = \dim V$. For $n = 1$, the statement is obvious. For $n > 1$, let $U \subset V$ be an $(n-1)$-dimensional

subspace containing $\operatorname{Im}\mathcal{N}$. Obviously, U is invariant with respect to $\mathcal{N}$. By induction,

$$U = U_1 \oplus \cdots \oplus U_k,$$

where $U_1, \ldots, U_k$ are cyclic subspaces. Pick a vector $e \in V \setminus U$. We have

$$\mathcal{N}e = u_1 + \cdots + u_k, \qquad u_i \in U_i.$$

If for some i,

$$u_i = \mathcal{N}v_i \in \mathcal{N}U_i, \qquad v_i \in U_i,$$

then we can replace e with $e - v_i$ and make $u_i = 0$. Therefore, we can assume that for every i, either $u_i = 0$ or $u_i \notin \mathcal{N}U_i$.

If $u_i = 0$ for all i, i.e., $\mathcal{N}e = 0$, then

$$V = \langle e \rangle \oplus U_1 \oplus \cdots \oplus U_k$$

is a decomposition of V into a direct sum of cyclic subspaces.

Now let $\mathcal{N}e \neq 0$. Clearly,

$$\operatorname{ht}\mathcal{N}e = \max_i \operatorname{ht} u_i.$$

Without loss of generality, we can assume that

$$\operatorname{ht}\mathcal{N}e = \operatorname{ht} u_1 = m.$$

Then $\operatorname{ht} e = m + 1$. Let us prove that

$$V = \langle e, \mathcal{N}e, \mathcal{N}^2e, \ldots, \mathcal{N}^m e \rangle \oplus U_2 \oplus \cdots \oplus U_k.$$

Since $u_1 \notin \mathcal{N}U_1$, we have $\dim U_1 = \operatorname{ht} u_1 = m$, thus

$$\dim V = \dim U + 1 = (m+1) + \dim U_2 + \cdots + \dim U_k.$$

Hence, it suffices to check that

$$\langle e, \mathcal{N}e, \mathcal{N}^2e, , \ldots, \mathcal{N}^m e \rangle \cap (U_2 \oplus \cdots \oplus U_k) = 0.$$

Assume that

$$\lambda_0 e + \lambda_1 \mathcal{N}e + \lambda_2 \mathcal{N}^2 e + \cdots + \lambda_m \mathcal{N}^m e \in U_2 \oplus \cdots \oplus U_m.$$

Since $e \notin U$, $\lambda_0 = 0$. By projecting the remaining terms onto U_1, we obtain

$$\lambda_1 u_1 + \lambda_2 \mathcal{N}u_1 + \cdots + \lambda_m \mathcal{N}^{m-1} u_1 = 0,$$

implying $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$.

It remains to prove the second claim of the theorem. Let

$$V = V_1 \oplus \cdots \oplus V_k$$

be a decomposition of the space V into a direct sum of cyclic subspaces of the operator $\mathcal{N}$. It is clear that

$$\operatorname{Ker}\mathcal{N} = \operatorname{Ker}\mathcal{N}|_{V_1} \oplus \cdots \oplus \operatorname{Ker}\mathcal{N}|_{V_k}.$$

Since

$$\dim \operatorname{Ker} \mathcal{N}|_{V_i} = 1$$

for all i, we have $\dim \operatorname{Ker} \mathcal{N} = k$. □

We resume our discussion of an arbitrary linear operator $\mathcal{A}$. Observe that restricted to a cyclic subspace of the nilpotent operator $\mathcal{N} = (\mathcal{A} - \lambda\mathcal{E})|_{V^\lambda(\mathcal{A})}$, the operator $\mathcal{A}$ has a matrix of the form

$$J(\lambda) = J(0) + \lambda E = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Such a matrix is called a *Jordan block* with the eigenvalue λ.

Definition 6.58. A *Jordan matrix* is a block-diagonal matrix

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & 0 \\ & & \ddots & \\ 0 & & & J_k \end{pmatrix},$$

where $J_1, J_2, \dots, J_k$ are Jordan blocks.

Combining Theorems 6.52 and 6.57, we obtain the following result.

Theorem 6.59. *If the characteristic polynomial $f_{\mathcal{A}}(t)$ splits into linear factors, then there exists a basis where the matrix of the operator $\mathcal{A}$ is Jordan.*

Such a matrix is called the *Jordan canonical form* of $\mathcal{A}$.

Corollary 6.60. *The matrix of any linear operator over the field of complex numbers reduces to a Jordan canonical form.*

The basis where $\mathcal{A}$ has a Jordan matrix is also called *Jordan*. The proof of Theorem 6.57 implies that in general there is a great freedom in selecting such a basis. However, the Jordan canonical form of a linear operator is determined uniquely up to a permutation of blocks. This will be shown in Section 9.3.

Obviously, in the matrix of $\mathcal{A}$ in the Jordan canonical form of $\mathcal{A}$, the sum of orders of Jordan blocks with the eigenvalue λ equals $\dim V^\lambda(\mathcal{A})$, which is the multiplicity of λ viewed as a root of the characteristic polynomial. The second part of Theorem 6.57 implies that the number of Jordan blocks with the eigenvalue λ equals $\dim V_\lambda(\mathcal{A})$.

Exercise 6.61. Prove that in the Jordan canonical form of $\mathcal{A}$, the maximum order of a Jordan block with the eigenvalue λ equals the height of the nilpotent operator $\mathcal{N} = (\mathcal{A} - \lambda\mathcal{E})|_{V^\lambda(\mathcal{A})}$.

Matrices A and B are called *similar* if there exists a nonsingular matrix C such that $B = C^{-1}AC$. Similar matrices can be regarded as matrices of the same linear operator in different bases. Corollary 6.52 can then be reformulated as saying that every complex matrix is similar to a Jordan one.

Exercise 6.62. Prove that a complex matrix is similar to its transpose.

6.5. Functions of a Linear Operator

Let $\mathcal{A}$ be a linear operator on an n-dimensional vector space V.

For any polynomial

$$f(t) = a_0t^m + a_1t^{m-1} + \cdots + a_{m-1}t + a_m \in K[t],$$

we can define its value at $\mathcal{A}$ as

$$f(\mathcal{A}) = a_0\mathcal{A}^m + a_1\mathcal{A}^{m-1} + \cdots + a_{m-1}\mathcal{A} + a_m\mathcal{E}.$$

It is clear that

$$(f+g)(\mathcal{A}) = f(\mathcal{A}) + g(\mathcal{A}), \qquad (fg)(\mathcal{A}) = f(\mathcal{A})g(\mathcal{A}). \tag{6.24}$$

Similarly, we can define $f(A)$ where A is a matrix. If the matrix of an operator $\mathcal{A}$ in some basis is A, then the matrix of the operator $f(\mathcal{A})$ in this basis is $f(A)$.

Since the space of all linear operators is finite-dimensional (as long as we stick to the agreement that the space V is finite-dimensional), there is only a finite number of linearly independent powers of $\mathcal{A}$. Therefore, there exists a polynomial f such that $f(\mathcal{A}) = 0$. Such a polynomial is called an *annihilating polynomial* of $\mathcal{A}$. An annihilating polynomial of the least degree is called a *minimal (annihilating) polynomial* of $\mathcal{A}$. We denote it $m_\mathcal{A}$.

Every annihilating polynomial f is divisible by $m_\mathcal{A}$. Indeed, if the remainder of division of f by $m_\mathcal{A}$ is nonzero, then it is an annihilating polynomial of degree less than $\deg m_\mathcal{A}$, contradicting the definition of a minimal polynomial. It follows, in particular, that a minimal polynomial is unique up to a constant factor. To define it uniquely, we assume that its leading coefficient is 1.

Exercise 6.63. Determine the minimal polynomials of the zero operator and the identity operator.

Similarly, we define the annihilating polynomials and the minimal polynomial of a matrix. The minimal polynomial of a linear operator equals the minimal polynomial of its matrix in any basis.

If the space V decomposes into a direct sum of invariant subspaces of an operator $\mathcal{A}$, then the minimal polynomial of it equals the least common multiple of the minimal polynomials of its restrictions to these subspaces. Knowing this fact, it is easy to find the minimal polynomial of a linear operator from its Jordan canonical form (if it has one, of course). The first step is to find the minimal polynomial of a Jordan block.

Lemma 6.64. *The minimal polynomial of the Jordan block of order m with the eigenvalue λ equals $(t-\lambda)^m$.*

Proof. Let $\mathcal{A}$ be the linear operator determined by this block. Then $\mathcal{N} = \mathcal{A} - \lambda\mathcal{E}$ is a nilpotent operator of height m, i.e.,

$$(\mathcal{A}-\lambda\mathcal{E})^m = 0, \qquad (\mathcal{A}-\lambda\mathcal{E})^{m-1} \neq 0.$$

It follows that $(t-\lambda)^m$ is an annihilating polynomial and none of its divisors is such. Therefore, $(t-\lambda)^m$ is the minimal polynomial. □

Now let $\mathcal{A}$ be a linear operator whose characteristic polynomial $f_{\mathcal{A}}$ splits into linear factors. Let $\lambda_1, \ldots, \lambda_s$ be all (distinct) roots of the polynomial $f_{\mathcal{A}}$. Lemma 6.64, together with the remark that precedes it, implies the following

Theorem 6.65. *The minimal polynomial of the operator $\mathcal{A}$ is*

$$m_{\mathcal{A}}(t) = \prod_{i=1}^{s} (t-\lambda_i)^{m_i},$$

where m_i is the maximal order of Jordan blocks with the eigenvalue λ_i in the Jordan canonical form of $\mathcal{A}$.

Corollary 6.66. *The Jordan canonical form of $\mathcal{A}$ is diagonal if and only if the minimal polynomial of $\mathcal{A}$ has no multiple roots.*

Example 6.67. Let $\mathcal{A}$ be a linear operator on a complex vector space such that $\mathcal{A}^m = \mathcal{E}$ for some natural m. Then $t^m - 1$ is an annihilating polynomial of $\mathcal{A}$. Since it has no multiple roots, the minimal polynomial of $\mathcal{A}$ has no multiple roots either. Thus, the Jordan canonical form of $\mathcal{A}$ is diagonal. Clearly, its diagonal entries (the eigenvalues of $\mathcal{A}$) are roots of unity of degree m.

Example 6.68. We will find here all linear operators $\mathcal{A}$ such that $\mathcal{A}^3 = \mathcal{A}^2$. This condition means that $t^3 - t^2$ is an annihilating polynomial of $\mathcal{A}$ or that, equivalently, the minimal polynomial of $\mathcal{A}$ divides $t^3 - t^2 = t^2(t-1)$. By Theorem 6.65, this holds if and only if the Jordan canonical form of $\mathcal{A}$ contains only blocks of the following types:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad (0), \qquad (1).$$

The number of blocks of each type can be arbitrary (in particular, zero) as long as the sum of their orders is n.

Corollary 6.69 (Cayley–Hamilton Theorem). $f_{\mathcal{A}}(\mathcal{A}) = 0$.

In particular, we conclude that for a linear operator $\mathcal{A}$ on a two-dimensional vector space,

$$\mathcal{A}^2 - (\operatorname{tr}\mathcal{A})\mathcal{A} + (\det\mathcal{A})\mathcal{E} = 0.$$

Of course, this can be deduced via direct calculation (do it!).

Remark 6.70. The Cayley–Hamilton theorem also holds without the assumption that the characteristic polynomial $f_{\mathcal{A}}$ splits into linear factors. This can be shown as follows. We will prove in Section 9.5 that there exists a field extension L of the field K where $f_{\mathcal{A}}$ splits into linear factors. Regarding the matrix A of the operator $\mathcal{A}$ as a matrix with entries from L, we can claim that it is annihilated by its characteristic polynomial (by the above corollary). But it is obvious that the characteristic polynomial of A does not depend on whether we treat it as a matrix with entries from K or with entries from L. Similarly, one can show that if the minimal polynomial of an operator $\mathcal{A}$ splits into linear factors over K, then its characteristic polynomial splits into linear factors over K as well.

Using Cayley–Hamilton theorem, we can reduce the calculation of the value of a polynomial f at a linear operator $\mathcal{A}$ to the calculation of the value of a polynomial of degree $< n$ at this operator. Namely, divide f by $f_{\mathcal{A}}$ with a remainder:

$$f = qf_{\mathcal{A}} + p, \qquad \deg p < n. \tag{6.25}$$

Then

$$f(\mathcal{A}) = p(\mathcal{A}).$$

Assume that $K = \mathbb{R}$ or $\mathbb{C}$ and that the polynomial $f_{\mathcal{A}}$ splits into linear factors (which is always true over $\mathbb{C}$). Let $\lambda_1, \dots, \lambda_s$ be all its (distinct) roots and $k_1, \dots, k_s$, their multiplicities, so that

$$k_1 + \cdots + k_s = n. \tag{6.26}$$

Then (6.25) implies that

$$f^{(j)}(\lambda_i) = p^{(j)}(\lambda_i) \qquad \text{for } i = 1, \dots, s,\ j = 0, 1, \dots, k_i - 1. \tag{6.27}$$

(We assume here that $f^{(0)} = f$ for every function f.) As the following proposition shows, equalities (6.27) determine the polynomial p uniquely.

Proposition 6.71. *Let $\lambda_1, \dots, \lambda_s \in K$ be distinct numbers and $k_1, \dots, k_s$, natural numbers satisfying condition* (6.26). *Denote by P_n the space of polynomials of degree $< n$. Then the correspondence $\varphi\colon P_n \to K^n$ sending a polynomial $p \in P_n$ to the collection*

$$(p^{(j)}(\lambda_i)\colon i = 1, \dots, s,\ j = 0, 1, \dots, k_i - 1)$$

is an isomorphism of vector spaces.

Proof. Clearly, φ is a linear map. Since $\dim P_n = \dim K^n = n$, it suffices to prove that $\operatorname{Ker}\varphi = 0$. But $\operatorname{Ker}\varphi$ must consist of polynomials for which every λ_i is a root of multiplicity $\geq k_i$, whereas a nonzero polynomial of degree $< n$ cannot have so many roots (counted with multiplicities). □

The problem of constructing a polynomial p of degree $< n$ given numbers $p^{(j)}(\lambda_i)$, $i = 1, \dots, s$; $j = 0, 1, \dots, k_i - 1$, is called the *interpolation problem* (*with multiple nodes*). In the case of simple nodes, i.e., when $k_1 = \cdots = k_s = 1$, the answer is provided by Lagrange's interpolation formula.

Example 6.72. Let us calculate A^m for

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 1 & -1 & -6 \\ -1 & 2 & 5 \end{pmatrix}.$$

We have

$$f_A(t) = \begin{vmatrix} t-1 & 0 & 3 \\ -1 & t+1 & 6 \\ 1 & -2 & t-5 \end{vmatrix} = t^3 - 5t^2 + 8t - 4 = (t-1)(t-2)^2.$$

The interpolation polynomial

$$p(t) = at^2 + bt + c$$

is determined by the following conditions:

$$\begin{aligned} p(1) &= a + b + c = 1, \\ p(2) &= 4a + 2b + c = 2^m, \\ p'(2) &= 4a + b = m \cdot 2^{m-1}, \end{aligned}$$

implying

$$\begin{aligned} a &= (m-2) \cdot 2^{m-1} + 1, \\ b &= -(3m-8) \cdot 2^{m-1} - 4, \\ c &= (2m-6) \cdot 2^{m-1} + 4. \end{aligned}$$

Therefore,

$$A^m = 2^{m-1}[(m-2)A^2-(3m-8)A+(2m-6)E] + A^2 - 4A + 4E$$
$$= 2^{m-1}\begin{pmatrix} 3m-6 & -6m+12 & -9m+12 \\ 3m-4 & -6m+8 & -9m+6 \\ -m & 2m & 3m+2 \end{pmatrix} + \begin{pmatrix} 4 & -6 & -6 \\ 2 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

The above discussion can be generalized from the case of polynomials to that of general analytic functions. In order to do this, we have to study topological properties of the algebra of linear operators.

Let V be a vector space over the field $K = \mathbb{R}$ or $\mathbb{C}$.

Definition 6.73. A *norm* on the space V is a function $\|\cdot\|: V \to \mathbb{R}$ such that

(i) $\|x\| > 0$ for $x \neq 0$;

(ii) $\|\lambda x\| = |\lambda|\|x\|$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$.

Let us give several examples of norms on K^n.

Example 6.74. $\|x\| = \max_i |x_i|$.

Example 6.75. Euclidean (Hermitian) norm $\|x\| = \sqrt{\sum_i |x_i|^2}$.

Example 6.76. $\|x\| = \sum_i |x_i|$.

Definition 6.77. A sequence consisting of vectors x_m *converges in norm* to a vector $x \in V$ if $\lim_{m\to\infty} \|x_m - x\| = 0$.

It is easy to see that convergence in any of the above norms means convergence in every coordinate. This is true for every norm on a finite-dimensional space as the following proposition demonstrates.

Proposition 6.78. *For any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a finite-dimensional space V, there exist positive constants a and b such that*

$$a \leq \frac{\|x\|_2}{\|x\|_1} \leq b \qquad \text{for all } x \in V,\ x \neq 0.$$

Proof. It suffices to compare every norm with a fixed one. Let $\|x\|_1 = \sum_i |x_i|$, where $x_1, \ldots, x_n$ are the coordinates of the vector x in some basis $\{e_1, \ldots, e_n\}$. Then

$$\|x\|_2 \leq \sum_i |x_i|\|e_i\|_2 \leq b\|x\|_1,$$

where $b = \max_i \|e_i\|_2$. The inequalities

$$|\|x + \Delta x\|_2 - \|x\|_2| \leq \|\Delta x\|_2 \leq b\|\Delta x\|_1$$

show that $\|\cdot\|_2$ is a continuous function in the coordinate topology. Let a be its minimum on the unit sphere $\|x\|_1 = 1$ in the sense of the first norm. Then $\|x\|_2 \geq a\|x\|_1$ for all $x \in V$. □

Remark 6.79. Generally speaking, on an infinite-dimensional space, different norms define different topologies. Check this, for instance, for the following norms on the space of continuous functions on the interval $[0, 1]$:

$$\|f\|_1 = \int_0^1 |f(x)|dx, \qquad \|f\|_2 = \max_{0\leq x\leq 1} |f(x)|.$$

Let V be a finite-dimensional vector space with a fixed norm $\|\cdot\|$.

Definition 6.80. A series $\sum_{m=1}^{\infty} x_m$ $(x_m \in V)$ *converges absolutely* if the numerical series $\sum_{m=1}^{\infty} \|x_m\|$ converges.

Just as for numerical series, we can prove the following two propositions:

Proposition 6.81. *Every absolutely converging series* $\sum_{m=1}^{\infty} x_m$ $(x_m \in V)$ *converges. Moreover,*

$$\left\| \sum_{m=1}^{\infty} x_m \right\| \leq \sum_{m=1}^{\infty} \|x_m\|.$$

Proposition 6.82. *The sum of an absolutely converging series does not change with any permutation of its terms.*

We can also define the norm on the space of linear operators on V.

Definition 6.83. The *norm of a linear operator* $\mathcal{A}$ is

$$\|\mathcal{A}\| = \max_{\|x\|=1} \|\mathcal{A}x\| = \max_{x\neq 0} \frac{\|\mathcal{A}x\|}{\|x\|}.$$

Proposition 6.84. *The above-defined function on the space of linear operators is indeed a norm. Moreover, it has the following property:*

$$\|\mathcal{A}\mathcal{B}\| \leq \|\mathcal{A}\| \, \|\mathcal{B}\|.$$

Proof. We have

$$\begin{aligned} \|\mathcal{A}+\mathcal{B}\| &= \max_{\|x\|=1} \|(\mathcal{A}+\mathcal{B})x\| = \max_{\|x\|=1} \|\mathcal{A}x + \mathcal{B}x\| \leq \max_{\|x\|=1} (\|\mathcal{A}x\| + \|\mathcal{B}x\|) \\ &\leq \max_{\|x\|=1} \|\mathcal{A}x\| + \max_{\|x\|=1} \|\mathcal{B}x\| = \|\mathcal{A}\| + \|\mathcal{B}\|. \end{aligned}$$

Other properties of a norm are obvious. Now,

$$\begin{aligned} \|\mathcal{A}\mathcal{B}\| &= \max_{x\neq 0} \frac{\|\mathcal{A}\mathcal{B}x\|}{\|x\|} = \max_{\mathcal{B}x\neq 0} \frac{\|\mathcal{A}\mathcal{B}x\|}{\|\mathcal{B}x\|} \cdot \frac{\|\mathcal{B}x\|}{\|x\|} \\ &\leq \max_{\mathcal{B}x\neq 0} \frac{\|\mathcal{A}\mathcal{B}x\|}{\|\mathcal{B}x\|} \cdot \max_{\mathcal{B}x\neq 0} \frac{\|\mathcal{B}x\|}{\|x\|} \leq \max_{y\neq 0} \frac{\|\mathcal{A}y\|}{\|y\|} \cdot \max_{x\neq 0} \frac{\|\mathcal{B}x\|}{\|x\|} = \|\mathcal{A}\| \, \|\mathcal{B}\|. \end{aligned}$$

□

Exercise 6.85. Determine explicitly the norm of a linear operator for each of the above three norms on the space K^n.

Clearly, a norm of a linear operator is not less than the absolute value of any of its eigenvalues.

Theorem 6.86. *Let a series* $\sum_{m=0}^{\infty} a_m t^m$ $(a_m \in K)$ *converge for* $|t| < R$. *Then the series*

$$f(\mathcal{A}) = \sum_{m=0}^{\infty} a_m \mathcal{A}^m \tag{6.28}$$

converges absolutely for any linear operator $\mathcal{A}$ *such that* $\|\mathcal{A}\| < R$.

Proof. It is known that convergence of the power series $f(t)$ for $|t| < R$ implies its absolute convergence on the same interval (disk). Since

$$\|a_m \mathcal{A}^m\| \leq |a_m| \, \|\mathcal{A}\|^m,$$

the series $f(\mathcal{A})$ converges absolutely for $\|\mathcal{A}\| < R$. □

Expression (6.28) is taken as the definition of how the function f is applied to a linear operator $\mathcal{A}$. Properties (6.24) are preserved here. Similarly, one defines the function of a matrix. As in the case of polynomials, if A is a matrix of an operator $\mathcal{A}$ in some basis, then $f(A)$ is a matrix of the operator $f(\mathcal{A})$ in the same basis.

Assume now that just as above, the characteristic polynomial $f_{\mathcal{A}}$ has roots $\lambda_1, \ldots, \lambda_s$ with multiplicities $k_1, \ldots, k_s$ such that $k_1 + \cdots + k_s = n$. If $\|\mathcal{A}\| < R$, then $|\lambda_i| < R$ for $i = 1, \ldots, s$.

Theorem 6.87. *Under the assumptions of Theorem 6.86, let* p *be a polynomial of degree* $< n$ *satisfying properties* (6.27). *Then* $f(\mathcal{A}) = p(\mathcal{A})$.

Proof. For any m, set

$$f_m(t) = \sum_{k=0}^{m} a_k t^k$$

and denote by p_m the polynomial of degree $< n$ satisfying properties (6.27) with f replaced by the polynomial f_m. We have $f_m(\mathcal{A}) = p_m(\mathcal{A})$. By Proposition 6.71, $\lim_{m\to\infty} p_m = p$. Thus,

$$f(\mathcal{A}) = \lim_{m\to\infty} f_m(\mathcal{A}) = \lim_{m\to\infty} p_m(\mathcal{A}) = p(\mathcal{A}).$$

□

In accordance with the above principle, for any linear operator $\mathcal{A}$ we define its *exponential* $e^{\mathcal{A}}$ $(= \exp \mathcal{A})$ as

$$e^{\mathcal{A}} = \mathcal{E} + \frac{\mathcal{A}}{1!} + \frac{\mathcal{A}^2}{2!} + \frac{\mathcal{A}^3}{3!} + \cdots. \tag{6.29}$$

We can multiply series with the use of Proposition 6.82. Just as in the case of numbers, we obtain the following

Theorem 6.88. $e^{\mathcal{A}+\mathcal{B}} = e^{\mathcal{A}}e^{\mathcal{B}}$ *whenever* $\mathcal{AB} = \mathcal{BA}$.

(When $\mathcal{AB} \neq \mathcal{BA}$, this does not hold, in general. One can say that this is the reason that the theory of Lie groups (see Chapter 12) exists.)

For a given $\mathcal{A}$, put

$$\mathcal{G}(t) = e^{t\mathcal{A}}, \qquad t \in K. \tag{6.30}$$

Obviously, $\mathcal{G}(0) = \mathcal{E}$. By Theorem 6.88,

$$\mathcal{G}(t+s) = \mathcal{G}(t)\mathcal{G}(s), \qquad \mathcal{G}(-t) = \mathcal{G}(t)^{-1}.$$

Therefore, the operators $\mathcal{G}(t)$ form a group. It is called the *one-parameter group* generated by the operator $\mathcal{A}$.

Example 6.89. Let $\mathcal{D}$ be the operator of differentiation on the space of polynomials of degree $< m$. Then

$$(e^{t\mathcal{D}}f)(x) = f(x) + \frac{f'(x)}{1!}t + \frac{f''(x)}{2!}t^2 + \cdots = f(x+t).$$

Example 6.90. $e^{t\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ (check this).

For an operator function of a real or complex variable, one defines the derivative in the usual way. Obviously, to differentiate an operator function, one must differentiate its matrix entries.

Theorem 6.91. $\mathcal{G}'(t) = \mathcal{G}(t)\mathcal{A} = \mathcal{A}\mathcal{G}(t)$.

Proof. Since

$$\mathcal{G}(t+\Delta t) = \mathcal{G}(t)\mathcal{G}(\Delta t) = \mathcal{G}(\Delta t)\mathcal{G}(t),$$

we have

$$\begin{aligned} \mathcal{G}'(t) &= \lim_{\Delta t \to 0} \frac{\mathcal{G}(t+\Delta t) - \mathcal{G}(t)}{\Delta t} \\ &= \mathcal{G}(t) \lim_{\Delta t \to 0} \frac{\mathcal{G}(\Delta t) - \mathcal{E}}{\Delta t} = \left(\lim_{\Delta t \to 0} \frac{\mathcal{G}(\Delta t) - \mathcal{E}}{\Delta t} \right) \mathcal{G}(t), \end{aligned}$$

and the proof, as in the case of the numerical exponential, comes down to calculating the following limit:

$$\lim_{t \to 0} \frac{e^{t\mathcal{A}} - \mathcal{E}}{t} = \mathcal{A}. \tag{6.31}$$

We have

$$\frac{e^{t\mathcal{A}} - \mathcal{E}}{t} = \mathcal{A}\left[\mathcal{E} + t\left(\frac{\mathcal{A}}{2!} + t\frac{\mathcal{A}^2}{3!} + \cdots\right)\right].$$

By Theorem 6.86, the series in parentheses converges absolutely for $|t| < 1$. Moreover, the norm of its sum is not greater than the sum of the numerical series

$$\frac{\|\mathcal{A}\|}{2!} + \frac{\|\mathcal{A}\|^2}{3!} + \frac{\|\mathcal{A}\|^3}{4!} + \cdots .$$

This verifies (6.31). □

Theorem 6.91 suggests the general form of a solution of a homogeneous system of linear ordinary differential equations with constant coefficients,

$$x_i'(t) = \sum_{j=1}^{n} a_{ij} x_j(t), \qquad i = 1, \ldots, n. \tag{6.32}$$

(Here $x_1(t), \ldots, x_n(t)$ are unknown functions of the variable t.) According to the general theory, system (6.32) has a unique solution satisfying the initial conditions

$$x_i(0) = x_{i0}, \qquad i = 1, \ldots, n. \tag{6.33}$$

In the vector form, the system (6.32) is

$$x'(t) = Ax(t), \tag{6.34}$$

where $x(t)$ is the column-vector with coordinates $x_i(t)$ and A is the matrix with entries a_{ij}. The initial condition (6.33) is thus rewritten as

$$x(0) = x_0, \tag{6.35}$$

where x_0 is the column-vector with coordinates x_{i0}. Then its solution is

$$x(t) = e^{tA} x_0. \tag{6.36}$$

The proof is a direct verification using Theorem 6.91.

Example 6.92. We will find here the solution of the following system of differential equations:

$$\begin{cases} x_1'(t) = x_1(t) - 3x_3(t), \\ x_2'(t) = x_1(t) - x_2(t) - 6x_3(t), \\ x_3'(t) = -x_1(t) + 2x_2(t) + 5x_3(t), \end{cases}$$

satisfying the initial conditions

$$x_1(0) = 1, \qquad x_2(0) = 1, \qquad x_3(0) = 0.$$

The matrix A of this system coincides with the matrix in Example 6.72. We have to calculate $f(A)$, where $f(u) = e^{tu}$ (here t acts as a constant). The

interpolation polynomial $p(u) = au^2 + bu + c$ is determined by the following conditions:

$$p(1) = a + b + c = e^t,$$
$$p(2) = 4a + 2b + c = e^{2t},$$
$$p'(2) = 4a + b = te^{2t},$$

implying

$$a = (t-1)e^{2t} + e^t,$$
$$b = -(3t-4)e^{2t} - 4e^t,$$
$$c = (2t-3)e^{2t} + 4e^t.$$

Therefore,

$$e^{tA} = e^{2t}[(t-1)A^2 - (3t-4)A + (2t-3)E] + e^t(A^2 - 4A + 4E)$$
$$= e^{2t}\begin{pmatrix} 3t-3 & -6t+6 & -9t+6 \\ 3t-2 & -6t+4 & -9t+3 \\ -t & 2t & 3t+1 \end{pmatrix} + e^t\begin{pmatrix} 4 & -6 & -6 \\ 2 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

The solution satisfying the given initial conditions is obtained by multiplication of the matrix e^{tA} by the column $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. We thus obtain the solution

$$x_1(t) = (-3t+3)e^{2t} - 2e^t,$$
$$x_2(t) = (-3t+2)e^{2t} - e^t,$$
$$x_3(t) = te^{2t}.$$

Affine and Projective Spaces

7.1. Affine Spaces

In elementary geometry, we deal not just with vectors but also with points (actually, most of the time we deal with points). Just as the axioms of a vector space reflect the basic properties of vectors considered in elementary geometry, the axioms of an affine space reflect the basic properties of points and vectors together.

In the "ordinary" Euclidean space of elementary geometry, one can define the operation of addition of a point and a vector. Namely, the sum of a point p and a vector x is the endpoint of a vector that starts at p and equals x. The properties of this operation lie at the foundation of the following definition.

Let V be a vector space over a field K.

Definition 7.1. The *affine space* associated to a vector space V is a set S with an operation of addition $S \times V \to S$ satisfying the following conditions:

(i) $p + (x + y) = (p + x) + y \qquad (p \in S,\ x, y \in V)$;

(ii) $p + 0 = p \qquad (p \in S,\ 0$ is the zero vector$)$;

(iii) for any $p, q \in S$ there exists a unique vector x such that $p + x = q$.

Elements of the set S are called *points*. The vector x in condition (iii) is called the *vector connecting points p and q* and is denoted $\overrightarrow{pq}$. Condition (i)

implies that

$$\overline{pq} + \overline{qr} = \overline{pr} \qquad \forall p, q, r \in S.$$

Every vector space V can be regarded as an affine one if we view vectors both as vectors and as points and define the operation of addition of a vector to a point as addition of vectors. Here the vector $\overline{pq}$ is the difference of vectors p and q.

Conversely, if we fix a point o (the "origin") in an affine space S, we can identify a point p with its *position vector* $\overline{op}$. Then addition of a vector to a point becomes just the addition of vectors. This identification of points with vectors is called the *vectorization* of an affine space. (Of course, it depends on the choice of an origin.)

The dimension of an affine space is defined as the dimension of the corresponding vector space.

A point o (the origin) together with a basis $\{e_1, \ldots, e_n\}$ of the space V is called a *frame* of the affine space S. Each frame is related to an *affine system of coordinates* in the space S. Namely, a point p gets the coordinates equal to those of the vector $\overline{op}$ in the basis $\{e_1, \ldots, e_n\}$. It is easy to see that

(i) coordinates of the point $p + x$ are equal to the sums of respective coordinates of the point p and the vector x;

(ii) coordinates of the vector $\overline{pq}$ are equal to the differences of respective coordinates of the points q and p.

Generally speaking, linear combinations of points are not defined in the affine space. However, some of them can be given a precise meaning. Namely, define the *barycentric linear combination* of points $p_1, \ldots, p_k \in S$ as a linear combination of the form $\sum_i \lambda_i p_i$, where $\sum_i \lambda_i = 1$, and take it to be equal to the point p such that

$$\overline{op} = \sum_i \lambda_i \overline{op_i},$$

where $o \in S$. By the condition $\sum_i \lambda_i = 1$, this definition does not depend on the choice of the point o. Indeed, let o' be another point. Then

$$\overline{o'p} = \overline{o'o} + \overline{op} = \sum_i \lambda_i (\overline{o'o} + \overline{op_i}) = \sum_i \lambda_i \overline{o'p_i}.$$

In particular, the *center of mass* of a system of points $p_1, \ldots, p_k \in S$ can be defined as

$$\text{cent}\,(p_1, \ldots, p_k) = \frac{1}{k}(p_1 + \cdots + p_k).$$

Exercise 7.2. Prove that in the ordinary Euclidean space

a) the barycentric combination $\lambda p + \mu q$ of two points p and q is the point dividing the interval pq in the ratio $\mu : \lambda$ (this point lies in this interval if $\lambda, \mu \geq 0$, and in its continuation otherwise);

b) the medians of a triangle intersect in the center of mass of the vertices.

Let $p_0, p_1, \ldots, p_n$ be points of an n-dimensional affine space S such that the vectors $\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n}$ are linearly independent. Then every point $p \in S$ can be uniquely presented as

$$p = \sum_{i=0}^{n} x_i p_i, \qquad \text{where} \quad \sum_{i=0}^{n} x_i = 1.$$

Indeed, this equality can be rewritten as

$$\overrightarrow{p_0p} = \sum_{i=1}^{n} x_i \overrightarrow{p_0p_i},$$

implying that we can (and should) take the coordinates of the vector $\overrightarrow{p_0p}$ in the basis $\{\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n}\}$ as $x_1, \ldots, x_n$. Then x_0 is determined as $x_0 = 1 - \sum_{i=1}^{n} x_i$.

The numbers $x_0, x_1, \ldots, x_n$ are called the *barycentric coordinates* of the point p with respect to $p_0, p_1, \ldots, p_n$.

The basic objects of elementary geometry are lines and planes. The following definition introduces these concepts in geometry of affine spaces.

Definition 7.3. A *plane* in an affine space S is a subset of the form

$$P = p_0 + U, \tag{7.1}$$

where p_0 is a point and U is a subspace of the space V.

The subspace U is uniquely determined as the collection of all vectors connecting the points on P and is called the *direction subspace* for P. The sum of a point on P and a vector in U lies in P and, with respect to this operation, the plane P is an affine space associated to the vector space U.

By definition, $\dim P = \dim U$. A zero-dimensional plane is a point. A one-dimensional plane is called a *line*. A plane of dimension $n - 1$ is called a *hyperplane*.

We can choose any point of the plane P as p_0 in (7.1).

If the intersection of two planes $P_1 = p_1 + U_1$ and $P_2 = p_2 + U_2$ is not empty, then this intersection is also a plane. Namely, if $p_0 \in P_1 \cap P_2$, then $P_1 \cap P_2 = (p_0 + U_1) \cap (p_0 + U_2) = p_0 + (U_1 \cap U_2)$.

For any subset $M \subset S$ and any point $p_0 \in M$, the plane

$$p_0 + \langle \overrightarrow{p_0 p} : p \in M \rangle$$

is the smallest plane that contains M (thus it does not depend on p_0). This plane is called the *affine hull* of M and is denoted $\operatorname{aff} M$. It can also be defined as the collection of all barycentric linear combinations of points from M.

Theorem 7.4. *Given any $k+1$ points of an affine space, there is a plane of dimension $\leq k$ passing through these points. Moreover, if these points are not contained in a plane of dimension $< k$, then there exists a unique k-dimensional plane passing through them.*

Proof. Let $p_0, p_1, \ldots, p_k \in S$. Then

$$P = p_0 + \langle \overrightarrow{p_0 p_1}, \ldots, \overrightarrow{p_0 p_k} \rangle$$

is a plane of dimension $\leq k$ through $p_0, p_1, \ldots, p_k$. If $\dim P = k$, the vectors $\overrightarrow{p_0 p_1}, \ldots, \overrightarrow{p_0 p_k}$ are linearly independent and P is the unique k-dimensional plane through $p_0, p_1, \ldots, p_k$. □

Points $p_0, p_1, \ldots, p_k \in S$ are called *affinely dependent* if they lie in a plane of dimension $< k$, and *affinely independent* otherwise. The proof of Theorem 7.4 implies that points $p_0, p_1, \ldots, p_k$ are affinely dependent if and only if the vectors $\overrightarrow{p_0 p_1}, \ldots, \overrightarrow{p_0 p_k}$ are linearly dependent. Also, by definition, the property of affine dependence or independence does not depend on the numbering of the points (in particular, on which one we choose as p_0).

Theorem 7.5. *Points $p_0, p_1, \ldots, p_k$ are affinely independent if and only if the rank of the matrix of their barycentric coordinates equals $k+1$.*

Proof. Let $x_{i0}, x_{i1}, \ldots, x_{in}$ be barycentric coordinates of the point p_i with respect to (affinely independent) points $q_0, q_1, \ldots, q_n$. Then $x_{i1}, \ldots, x_{in}$ are coordinates of the vector $\overrightarrow{q_0 p_i}$ in the basis $\{\overrightarrow{q_0 q_1}, \ldots, \overrightarrow{q_0 q_n}\}$.

The rank of the matrix

$$\begin{pmatrix} x_{00} & x_{01} & \ldots & x_{0n} \\ x_{10} & x_{11} & \ldots & x_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{k0} & x_{k1} & \ldots & x_{kn} \end{pmatrix} \tag{7.2}$$

does not change if we add to the first column the sum of all other columns. By doing so, we obtain the matrix

$$\begin{pmatrix} 1 & x_{01} & \ldots & x_{0n} \\ 0 & x_{11} - x_{01} & \ldots & x_{1n} - x_{0n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & x_{k1} - x_{01} & \ldots & x_{kn} - x_{0n} \end{pmatrix}.$$

Its rank is greater by 1 than the rank of the following matrix:

$$\begin{pmatrix} x_{11} - x_{01} & \dots & x_{1n} - x_{0n} \\ \dots\dots\dots\dots\dots\dots & & \\ x_{k1} - x_{01} & \dots & x_{kn} - x_{0n} \end{pmatrix}. \tag{7.3}$$

The entries of this matrix are coordinates of the vectors $\overrightarrow{p_0p_1}, \dots, \overrightarrow{p_0p_k}$ in the basis $\{\overrightarrow{q_0q_1}, \dots, \overrightarrow{q_0q_n}\}$.

Thus, the rank of the matrix (7.2) equals $k+1$ if and only if the rank of the matrix (7.3) equals k. The latter means precisely that the vectors $\overrightarrow{p_0p_1}, \dots, \overrightarrow{p_0p_k}$ are linearly independent, i.e., that the points $p_0, p_1, \dots, p_k$ are affinely independent. □

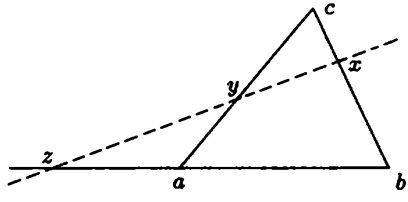


Figure 7.1

Example 7.6. Let points x, y, z lie on the sides bc, ca, ab of the triangle abc (Figure 7.1) or their continuations. Suppose they divide these sides in the ratio $\lambda : 1$, $\mu : 1$, $\nu : 1$, respectively. We will find under which condition on λ, μ, ν, the points x, y, z lie on the same line, i.e., are affinely dependent. By Exercise 7.2, the matrix of barycentric coordinates of x, y, z with respect to a, b, c is

$$\begin{pmatrix} 0 & \frac{1}{\lambda+1} & \frac{\lambda}{\lambda+1} \\ \frac{\mu}{\mu+1} & 0 & \frac{1}{\mu+1} \\ \frac{1}{\nu+1} & \frac{\nu}{\nu+1} & 0 \end{pmatrix}.$$

By Theorem 7.5, the points x, y, z lie on the same line if and only if the determinant of this matrix is zero, i.e., if

$$\lambda\mu\nu = -1.$$

This statement is known as *Menelaus's Theorem.*

Exercise 7.7. Using barycentric coordinates, prove *Ceva's Theorem*: in the notation of Example 7.6, the lines ax, by, cz intersect at one point if and only if

$$\lambda\mu\nu = 1$$

(Figure 7.2).

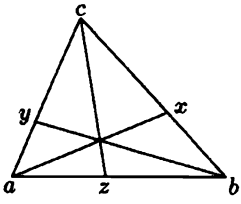


Figure 7.2

Theorem 7.8. *A nonempty subset $P \subset S$ is a plane if and only if for any two points $a, b \in P$, the line through a and b also lies on P.*

Proof. The "only if" part is obvious. Now let $P \subset S$ be a nonempty subset with the above property. Let $\{p_0, p_1, \ldots, p_k\}$ be the maximal affinely independent system of points of P. Then $P \subset \operatorname{aff}\{p_0, p_1, \ldots, p_k\}$. We will prove that $P = \operatorname{aff}\{p_0, p_1, \ldots, p_k\}$.

Let $p = \sum_i \lambda_i p_i$ be a barycentric linear combination of points $p_0, p_1, \ldots, p_k$. Let us prove that $p \in P$ by induction on the number l of nonzero coefficients $\lambda_0, \lambda_1, \ldots, \lambda_k$. For $l = 1$, p coincides with one of the points $p_0, p_1, \ldots, p_k$, so there is nothing to prove. Let $l > 1$. Without loss of generality, we may assume that $\lambda_k \neq 0$. Then

$$p = (1 - \lambda_k)\left(\sum_{i=0}^{k-1} \frac{\lambda_i}{1 - \lambda_k} p_i\right) + \lambda_k p_k,$$

i.e., p lies on the line through the points

$$p' = \sum_{i=0}^{k-1} \frac{\lambda_i}{1 - \lambda_k} p_i$$

and p_k. By induction, $p' \in P$. Therefore, $p \in P$ as well. □

Another point of view is to regard planes as sets of solutions of systems of linear equations.

Consider the following system of linear equations:

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \qquad i = 1, \ldots, m. \tag{7.4}$$

We interpret $x_1, \ldots, x_n$ as coordinates of points in an n-dimensional affine space S with respect to a frame $(o; e_1, \ldots, e_n)$. Then the solutions of system (7.4) can be viewed as points of the space S. Assume that this system is compatible and $p_0 \in S$ is one of its solutions. It is easy to see that a point

$p \in S$ is a solution of system (7.4) if and only if the coordinates of the vector $\overline{p_0p}$ satisfy the following system of homogeneous equations:

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \qquad i = 1, \ldots, m. \tag{7.5}$$

We know (see Theorem 2.63) that the solutions of system (7.5) form a subspace $U \subset V$ of dimension $n-r$, where r is the rank of the coefficient matrix (it is common for systems (7.4) and (7.5)). Therefore, the set of solutions of system (7.4) is the plane $P = p_0 + U$ of the same dimension. We have thus proved the following

Theorem 7.9. *The set of solutions of a compatible system of linear equations is a plane of dimension $n-r$, where n is the number of variables and r, the rank of its coefficient matrix.*

Conversely, let $P = p_0 + U$ be a plane. By Theorem 5.32, the subspace U is determined by a system of homogeneous linear equations. By replacing the free terms of these equations with the values that their left-hand sides assume at the point p_0, we obtain a system of linear equations that determines the plane P. Thus, we have proved the following

Theorem 7.10. *Every plane is the set of solutions of a system of linear equations.*

We will discuss now the relative position of two planes

$$P_1 = p_1 + U_1, \qquad P_2 = p_2 + U_2.$$

Theorem 7.11. *Planes P_1 and P_2 intersect if and only if*

$$\overline{p_1p_2} \in U_1 + U_2.$$

Proof. Planes P_1 and P_2 intersect if and only if there exist vectors $u_1 \in U_1$ and $u_2 \in U_2$ such that

$$p_1 + u_1 = p_2 + u_2.$$

This equality can be rewritten as

$$\overline{p_1p_2} = u_1 - u_2.$$

The existence of such vectors u_1, u_2 means that $\overline{p_1p_2} \in U_1 + U_2$. □

Planes P_1 and P_2 are called *parallel* if either $U_1 \subset U_2$ or $U_2 \subset U_1$ and *skew* if $P_1 \cap P_2 = \varnothing$ and $U_1 \cap U_2 = 0$.

Exercise 7.12. What is the least dimension of a space with two skew two-dimensional planes?

Exercise 7.13. Determine $\dim \operatorname{aff}(P_1 \cup P_2)$.

Consider now the class of functions on an affine space corresponding to the class of linear functions on a vector space.

Definition 7.14. An *affine-linear* function on an affine space S is a function $f: S \to K$ such that

$$f(p+x) = f(p) + \alpha(x), \qquad p \in S,\ x \in V, \tag{7.6}$$

where α is a linear function on the vector space V.

The function α is called the *differential* of f and is denoted df.

Let $o \in S$ be a fixed origin. By setting $p = o$ in (7.6), we express an affine-linear function in vectorized form as follows:

$$f(x) = \alpha(x) + b, \qquad b \in K, \tag{7.7}$$

where $b = f(o)$. This implies the following coordinate form of f:

$$f(x) = \sum_i a_i x_i + b. \tag{7.8}$$

Conversely, for any linear function $\alpha \in V^*$ and any number $b \in K$, the function f determined by (7.7) is an affine-linear function with the differential α. Indeed, let $p = o + y$; then by vectorization,

$$\begin{aligned} f(p+x) &= f(y+x) = \alpha(y+x) + b = \alpha(y) + \alpha(x) + b \\ &= f(y) + \alpha(x) = f(p) + \alpha(x). \end{aligned}$$

A particular case of affine-linear functions are constant functions. Their defining characteristic is the zero differential. If f is a nonconstant affine-linear function, then its level variety $f(p) = c$ is a hyperplane with the direction subspace determined by the equation $df(x) = 0$.

Affine-linear functions form an $(n+1)$-dimensional subspace (here $n = \dim S$) in the space of all functions on S. This follows, for instance, from the coordinate form (7.8).

We will prove now two statements about affine-linear functions that will be used in the next section.

Proposition 7.15. *Barycentric coordinates are affine-linear functions.*

Proof. Let $x_0, x_1, \ldots, x_n$ be the barycentric coordinates with respect to points $p_0, p_1, \ldots, p_n$. If we take p_0 as the origin and then vectorize S, $x_1, \ldots, x_n$ become the usual coordinates with respect to the basis $(\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n})$. Therefore, $x_1, \ldots, x_n$ are affine-linear functions. Since $x_0 = 1 - \sum_{i=1}^n x_i$, we obtain that x_0 is also an affine-linear function (this can also be shown by taking another point p_i as the origin). □

Proposition 7.16. *Let f be an affine-linear function. Then*

$$f\left(\sum_i \lambda_i p_i\right) = \sum_i \lambda_i f(p_i)$$

for any barycentric linear combination $\sum_i \lambda_i p_i$ of points $p_1, \ldots, p_k$.

Proof. Vectorize the space S. Then f can be written in the form (7.7), so that we obtain

$$f\left(\sum_i \lambda_i p_i\right) = \alpha\left(\sum_i \lambda_i p_i\right) + b = \sum_i \lambda_i(\alpha(p_i) + b) = \sum_i \lambda_i f(p_i).$$

□

By combining the axioms of a Euclidean vector space and the axioms of an affine space we can finally introduce the concept that encompasses all elementary geometry.

Definition 7.17. An affine space associated with a Euclidean vector space is called a *Euclidean affine space* (or simply a *Euclidean space* if the context is clear).

The *distance* ρ between two points in a Euclidean space is defined as

$$\rho(p, q) = |\overline{pq}|.$$

This notion satisfies the axioms of a metric space. In particular, the triangle inequality follows from inequality (5.26) for the length of a sum of vectors.

Exercise 7.18. Prove that the distance between two planes $P_1 = p_1 + U_1$ and $P_2 = p_2 + U_2$ in the Euclidean space is determined by the following formula:

$$\rho(P_1, P_2) = |\operatorname{ort}_{U_1+U_2} \overline{p_1 p_2}|.$$

Among all affine coordinate systems of a Euclidean space, those corresponding to orthonormal bases stand out. They are called *Cartesian coordinate systems*.

7.2. Convex Sets

Let S be an affine space over the field of real numbers and V, the associated vector space.

Definition 7.19. The (closed) *interval* connecting points $p, q \in S$ is the set

$$pq = \{\lambda p + (1 - \lambda)q : 0 \leq \lambda \leq 1\} \subset S.$$

Definition 7.20. A set $M \subset S$ is *convex* if for any two points $p, q \in S$, it contains the whole interval pq.

Obviously, the intersection of convex sets is convex. A plane is a convex set.

Definition 7.21. A *convex linear combination* of points in S is their barycentric linear combination with nonnegative coefficients.

Proposition 7.22. *For any points $p_0, p_1, \ldots, p_k$ in a convex set $M \subset S$, the set M also contains every convex linear combination $p = \sum_i \lambda_i p_i$.*

Proof. Use induction on the number of nonzero coefficients $\lambda_0, \lambda_1, \ldots, \lambda_k$ just as in the proof of Theorem 7.8 but with intervals instead of lines. □

Proposition 7.23. *For any set $M \subset S$, the set* conv M *of all convex linear combinations of points in M is convex.*

Proof. Let $p = \sum_i \lambda_i p_i$ and $q = \sum_i \mu_i q_i$ be convex linear combinations of points in M. Then for $0 \le \lambda \le 1$,

$$\lambda p + (1-\lambda) q = \sum_i \lambda \lambda_i p_i + \sum_i (1-\lambda) \mu_i q_i$$

is also a convex linear combination of points in M. □

The set conv M is the smallest convex set containing M. It is called the *convex hull* of M.

The convex hull of a system of affinely independent points $p_0, p_1, \ldots, p_n$ in an n-dimensional affine space is called the *n-dimensional simplex* with vertices $p_0, p_1, \ldots, p_n$. In other words, a simplex consists of points whose barycentric coordinates with respect to some points $p_0, p_1, \ldots, p_n$ are nonnegative. A zero-dimensional simplex is a point; a one-dimensional simplex, an interval; a two-dimensional simplex, a triangle; a three-dimensional simplex, a tetrahedron.

A point p in a set $M \subset S$ is called an *interior point* if there exists a neighborhood of p which is completely contained in M, and a *boundary point* otherwise. Obviously, all points of a simplex whose barycentric coordinates with respect to the vertices are all positive are interior (and vice versa).

Proposition 7.24. *A convex set M has interior points if and only if* aff $M = S$.

Proof. If aff $M = S$, then M contains a system of $n+1$ affinely independent points. But then M contains a simplex with vertices at these points, hence contains interior points. The converse is obvious. □

Exercise 7.25. Prove that the closure of a convex set is convex and that, moreover, every interior point of the closure is an interior point of the set itself.

A convex set that has interior points is called a *convex body*.

Proposition 7.26. *Let p be an interior point of a convex body M and q, any point of M. Then all points of the interval $\overline{pq}$ are interior points of M except, maybe, the point q.*

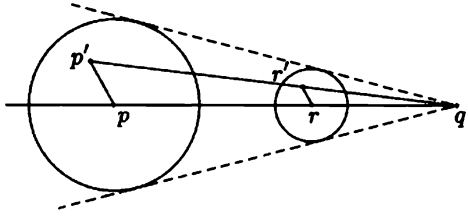


Figure 7.3

Proof. Consider the point

$$r = \lambda p + (1 - \lambda)q \quad (0 < \lambda \leq 1).$$

We have

$$p = \frac{1}{\lambda} r + \frac{\lambda - 1}{\lambda} q.$$

If a point r' is sufficiently close to r, then the point

$$p' = \frac{1}{\lambda} r' + \frac{\lambda - 1}{\lambda} q$$

is close to p, hence lies in M (Figure 7.3). Since

$$r' = \lambda p' + (1 - \lambda)q,$$

it follows that $r' \in M$. □

Corollary 7.27. *Interior points of a convex body form a convex set.*

Corollary 7.28. *Every point of a convex body is a limit of its interior points.*

Denote the set of interior points of a convex body M by M°. This is an open convex body.

By Proposition 7.24, any convex set $M \subset S$ is a convex body in aff M. Abusing the language, points of a convex set M that are interior within M regarded as a body in aff M are often called interior points of M.

For any nonconstant affine-linear function f on the set S (see Section 7.1), let

$$H_f = \{p \in S : f(p) = 0\},$$
$$H_f^+ = \{p \in S : f(p) \geq 0\},$$
$$H_f^- = \{p \in S : f(p) \leq 0\} \ (= H_{-f}^+).$$

The set H_f is a hyperplane. The sets H_f^+ and H_f^- are called (*closed*) *half-spaces* bounded by the hyperplane H_f. Proposition 7.16 implies that every half-space is a convex set. On the other hand, an interval connecting a point in H_f^+ to a point in H_f^- crosses the hyperplane H_f.

Definition 7.29. A hyperplane H_f is a *supporting hyperplane* of a closed convex body M if $M \subset H_f^+$ and H_f contains at least one (boundary) point of M. The half-space H_f^+ is then called the *supporting half-space* of M.

Proposition 7.30. *A hyperplane H that passes through a boundary point of a closed convex body M, is supporting if and only if $H \cap M^\circ = \varnothing$.*

Proof. If $H \cap M^\circ \neq \varnothing$, the points of M° (hence, also of M) lie on both sides of H. Conversely, let points of M lie on both sides of H. Then, as every point of M is a limit point for the set M°, there are points of M° on both sides of H too. The interval connecting two such points lies in M° and intersects H. Thus, $H \cap M^\circ \neq \varnothing$. □

The key theorem of the theory of convex sets is the following *separation theorem*:

Theorem 7.31. *For every boundary point of a closed convex body, there exists a supporting hyperplane passing through this point.*

Proof. Let p be a boundary point of a closed convex body M in an n-dimensional affine space. By induction on k, we prove that for $k \leq n-1$, there exists a k-dimensional plane through p that does not intersect M°. For $k = 0$, the point p is such a plane. Assume that we found a $(k-1)$-dimensional plane P satisfying the induction hypothesis. Pick any $(k+1)$-dimensional plane S' containing P and an interior point p_0 of M. Let us find our k-dimensional plane among the planes in S' containing P.

Consider a convex body $M' = M \cap S'$ in the space S'. Clearly, $M^\circ \cap S' \subset (M')^\circ$. Conversely, every point $r \in (M')^\circ$ is an interior point of the interval connecting p_0 to a point $q \in M' \subset M$ (Figure 7.4), hence it belongs to M°. Therefore,

$$(M')^\circ = M^\circ \cap S'.$$

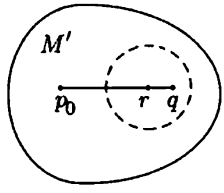


Figure 7.4

In particular, it follows that $P \cap (M')^\circ = \varnothing$ and it suffices to prove that S' contains a supporting hyperplane of the body M' that contains P. We change the notation and denote $S' = S$, $M' = M$, and $k + 1 = n$.

So, let P be an $(n-2)$-dimensional plane through the point p that does intersect M°. Let us prove that there exists a supporting hyperplane of M such that it contains P.

If a hyperplane H contains P, then P divides H into two half-planes (or, rather, two half-hyperplanes), say, H^+ and H^- (this notation is not to be confused with our notation for half-spaces). If none of the half-planes H^+ and H^- intersects M°, we are done. If both of them intersect M°, then so does P; hence, this does not happen.

Now assume that H^+ intersects M° while H^- does not. Let us start rotating the hyperplane H about P, clockwise. Clearly, after rotating just a little, the half-plane H^+ will still intersect M°. After rotating by π it will turn into the half-plane H^- which does not intersect M°. So, there exists a minimal angle of rotation at which H^+ no longer intersects M°. Denote the result of rotation of H through this angle by H_0.

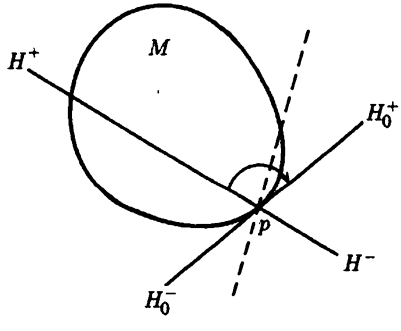


Figure 7.5

By construction, the hyperplane H_0^+ does not intersect the set M° but after even the slightest counterclockwise rotation, it does (Figure 7.5). On

the other hand, if the half-plane H_0^- intersected M°, it would still intersect M° after a small rotation. But, as we have already remarked, both halves of a hyperplane containing P cannot intersect M°.

Therefore, H_0^- does not intersect M°, hence H_0 is a supporting hyperplane of M. □

Remark 7.32. Actually we proved a stronger statement, namely that every plane that passes through p and does not intersect M° is contained in a supporting hyperplane.

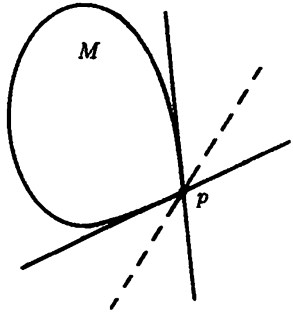


Figure 7.6

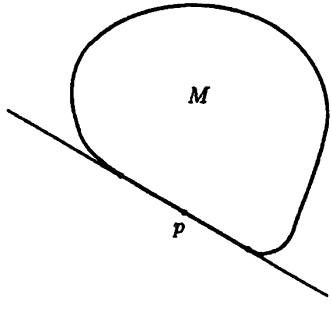


Figure 7.7

Remark 7.33. A boundary point p of a body M can belong to either a unique supporting hyperplane (as in Figure 7.5) or infinitely many such hyperplanes (as in Figure 7.6). A supporting hyperplane can also contain other points of M (as in Figure 7.7).

Theorem 7.34. *Every closed convex set M is an intersection of (perhaps infinitely many) half-spaces.*

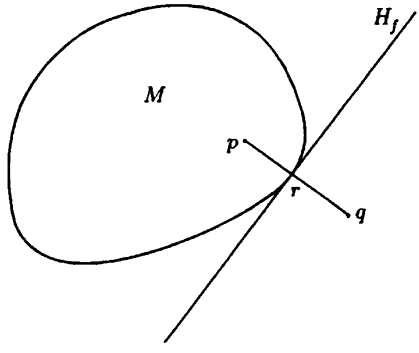


Figure 7.8

Proof. Observe that every hyperplane H_f is the intersection of the half-spaces H_f^+ and H_f^-. This implies that a plane of any dimension is also an intersection of half-spaces. Thus, it suffices to prove the theorem for a body M.

We will show that a closed convex body M is the intersection of its supporting half-spaces. Let $q \notin M$ and p be an interior point of M. The interval $\overline{pq}$ intersects the boundary of the body M at a point $r \neq q$. Denote a supporting hyperplane through r by H_f (Figure 7.8). Since $f(p) > 0$ and $f(r) = 0$, $f(q) < 0$, i.e., $q \notin H_f^+$. □

Definition 7.35. A *polyhedron* is the intersection of a finite number of half-spaces. A convex polyhedron which is also a body is called a *convex solid*.

In other words, a convex polyhedron is the set of solutions of a finite system of linear inequalities. Notice that a convex polyhedron does not have to be bounded. For instance, the space S itself is a convex polyhedron (as the intersection of the empty set of half-spaces). A convex polyhedron does not have to be a solid (though other texts sometimes require it).

Obviously, the intersection of a finite number of convex polyhedra is a convex polyhedron. Every plane is a convex polyhedron.

Example 7.36. A simplex with vertices $p_0, p_1, \ldots, p_n$ is a convex polyhedron since it is determined by linear inequalities $x_i \geq 0$, $i = 0, 1, \ldots, n$, where $x_0, x_1, \ldots, x_n$ are barycentric coordinates with respect to $p_0, p_1, \ldots, p_n$.

Example 7.37. A convex polyhedron determined by linear inequalities $0 \leq x_i \leq 1$, $i = 1, \ldots, n$, where $x_1, \ldots, x_n$ are affine coordinates with respect to some frame, is called an *n-dimensional parallelepiped*.

Definition 7.38. A point p of a convex set M is *extreme* if it is not an interior point of any interval in M.

Theorem 7.39. *A bounded closed convex set M is the convex hull of the set $E(M)$ of its extreme points.*

Proof. Let $\widetilde{M} = \operatorname{conv} E(M)$. Clearly, $\widetilde{M} \subset M$. We will prove by induction on $\dim S$ that $M \subset \widetilde{M}$. For $\dim S = 0$, there is nothing to prove. Let $\dim S > 0$ and $p \in M$. Let us prove that $p \in \widetilde{M}$. Assume that M is a body; otherwise we can use the induction hypothesis directly. Consider the following two cases.

Case 1. Let p be a boundary point. There is a supporting hyperplane H passing through p. The set $M \cap H$ is bounded, closed, and convex and all its extreme points are extreme points of M. By induction, $M \cap H$ is the convex hull of its extreme points. Thus, $p \in \widetilde{M}$.

Case 2. Let p be an interior point. Draw a line through p. Since the set M is bounded, this line intersects it by an interval qr that contains p. Points q and r are boundary points of M and, by the above, they belong to $\widetilde{M}$. Therefore, $p \in \widetilde{M}$. □

Theorem 7.40 (Minkowski–Weyl Theorem). *The following properties of a bounded set $M \subset S$ are equivalent:*

(i) M is a convex polyhedron;

(ii) M is a convex hull of a finite number of points.

Proof. (i) Let

$$M = \bigcap_{i=1}^{m} H_{f_i}^{+} \tag{7.9}$$

be a convex polyhedron. Let us prove that each extreme point of M is the only point in the intersection of some of the hyperplanes $H_{f_1}, \ldots, H_{f_m}$. This will imply that M has only a finite number of extreme points. By Theorem 7.39, M is their convex hull.

Let $p \in M$ be an extreme point. Define

$$J = \{j : f_j(p) = 0\} \subset \{1, \ldots, m\},$$
$$P = \{x \in S : f_j(x) = 0 \text{ for } j \in J\}.$$

Since $f_i(p) > 0$ for $i \notin J$, we see that p is an interior point of the convex polyhedron $M \cap P$ in the space P. But p is an extreme point of M, hence it is an extreme point of $M \cap P$. Thus, $\dim P = 0$, i.e., $P = \{p\}$.

(ii) Let $M = \operatorname{conv}\{p_1, \ldots, p_k\}$. We can assume that $\operatorname{aff} M = S$. Consider the following convex polyhedron in the space of affine-linear functions on S:

$$M^* = \left\{ f : f(p_i) \geq 0 \text{ for } i = 1, \ldots, k, \sum_{i=1}^{k} f(p_i) = 1 \right\}.$$

An affine-linear function on S is uniquely determined by its values at points $p_1, \ldots, p_k$. Since for a function from M^*, these values lie on the interval $[0, 1]$, M^* is a bounded polyhedron. By the above, it is the convex hull of a finite number of points, say, $f_1, \ldots, f_m$.

By Theorem 7.34, the set M is determined by linear inequalities (clearly M is closed). Thus,

$$M = \{p \in S : f(p) \geq 0 \ \forall f \in M^*\} = \{p \in S : f_i(p) \geq 0 \text{ for } i = 1, \ldots, m\}.$$

Therefore, M is a convex polyhedron. □

Definition 7.41. A *face* of a convex polyhedron M is a nonempty intersection of M with some of its supporting hyperplanes. (The polyhedron itself is also regarded as its own face being the intersection of itself with the empty set of supporting hyperplanes.)

A zero-dimensional face is called a *vertex*; a one-dimensional, an *edge*; an $(n-1)$-dimensional, a *hyperface* (here, $n = \dim \operatorname{aff} M$). Consider a polyhedron M determined by formula (7.9). The following theorem demonstrates that in order to find its faces, it suffices to consider only the hyperplanes $H_{f_1}, \ldots, H_{f_m}$.

Theorem 7.42. *Every face Γ of the polyhedron M is of the form*

$$\Gamma = M \cap \left(\bigcap_{j \in J} H_{f_j} \right), \tag{7.10}$$

where $J \subset \{1, \ldots, m\}$.

Proof. Let Γ' be a face of M. Let

$$J = \{j : \Gamma' \subset H_{f_j}\} \subset \{1, \ldots, m\}.$$

For any $i \notin J$, there exists a point $p_i \in \Gamma'$ such that $f_i(p_i) > 0$. Let p be the center of mass of the system of all these points. Then $f_i(p) > 0$ for every $i \notin J$.

Now, define Γ by formula (7.10). Let us prove that $\Gamma' = \Gamma$. Clearly, $\Gamma' \subset \Gamma$. It is also clear that p is an interior point of the face Γ. Hence, every supporting hyperplane passing through p contains Γ. Thus, $\Gamma' = \Gamma$. □

Therefore, if a convex polyhedron is determined by a system of linear inequalities, we can obtain its faces by replacing some of these inequalities with equalities (but in such a way that we do not get an empty set). One should keep in mind, though, that on some faces defined in such a way, some of the remaining inequalities can automatically turn into equalities too.

Example 7.43. The faces of the n-dimensional parallelepiped determined by inequalities $0 \le x_i \le 1$, $i = 1, \dots, n$, are obtained by setting some of the coordinates to 0 or 1. In particular, vertices are points such that all their coordinates are either 0 or 1.

Exercise 7.44. Describe the faces of the intersection of the n-dimensional parallelepiped $0 \le x_i \le 1$, $i = 1, \dots, n$, with the hyperplane $x_1 + \dots + x_n = \frac{n}{2}$.

Exercise 7.45. Determine the faces of the n-dimensional simplex.

Exercise 7.46. Prove that every face of the polyhedron $\operatorname{conv}\{p_1, \dots, p_k\}$ is a convex hull of a subcollection of the points $p_1, \dots, p_k$. In particular, any vertex of $\operatorname{conv}\{p_1, \dots, p_k\}$ is one of the points $p_1, \dots, p_k$ (but not all of them need to be vertices).

Remark 7.47. The study of the combinatorial structure of convex polyhedra is an interesting and important area of mathematics. Here are two results from this area:

(i) The sequence $(a_0, a_1, \dots, a_{n-1})$, where a_k is the number of k-dimensional faces of an n-dimensional bounded convex polyhedron, is called the *f-vector* of this polyhedron. What are the necessary and sufficient conditions for a sequence of n natural numbers to be the f-vector of some n-dimensional polyhedron? For $n = 3$, the conditions are as follows: $a_0 - a_1 + a_2 = 2$, $4 \le a_0,\ a_2 \le \frac{2a_1}{3}$ (Steinitz's Theorem). For a general n the answer is still unknown.

(ii) Points p and q in a convex polyhedron are called *neighboring* if the interval pq is an edge of this polyhedron. It is easy to see that tetrahedron is the only 3-dimensional convex polyhedron such that each of its two vertices are neighboring. The story is completely different in the 4-dimensional space. D. Gale showed that this space contains polyhedra with an arbitrary number of vertices such that each two of them are neighboring. For instance, let M be the convex hull of points

$$p_i = (t_i, t_i^2, t_i^3, t_i^4), \qquad i = 1, \dots, N,$$

where $t_1, \dots, t_N$ are distinct real numbers. Then

(i) every point p_i is a vertex of M (and these are all its vertices; see Exercise 7.46);

(ii) every interval p_ip_j $(i \neq j)$ is an edge of M.

Prove these two statements.

Proposition 7.48. *The extreme points of a convex polyhedron M are exactly its vertices.*

Proof. If a point p is an interior point of an interval in M, then a supporting hyperplane passing through p contains this interval. Hence, p is not a vertex. Conversely, if p is not a vertex, then it is an interior point of a face of positive dimension, hence is not extreme. □

Beside mathematics, the most important examples of the use of convex polyhedra are found in *linear programming*. The basic problem of linear programming is as follows: find the maximum (minimum) of a given affine-linear function on a given convex polyhedron. Clearly, to find the minimum of a function f is the same as to find the maximum of the function $-f$, so one can consider just one of these problems.

At the foundation of linear programming lies the following theorem:

Theorem 7.49. *The maximum of an affine-linear function on a bounded convex polyhedron M is attained at a vertex.*

Proof. By Theorem 7.39 and Proposition 7.48, every point p of a polyhedron M is a convex linear combination of its vertices $p_1, \ldots, p_k$:

$$p = \sum_{i=1}^{k} \lambda_i p_i, \qquad \sum_{i=1}^{k} \lambda_i = 1, \qquad \lambda_i \geq 0, \qquad i = 1, \ldots, k.$$

By Proposition 7.16,

$$f(p) = \sum_{i=1}^{k} \lambda_i f(p_i) \leq \max_i f(p_i)$$

and the theorem follows. □

The following two examples show how the need for linear programming arises naturally.

Example 7.50 (The Maximum Profit Problem). A company possesses resources $\mathrm{R}_1, \ldots, \mathrm{R}_m$ of amounts $b_1, \ldots, b_m$, respectively, and wants to produce products $\mathrm{P}_1, \ldots, \mathrm{P}_n$ of amounts $x_1, \ldots, x_n$, respectively. Let a_{ij} be the amount of the resource R_i needed to produce a unit of the product P_j and

let c_j be the selling price of a unit of the product P_j. Clearly, the following inequalities should hold:

$$\sum_{j=1}^{n} a_{ij}x_j \le b_i, \qquad i = 1, \dots, m, \quad x_j \ge 0, \quad j = 1, \dots, n.$$

They determine a convex polyhedron M in the n-dimensional space with coordinates $x_1, \dots, x_n$. To maximize the profit, one needs to find the point $(x_1, \dots, x_n) \in M$ where the linear function $\sum_{j=1}^{n} c_j x_j$ (the total selling price of the product produced) is maximal.

Example 7.51 (The Transportation Problem). Suppliers $A_1, \dots, A_m$ carry the amounts $a_1, \dots, a_m$, respectively, of a certain product. Customers $B_1, \dots, B_n$ need the amounts $b_1, \dots, b_n$, respectively, of the same product. It is also given that $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$. Let x_{ij} be the amount of product that is transported from A_i to B_j and c_{ij}, the cost to deliver a unit of the product from A_i to B_j. The following conditions must hold:

$$\sum_{j=1}^{n} x_{ij} = a_i, \quad \sum_{i=1}^{m} x_{ij} = b_j, \quad x_{ij} \ge 0.$$

They define a convex polyhedron in the mn-dimensional space with coordinates x_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$. The problem is to minimize the linear function $\sum_{i,j} c_{ij}x_{ij}$ (total cost of transportation) on this polyhedron.

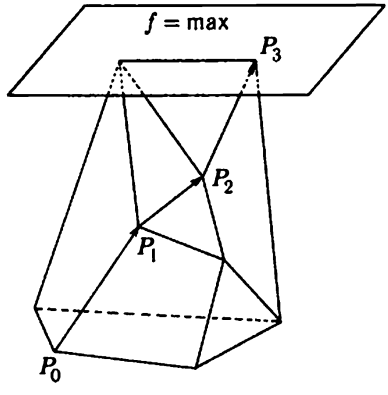


Figure 7.9

To solve a problem of linear programming, one commonly uses the *simplex method*. It consists in sliding by the edges of M in the direction of the increase of f, while possible. The movement ends at a vertex where the function f attains its maximum (see Figure 7.9).

7.3. Affine Transformations and Motions

Consider affine spaces S and S' associated with vector spaces V and V', respectively (over the same field).

Definition 7.52. An *affine map* from the space S to the space S' is a map $f\colon S \to S'$ such that

$$f(p+x) = f(p) + \varphi(x), \qquad p \in S,\ x \in V, \tag{7.11}$$

for some linear map φ from V to V' (independent of p and x).

In particular, affine-linear functions defined in Section 7.1 are simply affine maps from the space S to the field K (regarded as an affine line).

It follows from (7.11) that

$$\varphi(\overline{pq}) = \overline{f(p)f(q)}, \qquad p, q \in S. \tag{7.12}$$

Thus, f determines the linear map φ uniquely. The latter is called the *differential* of f and is denoted df.

Let us vectorize the spaces S and S' by taking points o and o', respectively, as origins. Put $p = o$ in (7.11). We thus obtain the following presentation of the affine map f in the vectorized form:

$$f(x) = \varphi(x) + b, \qquad b \in V', \tag{7.13}$$

where $b = \overline{o'f(o)}$. This implies the coordinate form of f:

$$y_i = \sum_{j=1}^{n} a_{ij}x_j + b_i, \qquad i = 1, \dots, m, \tag{7.14}$$

where $x_1, \dots, x_n$ are the coordinates of the point x and $y_1, \dots, y_m$ are the coordinates of $y = f(x)$.

Conversely, it is easy to check that for any linear map $\varphi\colon V \to V'$ and any vector $b \in V'$, the map defined by (7.13) is affine with the differential equal to φ.

Let S'' be another affine space and $g\colon S' \to S''$, an affine map.

Proposition 7.53. *The map $gf\colon S \to S''$ is affine. Also,*

$$d(gf) = dg \cdot df. \tag{7.15}$$

Proof. For $p \in S$, $x \in V$, we have

$$\begin{aligned}(gf)(p+x) &= g(f(p+x)) = g(f(p) + df(x)) \\ &= g(f(p)) + dg(df(x)) = (gf)(p) + (dg \cdot df)(x).\end{aligned}$$

□

For $K = \mathbb{R}$, the differential of an affine map is a particular case of a differential of a smooth map in analysis. Formula (7.15) is then a particular case of the formula for the differential of the product of smooth maps (i.e., the "composition function").

Proposition 7.54. *An affine map is bijective if and only if its differential is bijective.*

Proof. First, choose origins o and o' in the spaces S and S', so that $f(o) = o'$. Then the map f in its vectorized form coincides with its own differential. The proposition follows. □

A bijective affine transformation is called an *isomorphism* of affine spaces. Affine spaces are *isomorphic* if there exists an isomorphism between them.

Corollary 7.55. *Finite-dimensional affine spaces (over the same field) are isomorphic if and only if they have the same dimension.*

Obviously, an affine map $f: S \to S'$ sends a plane $P = p + U$ of the space S to the plane $f(P) = f(p) + df(U)$ of the space S'. If f is bijective, then $\dim f(P) = \dim P$.

Similarly to the proof of Proposition 7.16, one proves that

$$f\left(\sum_i \lambda_i p_i\right) = \sum_i \lambda_i f(p_i)$$

for any barycentric linear combination $\sum_i \lambda_i p_i$ of points $p_1, \ldots, p_k \in S$. In particular, an affine map sends the center of mass of a system of points to the center of mass of the system of images of these points.

An affine map from an affine space S to itself is called an *affine transformation*. Bijective affine transformations form a group called the *general affine group* of S and denoted $\mathrm{GA}(S)$. (This agrees with the definition in vector form in Section 4.2.)

By Proposition 7.53, the map

$$d: \mathrm{GA}(S) \to \mathrm{GL}(V)$$

is a group homomorphism. Its kernel is the *group of parallel translations*

$$t_a: p \mapsto p + a, \qquad a \in V.$$

We denote it $\mathrm{Tran}\, S$.

Proposition 7.56. *For any $f \in \mathrm{GA}(S)$ and $a \in V$,*

$$f t_a f^{-1} = t_{df(a)}. \tag{7.16}$$

Proof. Apply the transformation ft_af^{-1} to a point $q = f(p)$. We have

$$ft_af^{-1}(q) = ft_a(p) = f(p+a) = f(p) + df(a) = q + df(a).$$

□

Of course, it is not surprising that the map ft_af^{-1} is a parallel translation: the group $\operatorname{Tran} S \subset \mathrm{GA}(S)$ is a kernel of a homomorphism, so it is a normal subgroup.

If one is to fix the origin $o \in S$, so as to identify the affine space S with the associated vector space V, the group $\mathrm{GL}(V)$ becomes a subgroup of $\mathrm{GA}(S)$. This subgroup is just the stabilizer of the point o in the group $\mathrm{GA}(S)$. The vectorized form (7.13) of an affine transformation implies that every affine transformation $f \in \mathrm{GA}(S)$ has the following unique presentation:

$$f = t_b\varphi, \qquad \varphi \in \mathrm{GL}(V),\ b \in V. \tag{7.17}$$

Clearly the map $\varphi = df$ does not depend on the choice of the origin but, generally speaking, the vector $b = \overrightarrow{of(o)}$ does.

Exercise 7.57. Prove that under the change of origin $o' = o + a$ $(a \in V)$, the vector b gets replaced by the vector

$$b' = b + \varphi(a) - a. \tag{7.18}$$

Example 7.58. By Proposition 4.28, every motion of the Euclidean plane E^2 is a (bijective) affine transformation. The same is true for the Euclidean space E^3.

Example 7.59. A *homothety* with the center o and coefficient λ is an affine transformation defined as

$$f(o + x) = o + \lambda x.$$

Clearly, $df = \lambda\mathcal{E}$. Let us prove that every affine transformation f such that $df = \lambda\mathcal{E}$, where $\lambda \neq 1$, is a homothety with a center at some point. It suffices to prove that f has a fixed point. In the vectorized form,

$$f(x) = \lambda x + b, \qquad b \in V.$$

The equation $f(x) = x$ reduces to

$$(1 - \lambda)x = b,$$

hence, it has a (unique) solution.

A homothety with coefficient -1 is called a *central symmetry*.

Exercise 7.60. Prove that a product of homotheties with different centers and coefficients λ and μ is a homothety when $\lambda\mu \neq 1$ and a nontrivial parallel translation when $\lambda\mu = 1$ (but with $\lambda, \mu \neq 1$).

The group of affine transformations defines *affine geometry* in the sense that affine geometry studies those properties of geometric figures that are invariant under (bijective) affine transformations. Such transformations map a plane to a plane of the same dimension, and a barycentric linear combination of points to a barycentric linear combination of their images with the same coefficients. Thus, the notions of a plane and a barycentric combination belong to affine geometry (hence, so do the notions of parallel lines, a parallelogram, an interval, the center of an interval, the center of mass of a system of points, a convex set, a simplex, etc.). But, for instance, the notions of a square and a circle are not the notions of affine geometry, since an affine transformation can map a square to a parallelogram that is not a square, and a circle to an ellipse that is not a circle.

The following theorem shows that all simplices are equal in affine geometry (e.g., on the affine plane, all triangles are equal).

Theorem 7.61. *Let* $\{p_0, p_1, \dots, p_n\}$ *and* $\{q_0, q_1, \dots, q_n\}$ *be two systems of affinely independent points in an* n*-dimensional affine space* S*. Then there exists a unique affine transformation* f *that maps* p_i *to* q_i *for* $i = 0, 1, \dots, n$*.*

Proof. There exists a unique linear map φ of the space V that maps the basis $\{\overrightarrow{p_0p_1}, \dots, \overrightarrow{p_0p_n}\}$ to the basis $\{\overrightarrow{q_0q_1}, \dots, \overrightarrow{q_0q_n}\}$. If we vectorize S by taking p_0 as the origin, the affine transformation in question has the form

$$f(x) = \varphi(x) + \overrightarrow{p_0q_0}.$$

□

Exercise 7.62. Prove that in real affine geometry all parallelepipeds are equal.

Exercise 7.63. Let $P_1, P_2, P_1', P_2' \subset S$ be planes with direction subspaces U_1, U_2, U_1', U_2', respectively. Assume that $\dim P_1 = \dim P_1'$, $\dim P_2 = \dim P_2'$, $\dim U_1 \cap U_2 = \dim U_1' \cap U_2'$, and that the intersections $P_1 \cap P_2$ and $P_1' \cap P_2'$ are simultaneously empty or nonempty. Prove that there exists a transformation $f \in \mathrm{GA}(S)$ that maps P_1 to P_1' and P_2 to P_2'.

In affine geometry, there is no notion of distance between points because any pair of different points can be mapped into any other pair of different points by an affine transformation. However, affine transformations preserve the so-called ratio of a triple of points lying on the same line.

Consider points p_1, p_2, p_3 on the same line l. If $p_2 \neq p_3$, then $\overrightarrow{p_1p_3} = c\overrightarrow{p_3p_2}$ $(c \in K)$. The number c is called the (*simple*) *ratio* of the triple p_1, p_2, p_3 and is denoted (p_1, p_2, p_3). If $p_1 \neq p_2 = p_3$, then we set $(p_1, p_2, p_3) = \infty$. If $p_1 = p_2 = p_3$, then (p_1, p_2, p_3) is undefined. If $c = \frac{\lambda}{\mu}$, we say that the point p_3 divides the interval p_1p_2 in the ratio $\lambda : \mu$ (though the notion of an

interval itself is defined only in real geometry). For $\lambda + \mu = 1$, this means that

$$p_3 = \mu p_1 + \lambda p_2.$$

It is clear that the ratio of points p_1, p_2, p_3 is preserved under any affine transformation that does not contract the line l to a point (in particular, it is preserved under any bijective affine transformation).

Exercise 7.64. Determine how the ratio of a triple of points changes when these points are permuted. What is the maximum and the minimum number of the values that it may have?

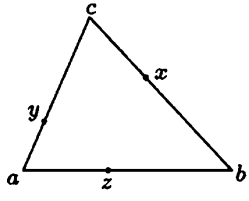


Figure 7.10

Exercise 7.65. Construct a triangle abc given points x, y, z on the sides bc, ca, ab (or their continuations) such that they divide the respective sides in the ratios $\lambda : 1$, $\mu : 1$, $\nu : 1$ (Figure 7.10). (*Hint*: consider the product of homotheties with centers at points x, y, z that map c to b, a to c, and b to a, respectively. See Example 7.6.)

Now let S be a Euclidean affine space associated to a Euclidean vector space V.

Definition 7.66. A *motion* of the space S is an affine transformation of S whose differential is an orthogonal operator. (In particular, it follows that every motion is bijective.)

Clearly, a motion preserves distances between points (for the definition, see Section 7.1). Conversely, a distance-preserving affine transformation is a motion.

Remark 7.67. Actually, one can show that every distance-preserving bijective transformation of S is an affine transformation, hence a motion.

Motions of a Euclidean space S form a group denoted $\operatorname{Isom} S$. A motion is called *proper* (or *orientation preserving*) if its differential belongs to $\mathrm{SO}(V)$ and *improper* (or *orientation reversing*) otherwise. Proper motions form a subgroup of index 2 in $\operatorname{Isom} S$; it is denoted $\operatorname{Isom}_+ S$ (cf. Example 4.114).

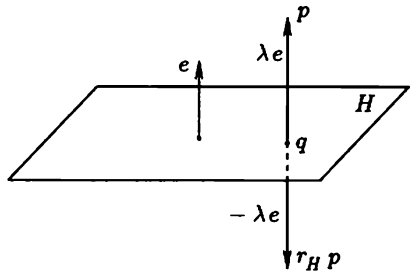


Figure 7.11

Example 7.68. An important example of an improper motion is the (orthogonal) *reflection* r_H through the hyperplane H. Let e be a unit vector which is orthogonal to H. Every point $p \in S$ can be uniquely presented as

$$p = q + \lambda e, \qquad q \in H.$$

By definition,

$$r_H p = q - \lambda e$$

(Figure 7.11).

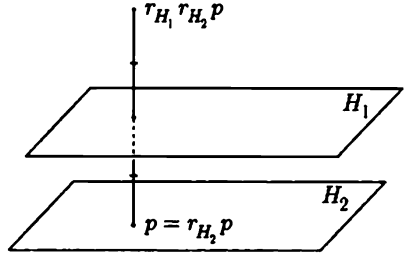


Figure 7.12

The differential of the reflection r_H is an (orthogonal) reflection through the direction subspace of the hyperplane H in the space V. Let H_1 and H_2 be hyperplanes. If they are parallel, $dr_{H_1} = dr_{H_2}$, hence

$$d(r_{H_1} r_{H_2}) = dr_{H_1} \cdot dr_{H_2} = \mathcal{E}.$$

In this case, $r_{H_1} r_{H_2}$ is the parallel translation along twice the common normal vector to H_1 and H_2 (Figure 7.12). On the other hand, if H_1 and H_2 intersect in an $(n-2)$-dimensional plane P, then $r_{H_1} r_{H_2}$ is the rotation through twice the angle between H_1 and H_2, i.e., the map that stabilizes every point of P and is the rotation through this angle in every two-dimensional plane orthogonal to P (cf. Example 6.41).

Exercise 7.69. Prove that the group $\operatorname{Isom} S$ is generated by reflections through hyperplanes.

If we choose an origin in the space S, every motion in S has a unique form (7.17), where $\varphi \in \mathrm{O}(V)$. But in general, the vector b depends on the choice of origin. The following theorem provides a canonical presentation of any motion.

Theorem 7.70. *For any motion f, there exists a uniquely determined plane $P = p_0 + U$ such that*

(i) $f(P) = P$ and $f|_P$ is a parallel translation (perhaps trivial);

(ii) df stabilizes no nonzero vectors in $U^\perp$.

Proof. If such a plane exists, its direction subspace coincides with the subspace of fixed vectors of the operator $\mathcal{A} = df$. Denote this subspace by U. Choose an origin and write f in the vectorized form

$$f(x) = \mathcal{A}x + a.$$

Let $a = b + \mathrm{c}$, $b \in U$, $\mathrm{c} \in U^\perp$. Since the operator $\mathcal{A} - \mathcal{E}$ is nonsingular on $U^\perp$, there exists a unique vector $x_0 \in U^\perp$ such that

$$\mathcal{A}x_0 + \mathrm{c} = x_0.$$

Let p_0 be the point corresponding to x_0. Then

$$f(p_0) = p_0 + b.$$

The plane $P = p_0 + U$ is the unique plane that satisfies conditions (i) and (ii). □

The plane P is called the *axis of the motion f*. A motion f is determined by its axis $P = p_0 + U$, the vector $b \in U$, and the orthogonal transformation $\mathcal{B} = \mathcal{A}_{U^\perp}$ of the space $U^\perp$ that has no fixed vectors. The description of orthogonal transformations implies that for proper motions, $\dim U^\perp$ is even, and for improper ones, odd.

The above theorem allows us to describe the motions of the Euclidean line, plane, and three-dimensional space in terms of elementary geometry. We denote the axis of a motion f by P.

Let f be a motion of the Euclidean line. Two cases are possible.

(i) $\dim P = 1$. Then f is a parallel translation.

(ii) $\dim P = 0$, i.e., P is a point. Then $\mathcal{B} = -\mathcal{E}$ and f is the reflection (symmetry) through the point P.

Let f be a motion of the Euclidean plane. Three cases are possible.

(i) $\dim P = 2$. Then f is a parallel translation.

(ii) $\dim P = 1$, i.e., P is a line. Then $\mathcal{B} = -\mathcal{E}$ and f is the reflection through the line P or a *glide reflection*, i.e., a composition of the reflection through P and a parallel translation along P.

(iii) $\dim P = 0$, i.e., P is a point. Then f is a (nontrivial) rotation about the point P.

Finally, let f be a motion of the three-dimensional Euclidean space. Four cases are possible.

(i) $\dim P = 3$. Then f is a parallel translation.

(ii) $\dim P = 2$. Then f is the reflection through P or a *glide reflection*, i.e., a composition of the reflection through P and a parallel translation by a vector parallel to P.

(iii) $\dim P = 1$. Then f is a (nontrivial) rotation about the line P or a *spiral motion*, i.e., a composition of a rotation about P and a parallel translation by a vector parallel to P.

(iv) $\dim P = 0$. Then f is a *mirror rotation*, i.e., a composition of a (nontrivial) rotation about a line and a reflection through a plane perpendicular to this line such that the plane and the line intersect at P.

Exercise 7.71. Describe the composition of rotations f and g of the Euclidean plane about different points. (*Hint*: calculate $d(fg)$.)

For every figure M of a Euclidean space S, we can define its *symmetry group*

$$\operatorname{Sym} M = \{f \in \operatorname{Isom} S : f(M) = M\}.$$

For instance, crystallographic groups appear as symmetry groups of crystallographic structures.

Observe that if the group $\operatorname{Sym} M$ contains improper motions, then the group

$$\operatorname{Sym}_+ M = \{f \in \operatorname{Isom}_+ S : f(M) = M\}$$

is its subgroup of order 2. It is the kernel of the homomorphism

$$\operatorname{Sym} M \to \{\pm 1\}, \qquad f \mapsto \det df.$$

If M is a bounded convex polyhedron, the group $\operatorname{Sym} M$ is finite since a motion mapping M into itself is uniquely determined by how it permutes the vertices of M (and there can be only a finite number of such permutations). Moreover, the group $\operatorname{Sym} M$ preserves the center of mass of the set of vertices of M, hence is actually a subgroup of the orthogonal group.

The most symmetric polyhedra are the so-called *regular polyhedra*.

Let M be a solid convex polyhedron in an n-dimensional Euclidean space. A *flag* of M is a collection of its faces $\{F_0, F_1, \dots, F_{n-1}\}$, where $\dim F_k = k$ and $F_0 \subset F_1 \subset \cdots \subset F_{n-1}$.

Definition 7.72. A convex polyhedron M is *regular* if for any two of its flags, there exists a motion $f \in \operatorname{Sym} M$ mapping the first to the second.

Since a motion $f \in \operatorname{Sym} M$ is obviously determined by where it maps one of the flags, the order of the symmetry group of a regular polyhedron is equal to the number of its flags.

Two-dimensional regular polyhedra are the ordinary *regular polygons*. Their symmetry groups were described in Example 4.20.

Three-dimensional regular polyhedra are *Platonic solids*, i.e., the regular tetrahedron T, cube K, octahedron O, dodecahedron D, and icosahedron I (see Figure 4.6). The cube and octahedron and the dodecahedron and icosahedron are the so-called dual polyhedra. The centers of faces of one of the dual polyhedra are the vertices of the other; this implies that their symmetry groups coincide. (The tetrahedron is dual to itself.)

By the above, the order of the symmetry group $\operatorname{Sym} P$ of a three-dimensional regular polyhedron P equals the number of its flags, i.e.,

$$|\operatorname{Sym} P| = (\text{number of vertices}) \times (\text{number of edges adjoint to each vertex}) \times 2.$$

Thus,

$$|\operatorname{Sym} T| = 24, \quad |\operatorname{Sym} K| = |\operatorname{Sym} O| = 48, \quad |\operatorname{Sym} D| = |\operatorname{Sym} I| = 120.$$

The order of the group $\operatorname{Sym}_+ P$ is half the order of $\operatorname{Sym} P$. It consists of rotations about the lines passing through the center of P and a boundary point that is either a vertex, or the midpoint of an edge, or the center of a face.

Exercise 7.73. List all elements of the symmetry group of the cube.

Just as we developed the group approach to Euclidean geometry, the same can be done in order to define *pseudo-Euclidean geometry*.

A real vector space V with a fixed symmetric bilinear function α of signature (k, l), where $k, l > 0$, $k + l = n = \dim V$, is called the *pseudo-Euclidean vector space* of signature (k, l). The group of α-preserving linear transformations of V is called the *pseudo-orthogonal group* and is denoted $\mathrm{O}(V, \alpha)$. In the basis where α assumes the normal form, the corresponding matrix group is denoted $\mathrm{O}_{k,l}$.

An affine space S associated to a pseudo-Euclidean vector space V is called a *pseudo-Euclidean affine space* of the corresponding signature. Its

group of motions is the group $\operatorname{Isom} S = d^{-1}(\mathrm{O}(V, \alpha))$. Pseudo-Euclidean geometry is the geometry defined by this group.

Spacetime of special relativity is the pseudo-Euclidean affine space of signature $(3, 1)$. It is called the *Minkowski space* and its group of motions, the *Poincaré group*. (The corresponding group of pseudo-orthogonal transformations is called the *Lorentz group*.)

Exercise 7.74. Describe the group $\mathrm{O}_{1,1}$. (*Hint*: use the coordinate system where the corresponding quadratic function has the form $q(x) = x_1 x_2$.)

Exercise 7.75. State and prove the "side-side-side triangle congruence theorem" for the pseudo-Euclidean plane.

7.4. Quadrics

Planes are the simplest objects of affine and Euclidean geometry. As we know, they are determined by systems of linear equations. A natural generalization of planes (which are also called *linear varieties*) is the so-called *algebraic varieties*. These are subsets of an affine space determined by systems of algebraic equations. Their study is the subject of algebraic geometry, which is a large mathematical field and is not covered in this book. We shall discuss briefly only several general problems of algebraic geometry in Chapter 9. Also, in this section, we will consider the next simple (after planes) type of algebraic varieties, the quadrics. These are determined by a single algebraic equation of the second degree. They include such objects of elementary geometry as circles and spheres.

Assume that $\operatorname{char} K \neq 2$.

Definition 7.76. An *affine-quadratic function* on an affine space S is a function $Q\colon S \to K$ such that its vectorized form is

$$Q(x) = q(x) + l(x) + c \tag{7.19}$$

for a quadratic function q, a linear function l, and a constant c.

Let $\dot{q}$ be the polarization of the quadratic function q, i.e., the corresponding symmetric bilinear function.

Lemma 7.77. *With a change of origin from o to $o' = o + a$ $(a \in V)$, the summands in (7.19) change as follows:*

$$q'(x) = q(x), \qquad l'(x) = 2\dot{q}(a, x) + l(x), \qquad c' = q(a) + l(a) + c. \tag{7.20}$$

Proof. We have

$$\begin{aligned} Q(o'+x) &= Q(o+a+x) = q(a+x) + l(a+x) + \mathrm{c} \\ &= q(a) + 2\dot{q}(a,x) + q(x) + l(a) + l(x) + \mathrm{c} \\ &= q(x) + (2\dot{q}(a,x) + l(x)) + (q(a) + l(a) + \mathrm{c}). \end{aligned}$$

□

In particular, the quadratic function q does not depend on the choice of origin.

In coordinate form, expression (7.19) becomes

$$Q(x) = \sum_{i,j} a_{ij}x_ix_j + \sum_i b_ix_i + \mathrm{c}, \qquad a_{ij} = a_{ji}. \tag{7.21}$$

Coefficients b_i and c have the following meaning:

$$\mathrm{c} = Q(o), \qquad b_i = \frac{\partial Q}{\partial x_i}(o). \tag{7.22}$$

The linear function

$$l(x) = \sum_i b_ix_i$$

is called the *differential* of Q at the point o and is denoted d_oQ. When $K = \mathbb{R}$, this agrees with the standard definition of the differential.

Definition 7.78. A point o is a *center* of an affine-quadratic function Q if

$$Q(o+x) = Q(o-x) \qquad \forall x \in V. \tag{7.23}$$

Clearly, o is a center of Q if and only if $d_oQ = 0$. Thus, the set of all centers of Q is determined by the following system of linear equations:

$$\frac{\partial Q}{\partial x_1} = \cdots = \frac{\partial Q}{\partial x_n} = 0. \tag{7.24}$$

This is either a plane of positive dimension or an empty set. It is easy to see that the coefficient matrix of system (7.24) is twice the matrix (a_{ij}) of q. Therefore, Q has a unique center if q is nondegenerate.

Let

$$X(Q) = \{p \in S : Q(p) = 0\}.$$

Definition 7.79. A *quadric* (or a *quadric hypersurface*) is the set $X(Q)$ for an affine-quadratic function Q unless $X(Q)$ is a plane or is empty.

A planar quadric is called a *conic* (or a *quadric curve*). A quadric in a three-dimensional space is also called a *quadric surface*.

Definition 7.80. A point o is a *center* of a quadric if this quadric is symmetric with respect to o, i.e., if the quadric contains the point $o - x$, $x \in V$, whenever it contains the point $o + x$. The center of a quadric that lies on the quadric itself is called a *vertex*.

A quadric is *central* if it has (at least one) center.

Obviously, every center of an affine-quadratic function Q is a center of the quadric $X(Q)$. Below we will see that the converse is also true.

Let us now show several easy geometric properties of quadrics.

Proposition 7.81. *If a line intersects a quadric in at least three distinct points, then it lies on the quadric entirely.*

Proof. Since any point can be chosen as the origin o, we can assume that our line passes through o. Let Q have vectorized form (7.19). Then the intersection of $X(Q)$ with the line $L = o + \langle x \rangle = \{o + tx : t \in K\}$, $x \in V$, is determined by the following condition:

$$Q(tx) = t^2 q(x) + tl(x) + c = 0. \tag{7.25}$$

This is a quadratic equation with respect to t. If all its coefficients are 0, then $L \subset X(Q)$. Otherwise, it has at most two roots and this implies that the intersection $L \cap X(Q)$ contains at most two points. □

Proposition 7.82. *If o is a vertex of a quadric X and X contains a point $p \neq o$, the quadric also contains the line op.*

Proof. Let $p = o + x$ for $x \in V$. Then X contains three distinct points of the line op: o, $o + x$, and $o - x$. Hence, it contains all of the line. □

A subset of an affine space such that together with points o and $p \neq o$, it contains the line op for every such p, is called a *cone* with the vertex at o. A quadric is called *conic* if it has a vertex.

Proposition 7.83. *Every quadric contains a point which is not its vertex.*

Proof. If all points of a quadric are vertices, then by Proposition 7.82, it contains every line passing through any two of its points. Hence, by Theorem 7.8, it is a plane. This contradicts the definition of a quadric. □

Clearly, any proportional affine-quadratic functions determine the same quadric. The converse is not so obvious and is the subject of the following

Theorem 7.84. *Let X be a quadric in an affine space over an infinite field K. If $X = X(Q_1) = X(Q_2)$ for affine-quadratic functions Q_1, Q_2, then these functions are proportional.*

Proof. Take as an origin a point o on X such that o is not a vertex of Q. Then, in the vectorized form,

$$Q_1(x) = q_1(x) + l_1(x), \qquad Q_2(x) = q_2(x) + l_2(x),$$

where $l_1, l_2 \neq 0$. Points of intersection of a line $o + \langle x \rangle$ with the quadric X are determined by either of the equations

$$t^2 q_1(x) + t l_1(x) = 0, \qquad t^2 q_2(x) + t l_2(x) = 0.$$

Since the solutions of these equations (with respect to t) must coincide, for $l_1(x)$ and $l_2(x) \neq 0$, we have

$$\frac{q_1(x)}{l_1(x)} = \frac{q_2(x)}{l_2(x)}.$$

Thus,

$$q_1(x) l_2(x) = q_2(x) l_1(x). \tag{7.26}$$

Multiplying the above equality by $l_1(x) l_2(x)$, we obtain

$$q_1(x) l_2(x) l_1(x) l_2(x) = q_2(x) l_1(x) l_1(x) l_2(x).$$

The latter equality holds for all x. However, since there are no zero divisors in the polynomial ring, we can cancel the common factor on both sides. Thus, we see that (7.26) holds for all x as well.

Assume that the linear functions l_1 and l_2 are not proportional. Then, in a suitable basis, $l_1(x) = x_1$, $l_2(x) = x_2$ and (7.26) can be rewritten as

$$q_1(x) x_2 = q_2(x) x_1.$$

Looking at factors on both sides of this equality, we see that

$$q_1(x) = l(x) x_1, \qquad q_2(x) = l(x) x_2$$

for a linear function $l(x)$. Hence,

$$Q_1(x) = (l(x) + 1) x_1, \qquad Q_2(x) = (l(x) + 1) x_2.$$

Since $X = X(Q_1)$, X contains the hyperplane $x_1 = 0$. Since $X = X(Q_2)$, we see that the function Q_2 is identically zero on this hyperplane. However, none of the factors $l(x) + 1$ and x_2 is identically zero there (the former is not zero even at the point 0). Since there are no zero divisors in the polynomial ring, we obtain a contradiction.

Therefore, $l_2 = \lambda l_1$, $\lambda \in K^*$. By (7.26), $q_2 = \lambda q_1$. Thus, $Q_2 = \lambda Q_1$. □

Corollary 7.85. *A center of a quadric $X(Q)$ is also a center of the function Q.*

Proof. If o is a center of $X(Q)$, then $X(Q) = X(\overline{Q})$ for

$$\overline{Q}(o+x) = Q(o-x).$$

Thus, $\overline{Q} = \lambda Q$, $\lambda \in K^*$. Comparing the terms of the second degree in the expressions for Q and $\overline{Q}$, we see that $\lambda = 1$. Hence, $\overline{Q} = Q$ and this implies that o is a center of the function Q. □

Corollary 7.86. *If a quadric $X(Q)$ is invariant under a parallel translation, then the function Q is invariant under this translation.*

Proof. If $X(Q)$ is mapped into itself under a parallel translation by a vector a, then $X(Q) = X(\overline{Q})$ for

$$\overline{Q}(p) = Q(p+a).$$

The rest of the proof repeats that of the previous corollary. □

Remark 7.87. A careful reading of the proof of Theorem 7.84, together with Remark 3.77, shows that it also holds for finite fields, the only exception being the field $\mathbb{Z}_3$. (Recall that we assumed that $\operatorname{char} K \neq 2$.) Over $\mathbb{Z}_3$, the following counterexample exists: consider equations $x_1^2 + x_1x_2 + 1 = 0$ and $x_2^2 + x_1x_2 + 1 = 0$. They determine the same quadric in $\mathbf{Z}_3^2$ (it consists of the points $(1,1)$ and $(-1,-1)$). However, both of the above corollaries remain valid over $\mathbb{Z}_3$.

Consider an affine-quadratic function Q in vectorized form (7.19). Set

$$\operatorname{Ker} Q = \operatorname{Ker} q \cap \operatorname{Ker} l \tag{7.27}$$

(here $\operatorname{Ker} q := \operatorname{Ker} \hat{q}$).

Proposition 7.88. *A function Q is invariant under a parallel translation by a vector a if and only if $a \in \operatorname{Ker} Q$.*

This implies, in particular, that $\operatorname{Ker} q \cap \operatorname{Ker} l$ does not depend on the choice of origin.

Proof. Invariance of Q under the parallel translation by a is equivalent to its preserving its form when the origin changes from o to $o' = o + a$. By Lemma 7.77, this holds if and only if $a \in \operatorname{Ker} Q$. □

Thus, if $U = \operatorname{Ker} Q \neq 0$, the quadric $X = X(Q)$ contains the plane $p+U$ for every point $p \in X$. Such a quadric is called a *cylindrical quadric* with a generator U. Choose the basis of the space V so that its last d vectors form a basis of the subspace U. Then the coordinate expression of Q does not contain the last d coordinates. The equation $Q = 0$ can be viewed as an equation of a quadric X_0 in the $(n-d)$-dimensional space. Then the first

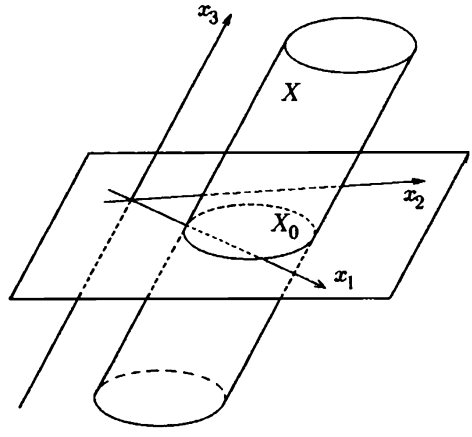


Figure 7.13

$n-d$ coordinates of a point on X are coordinates of a point on X_0 and the other coordinates are arbitrary (Figure 7.13).

Therefore, in order to describe all quadrics, it suffices to describe only the noncylindrical ones.

Proposition 7.89. *A noncylindrical quadric has at most one center.*

Proof. Assume that a quadric X has two centers o and o'. Denote by s and s' the central symmetries about o and o', respectively. Then $sX = s'X = X$, hence $ss'X = X$. Since

$$d(ss') = ds \cdot ds' = (-\mathcal{E})^2 = \mathcal{E},$$

ss' is a (nontrivial) parallel translation. Hence X is a cylindrical quadric. □

Noncylindrical quadrics can be arranged in three types.

I. *Nonconic central quadrics.*

We set the origin at the center of the quadric. Multiplying the equation of the quadric by an appropriate factor, we obtain

$$q(x_1, \ldots, x_n) = 1, \tag{7.28}$$

where q is a nondegenerate quadratic function.

II. *Conic quadrics.*

We set the origin at the vertex of the quadric. The equation of the quadric thus becomes

$$q(x_1, \ldots, x_n) = 0, \tag{7.29}$$

where q is a nondegenerate quadratic function. We can still multiply the equation by any number $\lambda \neq 0$.

III. *Noncentral quadrics.*

Since $\operatorname{Ker} q \cap \operatorname{Ker} l = 0$ but $\operatorname{Ker} q \neq 0$ (otherwise the quadric is central), $\dim \operatorname{Ker} q = 1$. Hence,

$$V = \operatorname{Ker} l \oplus \operatorname{Ker} q. \tag{7.30}$$

Choose an origin lying on the quadric and a basis of V that agrees with decomposition (7.30). The equation of the quadric becomes

$$u(x_1, \ldots, x_{n-1}) = x_n, \tag{7.31}$$

where $u = q|_{\operatorname{Ker} l}$ is a nondegenerate quadratic function in $n-1$ variables. We can also multiply this equation by any number $\lambda \neq 0$, as long as we also divide the last vector in the basis by λ.

The equation of the quadric can be simplified further by choosing a suitable basis in V. What can be achieved in this way depends on the field K (see Section 5.3). In particular, when $K = \mathbb{C}$ or $\mathbb{R}$, the quadratic function q can be reduced to its normal form.

Let us focus now on the case $K = \mathbb{R}$. Here the equation of a noncylindrical quadric reduces to one and only one of the following types:

I. *Nonconic central quadrics.*

$$x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2 = 1, \qquad 0 < k \leq n. \tag{7.32}$$

II. *Conic quadrics.*

$$x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2 = 0, \qquad \frac{n}{2} \leq k < n. \tag{7.33}$$

(The inequality $k \geq \frac{n}{2}$ can be satisfied by possible multiplication of the equation by -1.)

III. *Noncentral quadrics.*

$$x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_{n-1}^2 = x_n, \qquad \frac{n-1}{2} \leq k < n. \tag{7.34}$$

The above list can be interpreted as the classification of real quadrics up to an affine transformation. Indeed, if quadrics X_1 and X_2 have the same equation in the affine systems of coordinates determined by frames $\{o; e_1, \ldots, e_n\}$ and $\{o'; e_1', \ldots, e_n'\}$, respectively, then X_1 is mapped to X_2 by an affine transformation that maps the first frame to the second. Conversely, if an affine transformation f maps the quadric X_1 to the quadric X_2, then X_1 and X_2 have the same equation in the affine systems of coordinates determined by frames $\{o; e_1, \ldots, e_n\}$ and $\{f(o); df(e_1), \ldots, df(e_n)\}$, respectively.

Table 1

n	type	k	name
2	I	2	ellipse
		1	hyperbola
	II	1	two intersecting lines
	III	1	parabola
3	I	3	ellipsoid
		2	hyperboloid of one sheet
		1	hyperboloid of two sheets
	II	2	quadric cone
	III	2	elliptic paraboloid
		1	hyperbolic paraboloid

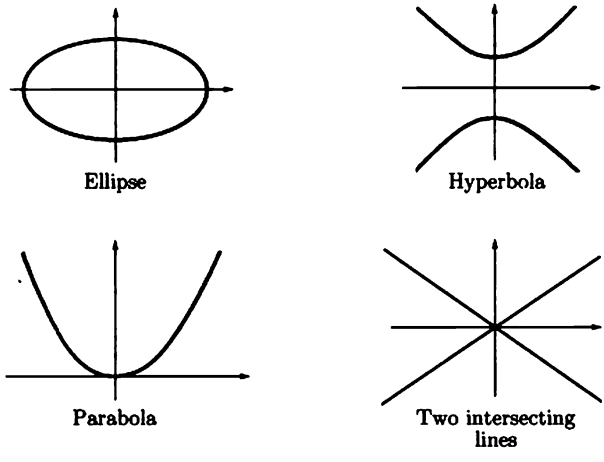
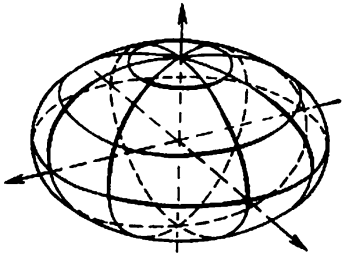


Figure 7.14

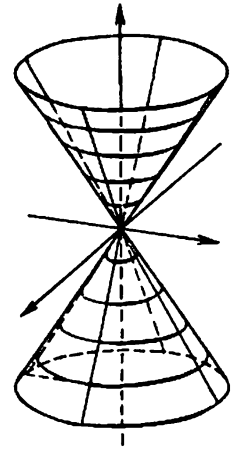
In particular, for $n = 2$ or 3, we obtain the well-known classes of real quadric curves and surfaces listed in Table 1 and shown in Figures 7.14 and 7.15, respectively.

In a space of arbitrary dimension, quadrics of type I are called *ellipsoids* for $k = n$ and *hyperboloids* for $k < n$. Quadrics of type II are called *quadric cones*. Quadrics of type III are called *elliptic paraboloids* for $k = n - 1$ and *hyperbolic paraboloids* for $k < n - 1$.

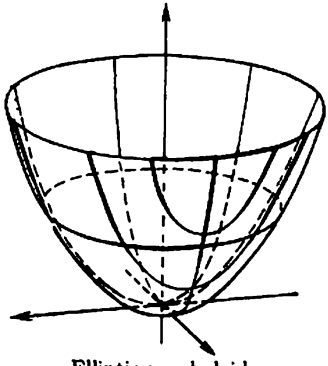
A real quadric X is a smooth hypersurface in a neighborhood of a point $p \in X$ if and only if $d_pQ \neq 0$, i.e., if p is not a vertex. In this case, the



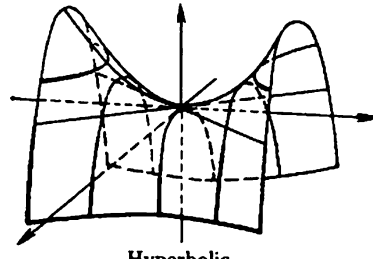
Ellipsoid



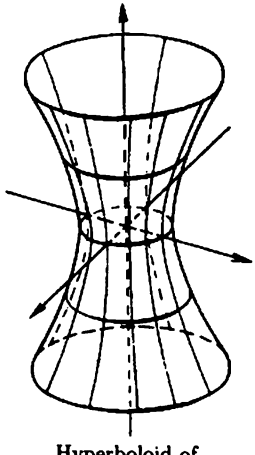
Cone



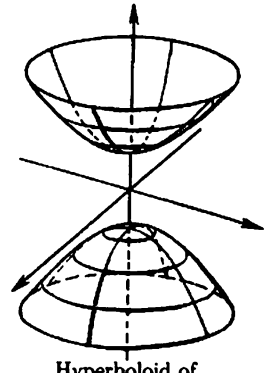
Elliptic paraboloid



Hyperbolic paraboloid



Hyperboloid of one sheet



Hyperboloid of two sheets

Figure 7.15

equation $d_pQ(x-p)=0$ determines the tangent hyperplane of X at p. In particular, nonconic quadrics are smooth everywhere.

Unlike generic hypersurfaces of higher degree, real (and complex) quadrics have a wonderful property: many affine symmetries.

Let X be a real quadric. Denote by $G(X)$ the group of all affine transformations mapping X to itself.

Theorem 7.90. *If X is a nonconic quadric, the group $G(X)$ acts transitively on X. If X is a conic quadric, the group $G(X)$ acts transitively on the complement of the set of vertices in X.*

Proof. If X is a cylindrical quadric with a generator U, then the group $G(X)$ contains the group of parallel translations by vectors in U, which acts transitively on every plane of the form $p+U$. Thus, in this case, it suffices to prove the theorem for the noncylindrical quadric X_0 in the space of smaller dimension (in the above notation).

Let X be an ellipsoid with the equation $q(x)=1$ (in vectorized form) for a positive definite quadratic function q. We can make the space V Euclidean by taking the polarization of q as the inner product. Then X becomes a unit sphere in this space and $G(X)$ at least contains the orthogonal group $\mathrm{O}(V)$. (In fact, they coincide but we do not need this stronger statement here.) Choose vectors x, x' in X. Then

$$V=\langle x\rangle\oplus\langle x\rangle^{\perp}=\langle x'\rangle\oplus\langle x'\rangle^{\perp}.$$

Consider a linear transformation $\varphi\in\mathrm{GL}(V)$ that sends x to x' and maps isomorphically the Euclidean space $\langle x\rangle^{\perp}$ to the Euclidean space $\langle x'\rangle^{\perp}$. Clearly $\varphi\in\mathrm{O}(V)$, so by construction $\varphi(x)=x'$.

The case of a hyperboloid X is similar. The only difference is that the inner product determined by the polarization of q makes V into a pseudo-Euclidean space of signature (k,l), $k+l=n$. Here the subspaces $\langle x\rangle^{\perp}$ and $\langle x'\rangle^{\perp}$ are pseudo-Euclidean of signature $(k-1,l)$, hence isomorphic.

Now let X be a quadric cone with the equation $q(x)=0$ (in vectorized form) for a quadratic function q of signature (k,l), $k+l=n$. As above, we make V into a pseudo-Euclidean space. For any nonzero vector $x\in X$, there exists a vector $y\in V$ such that $(x,y)\neq 0$. By normalizing vector y we may assume that $(x,y)=1$. Next, we can replace y with $y-\frac{1}{2}(y,y)x$ so that $(y,y)=0$ (the equality above still holds). Then, in the two-dimensional subspace $\langle x,y\rangle$, the matrix of the inner product is $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$. Thus, it is nondegenerate and of signature $(1,1)$. This implies that

$$V=\langle x,y\rangle\oplus\langle x,y\rangle^{\perp},$$

where $\langle x, y\rangle^{\perp}$ is a pseudo-Euclidean (or Euclidean) space of signature $(k-1, l-1)$. Likewise, for a different nonzero vector $x' \in X$, we obtain the decomposition

$$V = \langle x', y'\rangle \oplus \langle x', y'\rangle^{\perp}.$$

Consider a linear transformation $\varphi \in \mathrm{GL}(V)$ mapping x to x', y to y', and the subspace $\langle x, y\rangle^{\perp}$ to $\langle x', y'\rangle^{\perp}$ so that it is an isomorphism of pseudo-Euclidean spaces. Then $\varphi \in \mathrm{O}(V, q) \subset G(X)$ and by construction, $\varphi(x) = x'$.

Finally, let X be a paraboloid with equation (7.31) in the vectorized form. Any vector $x \in V$ can be presented in the form $x = y + te$, where $y \in \operatorname{Ker} l$, $t \in \mathbb{R}$, and e is a basis vector of the subspace $\operatorname{Ker} q$, so that $x \in X$ if and only if $u(y) = t$. For any $a \in \operatorname{Ker} l$, consider the affine transformation

$$f_a \colon y + te \mapsto y + a + (t + 2\dot{u}(a, y) + u(a))e.$$

If $u(y) = t$, then

$$u(y + a) = t + 2\dot{u}(a, y) + u(a)$$

and vice versa. This means that $f_a \in G(X)$. Obviously, transformations f_a, $a \in \operatorname{Ker} l$, form a group that acts transitively on X. □

Exercise 7.91. Prove that if X is a paraboloid determined by (7.31), then the group $G(X)$ acts transitively on the region $u(x_1, \ldots, x_{n-1}) < x_n$.

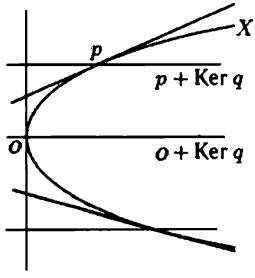


Figure 7.16

To every paraboloid $X = X(Q)$, we associate canonically one-dimensional subspace $\operatorname{Ker} q \subset V$ called the *special direction* of X. Since for any choice of origin, $\operatorname{Ker} q \not\subset \operatorname{Ker} l$, equation (7.25) has exactly one solution for $x \in \operatorname{Ker} q$. Therefore, a line in the special direction intersects the paraboloid at exactly one point. Moreover, for the same reason this intersection is transversal (Figure 7.16).

Exercise 7.92. Prove that for a nonspecial direction of a paraboloid X, there exists a line in this direction that does not intersect X.

Let us discuss now to which form we can reduce the equation of a quadric in a Euclidean space if we restrict ourselves to Cartesian coordinate systems. As in affine geometry, the problem reduces to the case of noncylindrical quadrics. Just as above, consider three types of such quadrics.

I. *Nonconic central quadrics.*

After reduction of the quadratic function q to principal axes (see Corollary 6.37), the equation of such a quadric in Cartesian coordinates reduces to the form

$$\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 = 1, \qquad \lambda_1, \ldots, \lambda_n \neq 0. \tag{7.35}$$

Factors $\lambda_1, \ldots, \lambda_n$ are determined uniquely up to a permutation.

II. *Conic quadrics.*

In Cartesian coordinates, the equation of such a quadric reduces to

$$\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 = 0, \qquad \lambda_1, \ldots, \lambda_n \neq 0. \tag{7.36}$$

Factors $\lambda_1, \ldots, \lambda_n$ are determined uniquely up to a permutation and simultaneous multiplication by $\lambda \neq 0$.

III. *Noncentral quadrics (paraboloids).*

Choose an origin and reduce the quadratic function q to principal axes. We thus obtain Cartesian coordinates in which the equation of the paraboloid is

$$\lambda_1 x_1^2 + \cdots + \lambda_{n-1} x_{n-1}^2 + b_1 x_1 + \cdots + b_{n-1} x_{n-1} + b_n x_n + c = 0$$
$$(\lambda_1, \ldots, \lambda_{n-1}, b_n \neq 0).$$

By changing the $x_1, \ldots, x_{n-1}$ coordinates of the origin, we can remove linear terms containing these coordinates (the free term will change as well). After this, by changing the x_n coordinate of the origin, we can remove the free term. Finally, after appropriately multiplying the equation of the paraboloid, we obtain

$$\lambda_1 x_1^2 + \cdots + \lambda_{n-1} x_{n-1}^2 = x_n, \qquad \lambda_1, \ldots, \lambda_{n-1} \neq 0. \tag{7.37}$$

Let us show that the choice of the origin such that the equation of the paraboloid assumes the form (7.37) is unique. We need to describe it in invariant terms.

Let $\{o; e_1, \ldots, e_n\}$ be the frame in which the equation of the paraboloid assumes the form (7.37). Then the special direction of this paraboloid is $\langle e_n \rangle$ and its tangent hyperplane at o is determined by the equation $x_n = 0$. Thus, if the basis $\{e_1, \ldots, e_n\}$ is orthonormal, the tangent hyperplane at o is orthogonal to the special direction. Such a point o is called a *vertex* of the paraboloid (even though this definition does not agree with Definition 7.80).

The line in the special direction passing through this point is called the *axis* of the paraboloid. Note that these notions are defined only for a paraboloid in a Euclidean space.

Proposition 7.93. *For every paraboloid in a Euclidean space, its vertex is unique.*

(In Figure 7.16, the point o is the vertex.)

Proof. Let p be a point on the paraboloid with coordinates $x_1, \dots, x_n$. Differentiating equation (7.37), we see that the coordinates of the normal vector to the paraboloid at p are

$$2\lambda_1 x_1, \dots, 2\lambda_{n-1} x_{n-1}, -1.$$

In order for p to be a vertex, it is necessary and sufficient that this vector be proportional to e_n. This happens if and only if $x_1 = \dots = x_{n-1} = 0$, i.e., $p = o$. □

Corollary 7.94. *Factors $\lambda_1, \dots, \lambda_{n-1}$ in equation (7.37) are determined up to a permutation and simultaneous multiplication by -1.*

Proof. As we have shown, the choice of the origin such that the equation of paraboloid assumes the form (7.37) is unique. The vector e_n is determined uniquely up to multiplication by -1 as the unit vector of the special direction. Such a multiplication results in the multiplication by -1 of all the left-hand side of (7.37). If e_n is fixed, we cannot multiply the equation by $\lambda \neq 1$ without changing the right-hand side. But then $\lambda_1, \dots, \lambda_{n-1}$ are determined uniquely (up to a permutation) as the eigenvalues of the symmetric operator corresponding to the function q. □

Above, we interpreted our results as the classification of quadrics up to affine transformations; here we can interpret our results as the classification of quadrics in a Euclidean space up to motions.

7.5. Projective Spaces

On a photo or a (realistic) painting of a flat locale, images of parallel lines intersect in general and images of equal intervals lying on the same line are not equal (Figure 7.17). This means that the transfer of a landscape onto the picture's plane is not affine. The same can be said regarding images that appear on the retina. In both cases we actually deal with central projections.

Another real-life example of a central projection is the light spot that a lamp with a round shade makes on the floor. When the shade is pointing straight down, the boundary of the spot is circular, just as the shade itself. But if we begin to rotate the shade about a horizontal axis, the circle turns

into an ellipse that stretches further and further. Finally, when its farthest end reaches infinity, it turns into a parabola. If we rotate the shade more, the parabola "opens up" and turns into a branch of a hyperbola (if we put another lamp on the other side of the shade, we would see the other branch of this hyperbola). So, the rim of the shade projects onto the floor either as an ellipse or as a parabola or as a hyperbola.

Notice one more thing. On a picture of a flat landscape, images of lines intersect at the point that is not the image of any point in the landscape itself (otherwise the lines would not be parallel). On the other hand, when the boundary of the light spot becomes a parabola, the image of the uppermost point of the rim disappears into infinity. We see that central projection is more than simply nonaffine, it is also nonsurjective and is not defined everywhere.

In order to study central projection, it is useful to consider the set called the projective plane. Its "points" are lines passing through the center of projection and a point at which such a line intersects the plane of projection is the image of the corresponding "point." Note that "points" corresponding

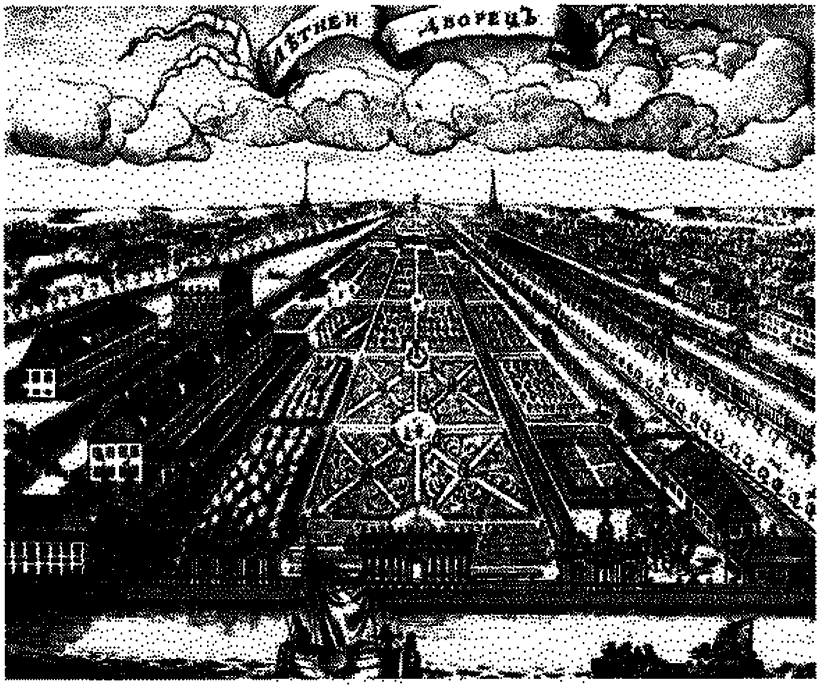


Figure 7.17. A. Zubov, Summer Gardens (St. Petersburg, Russia), engraving.

to lines that are parallel to the plane of projection have no images. (They will if we choose to project onto another plane.) They are called "points at infinity" with respect to the given plane of projection.

Also, the set of "points" corresponding to all lines on a plane that passes through the center of projection is quite naturally called a "line" of the projective plane. On the plane of projection such a "line"—with its "point at infinity" removed—is represented by a regular line. The only exception is the "line" corresponding to the plane parallel to the plane of projection. It consists of "points at infinity" and is not represented at all. This "line" is called the "line at infinity" with respect to the given plane of projection.

This construction can be interpreted as the addition of the "line at infinity" formed by the "points at infinity" to the affine plane. Here we add the same "point at infinity" to all lines from the sheaf of parallel lines on the affine plane. In the resulting "plane," any two lines intersect.

Generalizing this approach to any dimension and any field, we arrive at the following definitions.

Definition 7.95. An n-dimensional *projective space* PV over a field K is the set of one-dimensional subspaces of an $(n+1)$-dimensional vector space V over K. For every $k+1$-dimensional subspace $U \subset V$, the subset $PU \subset PV$ is called a k-dimensional *plane* of the space PV.

In particular, zero-dimensional planes are the points of PV, one-dimensional planes are called *lines* and $(n-1)$-dimensional planes are called *hyperplanes*.

Obviously, the intersection of several planes is a plane (as long as it is not empty).

The space PK^{n+1} built on the space of rows K^{n+1} is sometimes denoted KP^n.

Given a nonzero vector $x \in V$, we denote by $\widehat{x}$ the one-dimensional space $\langle x \rangle$ regarded as a point of the space PV.

Let S be a hyperplane of the space V that does not pass through the origin; let V_S be its direction subspace. Define the map

$$\varphi_S \colon PV \setminus PV_S \to S$$

that sends a point $\widehat{x} \in PV \setminus PV_S$ ($x \in V \setminus V_S$) to the intersection of the line $\langle x \rangle$ with S (see Figure 7.18).

Definition 7.96. The hyperplane S together with the map φ_S is an *affine chart* of the space PV. Points of the hyperplane PV_S of V are called *points at infinity* with respect to the affine chart S.

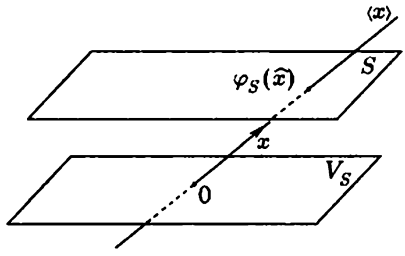


Figure 7.18

Remark 7.97. The term *affine chart* agrees with the usual meaning of the word *chart*. Just as a geographical chart (i.e., a map) is a projection of a piece of the Earth surface onto a piece of paper, an affine chart is a projection of a piece of the projective space onto an affine space.

Remark 7.98. Identifying points of the projective space with their images on an affine chart, we sometimes speak of an affine chart as if it were a piece of the projective space. Likewise, we can also say that we get the projective space by adding points at infinity to an affine chart.

Every k-dimensional plane of the space PV that does not completely lie in PV_S is depicted as a k-dimensional plane on the affine chart S. Planes that lie in PV_S are called *planes at infinity* with respect to S.

The *homogeneous coordinates* of a point $\widehat{x} \in PV$ are coordinates of x in a basis of V. Homogeneous coordinates of a point are defined up to a multiplication by $\lambda \neq 0$. In this they differ from coordinates in the usual meaning of this word; moreover, they cannot be simultaneously zero. A point with homogeneous coordinates $x_0, x_1, \ldots, x_n$ is denoted $(x_0 : x_1 : \cdots : x_n)$.

The *nonhomogeneous* coordinates of a point in PV are the affine coordinates of its image in an affine chart. Unlike the homogeneous coordinates, the nonhomogeneous coordinates of a point are defined uniquely; however, not every point has nonhomogeneous coordinates, namely, they are not defined for points at infinity with respect to the given affine chart.

We will now find a relation between homogeneous and nonhomogeneous coordinates. Let $\{e_0, e_1, \ldots, e_n\}$ be a basis of V. Consider the affine chart

$$S_0 = e_0 + \langle e_1, \ldots, e_n \rangle \tag{7.38}$$

(see Figure 7.19). The image of a point $\widehat{x} = (x_0 : x_1 : \cdots : x_n)$ on S_0 is the point

$$e_0 + \frac{x_1}{x_0} e_1 + \cdots + \frac{x_n}{x_0} e_n$$

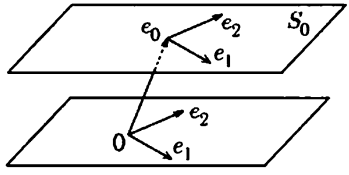


Figure 7.19

whose affine coordinates in the frame $\{e_0; e_1, \ldots, e_n\}$ are $\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}$. Therefore, for the given affine chart and frame, the nonhomogeneous coordinates of the point $(x_0 : x_1 : \cdots : x_n)$ are the ratios $\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}$. The points with $x_0 = 0$ are points at infinity with respect to S_0.

Similarly, nonhomogeneous coordinates of the point $\widehat{x}$ on the affine chart

$$S_i = e_i + \langle e_0, e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n \rangle \tag{7.39}$$

are the ratios $\frac{x_0}{x_i}, \frac{x_1}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i}$. The points with $x_i = 0$ are points at infinity with respect to S_i.

Observe that the charts $S_0, S_1, \ldots, S_n$ form an *atlas* in the sense that they cover all of PV.

Exercise 7.99. Prove that no atlas of PV contains less than $n + 1$ charts.

Exercise 7.100. Let $y_1, \ldots, y_n$ be nonhomogeneous coordinates of the image of a point $\widehat{x} \in PV$ on the chart S_0. Find its nonhomogeneous coordinates on the chart S_1.

Theorem 7.101. *For any $k + 1$ points of a projective space, there exists a plane of dimension $\leq k$ passing through all of them. Moreover, if these points are not contained in a plane of dimension $< k$, there is a unique k-dimensional plane that passes through them.*

Proof. The translation of this theorem to the language of vector spaces is the following obvious statement: any set of $k + 1$ vectors is contained in a subspace of dimension $\leq k + 1$; if they do not lie in a subspace of dimension $< k + 1$, then they lie in a unique subspace of dimension $k + 1$. □

Theorem 7.102. *Let Π_1 and Π_2 be two planes of an n-dimensional projective space. If $\dim \Pi_1 + \dim \Pi_2 \geq n$, then $\Pi_1 \cap \Pi_2 \neq \varnothing$ and*

$$\dim(\Pi_1 \cap \Pi_2) \geq \dim \Pi_1 + \dim \Pi_2 - n. \tag{7.40}$$

For instance, every two lines on a projective plane intersect.

Proof. If $\Pi_1 = PU_1$ and $\Pi_2 = PU_2$, then

$$\dim U_1 + \dim U_2 = \dim \Pi_1 + \dim \Pi_2 + 2 \geq n + 2 > \dim V.$$

Therefore, $U_1 \cap U_2 \neq 0$ and hence, $\Pi_1 \cap \Pi_2 = P(U_1 \cap U_2) \neq \varnothing$. More precisely,

$$\dim(U_1 \cap U_2) \geq \dim U_1 + \dim U_2 - \dim V$$

and this implies (7.40). □

Every nonsingular linear operator $\mathcal{A} \in \mathrm{GL}(V)$ maps one-dimensional subspaces into one-dimensional subspaces, hence defines a bijective map $\widehat{\mathcal{A}}$ of the space PV.

Definition 7.103. A *projective transformation* of the space PV is a transformation of the form $\widehat{\mathcal{A}}$, $\mathcal{A} \in \mathrm{GL}(V)$.

Obviously, a projective transformation maps a plane in PV to another plane of the same dimension.

The map $\mathcal{A} \mapsto \widehat{\mathcal{A}}$ is a homomorphism of the group $\mathrm{GL}(V)$ to the group of transformations of the space PV. Its image is the group of all projective transformations of PV, also called the *general projective group* of PV; it is denoted $\mathrm{PGL}(V)$.

Lemma 7.104. *The kernel of the homomorphism $\mathcal{A} \mapsto \widehat{\mathcal{A}}$ is the group of scalar operators $\lambda\mathcal{E}$, $\lambda \in K^*$.*

Proof. If the operator $\mathcal{A}$ maps every one-dimensional space into itself, then every nonzero vector is its eigenvector. But clearly, the sum of eigenvectors with different eigenvalues is not an eigenvector. Therefore, all eigenvalues of $\mathcal{A}$ are the same, hence $\mathcal{A}$ is a scalar operator. □

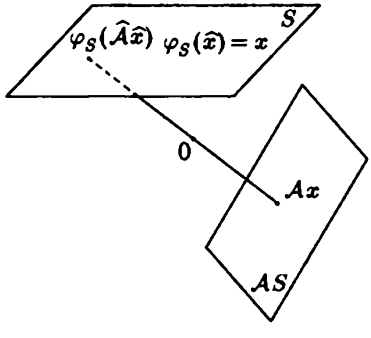


Figure 7.20

Thus we obtain

$$\mathrm{PGL}(V) \simeq \mathrm{GL}(V)/\{\lambda\mathcal{E} \colon \lambda \in K^*\}.$$

Let us study how a projective transformation $\widehat{\mathcal{A}}$ acts on an affine chart S. The operator $\mathcal{A}$ acts as an affine map from the hyperplane S to the hyperplane $\mathcal{A}S$. The image of a point $\widehat{\mathcal{A}}\widehat{x} = \widehat{\mathcal{A}x}$, $x \in S$, on the chart S is the central projection (with the center at the origin) of the point $\mathcal{A}x \in \mathcal{A}S$ onto S (Figure 7.20). Thus, we can say that from the affine chart's point of view, a projective transformation is a composition of an affine map and a central projection.

In coordinate form, the above discussion looks as follows. Let $A = (a_{ij})_{i,j=0}^{n}$ be the matrix of $\mathcal{A}$ in the basis $\{e_0, e_1, \dots, e_n\}$. Consider nonhomogeneous coordinates of the space PV determined by the frame $\{e_0; e_1, \dots, e_n\}$ of the affine chart S_0 (see (7.38)). Choose

$$x = e_0 + x_1 e_1 + \cdots + x_n e_n$$

so that the point $\widehat{x} \in PV$ have nonhomogeneous coordinates $x_1, \dots, x_n$. Let $y_1, \dots, y_n$ be the nonhomogeneous coordinates of its image. Then

$$y_i = \frac{a_{i0} + \sum_{j=1}^{n} a_{ij} x_j}{a_{00} + \sum_{j=1}^{n} a_{0j} x_j}, \qquad i = 1, \dots, n. \tag{7.41}$$

For example, a projective transformation of a line is a linear fractional transformation

$$y = \frac{ax + b}{cx + d}, \qquad ad - bc \neq 0. \tag{7.42}$$

(When $c \neq 0$, the point $-d/c$ is mapped to the point at infinity and the point at infinity is mapped to the point a/c.)

If $\mathcal{A}S = S$, the transformation $\widehat{\mathcal{A}}$ acts on a chart S as an affine transformation. The following lemma demonstrates that every affine transformation of the space S is obtained in this way.

Lemma 7.105. *Every affine transformation of a hyperplane $S \subset V$ that does not pass through the origin extends uniquely to a linear transformation of V.*

Proof. A frame $\{e_0; e_1, \dots, e_n\}$ of the hyperplane S is also a basis of V (see Figure 7.19). The extension of an affine transformation f of S is a linear transformation of V mapping the basis $\{e_0, e_1, \dots, e_n\}$ to the basis $\{f(e_0), df(e_1), \dots, df(e_n)\}$. □

Considering the affine space S as a piece of the projective space PV, we can say that the group $\mathrm{GA}(S)$ is a subgroup of $\mathrm{PGL}(V)$.

Exercise 7.106. Prove that for every projective transformation of a complex projective space, there exists an affine chart where it acts as an affine transformation.

Geometry defined by the group of projective transformations is called *projective geometry*. In comparison with Theorem 7.61, the following theorem shows how rich is the group of projective transformations relative to the group of affine transformations.

A system of $n+2$ points of an n-dimensional projective plane is called a *system of points in general position* if no $n+1$ of them lie on the same hyperplane.

Theorem 7.107. *Let $\{p_0, p_1, \dots, p_{n+1}\}$ and $\{q_0, q_1, \dots, q_{n+1}\}$ be two systems of points in general position of an n-dimensional projective space PV. Then there exists a unique projective transformation that maps p_i to q_i for $i = 0, 1, \dots, n+1$.*

Proof. Let $p_i = \widehat{e_i}$, $q_i = \widehat{f_i}$, where e_i, f_i, $i = 0, 1, \dots, n+1$, are nonzero vectors of the space V. The condition of the theorem means that $\{e_0, e_1, \dots, e_n\}$ (respectively, $\{f_0, f_1, \dots, f_n\}$) is a basis of V and all coordinates of the vector e_{n+1} (respectively, f_{n+1}) in this basis are nonzero. By normalizing vectors $e_0, e_1, \dots, e_n$ (respectively, $f_0, f_1, \dots, f_n$) in a particular way, we can obtain $e_{n+1} = e_0 + e_1 + \dots + e_n$ (respectively, $f_{n+1} = f_0 + f_1 + \dots + f_n$). With these conditions, let $\mathcal{A}$ be the linear operator that maps the basis $\{e_0, e_1, \dots, e_n\}$ to $\{f_0, f_1, \dots, f_n\}$. Then $\mathcal{A}e_{n+1} = f_{n+1}$ and $\widehat{\mathcal{A}}$ is the unique projective transformation that satisfies the condition of the theorem. □

In particular, any three distinct points on a projective line can be mapped to any other three distinct points by a projective transformation. For this reason, not only projective geometry does not have the notion of the distance between two points but it also lacks the notion of the ratio of a triple of points on a line, which affine geometry has. However, there exists an invariant of a quadruple of points on a line.

Namely, let p_1, p_2, p_3, p_4 be points on a line $PU \subset PV$. Choose a basis $\{e_1, e_2\}$ in the space U and, for any two vectors $u, v \in U$, denote by $\det(u, v)$ the determinant of the matrix formed by coordinates of these vectors in this basis. Let $p_i = \widehat{u_i}$, $i = 1, 2, 3, 4$. It is easy to see that the expression

$$(p_1, p_2; p_3, p_4) = \frac{\det(u_1, u_3)}{\det(u_3, u_2)} : \frac{\det(u_1, u_4)}{\det(u_4, u_2)} \tag{7.43}$$

depends neither on the normalization of vectors u_i nor on the choice of $\{e_1, e_2\}$ in U. It is called the *cross-ratio* of the four points p_1, p_2, p_3, p_4.

Let L be an affine chart of the line PU. Choose the basis $\{e_1, e_2\}$ so that $L = e_2 + \langle e_1 \rangle$ and let $u_i = e_2 + x_i e_1$. Then x_i is the nonhomogeneous coordinate of the point p_i on the chart L (Figure 7.21) and

$$\det(u_i, u_j) = \begin{vmatrix} x_i & 1 \\ x_j & 1 \end{vmatrix} = x_i - x_j.$$

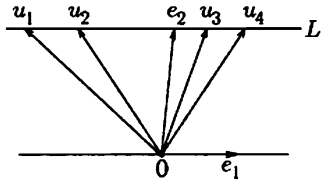


Figure 7.21

Therefore,

$$(p_1, p_2; p_3, p_4) = \frac{x_1 - x_3}{x_3 - x_2} : \frac{x_1 - x_4}{x_4 - x_2}. \tag{7.44}$$

Let us emphasize that since the cross-ratio can be defined by formula (7.43), expression (7.44) depends neither on the choice of the affine chart nor on the choice of coordinate on it.

Remark 7.108. The cross-ratio is assumed to be determined if no three of the points p_1, p_2, p_3, p_4 are the same. Under this condition, if $p_2 = p_3$ or $p_1 = p_4$, the value of the cross-ratio is taken to be ∞.

Exercise 7.109. Determine what happens to the cross-ratio $(p_1, p_2; p_3, p_4) = \delta$ when points p_1, p_2, p_3, p_4 are permuted. Prove that the expression

$$\frac{(\delta^2 - \delta + 1)^3}{\delta^2(\delta - 1)^2}$$

does not change under any permutation.

Exercise 7.110. Study the images of four square flower beds along the central alley in Figure 7.17 and show that the engraver seriously distorted the perspective. (*Hint*: compare the cross-ratio of the three equidistant points of the central alley determined by the flower beds and the alley's point at infinity to the cross-ratio of their images on the engraving.)

Since the cross-ratio was defined in terms invariant under any linear transformation of V, it is preserved by every projective transformation.

We now turn to the projective theory of quadrics. As we will see, it is much simpler than its affine counterpart. This is one of the manifestations of projective geometry's perfection that fascinated 19th century mathematicians so much—they even believed that all geometries must be deduced from it.

We call a subset of a vector space V a *cone* if it is invariant under multiplication by numbers, i.e., if together with a given vector, it contains all vectors proportional to it. (In the sense of the definition in Section 7.4, this is the same as a cone with vertex at the origin.) In particular, a quadric

$X \subset V$ is a cone in the above sense if and only if $X = X(Q)$, where Q is a quadratic function on the space V. Such quadrics are called *quadratic cones* (here we deviate slightly from the terminology in Section 7.4).

For any cone $X \subset V$, call the subset PX of the space PV formed by all one-dimensional spaces in X the *projectivization* of X. Clearly, the image of PX on an affine chart S is the intersection $X \cap S$.

Definition 7.111. A *quadric* in the space PV is a projectivization of a quadratic cone from the space V.

In other words, this is a subset of the form $PX(Q)$, where Q is a quadratic function on the space V, as long as this subset is not empty and is not a plane. The image of a projective quadric on an affine chart, as long as it is not empty and is not a plane, is an affine quadric; however, the type of the quadric depends on the chart. (Recall that light spot from a lamp with a shade.)

A projective quadric $PX(Q)$ is called *nondegenerate* if the quadratic function Q is nondegenerate.

Remark 7.112. Using ideas from the proof of Theorem 7.84, it is not difficult to show that whenever the field K contains more than five elements, the intersection $X(Q) \cap S$ is never empty. Also, $X(Q) \cap S$ can become a hyperplane if and only if Q is the product of two linear functions (and hence, $PX(Q)$ is the union of two hyperplanes).

In homogeneous coordinates, the equation of a projective quadric $PX(Q)$ is

$$Q(x_0, x_1, \dots, x_n) = \sum_{i,j=0}^{n} a_{ij} x_i x_j = 0, \qquad a_{ij} = a_{ji}. \tag{7.45}$$

On the affine chart S_0, it is described in affine coordinates as

$$Q(1, x_1, \dots, x_n) = 0 \tag{7.46}$$

and its intersection with the hyperplane at infinity (with respect to S_0) is described by the equation

$$Q(0, x_1, \dots, x_n) = 0 \tag{7.47}$$

in homogeneous coordinates.

Observe that every quadric X on the chart S_0 (hence, on any affine chart) depicts a projective quadric $\overline{X}$. The equation of $\overline{X}$ in homogeneous coordinates is obtained from the equation of X by entering x_0 into every linear term and x_0^2 into the free term. When applicable, Theorem 7.84 implies that the quadric $\overline{X}$ is uniquely determined by the quadric X.

Example 7.113. Consider the conic $C \subset \mathbb{R}P^2$ defined by the following equation in homogeneous coordinates:

$$x_0^2 - x_1^2 - x_2^2 = 0.$$

Its image on the affine chart S_0 is the ellipse

$$x_1^2 + x_2^2 = 1$$

and there are no points at infinity on C with respect to S_0. On the affine chart $x_0 - x_2 = 1$, the same conic is depicted as the parabola

$$y = x_1^2,$$

where $y = x_0 + x_2$. In this case, there exists one point at infinity: $(1 : 0 : 1)$. Finally, on the affine chart S_2, the conic C is depicted as the hyperbola

$$x_0^2 - x_1^2 = 1$$

and there are two points at infinity, $(1 : 1 : 0)$ and $(1 : (-1) : 0)$.

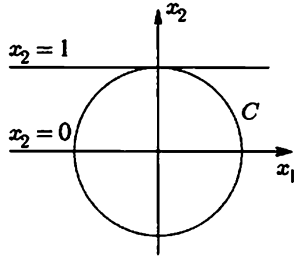


Figure 7.22

All this is well observed on the chart S_0, where the image of the line $x_0 - x_2 = 0$, which is the line at infinity with respect to the chart $x_0 - x_2 = 1$, has the equation $x_2 = 1$ and touches the image of the conic at one point. In turn, the image of the line $x_2 = 0$, which is the line at infinity with respect to S_2, intersects the image of the conic at two points (Figure 7.22). Therefore, we can claim that a parabola is tangent to the line at infinity and a hyperbola intersects it at two points. It is not difficult to see that the point at infinity that lies on the parabola corresponds to its special direction (see Section 7.4) and the points at infinity that lie on the hyperbola correspond to its asymptotes.

Exercise 7.114. Prove that every paraboloid in a real affine space is tangent to the hyperplane at infinity.

If a quadratic function Q is degenerate and the one-dimensional space $\langle x_0 \rangle$ lies in its kernel, then for any one-dimensional subspace $\langle x \rangle$ contained in a cone $X(Q)$, this cone also contains the two-dimensional space $\langle x, x_0 \rangle$. This implies that for every point $\widehat{x} \neq \widehat{x_0}$, the quadric $PX(Q)$ also contains the line $\widehat{x}\widehat{x_0}$, i.e., that it is a cone with the vertex $\widehat{x_0}$.

Its image on an affine chart is either a cone or a cylinder depending on whether the point $\widehat{x_0}$ lies in this chart or not. (So, the difference between cones and cylinders disappears in projective geometry.)

In the cases $K = \mathbb{C}$ or $\mathbb{R}$, we can choose a basis of V such that the quadratic function Q reduces to its normal form. It follows that the equation of a nondegenerate quadric in a complex projective space always reduces to the form

$$x_0^2 + x_1^2 + \cdots + x_n^2 = 0. \tag{7.48}$$

In a real projective space, it always reduces to the form

$$x_0^2 + x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2 = 0, \qquad \frac{n-1}{2} \leq k < n. \tag{7.49}$$

(The inequality $k \geq \frac{n-1}{2}$ appears because we can always multiply the equation by -1.)

We see that all nondegenerate complex quadrics are projectively equivalent and that the set of nondegenerate real quadrics splits into $\left[\frac{n-1}{2}\right] + 1$ classes with respect to projective equivalence.

That quadrics determined by equation (7.49) are not projectively equivalent for different k follows from Theorem 7.84 and the Law of Inertia. However, the differences also show up in the geometric structure of such quadrics. The following theorem describes one such difference.

Theorem 7.115. *The maximum dimension of a plane contained in a real quadric with equation* (7.49) *equals* $n - k - 1$.

Proof. Obviously, a k-dimensional plane Π_0 determined by equations

$$x_{k+1} = \cdots = x_n = 0$$

does not intersect quadric (7.49). Since each plane of dimension $\geq n - k$ intersects Π_0, none of them can lie in the quadric.

On the other hand, by changing the basis, we can rewrite equation (7.49) as

$$y_0 y_{k+1} + y_1 y_{k+2} + \cdots + y_{n-k-1} y_n + y_{n-k}^2 + \cdots + y_k^2 = 0.$$

This shows that the $(n - k - 1)$-dimensional plane

$$y_0 = y_1 = \cdots = y_k = 0$$

lies in the quadric. □

In particular, quadric (7.49) does not contain straight lines if and only if $k = n - 1$. Such a quadric is called an *oval*. One of its affine images is an ellipsoid. When $k < n - 1$, the quadric is called *ruled*.

Table 2

n	k	name	affine image	part at infinity
2	1	conic	ellipse	$\varnothing$
			parabola	point
			hyperbola	two points
3	2	oval quadric	ellipsoid	$\varnothing$
			elliptic paraboloid	point
			hyperboloid of two sheets	conic
	1	ruled quadric	hyperboloid of one sheet	conic
			hyperbolic paraboloid	two lines

We list nondegenerate quadrics in $\mathbb{R}P^2$ and $\mathbb{R}P^3$ together with their affine images in Table 2. In each case we also describe the part at infinity with respect to the given affine chart.

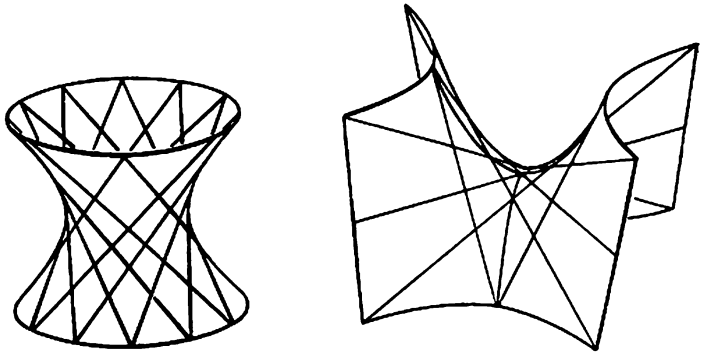


Figure 7.23

Observe that a ruled quadric in $\mathbb{R}P^3$ is "woven" from two families of lines; see Figure 7.23 for its affine image.

The following theorem is closely related to Theorem 7.90 and is also its projective analogue.

Theorem 7.116. *For any nondegenerate real projective quadric PX, the group $G(PX)$ of X-preserving projective transformations acts transitively on X.*

Proof. Nondegeneracy implies that the zero vector is the only vertex of the cone X. Thus, applied to X, Theorem 7.90 implies that the group of X-preserving linear transformations acts transitively on the set of its nonzero vectors. The theorem follows by projectivization. □

Exercise 7.117. Prove that if PX is an oval quadric, the group $G(PX)$ acts transitively on the set of (ordered) triples of different points of PX.

In the case of an oval quadric PX, the group $G(PX)$ can serve as a foundation for the construction of *conformal* and *Lobachevsky* geometries. Namely, conformal geometry is realized on the quadric PX itself while Lobachevsky geometry is realized on its interior. In both cases, the geometry-defining group in the sense of Section 4.2 is the group $G(PX)$ (but it acts on different sets).

Exercise 7.118. Prove that the group $G(PX)$ acts transitively on the interior of an oval quadric PX (cf. Exercise 7.91).

Chapter 8

Tensor Algebra

Tensor algebra is rather a language than a substantial theory. Yet this language is very useful and, moreover, absolutely indispensable. In particular, it allows us to give a uniform description of all objects of linear algebra and even arrange them in one algebraic structure.

8.1. Tensor Product of Vector Spaces

We begin by introducing a general concept that encompasses many of the notions we considered in previous chapters.

Let $V_1, \dots, V_p$ and U be vector spaces over a field K. The map

$$\varphi\colon V_1 \times \cdots \times V_p \longrightarrow U \tag{8.1}$$

is called *multilinear* (or, rather, *p-linear*) if it is linear in each of the p arguments when the other arguments are fixed. Such maps form a vector space, which is itself a subspace in the space of all maps from $V_1 \times \cdots \times V_p$ to U. We denote this subspace as $\operatorname{Hom}(V_1, \dots, V_p; U)$.

If the spaces $V_1, \dots, V_p$ and U are finite-dimensional, then so is the space $\operatorname{Hom}(V_1, \dots, V_p; U)$. More precisely,

$$\dim \operatorname{Hom}(V_1, \dots, V_p; U) = \dim V_1 \cdots \dim V_p \cdot \dim U,$$

since a multilinear map (8.1) is determined by the images of the basis vectors of the spaces $V_1, \dots, V_p$, which, in turn, are determined by their coordinates in a basis of U.

When $U = K$, we obtain the space $\operatorname{Hom}(V_1, \dots, V_p; K)$ of *multilinear functions* on $V_1 \times \cdots \times V_p$. In particular, $\operatorname{Hom}(V; K)$ is the dual space V^* of V.

The tensor products of vector spaces V and W arises naturally when we consider bilinear mappings $\varphi\colon V \times W \to U$. We will see that one of them is "universal;" in some sense, it describes all others. The corresponding space U is called the tensor product of V and W.

Proposition 8.1. *Let V and W be vector spaces with bases $\{e_i\colon i \in I\}$ and $\{f_j\colon j \in J\}$, respectively. The following properties of a bilinear map $\varphi\colon V \times W \to U$ are equivalent:*

(i) vectors $\varphi(e_i, f_j)$, $i \in I, j \in J$, form a basis of U;

(ii) every vector $z \in U$ decomposes uniquely as $z = \sum_i \varphi(e_i, y_i)$, $y_i \in W$;

(iii) every vector $z \in U$ decomposes uniquely as $z = \sum_j \varphi(x_j, f_j)$, $x_j \in V$.

(When these spaces are infinite-dimensional, it is assumed that the sums above are finite.)

Proof. If $z = \sum_{i,j} z_{ij}\varphi(e_i, f_j)$, then $z = \sum_i \varphi(e_i, y_i)$ for $y_i = \sum_j z_{ij} f_j$ and vice versa. Equivalence of properties (i) and (ii) follows. Similarly, we prove the equivalence of (i) and (iii). □

Corollary 8.2. *If property (i) holds for some bases of V and W, then it holds for any bases.*

Definition 8.3. A *tensor product* of vector spaces V and W is a vector space T with a bilinear map

$$\otimes\colon V \times W \to T, \qquad (x, y) \mapsto x \otimes y$$

that satisfies the following condition: if $\{e_i\colon i \in I\}$ and $\{f_j\colon j \in J\}$ are bases of V and W, respectively, then $\{e_i \otimes f_j\colon i \in I, j \in J\}$ is a basis of T.

The above discussion implies that this condition does not depend on the choice of bases of V and W.

Obviously, a tensor product exists for any two spaces V and W. Indeed, consider the vector space T with a basis $\{t_{ij}\colon i \in I, j \in J\}$ and define the bilinear map $\otimes\colon V \times W \to T$ so that $e_i \otimes f_j = t_{ij}$ for the basis vectors e_i, f_j.

A tensor product is unique in the following sense: if $(T_1, \otimes_1)$ and $(T_2, \otimes_2)$ are two tensor products of V and W, then there exists a (unique) isomorphism

$$\psi\colon T_1 \to T_2$$

such that

$$\psi(x \otimes_1 y) = x \otimes_2 y \tag{8.2}$$

for any $x \in V$, $y \in W$. Indeed, for basis vectors, the isomorphism in question can be constructed as

$$\psi(e_i \otimes_1 f_j) = e_i \otimes_2 f_j.$$

By linearity, (8.2) then holds for any $x \in V$, $y \in W$.

The tensor product of vector spaces V and W is denoted as $V \otimes W$ or, if we need to emphasize the base field, $V \otimes_K W$. It follows from the definition that in the finite-dimensional case,

$$\dim(V \otimes W) = \dim V \cdot \dim W. \tag{8.3}$$

Example 8.4. Consider the bilinear map

$$\otimes : K[x] \times K[y] \to K[x, y]$$

defined as

$$(f \otimes g)(x, y) = f(x)g(y).$$

The products $x^i \otimes y^j = x^i y^j$ $(i, j = 0, 1, 2, \dots)$ form a basis of $K[x, y]$, hence $K[x, y] = K[x] \otimes K[y]$. Similarly,

$$K[x_1, \dots, x_m, y_1, \dots, y_n] = K[x_1, \dots, x_m] \otimes K[y_1, \dots, y_n]. \tag{8.4}$$

In the following two examples, V and W are finite-dimensional vector spaces with bases $\{e_1, \dots, e_n\}$ and $\{f_1, \dots, f_m\}$. The dual bases of V^* and W^* are denoted $\{\varepsilon_1, \dots, \varepsilon_n\}$ and $\{\theta_1, \dots, \theta_m\}$, respectively.

Example 8.5. For any $\alpha \in V^*$ and $y \in W$, define the linear map $\alpha \otimes y$ from V to W as

$$(\alpha \otimes y)(x) = \alpha(x)y. \tag{8.5}$$

We have thus constructed a bilinear map

$$\otimes : V^* \times W \to \mathrm{Hom}(V; W).$$

It is easy to see that $\varepsilon_i \otimes f_j$ is a linear map determined by the matrix E_{ji}. Since these matrices comprise a basis of the space of all $m \times n$ matrices,

$$\mathrm{Hom}(V; W) = V^* \otimes W. \tag{8.6}$$

Example 8.6. For any $\alpha \in V^*$ and $\beta \in W^*$, define the bilinear function $\alpha \otimes \beta$ on $V \times W$ as

$$(\alpha \otimes \beta)(x, y) = \alpha(x)\beta(y). \tag{8.7}$$

We thus obtain a bilinear map

$$\otimes : V^* \otimes W^* \to \mathrm{Hom}(V, W; K).$$

Here $(\varepsilon_i \otimes \theta_j)(x, y) = x_i y_j$, where $x_1, \dots, x_n$ and $y_1, \dots y_m$ are coordinates of vectors x and y, respectively. Since every bilinear function γ on $V \times W$

decomposes uniquely as $\gamma(x,y) = \sum_{i,j} c_{ij}x_iy_j$, the functions $\varepsilon_i \otimes \theta_j$ form a basis of the space $\mathrm{Hom}(V,W;K)$. Therefore,

$$\mathrm{Hom}(V,W;K) = V^* \otimes W^*. \tag{8.8}$$

Universality of the tensor product that we mentioned in the beginning of this section can be expressed as follows:

Proposition 8.7. *For any bilinear map $\varphi\colon V \times W \to U$, there exists a unique linear map $\psi\colon V \otimes W \to U$ such that*

$$\varphi(x,y) = \psi(x \otimes y) \tag{8.9}$$

for any $x \in V$, $y \in W$.

Proof. On the basis vectors of $V \otimes W$, this linear map is determined as

$$\psi(e_i \otimes f_j) = \varphi(e_i, f_j).$$

□

Every element $z \in V \otimes W$ decomposes uniquely as

$$z = \sum_{i,j} z_{ij} e_i \otimes f_j, \qquad z_{ij} \in K. \tag{8.10}$$

The numbers z_{ij} are called *coordinates* of z with respect to the given bases of V and W. In particular, in the finite-dimensional case z is described by the $m \times n$ matrix (z_{ij}) (here $m = \dim V$, $n = \dim W$).

An element $z \in V \otimes W$ is called *decomposable* if it decomposes as

$$z = x \otimes y, \qquad x \in V,\ y \in W. \tag{8.11}$$

Clearly, if $x = \sum_i x_i e_i$, $y = \sum_j y_j f_j$, then $z_{ij} = x_i y_j$. In the finite-dimensional case this means that $\mathrm{rk}(z_{ij}) \le 1$. Thus decomposable elements comprise a very small part of the space $V \otimes W$ (unless V or W is one-dimensional); however, they span all of it.

Exercise 8.8. Prove that decomposition (8.11) of a nonzero decomposable element $z \in V \otimes W$ is unique up to replacements $x \mapsto \lambda x$, $y \mapsto \lambda^{-1}y$, $\lambda \in K^*$.

Proposition 8.1 suggests other—often useful—decompositions of an element of a tensor product. Namely, every element $z \in V \otimes W$ decomposes uniquely as

$$z = \sum_i e_i \otimes y_i, \qquad y_i \in W, \tag{8.12}$$

or as

$$z = \sum_j x_j \otimes f_j, \qquad x_j \in V. \tag{8.13}$$

Exercise 8.9. Prove that every element $z \in V \otimes W$ decomposes as

$$z = \sum_{k=1}^{r} v_k \otimes w_k, \tag{8.14}$$

where the vectors $v_1, \dots, v_r \in V$, as well as $w_1, \dots, w_r \in W$, are linearly independent. This decomposition is unique up to replacements

$$v_k \mapsto \sum_l a_{kl} v_l, \qquad w_k \mapsto \sum_l b_{kl} w_l,$$

where $A = (a_{kl})$ and $B = (b_{kl})$ are nonsingular square matrices of order r such that $A^\top B = E$. Here r is the rank of the coordinate matrix of z.

An important example of a tensor product is *base field extension*. We discussed its simplest case, the complexification of a real vector space, in Section 6.2.

Let V be a vector space over a field K. Let L be an extension of K, i.e., a field that contains K as a subfield. We can consider L as a vector space over K and thus form the tensor product

$$V(L) = L \otimes V.$$

By definition this is a vector space over K. However, we can make it into a vector space over L by defining multiplication by elements of L as

$$\lambda(\mu \otimes v) = \lambda\mu \otimes v, \qquad \lambda, \mu \in L,\ v \in V.$$

We can view the original space V as embedded into $V(L)$ by identifying a vector $v \in V$ with the vector $1 \otimes v \in V(L)$. With this identification, $\lambda \otimes v = \lambda v$. Consider the decomposition of elements of $V(L)$ in a basis of the second component in the tensor product. We obtain that every basis of V over K is also a basis of $V(L)$ over L. However, base field extension is useful precisely because there exist other bases of $V(L)$ where certain objects (e.g., linear operators) look simpler.

On the other hand, if $\{\theta_i : i \in I\}$ is a basis of L over K, every vector of $V(L)$ decomposes uniquely as $\sum_i \theta_i v_i$, where v_i, $i \in I$, are some vectors of V (such that only finitely many of them are nonzero). For instance, every vector of the complexification $V(\mathbb{C})$ of a real vector space V decomposes uniquely as $x + iy$ for $x, y \in V$.

In some sense, the operation of tensor product on vector spaces is commutative and associative. Namely, for any two vector spaces V and W, there exists an isomorphism

$$V \otimes W \xrightarrow{\sim} W \otimes V \tag{8.15}$$

that maps $x \otimes y$ ($x \in V$, $y \in W$) to $y \otimes x$. It is defined by the condition that it maps any basis vector $e_i \otimes f_j$ of $V \otimes W$ to the basis vector $f_j \otimes e_i$ of $W \otimes V$. Similarly, for any three space U, V, W there exists an isomorphism

$$(U \otimes V) \otimes W \xrightarrow{\sim} U \otimes (V \otimes W) \tag{8.16}$$

mapping $(x \otimes y) \otimes z$ ($x \in U$, $y \in V$, $z \in W$) to $x \otimes (y \otimes z)$.

By identifying spaces $(U \otimes V) \otimes W$ and $U \otimes (V \otimes W)$ via isomorphism (8.16), we can write tensor products of any (finite) number of vector spaces $V_1, V_2, \ldots, V_p$ without parentheses. Induction on p shows that tensor products of basis vectors of $V_1, \ldots, V_p$ form a basis of the space $V_1 \otimes \cdots \otimes V_p$. On the other hand, this property can be taken as the definition of $V_1 \otimes \cdots \otimes V_p$, i.e., one can define the tensor product of several vector spaces just as we did it for two of them (one should only replace a bilinear map with a p-linear one).

In view of Proposition 8.7, there exists an isomorphism

$$\operatorname{Hom}(V \otimes W; U) \xrightarrow{\sim} \operatorname{Hom}(V, W; U) \tag{8.17}$$

that sends a linear map $\psi\colon V \otimes W \to U$ to the bilinear map $\varphi\colon V \times W \to U$ determined by (8.9). In particular, for $U = K$ we obtain

$$(V \otimes W)^* \xrightarrow{\sim} \operatorname{Hom}(V, W; K). \tag{8.18}$$

We can obviously generalize Proposition 8.7 to the case of any (finite) number of vector spaces (instead of just two of them). Therefore, we obtain an isomorphism

$$\operatorname{Hom}(V_1 \otimes \cdots \otimes V_p; U) \xrightarrow{\sim} \operatorname{Hom}(V_1, \ldots, V_p; U) \tag{8.19}$$

that sends a linear map $\psi\colon V_1 \otimes \cdots \otimes V_p \to U$ to the p-linear map $\varphi\colon V_1 \times \cdots \times V_p \to U$ defined as

$$\varphi(x_1, \ldots, x_p) = \psi(x_1 \otimes \cdots \otimes x_p). \tag{8.20}$$

In particular, for $U = K$ we obtain an isomorphism

$$(V_1 \otimes \cdots \otimes V_p)^* \xrightarrow{\sim} \operatorname{Hom}(V_1, \ldots, V_p; K). \tag{8.21}$$

Elements of the type $x_1 \otimes \cdots \otimes x_p$ are called *decomposable elements* of the tensor product $V_1 \otimes \cdots \otimes V_p$. The existence of canonical isomorphism (8.19) is equivalent to the following *fundamental principle of tensor algebra*: for any p-linear map $\varphi\colon V_1 \times \cdots \times V_p \to U$, there exists a unique linear map $\psi\colon V_1 \otimes \cdots \otimes V_p \to U$ satisfying condition (8.20). This allows us to construct linear maps of a tensor product by defining them on indecomposable elements.

There exist other important canonical isomorphisms of tensor products of finite-dimensional spaces.

First of all, we can generalize Example 8.6 to any number of finite-dimensional vector spaces $V_1, \dots, V_p$. Let

$$(\alpha_1 \otimes \cdots \otimes \alpha_p)(x_1, \dots, x_p) = \alpha_1(x_1) \cdots \alpha_p(x_p) \tag{8.22}$$

for $\alpha_1 \in V_1^*, \dots, \alpha_p \in V_p^*$. Then,

$$\operatorname{Hom}(V_1, \dots, V_p; K) = V_1^* \otimes \cdots \otimes V_p^*. \tag{8.23}$$

Together with isomorphism (8.21), this establishes the following isomorphisms:

$$V_1^* \otimes \cdots \otimes V_p^* \xrightarrow{\sim} (V_1 \otimes \cdots \otimes V_p)^*. \tag{8.24}$$

Combining equality (8.6) with isomorphisms (8.19) and (8.24), we obtain the isomorphism

$$V_1^* \otimes \cdots \otimes V_p^* \otimes U \xrightarrow{\sim} \operatorname{Hom}(V_1, \dots, V_p; U) \tag{8.25}$$

for any finite-dimensional spaces $V_1, \dots, V_p$, and U.

In view of the canonical isomorphisms above, we can identify the respective spaces, i.e., assume that $V \otimes W = W \otimes V$, $(U \otimes V) \otimes W = U \otimes (V \otimes W)$, $V^* \otimes W^* \otimes U = \operatorname{Hom}(V, W; U)$ (for finite-dimensional spaces), etc.

For any linear operators $\mathcal{A} \in \mathrm{L}(V)$ and $\mathcal{B} \in \mathrm{L}(W)$, one constructs the linear operator $\mathcal{A} \otimes \mathcal{B} \in \mathrm{L}(V \otimes W)$ by defining it as

$$(\mathcal{A} \otimes \mathcal{B})(x \otimes y) = \mathcal{A}x \otimes \mathcal{B}y \tag{8.26}$$

on a decomposable element $x \otimes y$. The operator $\mathcal{A} \otimes \mathcal{B}$ is called the *tensor product* of operators $\mathcal{A}$ and $\mathcal{B}$.

Let $A = (a_{ij})$ be the matrix of the operator $\mathcal{A}$ in a basis $\{e_1, \dots, e_n\}$ of V and $B = (b_{kl})$, the matrix of the operator $\mathcal{B}$ in a basis $\{f_1, \dots, f_m\}$ of W. Then the matrix of the operator $\mathcal{A} \otimes \mathcal{B}$ in the basis $\{e_1 \otimes f_1,\ e_1 \otimes f_2, \dots, e_1 \otimes f_m,\ e_2 \otimes f_1,\ e_2 \otimes f_2, \dots, e_2 \otimes f_m, \dots, e_n \otimes f_1,\ e_n \otimes f_2, \dots, e_n \otimes f_m\}$ of the space $V \otimes W$ is

$$\begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \dots & \dots & \dots & \dots \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{pmatrix}. \tag{8.27}$$

It is called the *tensor product* of matrices A and B and is denoted $A \otimes B$.

It is easy to see that $\operatorname{tr} A \otimes B = \operatorname{tr} A \cdot \operatorname{tr} B$, hence

$$\operatorname{tr} \mathcal{A} \otimes \mathcal{B} = \operatorname{tr} \mathcal{A} \cdot \operatorname{tr} \mathcal{B}. \tag{8.28}$$

Exercise 8.10. Prove that

$$\det \mathcal{A} \otimes \mathcal{B} = (\det \mathcal{A})^m (\det \mathcal{B})^n. \tag{8.29}$$

Exercise 8.11. Assume that the characteristic polynomial of $\mathcal{A}$ has n roots $\lambda_1, \ldots, \lambda_n$ (counted with multiplicities). Assume as well that the characteristic polynomial of $\mathcal{B}$ has m roots $\mu_1, \ldots, \mu_m$. Prove that the characteristic polynomial of $\mathcal{A} \otimes \mathcal{B}$ has nm roots $\lambda_i \mu_j$, $i = 1, \ldots, n$; $j = 1, \ldots, m$. Deduce from this formulas (8.28) and (8.29).

Exercise 8.12. Prove that (for finite-dimensional spaces V and W) the space $\mathrm{L}(V \otimes W)$ is the tensor product of the spaces $\mathrm{L}(V)$ and $\mathrm{L}(W)$ with respect to the tensor product of linear operators defined above.

Similarly, one can define a tensor product of several linear operators.

8.2. Tensor Algebra of a Vector Space

In this section, V always stands for an n-dimensional vector space.

The space

$$T_q^p(V) = \underbrace{V \otimes \cdots \otimes V}_{p} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q}$$

is called the space of *tensors of type* (p, q) on V. (The space $T_0^0(V)$ is assumed to be equal to K.) Clearly, $\dim T_q^p(V) = n^{p+q}$. Also, $T_0^1(V) = V$, $T_1^0(V) = V^*$ and, more generally,

$$T_q^0(V) = \mathrm{Hom}(\underbrace{V, \ldots, V}_{q}; K), \tag{8.30}$$

$$T_q^1(V) = \mathrm{Hom}(\underbrace{V, \ldots, V}_{q}; V). \tag{8.31}$$

In particular, tensors of type $(0, 2)$ are bilinear functions; tensors of type $(1, 1)$, linear operators; tensors of type $(1, 2)$, bilinear operations (algebraic structures) on V.

Tensor multiplication determines a bilinear operation

$$\otimes\colon T_q^p(V) \times T_s^r(V) \to T_{q+s}^{p+r}(V),$$

such that

$$(x_1 \otimes \cdots \otimes x_p \otimes \alpha_1 \otimes \cdots \otimes \alpha_q) \otimes (x_{p+1} \otimes \cdots \otimes x_{p+r} \otimes \alpha_{q+1} \otimes \cdots \otimes \alpha_{q+s}) = x_1 \otimes \cdots \otimes x_{p+r} \otimes \alpha_1 \otimes \cdots \otimes \alpha_{q+s}.$$

Example 8.13. The space

$$T_2^2(V) = V \otimes V \otimes V^* \otimes V^* = (V \otimes V) \otimes (V \otimes V)^*$$

can be identified with the space $\mathrm{L}(V \otimes V)$. Then the tensor multiplication

$$T_1^1(V) \times T_1^1(V) \to T_2^2(V)$$

coincides with the tensor multiplication of linear operators in the sense of Section 8.1. Indeed, bilinearity of both multiplications implies that it suffices to check that they coincide for decomposable linear operators. Let $\mathcal{A} = u \otimes \alpha$, $\mathcal{B} = v \otimes \beta$, $u, v \in V$, $\alpha, \beta \in V^*$, and let $\mathcal{A} \otimes \mathcal{B}$ be the tensor product of the operators $\mathcal{A}$ and $\mathcal{B}$ in the sense of Section 8.1. Given that $(\alpha \otimes \beta)(x \otimes y) = \alpha(x)\beta(y)$ (see Example 8.6 and the definition of isomorphism (8.18)), we obtain that

$$\begin{aligned}(\mathcal{A} \otimes \mathcal{B})(x \otimes y) &= \mathcal{A}x \otimes \mathcal{B}y = \alpha(x)\beta(y)u \otimes v \\ &= ((\alpha \otimes \beta)(x \otimes y))u \otimes v = ((u \otimes v) \otimes (\alpha \otimes \beta))(x \otimes y).\end{aligned}$$

Thus,

$$\mathcal{A} \otimes \mathcal{B} = u \otimes v \otimes \alpha \otimes \beta,$$

as required.

Another important operation on tensors is *contraction*. This is a linear map

$$T_q^p(V) \longrightarrow T_{q-1}^{p-1}(V), \qquad p, q > 0,$$

defined as follows. Consider the map

$$\underbrace{V \times \cdots \times V}_{p} \times \underbrace{V^* \times \cdots \times V^*}_{q} \longrightarrow T_{q-1}^{p-1}(V),$$

$$(x_1, \ldots, x_p, \alpha_1, \ldots, \alpha_q) \mapsto \alpha_1(x_1)(x_2 \otimes \cdots \otimes x_p \otimes \alpha_2 \otimes \cdots \otimes \alpha_q).$$

It is clearly multilinear. Thus, there exists a linear map $T_q^p(V) \longrightarrow T_{q-1}^{p-1}(V)$ such that

$$x_1 \otimes \cdots \otimes x_p \otimes \alpha_1 \otimes \cdots \otimes \alpha_q \mapsto \alpha_1(x_1)(x_2 \otimes \cdots \otimes x_p \otimes \alpha_2 \otimes \cdots \otimes \alpha_q).$$

This is exactly the contraction.

In this definition we contract the first components in the products $\underbrace{V \otimes \cdots \otimes V}_{p}$ and $\underbrace{V^* \otimes \cdots \otimes V^*}_{q}$ whose tensor product is $T_q^p(V)$. One can similarly define a contraction in any pair of components.

Example 8.14. The contraction of a linear operator (i.e., a $(1,1)$-tensor) is its trace. Indeed, by linearity it suffices to verify this statement for decomposable operators, i.e., for operators of the form $x \otimes \alpha$, $x \in V, \alpha \in V^*$. Such an operator is zero on the $(n-1)$-dimensional subspace $\operatorname{Ker} \alpha$ and acts as the multiplication by $\alpha(x)$ on the quotient space $V/\operatorname{Ker} \alpha$. Hence, its trace is $\alpha(x)$, precisely its contraction.

Example 8.15. The contraction of the product of a linear operator $\mathcal{A}$ and a vector x in the second V-component and the first (and only) V^*-component is the vector $\mathcal{A}x$. Indeed, for a decomposable operator $\mathcal{A} = u \otimes \alpha$, the result of this contraction is $\alpha(x)u$, which is precisely $\mathcal{A}x$.

Example 8.16. The contraction of the product of linear operators $\mathcal{A}$ and $\mathcal{B}$ in the second V component and the first V^* component equals the usual product $\mathcal{AB}$ of these operators. Indeed, for decomposable operators $\mathcal{A} = u \otimes \alpha$ and $\mathcal{B} = v \otimes \beta$, the result of the contraction is the operator $\alpha(v)u \otimes \beta$ that maps a vector $x \in V$ to $\alpha(v)\beta(x)u$. On the other hand,

$$\mathcal{AB}x = \beta(x)\mathcal{A}v = \alpha(v)\beta(x)u.$$

Example 8.17. Formula (8.7) implies that the contraction of the tensor product of a bilinear function α and two vectors x and y in two pairs of components equals $\alpha(x, y)$ or $\alpha(y, x)$ (depending on how the contracted components are combined).

A contraction of the tensor product of tensors T and U is often called the contraction of T with U.

Let $\{e_i\}$ be a basis of the space V and $\{\varepsilon_j\}$, the dual basis of V^*. Then $\{e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \varepsilon_{j_1} \otimes \cdots \otimes \varepsilon_{j_q}\}$ is the basis of the space $T_q^p(V)$. Every tensor T of the type (p, q) can be expressed in this basis as

$$T = \sum_{i_1,\ldots,i_p,j_1,\ldots,j_q} T_{i_1\ldots i_p j_1\ldots j_q} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \varepsilon_{j_1} \otimes \cdots \otimes \varepsilon_{j_q}.$$

The coefficients $T_{i_1\ldots i_p j_1\ldots j_q}$ are called the *coordinates* of T in the basis $\{e_i\}$ of V.

Example 8.18. The coordinates of a linear operator regarded as a $(1,1)$-tensor are exactly the matrix entries of this operator. Indeed, if $\mathcal{A} = \sum_{i,j} \mathcal{A}_{ij} e_i \otimes \varepsilon_j$, then $\mathcal{A}e_j = \sum_i \mathcal{A}_{ij} e_i$.

Example 8.19. Similarly, coordinates of a bilinear function regarded as a $(0,2)$-tensor are the matrix entries of this function.

There exists another notation for tensors, originally introduced by Einstein. Here both subscripts and superscripts are used: basis vectors of V are indexed by the subscripts and those of V^*, by the superscripts. The corresponding indices of tensor coordinates are superscripts and subscripts, respectively. If the same index appears in a product twice, once as a subscript and once as a superscript (no other repetitions are allowed), it is assumed that we sum over this index. So in Einstein's notations, the above formula looks like

$$T = T^{i_1\ldots i_p}_{j_1\ldots j_q} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \epsilon^{j_1} \otimes \cdots \otimes \epsilon^{j_q}. \tag{8.32}$$

Example 8.20. Consider a linear operator $\mathcal{A}$ acting on a vector x. In Einstein's notations, coordinates of the image of x are

$$(\mathcal{A}x)^i = \mathcal{A}^i_j x^j.$$

Example 8.21. The product of linear operators $\mathcal{A}$ and $\mathcal{B}$ is given by the formula

$$(\mathcal{AB})^i_k = \mathcal{A}^i_j \mathcal{B}^j_k.$$

The coordinates of the tensor product $T \otimes U$ of tensors $T \in T^p_q(V)$ and $U \in T^r_s(V)$ are the products of coordinates of the factors:

$$(T \otimes U)^{i_1 \dots i_{p+r}}_{j_1 \dots j_{q+s}} = T^{i_1 \dots i_p}_{j_1 \dots j_q} U^{i_{p+1} \dots i_{p+r}}_{j_{q+1} \dots j_{q+s}}.$$

The coordinates of the contraction S of a tensor $T \in T^p_q(V)$ in the first pair of components (also called the contraction in the first pair of indices) are

$$S^{i_2 \dots i_p}_{j_2 \dots j_q} = T^{k i_2 \dots i_p}_{k j_2 \dots j_q}.$$

This follows from equality (8.32) because the contraction of the product $e_{i_1} \otimes \dots \otimes e_{i_r} \otimes \varepsilon^{j_1} \otimes \dots \otimes \varepsilon^{j_q}$ equals $\delta^{j_1}_{i_1} e_{i_2} \otimes \dots \otimes e_{i_r} \otimes \varepsilon^{j_2} \otimes \dots \otimes \varepsilon^{j_q}$ (here δ^j_i is the Kronecker symbol). Similarly, one finds the coordinates of a contraction of T in any pair of indices.

On a Euclidean vector space V, there exists a special tensor $g \in T^0_2(V)$ that determines the inner product. It is called the *metric tensor* of V. The contraction of the metric tensor with any tensor $T \in T^p_q(V)$ in any index of g and the first superscript of T is the tensor $\widetilde{T} \in T^{p-1}_{q+1}(V)$ with coordinates

$$\widetilde{T}^{i_2 \dots i_p}_{j j_1 \dots j_q} = g_{jk} T^{k i_2 \dots i_p}_{j_1 \dots j_q}.$$

The transition from the tensor T to the tensor $\widetilde{T}$ is called the *lowering of* the first superscript of the tensor T. We define the lowering of any other superscript similarly.

In an orthonormal basis of V, $g_{jk} = \delta_{jk}$, hence

$$\widetilde{T}^{i_2 \dots i_p}_{j j_1 \dots j_q} = T^{j i_2 \dots i_p}_{j_1 \dots j_q}.$$

This implies, first of all, that the operation of lowering a superscript is invertible. Its inverse is called the *raising of a subscript*. Second, it follows that if we restrict ourselves to orthonormal bases, there is no difference between subscripts and superscripts of tensors in a Euclidean space.

Example 8.22. When lowering the index of a vector $u \in V$, we obtain a linear function

$$\alpha(x) = g_{jk} x^j u^k = (x, u).$$

We thus establish again the canonical isomorphism between the Euclidean space V and its dual space V^*.

Example 8.23. When lowering the index of a linear operator $\mathcal{A}$, we obtain a bilinear function

$$\alpha(x, y) = g_{jk} x^j \mathcal{A}^k_l y^l = (x, \mathcal{A}y).$$

This establishes the canonical isomorphism between the space of linear operators and the space of bilinear functions on a Euclidean space (we already described this isomorphism in Section 6.3).

Tensors of type $(p,0)$ are called *contravariant tensors* of rank p. We denote

$$T^p(V) = T_0^p(V).$$

The spaces $T^0(V) = K$, $T^1(V) = V$, $T^2(V), \ldots$ can be arranged into an algebra. In order to do this, we need the notion of an external direct sum of vector spaces.

We already encountered the decomposition of a vector space into a direct sum of subspaces in Section 5.1. The corresponding definition can be stated as follows:

Definition 8.24. A vector space V decomposes into a *direct sum* of subspaces $V_1, \ldots, V_k$ if every element $x \in V$ decomposes uniquely as $x = x_1 + \cdots + x_k$ for $x_i \in V_i$. This is written as

$$V = V_1 \oplus \cdots \oplus V_k.$$

For subspaces V_1, V_2, the uniqueness of the decomposition of $x \in V$ as $x = x_1 + x_2$, $x_1 \in V_1$, $x_2 \in V_2$, is equivalent to $V_1 \cap V_2 = 0$.

There exists another approach to the concept of a direct sum where we do not assume beforehand that spaces $V_1, \ldots, V_k$ are embedded into some common space.

Definition 8.25. The *direct sum* of vector spaces $V_1, \ldots, V_k$ is the vector space $V_1 \oplus \cdots \oplus V_k$ formed by all sequences $(x_1, \ldots, x_k)$, where $x_i \in V_i$, with componentwise operations of addition and multiplication by elements of the base field.

So, operations on $V_1 \oplus \cdots \oplus V_k$ are determined as follows:

$$(x_1, \ldots, x_k) + (y_1, \ldots, y_k) = (x_1 + y_1, \ldots, x_k + y_k),$$
$$\lambda(x_1, \ldots, x_k) = (x_1, \ldots, x_k).$$

That we indeed get a vector space in this way is obvious. In particular, its zero is the sequence $(0, \ldots, 0)$.

A direct sum in the sense of Definition 8.24 is called *internal* and in the sense of Definition 8.25, *external*. However, these two notions are closely related.

Namely, consider sequences of the form $(0, \ldots, x, \ldots, 0)$, where $x \in V_i$ is in the ith place. Operations on such sequences reduce to respective operations on the ith component. By identifying an element $x \in V$ with such a sequence, we obtain an embedding of the space V_i as a subspace of the space

$V_1 \oplus \cdots \oplus V_k$. Moreover, every element of $V_1 \oplus \cdots \oplus V_k$ decomposes uniquely into a sum of elements of these subspaces. Hence, the space $V_1 \oplus \cdots \oplus V_k$ is the direct sum of its subspaces $V_1, \ldots, V_k$. With this identification, an element $(x_1, \ldots, x_k)$ of the external direct sum $V_1 \oplus \cdots \oplus V_k$ is commonly written as $x_1 + \cdots + x_k$.

Conversely, let a vector space V decompose into a direct sum of its subspaces $V_1, \ldots, V_k$. Form the external direct sum $V_1 \oplus \cdots \oplus V_k$. Then the map

$$V_1 \oplus \cdots \oplus V_k \to V, \quad (x_1, \ldots, x_k) \mapsto x_1 + \cdots + x_k,$$

is an isomorphism of vector spaces.

The above discussion can be generalized to the case of an infinite number of components $V_1, V_2, \ldots$ as long as we consider only *finitary* sequences $(x_1, x_2, \ldots)$, $x_i \in V_i$, i.e., sequences with only a finite number of nonzero terms.

Now we can describe the construction of the tensor algebra. Consider the following infinite direct sum:

$$T(V) = \bigoplus_{p=0}^{\infty} T^p(V). \tag{8.33}$$

Since

$$T^p(V) \otimes T^q(V) \subset T^{p+q}(V),$$

the tensor product defines the structure of a graded algebra on $T(V)$. This algebra is called the *tensor algebra* of V. Note that it is associative (but not commutative) and has a unity, which is the unity of the field $K = T^0(V)$.

Similarly, tensors of type $(0, p)$ are called *covariant tensors* of rank p. Denote $T_p(V) = T_p^0(V)$. The algebra

$$T_*(V) = \bigoplus_{p=0}^{\infty} T_p(V)$$

is called the *algebra of multilinear functions* on V. The tensor product of multilinear functions has a simple interpretation. Namely, the values of a $(p+q)$-linear function $\alpha \otimes \beta$, $\alpha \in T_p(V)$, $\beta \in T_q(V)$, are determined as follows:

$$(\alpha \otimes \beta)(x_1, \ldots, x_{p+q}) = \alpha(x_1, \ldots, x_p)\beta(x_{p+1}, \ldots, x_{p+q}). \tag{8.34}$$

Indeed, by linearity, it suffices to verify that this formula holds for $\alpha = \alpha_1 \otimes \cdots \otimes \alpha_p$ $(\alpha_1, \ldots, \alpha_p \in V^*)$ and $\beta = \beta_1 \otimes \cdots \otimes \beta_q$ $(\beta_1, \ldots, \beta_q \in V^*)$ but then it easily follows from (8.22).

On the other hand, since

$$T_p(V) = T^p(V^*),$$

the algebra of covariant tensors can be regarded as the tensor algebra of the space V^*.

By the fundamental principle of tensor algebra (see Section 8.1), every p-linear map

$$\varphi\colon \underbrace{V \times \cdots \times V}_{p} \to U \tag{8.35}$$

"passes through" $T^p(V)$ in the sense that there exists a (unique) linear map

$$\psi\colon T^p(V) \to U \tag{8.36}$$

such that

$$\varphi(x_1, \ldots, x_p) = \psi(x_1 \otimes \cdots \otimes x_p) \tag{8.37}$$

for any $x_1, \ldots, x_p \in V$. When $U = K$, this establishes the isomorphism

$$T_p(V) \xrightarrow{\sim} (T^p(V))^* \tag{8.38}$$

(a special case of isomorphism (8.21)).

If we consider only symmetric or skew-symmetric multilinear maps, we arrive at the notions of the symmetric or the exterior power of V, respectively. These are discussed in the next two sections.

8.3. Symmetric Algebra

Definition 8.26. A multilinear map (8.35) is *symmetric* if

$$\varphi(x_{i_1}, \ldots, x_{i_p}) = \varphi(x_1, \ldots, x_p)$$

for any permutation $(i_1, \ldots, i_p)$ of indices $1, \ldots, p$.

Clearly, it suffices to consider only permutations of two indices.

When $U = K$, this definition turns into that of a *symmetric multilinear function.*

Let $\{e_1, \ldots, e_n\}$ be a basis of a space V.

Definition 8.27. A pth symmetric power of V is a vector space S together with a symmetric p-linear map

$$\underbrace{V \times \cdots \times V}_{p} \to S, \quad (x_1, \ldots, x_p) \mapsto x_1 \vee \cdots \vee x_p \tag{8.39}$$

such that the vectors $e_{i_1} \vee \cdots \vee e_{i_p}$, $i_1 \le \cdots \le i_p$, form a basis of S.

Notice that the expression $x_1 \vee \cdots \vee x_p$ in (8.39) is one (inseparable) symbol that denotes the image of the element $(x_1, \ldots, x_p)$.

This definition does not depend on the choice of a basis of V. Indeed, if $\{e'_1, \ldots, e'_n\}$ is another basis, the vectors $e'_{j_1} \vee \cdots \vee e'_{j_p}$, $j_1 \le \cdots \le j_p$, also

form a basis of the space S: the number of these vectors is the same as the number of the vectors $e_{i_1} \vee \cdots \vee e_{i_p}$, $i_1 \le \cdots \le i_p$, and the latter can be expressed as the linear combinations of the former.

A symmetric power exists: it suffices to consider a vector space S with the basis $\{s_{i_1 \dots i_p} : i_1 \le \cdots \le i_p\}$ and define the p-linear map (8.39) on the basis vectors of V as $e_{i_1} \vee \cdots \vee e_{i_p} = s_{j_1 \dots j_p}$, where the nondecreasing sequence $j_1, \dots, j_p$ is an arrangement of $i_1, \dots, i_p$.

The symmetric power is unique in the following sense: if $(S_1, \vee_1)$ and $(S_2, \vee_2)$ are two p-linear symmetric powers of V, then there exists a (unique) isomorphism $\psi : S_1 \to S_2$ such that

$$\psi(x_1 \vee_1 \cdots \vee_1 x_p) = x_1 \vee_2 \cdots \vee_2 x_p$$

for any $x_1, \dots, x_p \in V$. We construct this isomorphism by first defining it for the basis vectors:

$$\psi(e_{i_1} \vee_1 \cdots \vee_1 e_{i_p}) = e_{i_1} \vee_2 \cdots \vee_2 e_{i_p}, \qquad i_1 \le \cdots \le i_p.$$

The symmetric power of a space V is denoted $S^p(V)$.

The following proposition describes the universality property of the symmetric power, which is similar to that of the tensor product (see Proposition 8.7).

Proposition 8.28. *For any symmetric p-linear map* (8.35), *there exists a unique linear map $\psi : S^p(V) \to U$ such that*

$$\varphi(x_1, \dots, x_p) = \psi(x_1 \vee \cdots \vee x_p) \tag{8.40}$$

for any $x_1, \dots, x_p \in V$.

Proof. Such a linear map is determined on the basis vectors of $S^p(V)$ as

$$\psi(e_{i_1} \vee \cdots \vee e_{i_p}) = \varphi(e_{i_1}, \dots, e_{i_p}), \qquad i_1 \le \cdots \le i_p.$$

Because the map φ is symmetric, this formula holds for every $i_1, \dots, i_p$. Then (8.40) follows by linearity. □

Elements of the form $x_1 \vee \cdots \vee x_p$, $x_1, \dots, x_p \in V$, of the symmetric power $S^p(V)$ are called *decomposable*. Proposition 8.28 implies that in order to define a linear map of $S^p(V)$, it suffices to define it on decomposable elements so that it is multilinear and symmetric with respect to $x_1, \dots, x_p$.

In particular, there exists a bilinear map

$$\vee : S^p(V) \times S^q(V) \to S^{p+q}(V)$$

determined on decomposable elements as

$$(x_1 \vee \cdots \vee x_p) \vee (x_{p+1} \vee \cdots \vee x_{p+q}) = x_1 \vee \cdots \vee x_{p+q}. \tag{8.41}$$

Consider the direct sum

$$S(V) = \bigoplus_{p=0}^{\infty} S^p(V).$$

The operation $\vee$ defined above turns $S(V)$ into a graded algebra. It is called the *symmetric algebra* of the space V. Clearly, it is associative, commutative, and has a unity (which is the unity of the field $K = S^0V$). Definition (8.41) implies that every decomposable element $x_1 \vee \cdots \vee x_p \in S^p(V)$ coincides with the product of elements $x_1, \ldots, x_p$ in the algebra $S(V)$.

The symmetric algebra of a vector space is actually a familiar object. We recognize it as the polynomial algebra. Namely, identify each product $e_{i_1} \vee \cdots \vee e_{i_p}$, $i_1 \leq \cdots \leq i_p$, with the monomial $u_{i_1} \cdots u_{i_p}$ in variables $u_1, \ldots, u_n$; we thus obtain an isomorphism between the algebra $S(V)$ and the algebra $K[u_1, \ldots, u_n]$.

Moreover, if one regards $e_1, \ldots, e_n$ as coordinate functions on the dual space V^*, then every element of the algebra $S(V)$ determines a function on V^* (as a polynomial in $e_1, \ldots, e_n$). We can thus say that the algebra $S(V)$ is the algebra of polynomials on V^* (even though over a finite field, its elements cannot be identified with the functions they define). Likewise, the algebra $S(V^*)$ is the polynomial algebra on V.

If $\operatorname{char} K = 0$, the space $S^p(V)$ can be identified with the subspace of symmetric tensors in $T^p(V)$.

Namely, for every permutation $\sigma \in S_p$, define a linear map $T \mapsto T^\sigma$ of the space $T^p(V)$ on decomposable elements as follows:

$$(x_1 \otimes \cdots \otimes x_p)^\sigma = x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(p)}. \tag{8.42}$$

Observe that

$$\begin{aligned}((x_1 \otimes \cdots \otimes x_p)^\sigma)^\tau &= (x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(p)})^\tau \\ &= x_{\sigma\tau(1)} \otimes \cdots \otimes x_{\sigma\tau(p)} = (x_1 \otimes \cdots \otimes x_p)^{\sigma\tau},\end{aligned}$$

hence,

$$(T^\sigma)^\tau = T^{\sigma\tau} \tag{8.43}$$

for every tensor $T \in T^p(V)$.

A tensor $T \in T^p(V)$ is called *symmetric* if $T^\sigma = T$ for any permutation $\sigma \in S_p$. Symmetric tensors form a subspace of $T^p(V)$; denote it by $ST^p(V)$.

Assume that $\operatorname{char} K = 0$. Then we can define the operator Sym of *symmetrization* on the space $T^p(V)$ as

$$\operatorname{Sym} T = \frac{1}{p!} \sum_{\sigma \in S_p} T^\sigma. \tag{8.44}$$

Clearly, $\operatorname{Sym} T \in ST^p(V)$ for any T and $\operatorname{Sym} T = T$ whenever $T \in ST^p(V)$. This means that Sym is a projection onto $ST^p(V)$.

Proposition 8.29. *If* $\operatorname{char} K = 0$, *there exists an isomorphism* $\mu : S^p(V) \to ST^p(V)$ *such that*

$$\mu(x_1 \vee \cdots \vee x_p) = \operatorname{Sym}(x_1 \otimes \cdots \otimes x_p). \tag{8.45}$$

Proof. Since the right-hand side of (8.45) is multilinear and symmetric with respect to $x_1, \ldots, x_p$, there exists a linear map μ: $S^p(V) \to ST^p(V)$ satisfying condition (8.45). It maps a basis vector $e_{i_1} \vee \cdots \vee e_{i_p}$, $i_1 \leq \cdots \leq i_p$, of $S^p(V)$ to the tensor $\operatorname{Sym}(e_{i_1} \otimes \cdots \otimes e_{i_8})$, $i_1 \leq \cdots \leq i_p$.

Let us look at the decomposition of a symmetric tensor in basis vectors $e_{i_1} \otimes \cdots \otimes e_{i_p}$ of the space $T^p(V)$. We notice that the basis vectors with the same indices (but different order) have the same coefficients. Thus, the tensors $\operatorname{Sym}(e_{i_1} \otimes \cdots \otimes e_{i_p})$ with $i_1 \leq \cdots \leq i_p$ form a basis of the space $ST^p(V)$.

Hence, μ is an isomorphism. □

By means of this isomorphism, $S(V)$ can be identified with $ST^p(V)$.

Observe that the subspace

$$ST(V) = \bigoplus_{p=0}^{\infty} ST^p(V) \subset T(V)$$

is not a subalgebra of $T(V)$; however, its identification with $S(V)$ allows us to endow it with an algebra structure. Multiplication in this algebra is defined as

$$T \vee U = \operatorname{Sym}(T \otimes U). \tag{8.46}$$

Let us apply the above construction to the dual space V^*. We denote

$$S_p(V) = S^p(V^*), \quad ST_p(V) = ST^p(V^*).$$

The space $ST_p(V)$ is nothing but the space of symmetric p-linear functions on V. Symmetrization works as follows:

$$(\operatorname{Sym}\alpha)(x_1, \ldots, x_p) = \frac{1}{p!} \sum_{\sigma \in S_p} \alpha(x_{\sigma(1)}, \ldots, x_{\sigma(p)}). \tag{8.47}$$

To each symmetric p-linear function $\alpha \in ST_p(V)$, we associated a polynomial $f_\alpha \in S_p(V)$ defined as

$$f_\alpha(x) = \alpha(x, \ldots, x) \tag{8.48}$$

(just as in Section 5.3, where to every bilinear function we associated a quadratic function).

Proposition 8.30. *If* $\operatorname{char} K = 0$, *then the map*

$$ST_p(V) \to S_p(V), \quad \alpha \mapsto f_\alpha \tag{8.49}$$

is a vector space isomorphism whose inverse $\mu\colon S_p(V) \to ST_p(V)$ *was defined in Proposition 8.29.*

Proof. It suffices to consider symmetric p-linear functions of the form

$$\alpha = \operatorname{Sym}(\alpha_1 \otimes \cdots \otimes \alpha_p) = \mu(\alpha_1 \vee \cdots \vee \alpha_p),$$

where $\alpha_1, \ldots, \alpha_p \in V^*$. For every such function, we have

$$f_\alpha(x) = \alpha_1(x) \cdots \alpha_p(x) = (\alpha_1 \vee \cdots \vee \alpha_p)(x),$$

implying

$$f_\alpha = \alpha_1 \vee \cdots \vee \alpha_p = \mu^{-1}(\alpha),$$

as required. □

The symmetric multilinear function α is called the *polarization* of the polynomial f_α.

Example 8.31. The polarization of the polynomial

$$f(x) = x_1^3 + x_2^2 x_3$$

is the symmetric trilinear function

$$\alpha(x, y, z) = x_1 y_1 z_1 + \frac{1}{3}(x_3 y_2 z_2 + x_2 y_3 z_2 + x_2 y_2 z_3).$$

(Here x, y, z are the vectors in a three-dimensional space with coordinates x_i, y_i, z_i, $i = 1, 2, 3$.)

Remark 8.32. When the base field has a positive characteristic, map (8.49) is not generally an isomorphism. For instance, the bilinear function $\alpha(x, y) = x_1 y_2 + x_2 y_1$ corresponds to the zero quadratic function over a field of characteristic 2. Also, over such a field, the quadratic function $f(x) = x_1 x_2$ does not correspond to any symmetric bilinear function.

Remark 8.33. Formula (8.48) associates to any (not necessarily symmetric) p-linear function a homogeneous polynomial of degree p. However, the linear map $T_p(V) \to S_p(V)$ defined in this way is not an isomorphism for any $p > 1$.

Multiplication in the algebra

$$ST_*(V) = \bigoplus_{p=0}^{\infty} ST_p(V)$$

of symmetric multilinear functions corresponding to multiplication in the algebra

$$S_*(V) = \bigoplus_{p=0}^{\infty} S_p(V)$$

looks as follows:

$$(8.50)\quad (\alpha \vee \beta)(x_1, \dots, x_{p+q}) = \frac{p!q!}{(p+q)!} \sum_{(i_1,\dots,i_p | i_{p+1},\dots,i_{p+q})} \alpha(x_{i_1}, \dots, x_{i_p})\beta(x_{i_{p+1}}, \dots, x_{i_{p+q}}),$$

where we sum over all partitions $(i_1, \dots, i_p \mid i_{p+1}, \dots, i_{p+q})$ of the set $\{1, \dots, p+q\}$ into two groups of p and q elements, respectively (the order of indices within each group does not matter). This follows from formulas (8.46), (8.34), and (8.47) and symmetry the of functions α and β.

The symmetric product of p linear functions $\alpha_1, \dots, \alpha_p \in V^*$ is given by the formula

$$(8.51)\qquad (\alpha_1 \vee \dots \vee \alpha_p)(x_1, \dots, x_p) = \frac{1}{p!} \operatorname{per}(\alpha_i(x_j)),$$

where per A is the *permanent* of the square matrix A. The permanent is defined similarly to the determinant with the only difference that all terms, corresponding to even or odd permutations, are added with the plus sign.

Remark 8.34. For a field of positive characteristic, formula (8.50) does not make sense. The situation improves if we remove the coefficient before the sum. Such an operation remains associative and commutative but this algebra is no longer isomorphic to $S_*(V)$.

Just as we defined the tensor product of linear operators, it is possible to define the *symmetric power* $S^p\mathcal{A}$ of a linear operator $\mathcal{A}$. This is a linear operator on the space $S^p(V)$ that acts on decomposable elements as

$$(8.52)\qquad (S^p\mathcal{A})(x_1 \vee \dots \vee x_p) = \mathcal{A}x_1 \vee \dots \vee \mathcal{A}x_p.$$

If we identify the space $S^p(V)$ with the space $ST^p(V)$ of symmetric tensors (in the case $\operatorname{char} K = 0$), the operator $S^p\mathcal{A}$ becomes simply the restriction of the pth tensor power of $\mathcal{A}$ to the invariant subspace $ST^p(V) \subset T^p(V)$.

Example 8.35. In representation theory of groups (see Section 11.4), it is sometimes necessary to know the trace of the symmetric square $S^2\mathcal{A}$ of an operator $\mathcal{A}$. Let $\{e_1, \dots, e_n\}$ be a basis of V. Then vectors $e_i \vee e_j$, $i \le j$, comprise a basis of the space $S^2(V)$. We have (in Einstein's notation)

$$\begin{aligned}(S^2\mathcal{A})(e_i \vee e_j) &= \mathcal{A}e_i \vee \mathcal{A}e_j = \mathcal{A}_i^k e_k \vee \mathcal{A}_j^l e_l \\ &= \mathcal{A}_i^k \mathcal{A}_j^l e_k \vee e_l = \frac{1}{2}\left(\mathcal{A}_i^k \mathcal{A}_j^l + \mathcal{A}_i^l \mathcal{A}_j^k\right) e_k \vee e_l.\end{aligned}$$

Therefore,

$$\operatorname{tr} S^2\mathcal{A} = \frac{1}{2}\left(\mathcal{A}_i^i\mathcal{A}_j^j + \mathcal{A}_i^j\mathcal{A}_j^i\right) = \frac{1}{2}\left((\operatorname{tr}\mathcal{A})^2 + \operatorname{tr}\mathcal{A}^2\right). \tag{8.53}$$

Exercise 8.36. Assume that the characteristic polynomial of an operator $\mathcal{A}$ has n roots $\lambda_1, \dots, \lambda_n$ (counted with multiplicities). Prove that the characteristic polynomial of the operator $S^2\mathcal{A}$ has $n(n+1)/2$ roots $\lambda_i\lambda_j$, $1 \le i \le j \le n$. Deduce from this formula (8.53).

8.4. Grassmann Algebra

The Grassmann or the exterior algebra is constructed like the symmetric algebra but with skew-symmetry replacing symmetry. We will assume here that $\operatorname{char} K \neq 2$. (The case $\operatorname{char} K = 2$ can be included in the general scheme but then we ought to treat it carefully.)

Definition 8.37. A multilinear map (8.35) is *skew-symmetric* if

$$\varphi(x_{i_1}, \dots, x_{i_p}) = \operatorname{sign}(i_1, \dots, i_p)\varphi(x_1, \dots, x_p)$$

for any arrangement $(i_1, \dots, i_p)$ of indices $1, \dots, p$.

Clearly, it suffices to consider only permutations of two arguments (and require that after such a permutation, the image be multiplied by -1). It is also clear that if φ is a skew-symmetric p-linear map, then $\varphi(x_1, \dots, x_p) = 0$ whenever some of the vectors $x_1, \dots, x_p$ are the same.

When $U = K$, this definition becomes that of a *skew-symmetric multilinear function*.

Let $\{e_1, \dots, e_n\}$ be a basis of the space V.

Definition 8.38. A pth *exterior power* of V is a vector space Λ together with a p-linear map

$$\underbrace{V \times \cdots \times V}_{p} \to \Lambda, \quad (x_1, \dots, x_p) \mapsto x_1 \wedge \cdots \wedge x_p \tag{8.54}$$

such that the vectors $e_{i_1} \wedge \cdots \wedge e_{i_p}$, $i_1 < \cdots < i_p$, form a basis of Λ.

For the same reasons as in the definition of a symmetric power, this definition does not depend on the choice of a basis of V.

Just as a symmetric power, an exterior power exists and is unique. It is denoted $\Lambda^p V$.

Its definition implies that

$$\dim \Lambda^p(V) = \binom{n}{p} = \frac{n(n-1)\cdots(n-p+1)}{p!}.$$

In particular, $\Lambda^p(V) = 0$ for $p > n$.

The elements of the space $\Lambda^p(V)$ are called *multivectors* or, rather, *p-vectors*. In particular, 1-vectors are just vectors of V; 2-vectors are called *bivectors*; 3-vectors are called *trivectors*.

The following proposition explains the universality of the exterior product; its proof is similar to that of Proposition 8.28.

Proposition 8.39. *For any skew-symmetric p-linear map* (8.35), *there exists a unique liner map* $\psi\colon \Lambda^p(V) \to U$ *such that*

$$\varphi(x_1, \dots, x_p) = \psi(x_1 \wedge \dots \wedge x_p) \tag{8.55}$$

for any $x_1, \dots, x_p \in V$.

Multivectors of the form $x_1 \wedge \dots \wedge x_p$, $x_1, \dots, x_p \in V$, are called *decomposable*.

There exists a bilinear map

$$\wedge\colon \Lambda^p(V) \times \Lambda^q(V) \to \Lambda^{p+q}(V)$$

defined on decomposable multivectors as

$$(x_1 \wedge \dots \wedge x_p) \wedge (x_{p+1} \wedge \dots \wedge x_{p+q}) = x_1 \wedge \dots \wedge x_{p+q}. \tag{8.56}$$

Consider the direct sum

$$\Lambda(V) = \bigoplus_{p=0}^{\infty} \Lambda^p(V).$$

Operation $\wedge$ turns $\Lambda(V)$ into a graded algebra called the *Grassmann* or the *exterior algebra* of the space V. It is associative and has a unity but it is not commutative. However, it possesses a property that replaces commutativity, namely

$$u \wedge v = (-1)^{pq} v \wedge u \quad \text{for} \quad u \in \Lambda^p(V),\ v \in \Lambda^q(V).$$

Graded algebras satisfying this property are called *supercommutative*. (Supercommutativity lies at the foundation of the so-called supermathematics.)

Every decomposable multivector $x_1 \wedge \dots \wedge x_p \in \Lambda^p(V)$ coincides with the product of vectors $x_1, \dots, x_p$ in the algebra $\Lambda(V)$.

Unlike the symmetric algebra, the exterior algebra is finite-dimensional. More precisely, since its basis vectors $e_{i_1} \wedge \dots \wedge e_{i_p}$, $i_1 < \dots < i_p$, are in one-to-one correspondence with subsets of the set $\{1, \dots, n\}$,

$$\dim \Lambda(V) = 2^n.$$

The space $\Lambda^p(V)$ can be identified with the subspace of skew-symmetric tensors in $T^p(V)$.

Namely, a tensor $T \in T^p(V)$ is called *skew-symmetric* if $T^\sigma = (\operatorname{sign}\sigma)T$ for any permutation $\sigma \in S_p$. Skew-symmetric tensors form a subspace of $T^p(V)$; denote it $\Lambda T^p(V)$.

Assume that $\operatorname{char} K = 0$. Then we can define the operator Alt of *alternation* on the space $T^p(V)$:

$$\operatorname{Alt} T = \frac{1}{p!} \sum_{\sigma \in S_p} (\operatorname{sign} \sigma) T^\sigma. \tag{8.57}$$

This is a projection onto $\Lambda T^p(V)$.

Proposition 8.40. *For* $\operatorname{char} K = 0$, *there exists an isomorphism* $\mu\colon \Lambda^p(V) \to \Lambda T^p(V)$ *such that*

$$\mu(x_1 \wedge \cdots \wedge x_p) = \operatorname{Alt}(x_1 \otimes \cdots \otimes x_p). \tag{8.58}$$

Proof. The proof is similar to that of Proposition 8.29. We should only take into account that in the decomposition of a skew-symmetric tensor in the basis $e_{i_1} \otimes \cdots \otimes e_{i_p}$ of $T^p(V)$, a basis vector appears with a nonzero coefficient only if all of its indices are distinct. □

By means of this isomorphism, we can identify the space $\Lambda^p(V)$ with $\Lambda T^p(V)$.

Exercise 8.41. Prove that

$$T^2(V) = ST^2(V) \oplus \Lambda T^2(V),$$

but if $\dim V > 1$, then $T^p(V) \neq ST^p(V) + \Lambda T^p(V)$ for $p > 2$.

The subspace

$$\Lambda T(V) = \bigoplus_{p=0}^{n} \Lambda T^p(V) \subset T(V)$$

is not a subalgebra of $T(V)$, but after identifying it with $\Lambda(V)$, we can endow it with an algebra structure. In particular, its multiplication operation is described as

$$T \wedge U = \operatorname{Alt}(T \otimes U).$$

Now we apply the above discussion to the dual space. We denote

$$\Lambda_p(V) = \Lambda^p(V^*), \quad \Lambda T_p(V) = \Lambda T^p(V^*).$$

The subspace $\Lambda T_p(V)$ is nothing but the space of skew-symmetric p-linear functions on V. The operation of alternation is

$$(\operatorname{Alt} \alpha)(x_1, \ldots, x_p) = \frac{1}{p!} \sum_{\sigma \in S_p} (\operatorname{sign} \sigma) \alpha(x_{\sigma(1)}, \ldots, x_{\sigma(p)}). \tag{8.59}$$

Multiplication in the algebra

$$\Lambda T_*(V) = \bigoplus_{p=0}^{\infty} \Lambda T_p(V)$$

of skew-symmetric multilinear functions that arises from multiplication in the algebra

$$\Lambda_*(V) = \bigoplus_{p=0}^{\infty} \Lambda_p(V)$$

is given by the formula

$$(8.60)\quad (\alpha \wedge \beta)(x_1, \dots, x_{p+q}) = \frac{p!q!}{(p+q)!} \sum_{(i_1,\dots,i_p|i_{p+1},\dots,i_{p+q})} \operatorname{sign}(i_1, \dots, i_{p+q}) \alpha(x_{i_1}, \dots, x_{i_p})(x_{i_{p+1}}, \dots, x_{i_{p+q}}),$$

where, as in (8.50), we sum over all partitions $(i_1, \dots, i_p \mid i_{p+1}, \dots, i_{p+q})$ of the set $\{1, \dots, p+q\}$ into two subset of p and q elements, respectively. The product $\alpha \wedge \beta$ is called the *exterior product* of the functions α and β.

The exterior product of p linear functions $\alpha_1, \dots, \alpha_p \in V^*$ is given by the formula

$$(8.61)\qquad (\alpha_1 \wedge \cdots \wedge \alpha_p)(x_1, \dots, x_p) = \frac{1}{p!} \det(\alpha_i(x_j)).$$

Remark 8.42. In the case of a field of positive characteristic, formula (8.60) does not make sense. However, if we remove the coefficient before the sum, we would still obtain an algebra isomorphic to $\Lambda_*(V)$. Sometimes such a definition of the exterior product is assumed also for the case of zero characteristic.

Similarly to the symmetric power of a linear operator, one defines the *exterior power* $\Lambda^p \mathcal{A}$ of a linear operator $\mathcal{A}$.

Exercise 8.43. Prove that

$$(8.62)\qquad \operatorname{tr} \Lambda^2 \mathcal{A} = \frac{1}{2} \left((\operatorname{tr} \mathcal{A})^2 - \operatorname{tr} \mathcal{A}^2 \right).$$

Whereas the symmetric algebra only provides a different approach to the polynomial algebra, the Grassmann algebra is explicitly featured here for the first time in this course. However, we have already encountered it in disguise when we discussed determinants. Applications of the Grassmann algebra given below can be viewed as a development of the theory of determinants.

Consider an n-dimensional vector space V over a field K of characteristic $\neq 2$.

Theorem 8.44. *(i) A system of vectors $\{a_1, \dots, a_p\}$ of V is linearly dependent if and only if $a_1 \wedge \cdots \wedge a_p = 0$.*

(ii) Assume that in each of the systems $\{a_1, \dots, a_p\}$ and $\{b_1, \dots, b_p\}$ vectors are linearly independent. Then $\langle a_1, \dots, a_p \rangle = \langle b_1, \dots, b_p \rangle$ if and only if the p-vectors $a_1 \wedge \cdots \wedge a_p$ and $b_1 \wedge \cdots \wedge b_p$ are proportional.

Proof. (i) If vectors $a_1, \ldots, a_p$ are linearly dependent, then one of them can be expressed in terms of others. For instance, let

$$a_p = \sum_{i=1}^{p-1} \lambda_i a_i.$$

Then

$$a_1 \wedge \cdots \wedge a_{p-1} \wedge a_p = \sum_{i=1}^{p-1} \lambda_i a_1 \wedge \cdots \wedge a_{p-1} \wedge a_i = 0.$$

If vectors $a_1, \ldots, a_p$ are linearly independent, then they are part of some basis of V. Then, according to the definition of the exterior product, the p-vector $a_1 \wedge \cdots \wedge a_p$ is a basis vector of the space $\Lambda^p(V)$. Hence, it is nonzero.

(ii) If $\langle a_1, \ldots, a_p \rangle = \langle b_1, \ldots, b_p \rangle$, then the vectors $b_1, \ldots, b_p$ can be expressed in terms of vectors $a_1, \ldots, a_p$. Thus the p-vector $b_1 \wedge \cdots \wedge b_p$ can be expressed as a linear combination of p-vectors of the form $a_{i_1} \wedge \cdots \wedge a_{i_p}$. However,

$$a_{i_1} \wedge \cdots \wedge a_{i_p} = \begin{cases} \pm a_1 \wedge \cdots \wedge a_p & \text{for different } i_1, \ldots, i_p, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, $b_1 \wedge \cdots \wedge b_p = \lambda a_1 \wedge \cdots \wedge a_p$.

If $\langle a_1, \ldots, a_p \rangle \neq \langle b_1, \ldots, b_p \rangle$, there exists a basis $\{e_1, \ldots, e_n\}$ of V such that

$$\langle a_1, \ldots, a_p \rangle = \langle e_1, \ldots, e_p \rangle, \ \langle b_1, \ldots, b_p \rangle = \langle e_{d+1}, \ldots, e_{d+p} \rangle \quad (0 < d \leq p).$$

We know that the p-vector $a_1 \wedge \cdots \wedge a_p$ is proportional to $e_1 \wedge \cdots \wedge e_p$ and the p-vector $b_1 \wedge \cdots \wedge b_p$ is proportional to $e_{d+1} \wedge \cdots \wedge e_{d+p}$. But the p-vectors $e_1 \wedge \cdots \wedge e_p$ and $e_{d+1} \wedge \cdots \wedge e_{d+p}$ are not proportional to each other: according to the definition of the exterior power, they are distinct basis vectors of the space $\Lambda^p(V)$. Hence, the p-vectors $a_1 \wedge \cdots \wedge a_p$ and $b_1 \wedge \cdots \wedge b_p$ are not proportional either. □

The set of decomposable p-vectors is called the *Grassmann cone*. The projectivization of this cone is called the *Grassmann variety* and is denoted $\mathrm{Gr}_p(V)$. By Theorem 8.44, the points of $\mathrm{Gr}_p(V)$ are in one-to-one correspondence with the p-dimensional subspaces of V.

Fix a basis $\{e_1, \ldots, e_n\}$ of the space V and let $\{a_1, \ldots, a_p\}$ be a basis of a subspace U. Let us find the coordinates of the p-vector $a_1 \wedge \cdots \wedge a_p$ in the basis of the space $\Lambda^p(V)$ formed by the products $e_{i_1} \wedge \cdots \wedge e_{i_p}$, $i_1 < \cdots < i_p$.

Let $A = (a_{ij})$ be a $p \times n$ matrix of coordinates of vectors $a_1, \ldots, a_p$ in the basis $\{e_1, \ldots, e_n\}$, i.e.,

$$a_i = \sum_j a_{ij} e_j, \qquad i = 1, \ldots, p.$$

We have

$$a_1 \wedge \cdots \wedge a_p = \sum_{i_1, \ldots, i_p} a_{1i_1} \cdots a_{pi_p} e_{i_1} \wedge \cdots \wedge e_{i_p}.$$

If some of the indices $i_1, \ldots, i_p$ are the same, then $e_{i_1} \wedge \cdots \wedge e_{i_p} = 0$. If they are all different, we can permute factors in $e_{i_1} \wedge \cdots \wedge e_{i_p}$ so that their indices form an increasing sequence; in the process the whole product gets multiplied by $(-1)^s$, where s is the number of inversions in the sequence $(i_1, \ldots, i_p)$. It follows that

$$a_1 \wedge \cdots \wedge a_p = \sum_{i_1 < \cdots < i_p} M_{i_1 \ldots i_p} e_{i_1} \wedge \cdots \wedge e_{i_p}, \tag{8.63}$$

where $M_{i_1 \ldots i_p}$ is the minor of A of order p formed by the columns with indices $i_1, \ldots, i_p$.

By Theorem 8.44, the values of $M_{i_1 \ldots i_k}$ determine the subspace U uniquely. They are called the *Plücker coordinates* of U. These are actually the homogeneous coordinates of the corresponding point of the projective space $P\Lambda^p(V)$ and are defined up to a multiplication by a number $c \neq 0$. Furthermore, since the decomposable p-vectors form just a part of the space $\Lambda^p(V)$, the Plücker coordinates of a subspace are not arbitrary; there exist certain relations between them (see the next theorem).

In order to express these relations better, we accept the following convention: given a collection of numbers $\mu_{i_1 \ldots i_p}$, $i_1 < \cdots < i_p$, we assume that $\mu_{i_1 \ldots i_p}$ are also defined for *any* $i_1, \ldots, i_p$ in such a way that after two indices are interchanged, $\mu_{i_1 \ldots i_p}$ gets multiplied by -1 (so, if two indices are the same, it is zero). In particular, for any $i_1, \ldots, i_p$, $M_{i_1 \ldots i_p}$ is then equal to the determinant of a matrix of order p formed by the columns of A with indices $i_1, \ldots, i_p$ (in this order).

Theorem 8.45. *Numbers $\mu_{i_1 \ldots i_p}$ are the Plücker coordinates of some p-dimensional subspace $U \subset V$ if and only if they are not simultaneously zero and if for any $i_1, \ldots, i_{p+1}, j_1, \ldots, j_{p-1}$, the following relation holds:*

$$\sum_{k=1}^{p+1} (-1)^k \mu_{i_1 \ldots \widehat{i_k} \ldots i_{p+1}} \mu_{i_k j_1 \ldots j_{p-1}} = 0 \tag{8.64}$$

(here the symbol $\widehat{\ }$ shows that the marked index is absent).

Relations (8.64) are called the *Plücker relations*.

Remark 8.46. Since the left-hand side of relation (8.64) is skew-symmetric in $i_1, \ldots, i_{p+1}$ and $j_1, \ldots, j_{p-1}$, we may assume that $i_1 < \cdots < i_{p+1}$ and $j_1 < \cdots < j_{p-1}$.

Proof. Let us prove that relations (8.64) hold for the Plücker coordinates $M_{i_1 \ldots i_p}$ of a p-dimensional subspace $U \subset V$. Expanding the determinant $M_{i_k j_1 \ldots j_{p-1}}$ along the first column, we obtain

$$M_{i_k j_1 \ldots j_{p-1}} = \sum_s a_{s i_k} N_s,$$

where N_s does not depend on k. Thus, it suffices to prove that

$$\sum_{k=1}^{p+1} (-1)^k M_{i_1 \ldots \hat{i}_k \ldots i_{p+1}} a_{s i_k} = 0 \tag{8.65}$$

for all s. Add the sth column of A to A to obtain a $(p+1) \times n$ matrix. Denote it by A_s. Then the left-hand side of (8.65) is, up to a sign, the expansion of the determinant of a matrix formed by the columns of A_s with indices $i_1, \ldots, i_{p+1}$ along the last line. Since two columns of the matrix A_s are the same, this determinant is zero.

Conversely, assume that $\mu_{i_1 \ldots i_p}$ are not simultaneously zero and that they satisfy relations (8.64). Let us prove that there exists a $p \times n$ matrix A such that

$$\mu_{i_1 \ldots i_p} = M_{i_1 \ldots i_p} \tag{8.66}$$

for any $i_1, \ldots, i_p$ (here the meaning of $M_{i_1 \ldots i_p}$ is as above).

Without loss of generality, we may assume that $\mu_{1 \ldots p} = 1$. We are looking for a matrix A of the form

$$A = \begin{pmatrix} 1 & 0 & \ldots & 0 & a_{1,p+1} & \ldots & a_{1n} \\ 0 & 1 & \ldots & 0 & a_{2,p+1} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & a_{p,p+1} & \ldots & a_{pn} \end{pmatrix}$$

satisfying (8.66). Then, for $j > p$

$$M_{1 \ldots \hat{i} \ldots p j} = (-1)^{p-i} a_{ij}.$$

Thus, we must set

$$a_{ij} = (-1)^{p-i} \mu_{1 \ldots \hat{i} \ldots p j};$$

then equality (8.66) holds whenever the set $\{i_1, \ldots, i_p\}$ differs from the set $\{1, \ldots, p\}$ in no more than one element.

Now it remains to prove that equality (8.66) holds if the set $\{i_1, \ldots, i_p\}$ differs from the set $\{1, \ldots, p\}$ in m elements for any m. We use induction

on m. Let $i_1 \notin \{1, \dots, p\}$. The following condition holds:

$$\mu_{i_1 i_2 \dots i_p} = \mu_{1 \dots p} \mu_{i_1 i_2 \dots i_p} = \sum_{k=1}^{p} (-1)^{k+1} \mu_{i_1 1 \dots \hat{k} \dots p} \mu_{k i_2 \dots i_p}. \tag{8.67}$$

On the other hand, it follows from the first part of the theorem that the same condition holds for minors of A:

$$M_{i_1 i_2 \dots i_p} = \sum_{k=1}^{p} (-1)^{k+1} M_{i_1 1 \dots \hat{k} \dots p} M_{k i_2 \dots i_p}. \tag{8.68}$$

By the induction hypothesis, the right-hand sides of (8.67) and (8.68) coincide. Therefore, $\mu_{i_1 \dots i_p} = M_{i_1 \dots i_p}$. □

Example 8.47. For $n = 4$, $p = 2$, the Plücker relations reduce to the following one:

$$\mu_{12}\mu_{34} + \mu_{23}\mu_{14} + \mu_{31}\mu_{24} = 0. \tag{8.69}$$

Example 8.48. For $p = n - 1$, there are no nontrivial Plücker relations. Hence, every $(n-1)$-vector is decomposable.

Exercise 8.49. Prove that for $p \le q$, there exists a bilinear map

$$\varphi\colon \Lambda^p(V) \times \Lambda^q(V^*) \to \Lambda^{q-p}(V^*)$$

defined on decomposable elements as

$$\varphi(x_1 \wedge \dots \wedge x_p, \alpha_1 \wedge \dots \wedge \alpha_q) = \sum_{i_1, \dots, i_p} \operatorname{sign}(i_1, \dots, i_p, j_1, \dots, j_{q-p}) \alpha_{i_1}(x_1) \cdots \alpha_{i_p}(x_p) \alpha_{j_1} \wedge \dots \wedge \alpha_{j_{q-p}}$$

(here we sum over all different $i_1, \dots, i_p$, and $\{j_1, \dots, j_{q-p}\}$ is the complement of $\{i_1, \dots, i_p\}$ in the set $\{1, \dots, q\}$ with an arbitrary ordering).

Exercise 8.50. Prove that for a nonzero element δ of $\Lambda^n(V^*)$, the map

$$\Lambda^p(V) \to \Lambda^{n-p}(V^*), \qquad u \mapsto \varphi(u, \delta),$$

where φ is the bilinear map from Exercise 8.41, is an isomorphism that maps decomposable elements into decomposable elements. Deduce from this that every $(n-1)$-vector is decomposable (cf. Example 8.48).

Another application of the Grassmann algebra is the construction of the so-called Pfaffian of a skew-symmetric matrix of even order.

Suppose $n = 2m$ and $A = (a_{ij})$ is a skew-symmetric matrix of order n. Consider the bivector

$$a = \sum_{i<j} a_{ij}(e_i \wedge e_j) = \frac{1}{2} \sum_{i,j} a_{ij}(e_i \wedge e_j),$$

where $\{e_1, \dots, e_n\}$ is a fixed basis of the space V. Let us compute the mth power of a in the algebra $\Lambda(V)$:

$$a^m = \underbrace{a \wedge \dots \wedge a}_{m} = \frac{1}{2^m} \sum_{i_1, \dots, i_n} a_{i_1 i_2} \cdots a_{i_{n-1} i_n} e_{i_1} \wedge \dots \wedge e_{i_n}$$

$$= \frac{1}{2^m} \left(\sum_{(i_1, \dots, i_n)} \operatorname{sign}(i_1, \dots, i_n) a_{i_1 i_2} \cdots a_{i_{n-1} i_n} \right) e_1 \wedge \dots \wedge e_n,$$

where the latter sum is taken over all permutations $(i_1, \dots, i_n)$ of indices $1, \dots, n$. The summands that differ only by the order of pairs $(i_1, i_2), \dots, (i_{n-1}, i_n)$ and the order of elements in each pair are equal. Therefore,

$$a^m = m! \left(\sum_{(i_1 i_2 | \cdots | i_{n-1} i_n)} \operatorname{sign}(i_1, \dots, i_n) a_{i_1 i_2} \cdots a_{i_{n-1} i_n} \right) e_1 \wedge \dots \wedge e_n, \tag{8.70}$$

where the sum is taken over all partitions of the set $\{1, \dots, n\}$ into pairs $(i_1, i_2), \dots, (i_{n-1}, i_n)$ (the order of pairs and the order of indices within pairs are chosen arbitrarily).

The expression

$$\operatorname{pf} A = \sum_{(i_1 i_2 | \cdots | i_{n-1} i_n)} \operatorname{sign}(i_1, \dots, i_n) a_{i_1 i_2} \cdots a_{i_{n-1} i_n} \tag{8.71}$$

is called the *Pfaffian* of A. Formula (8.70) can be rewritten as

$$a^m = m! \, (\operatorname{pf} A) e_1 \wedge \dots \wedge e_n. \tag{8.72}$$

It remains valid when vectors $e_1, \dots, e_n$ are linearly dependent: then it says that $a^m = 0$.

Theorem 8.51. *(i)* $\operatorname{pf} CAC^{\mathsf{T}} = \det C \cdot \operatorname{pf} A$ *for any matrix* C *of order* n.
(ii) $(\operatorname{pf} A)^2 = \det A$.

Proof. (i) We will prove this formula first under the assumption that $\operatorname{char} K = 0$. Let $\{e'_1, \dots, e'_n\}$ be a basis of V and

$$(e_1, \dots, e_n) = (e'_1, \dots, e'_n) C.$$

Let us express the bivector $a = \frac{1}{2} \sum_{i,j} a_{ij} e_i \wedge e_j$ in terms of $e'_1, \dots, e'_n$. Put $C = (c_{ij})$; then

$$a = \frac{1}{2} \sum_{i,j,k,l} a_{ij} c_{ki} c_{lj} e'_k \wedge e'_l = \frac{1}{2} \sum_{k,l} a'_{kl} e'_k \wedge e'_l,$$

where

$$a'_{kl} = \sum_{i,j} a_{ij} c_{ki} c_{lj}.$$

Set $A' = (a'_{kl})$; then

$$A' = CAC^{\mathsf{T}}.$$

Thus,

$$a^m = m!\,(\mathrm{pf}\,A)e_1 \wedge \cdots \wedge e_n = m!\,(\mathrm{pf}\,A')e'_1 \wedge \cdots \wedge e'_n.$$

On the other hand,

$$e_1 \wedge \cdots \wedge e_n = (\det C)e'_1 \wedge \cdots \wedge e'_n$$

(cf. expression (8.63)). Therefore,

$$\mathrm{pf}\,A \cdot \det C = \mathrm{pf}\,A',$$

as required.

The equality that we have just proved can be viewed as an identity in the ring of polynomials over $\mathbb{Z}$ in entries of matrices A and C. Reduction modulo p shows that it still holds over the field $\mathbb{Z}_p$, hence over any field of characteristic p.

(ii) By Theorem 5.59, there exists a nonsingular matrix C such that

$$A = CFC^{\mathsf{T}},$$

where F is a matrix of the following form:

$$F = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & & & & \\ & \ddots & & 0 & \\ & & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & & \\ & 0 & & 0 & \\ & & & & \ddots \\ & & & & 0 \end{pmatrix}.$$

It is easy to see that

$$\det F = \mathrm{pf}\,F = \begin{cases} 1 & \text{for } k = m, \\ 0 & \text{for } k < m, \end{cases}$$

but in either case $\det F = (\mathrm{pf}\,F)^2$. According to the first part of the theorem,

$$\mathrm{pf}\,A = \det C \cdot \mathrm{pf}\,F.$$

On the other hand,

$$\det A = (\det C)^2 \det F.$$

Therefore, $\det A = (\mathrm{pf}\,A)^2$. □

Example 8.52. For a skew-symmetric matrix A of order 4, we have

$$\operatorname{pf} A = a_{12}a_{34} + a_{23}a_{14} + a_{31}a_{24}.$$

Comparing this formula with formula (8.69), we see that the bivector $a = \sum_{i<j} a_{ij} e_i \wedge e_j$ is decomposable if and only if $\operatorname{pf} A = 0$. By Theorem 8.51, this condition is equivalent to $\det A = 0$. Since the rank of a skew-symmetric matrix is always even, the latter condition holds if and only if $\operatorname{rk} A \leq 2$. On the other hand, one can easily prove directly that the bivector a is decomposable if and only if $\operatorname{rk} A \leq 2$.

Commutative Algebra

Rings (and, in particular, fields) and groups are the most important algebraic structures, for which a substantial theory exists. In this chapter, we develop the subjects of abelian groups and commutative associative rings that we already touched upon in Chapters 1 and 3. Some general definitions and the simplest facts in Sections 9.2 and 9.3 hold for more general kinds of rings.

9.1. Abelian Groups

Abelian groups are to some extent similar to vector spaces, an object with which you are already well acquainted. At any rate, the notion of linear dependence also plays an important role in the theory of abelian groups.

Recall that elements of an additive abelian group can be multiplied by integers (this corresponds to taking an integer power of an element in a multiplicative group). This operation has the same properties as the operation of multiplication of vectors by elements of the base field.

Namely, let A be an additive abelian group. Then it is easy to check that for any $a, b \in A$, $k, l \in \mathbb{Z}$,

$$k(a+b) = ka + kb, \tag{9.1}$$

$$(k+l)a = ka + la, \tag{9.2}$$

$$(kl)a = k(la) \tag{9.3}$$

(in the multiplicative version, property (9.2) was proven in Section 4.3). Properties (9.1) and (9.2) imply similar properties for subtraction:

$$k(a-b) = ka - kb, \quad (k-l)a = ka - la.$$

For any subset $S \subset A$, the collection of all linear combinations

$$k_1a_1 + \cdots + k_na_n, \qquad a_i \in S,\ k_i \in \mathbb{Z},$$

is the smallest subgroup of the group A that contains S. It is called the *subgroup generated by* S and is denoted $\langle S \rangle$. If $\langle S \rangle = A$, then we say that A is *generated* by S or that S is a *generating set* of A. (This agrees with notions introduced in Section 4.4 for arbitrary groups.) An abelian group that has a finite generating set is called *finitely generated*. Finitely generated abelian groups are similar to finite-dimensional vector spaces.

A system $\{a_1, \ldots, a_n\}$ of elements of a group A is called *linearly independent* if $k_1a_1 + \cdots + k_na_n = 0$ only for $k_1 = \cdots = k_n = 0$. A system of linearly independent elements that generates A is called a *basis*.

Every finite-dimensional vector space has a basis, but not every finitely generated abelian group has one. For instance, the group $\mathbb{Z}_n$ is generated by one element, but it has no basis since every element $a \in \mathbb{Z}_n$ satisfies the nontrivial relation $na = 0$.

Definition 9.1. A finitely generated abelian group is *free* if it has a basis.

Analogues of some statements concerning vector spaces (see Section 2.2) hold for free abelian groups.

Theorem 9.2. *All bases of a free abelian group L contain the same number of elements.*

Proof. Let $\{e_1, \ldots, e_n\}$ and $\{e'_1, \ldots, e'_m\}$ be bases of the group L. Assume that $m > n$. We have

$$(e'_1, \ldots, e'_m) = (e_1, \ldots, e_n)C,$$

where C is an $n \times m$ integral matrix (i.e., a matrix with integer entries). By Proposition 2.25, the columns of C are linearly dependent as elements of the space $\mathbb{Q}^n$. It follows that there exists a nontrivial linear dependence with integer coefficients between them; clearly, $e'_1, \ldots, e'_m$ are subject to the same dependence and this is impossible. □

The number of elements in a basis of a free abelian group L is called its *rank* and is denoted $\operatorname{rk} L$. Obviously, every free abelian group of rank n is isomorphic to the group $\mathbb{Z}^n$ of integer rows of length n.

Remark 9.3. The zero group is regarded as a free abelian group of rank 0.

We will now describe all bases of a free abelian group L. Let $\{e_1, \ldots, e_n\}$ be a basis of L and let $e'_1, \ldots, e'_n$ be some elements of L. We have

$$(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)C, \tag{9.4}$$

where C is a square integral matrix of order n.

Theorem 9.4. *Elements $e'_1, \dots, e'_n$ form a basis of L if and only if* $\det C = \pm 1$.

Proof. If $\det C = \pm 1$, the matrix C^{-1} is integral, thus

$$(e_1, \dots, e_n) = (e'_1, \dots, e'_n)C^{-1}.$$

This implies that the elements $e'_1, \dots, e'_n$ generate L and, since C is nonsingular, they are linearly independent.

Conversely, let $\{e'_1, \dots, e'_n\}$ be a basis of L. Then

$$(e_1, \dots, e_n) = (e'_1, \dots, e'_n)D \tag{9.5}$$

for an integral matrix D. By (9.4) and (9.5), we see that $CD = E$, hence $(\det C)(\det D) = 1$. Since both $\det C$ and $\det D$ are integers, $\det C = \pm 1$. □

If one is to match free abelian groups with vector spaces, then subspaces should be matched with subgroups. This is partially justified by the following theorem.

Theorem 9.5. *Every subgroup N of a free abelian group L of rank n is a free abelian group of rank $\leq n$.*

Proof. We use induction on n. For $n = 0$, there is nothing to prove.

For $n > 0$, let $\{e_1, \dots, e_n\}$ be a basis of L. Consider the subgroup $L_1 = \langle e_1, \dots, e_{n-1} \rangle \subset L$. This is a free abelian group of rank $n - 1$. By the induction hypothesis, the subgroup $N_1 = N \cap L_1$ is a free abelian subgroup of rank $m \leq n - 1$. Let $\{f_1, \dots, f_m\}$ be its basis.

Consider the last coordinates of all elements of N in the basis $\{e_1, \dots, e_n\}$ of L. They form a subgroup of the group $\mathbb{Z}$, which, by Theorem 4.50, has the form $k\mathbb{Z}$ for some $k \in \mathbb{Z}_+$. If $k = 0$, $N = N_1$ and we are done. If $k > 0$, let f_{m+1} be an element of N whose last coordinate is k. Then $\{f_1, \dots, f_m, f_{m+1}\}$ is a basis of N, which completes the proof. □

The analogy between subgroups of a free abelian group and subspaces of a vector space is not complete. Unlike a vector space, a free abelian group of rank $n > 0$ contains subgroups of the same rank that do not coincide with the whole group. For instance, the subgroup $m\mathbb{Z} \subset \mathbb{Z}$, $m > 0$, has rank 1, just as the whole group $\mathbb{Z}$.

However, the connection between free abelian groups and vector spaces goes beyond the above analogy. A free abelian group of rank n can be embedded as a subgroup into an n-dimensional Euclidean vector space E^n. Namely, let $\{e_1, \dots, e_n\}$ be a basis of E^n. Then the subgroup generated by vectors $e_1, \dots, e_n$ (i.e., the set of vectors with integer coordinates in the

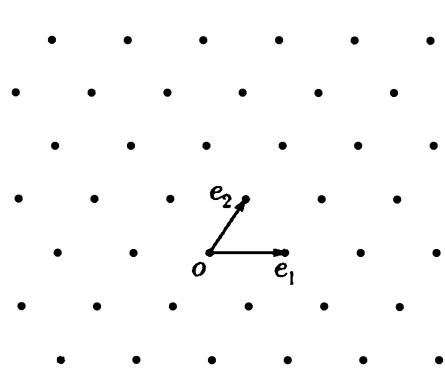


Figure 9.1

basis $\{e_1, \ldots, e_n\}$) is a free abelian group of rank n. This geometric picture (Figure 9.1) helps to understand free abelian groups better.

A subgroup $L \subset E^n$ obtained as above is called a *lattice* in E^n.

Exercise 9.6. A parallelepiped $P(e_1, \ldots, e_n)$ on a basis $\{e_1, \ldots, e_n\}$ of a lattice $L \subset E^n$ is called a *fundamental parallelepiped* of this lattice. Prove that its volume does not depend on the choice of basis of L.

There is also an axiomatic description of lattices in E^n which uses topology of this space.

Definition 9.7. A subgroup $L \subset E^n$ is *discrete* if every bounded subset of E^n contains a finite number of elements of L.

Obviously, every lattice is discrete. More generally, a subgroup generated by a linearly independent system of vectors (i.e., a lattice in a subspace of E^n) is discrete.

Exercise 9.8. Prove that a subgroup $L \subset E^n$ is discrete if and only if its intersection with a neighborhood of zero consists only of zero itself.

Theorem 9.9. *Every discrete subgroup $L \subset E^n$ is generated by a linearly independent system of vectors of E^n.*

Proof. Let $U \subset E^n$ be the linear span of the subgroup L. Clearly, L is a discrete subgroup of U. Thus, by switching from E^n to U, we may assume that the linear span of L is all of the space itself.

In this case, the subgroup L contains a basis $\{e_1, \ldots, e_n\}$ of E^n. Consider the lattice L_0 generated by this basis in E^n. In every coset of L by L_0, there is a vector from the parallelepiped $P(e_1, \ldots, e_n)$. Since the intersection $L \cap P(e_1, \ldots, e_n)$ is finite, the index $|L : L_0|$ is finite. Denote it by d. Then

$dx \in L_0$ for any $x \in L$ (see Corollary 4.68). Therefore,

$$L_0 \subset L \subset d^{-1}L_0. \tag{9.6}$$

Observe that $d^{-1}L_0$ is a lattice in E^n generated by the basis $\{d^{-1}e_1, \ldots, d^{-1}e_n\}$. By Theorem 9.5, it follows from (9.6) that L is a free group; moreover,

$$n = \operatorname{rk} L_0 \leq \operatorname{rk} L \leq \operatorname{rk} d^{-1}L_0 = n,$$

thus, $\operatorname{rk} L = n$. It follows that any basis of L is also a basis of E^n. This means that L is a lattice in E^n. □

Corollary 9.10. *A discrete subgroup $L \subset E^n$ whose linear span coincides with E^n, is a lattice in E^n.*

Example 9.11. Lattices in E^3 play an important role in crystallography. The defining feature of a crystal structure is the periodic repetition of the configuration of atoms in all three dimensions (see Figure 4.2 in Section 4.2). More explicitly, let Γ be the symmetry group of a crystal structure (which we extend to all of the space). Denote by L the group of all vectors a such that the parallel translation t_a belongs to Γ. The above implies that L generates all of E^3 (as a vector space). On the other hand, since there exist only finitely many atoms of the crystal structure in a bounded part of the space, the group L is a discrete subgroup of E^n. Hence, L is a lattice in E^3.

Usually, the group Γ contains other motions apart from parallel translations. They are the ones that determine the symmetry of real crystals in nature. That is, the symmetry group G of a crystal whose structure has symmetries described by Γ coincides with the group $d\Gamma$ of linear parts of motions in Γ.

Using the above description of the group of parallel translations contained in Γ, we can learn something about G. Namely, for each $\gamma \in \Gamma$ and every $a \in L$, we have

$$t_{d\gamma(a)} = \gamma t_a \gamma^{-1} \in \Gamma$$

(see (4.2)). Therefore, every transformation $g \in G$ preserves the lattice L, and hence its matrix in a basis of the lattice is integral. We obtain $\operatorname{tr} g \in \mathbb{Z}$. On the other hand, if g is a rotation through an angle α about an axis, then $\operatorname{tr} g = 2\cos\alpha + 1$. Hence, $2\cos\alpha \in \mathbb{Z}$. This implies that

$$\alpha \in \left\{0, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \pi\right\}.$$

In particular, unlike flowers and some lower animal species, crystals cannot have a rotational symmetry of the fifth order.

We will now provide a more precise description of subgroups of free abelian groups. The key role will be played by an auxiliary statement about integral matrices.

Definition 9.12. An *integral elementary row transformation* of a matrix is a transformation of one of the following three types:

(i) adding a row multiplied by an integer to another row;

(ii) interchanging two rows;

(iii) multiplying a row by -1.

An *integral elementary column transformation* is defined similarly.

A rectangular $n \times m$ matrix $C = (c_{ij})$ is called *diagonal* if $c_{ij} = 0$ for $i \neq j$ and $c_{ii} = u_i$ for $i = 1, \ldots, p$, $p = \min\{n, m\}$. The notation is $\operatorname{diag}(u_1, \ldots, u_p)$.

Proposition 9.13. *Every integral rectangular matrix $C = (c_{ij})$ can be reduced by integral elementary row and column transformations to the form $\operatorname{diag}(u_1, \ldots, u_p)$, where $u_1, \ldots, u_p \geq 0$ and $u_i | u_{i+1}$, $i = 1, \ldots, p-1$.*

Proof. If $C = 0$, there is nothing to prove. If $C \neq 0$ but $c_{11} = 0$, we can get $c_{11} \neq 0$ by interchanging rows and columns. We can also obtain $c_{11} > 0$ by multiplying the first row by -1, if necessary. Let us now try to reduce c_{11} by integral elementary row and column transformations.

If an entry c_{i1}, $i \geq 2$, is not divisible by c_{11}, divide it by c_{11} with a remainder:

$$c_{i1} = qc_{11} + r, \qquad 0 < r < c_{11}.$$

Subtract the first row multiplied by q from the ith row and then interchange the ith and the first rows. We thus reduced c_{11}. Similarly, if an entry c_{1j}, $j \geq 2$, is not divisible by c_{11}, we can reduce c_{11} by applying column transformations.

If all entries in the first row and the first column are divisible by c_{11} but an entry c_{ij}, $i, j \geq 2$, is not divisible by c_{11}, we act as follows. Subtract the first row multiplied by an appropriate number from the ith row in order to obtain $c_{i1} = 0$ (c_{ij} remains nondivisible by c_{11}). Then add the ith row to the first one. The entry c_{11} remains the same, but c_{1j} is no longer divisible by c_{11} and we can reduce c_{11} as above.

Following the above procedure, we finally obtain a matrix where all entries are divisible by c_{11}. By subtracting the appropriate multiples of the first row from all other rows and the appropriate multiplies of the first column from all other columns, we obtain a matrix of the form

$$\begin{pmatrix} u_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & C_1 & \\ 0 & & & \end{pmatrix},$$

where all entries of the matrix C_1 are divisible by u_1. The latter property is preserved under any integral elementary row or column transformation of C_1.

Applying the same procedure to the matrix C_1, etc., we will finally reduce C to the required form. □

For 2×1 or 1×2 matrices, the procedure in the above proof is simply the Euclidean algorithm that produces the greatest common divisor of two integers.

Example 9.14. To illustrate the procedure in general, let us consider the following example:

$$\begin{pmatrix} 2 & 6 & 2 \\ 2 & 3 & 4 \\ 4 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 2 & 6 & 2 \\ 0 & -3 & 2 \\ 4 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 2 & 3 & 4 \\ 0 & -3 & 2 \\ 4 & 2 & 4 \end{pmatrix}$$
$$\to \begin{pmatrix} 2 & 1 & 4 \\ 0 & -3 & 2 \\ 4 & -2 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 4 \\ -3 & 0 & 2 \\ -2 & 4 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 14 \\ 0 & 8 & 12 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 14 \\ 0 & 2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & 6 & 14 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 20 \end{pmatrix}.$$

Here all entries in the first row and the first column were divisible by $c_{11} = 2$ from the beginning; however, $c_{22} = 3$ was not divisible by c_{11}. This is why we subtracted the first row from the second and added the resulting second row to the first one. Of course, using particular properties of the original matrix, one could obtain the same result faster.

Exercise 9.15. Prove that $u_i = d_i/d_{i-1}$, where d_i is the greatest common divisor of the minors of order i of the original matrix C (d_0 is assumed to be equal to 1).

Remark 9.16. The above exercise implies that the numbers $u_1, \ldots, u_p$ are uniquely determined by C. If one does not require that $u_i | u_{i+1}$, the reduction to the diagonal form simplifies somewhat but, in general, the diagonal form is then no longer determined uniquely.

Theorem 9.17. *For any subgroup N of a free abelian group L of rank n, there exists a basis $\{e_1, \ldots, e_n\}$ of L and natural numbers $u_1, \ldots, u_m$, $m \le n$, such that $\{u_1 e_1, \ldots, u_m e_m\}$ is a basis of the group N and $u_i | u_{i+1}$ for $i = 1, \ldots, m-1$.*

Proof. By Theorem 9.5, N is a free abelian group of rank $m \leq n$. Let $\{e_1, \ldots, e_n\}$ be a basis of L and $\{f_1, \ldots, f_m\}$, a basis of N. Then

$$(f_1, \ldots, f_m) = (e_1, \ldots, e_n)C,$$

where C is an integral $n \times m$ matrix of rank m. We can perform the following "elementary" transformations on the bases of L and N:

(i) adding a basis element multiplied by an integer to another basis element;

(ii) interchanging two basis elements;

(iii) multiplying a basis element by -1.

The elementary transformations of the basis of L induce integral elementary row transformations of the matrix C, and elementary transformations of the basis of N induce integral elementary column transformations of this matrix. By Proposition 9.13, using these transformations we can reduce C to

$$C = \operatorname{diag}(u_1, \ldots, u_m),$$

where $u_1, \ldots, u_m > 0$ and $u_i | u_{i+1}$ for $i = 1, \ldots, m-1$. (Since $\operatorname{rk} C = m$, there are no zeros among $u_1, \ldots, u_m$.) But this means exactly that the bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ of L and N that we obtain after applying these particular transformations are related by the following equations:

$$f_i = u_i e_i, \qquad i = 1, \ldots, m.$$

□

A basis $\{e_1, \ldots, e_n\}$ of the group L that satisfies the assumptions of Theorem 9.17 is not unique. However, as we will see below, the numbers $u_1, \ldots, u_m$ are determined uniquely. They are called *invariant factors* of the subgroup $N \subset L$.

Exercise 9.18. Prove that when $m = n$, the index $|L : N|$ is finite and equal to the product of invariant factors.

Exercise 9.19. Let L be a lattice in E^n and N, its sublattice. Prove that the index $|L : N|$ equals the ratio of the volumes of the fundamental parallelepipeds of the lattices L and N.

Figure 9.2 illustrates Theorem 9.17: dots stand for elements of the lattice $L \subset E^2$ and encircled dots, for elements of the sublattice N. Vectors e_1 and e_2 form a basis of L that satisfies the assumptions of the theorem; here $u_1 = 1$, $u_2 = 4$.

Let us study now the structure of an arbitrary finitely generated abelian subgroup. For this, we need the notion of a direct sum of abelian groups.

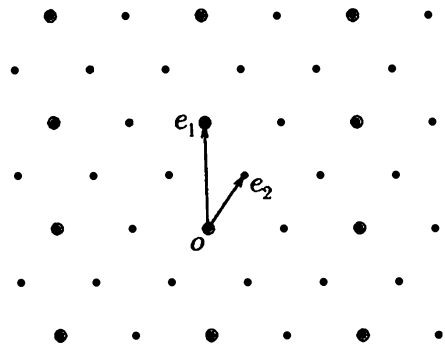


Figure 9.2

Definition 9.20. An (additive) abelian group decomposes into a *direct sum* of subgroups $A_1, \ldots, A_k$ if every element $a \in A$ decomposes uniquely as $a = a_1 + \cdots + a_k$ with $a_i \in A_i$. We denote this as

$$A = A_1 \oplus \cdots \oplus A_k.$$

In the case of two subgroups A_1 and A_2, the uniqueness of the decomposition of every element $a \in A$ in the form $a = a_1 + a_2$, $a_1 \in A_1, a_2 \in A_2$, is equivalent to the condition $A_1 \cap A_2 = 0$.

Definition 9.21. The *direct sum* of (additive) abelian groups $A_1, \ldots, A_k$ is the group $A_1 \oplus \cdots \oplus A_k$ that consists of all sequences $(a_1, \ldots, a_k)$, $a_i \in A_i$, with componentwise addition.

For instance, $\underbrace{\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}}_{n} = \mathbb{Z}^n$. Notice that if the groups $A_1, \ldots, A_k$ are finite, then

$$|A_1 \oplus \cdots \oplus A_k| = |A_1| \cdots |A_k|.$$

The direct sum in the sense of Definition 9.20 is called *internal* and in the sense of Definition 9.21, *external*. These two notions are related just as in the case of vector spaces (see Section 8.2).

In the case of multiplicative abelian groups $G_1, \ldots, G_k$, one usually speaks about a *direct product* and denotes it as $G_1 \times \cdots \times G_k$. This agrees with the general definition of a direct product of groups that is given in Section 10.1.

First consider the decomposition of cyclic groups into a direct sum of (cyclic) subgroups.

Recall that every infinite cyclic group is isomorphic to the additive group $\mathbb{Z}$ and every finite cyclic group of order n is isomorphic to the additive group $\mathbb{Z}_n$ of residue classes modulo n.

Exercise 9.22. Prove that the group $\mathbf{Z}$ cannot decompose into a direct sum of two nonzero subgroups.

Proposition 9.23. *If $n = kl$ with $(k, l) = 1$, then*

$$\mathbf{Z}_n \simeq \mathbf{Z}_k \oplus \mathbf{Z}_l. \tag{9.7}$$

Proof. Since $|\mathbf{Z}_k \oplus \mathbf{Z}_l| = kl = n$, it suffices to produce an element of order n in the group $\mathbf{Z}_k \oplus \mathbf{Z}_l$. For example, such is the element $([1]_k, [1]_l)$. □

Exercise 9.24. Determine preimages of the elements $[1]_3 \in \mathbf{Z}_3$ and $[1]_5 \in \mathbf{Z}_5$ under the isomorphism $\mathbf{Z}_{15} \xrightarrow{\sim} \mathbf{Z}_5 \oplus \mathbf{Z}_3$ that maps $[1]_{15}$ into $([1]_3, [1]_5)$.

Corollary 9.25. *If $n = p_1^{k_1} \cdots p_s^{k_s}$ is the factorization of n into prime numbers, then*

$$\mathbf{Z}_n \simeq \mathbf{Z}_{p_1^{k_1}} \oplus \cdots \oplus \mathbf{Z}_{p_s^{k_s}}. \tag{9.8}$$

Definition 9.26. A *primary group* or a *p-group* is a finite group whose order is a power of a prime number p.

Thus, every finite cyclic group decomposes into a direct sum of primary cyclic groups.

Exercise 9.27. Prove that a primary cyclic group cannot decompose into a direct sum of two nonzero subgroups.

Theorem 9.28. *Every finitely generated abelian group A decomposes into a direct sum of primary cyclic and infinite cyclic subgroups. Moreover, the collection of orders of these subgroups is determined uniquely.*

Proof. Let $\{a_1, \ldots, a_n\}$ be a generating set of A. Consider the following homomorphism

$$\varphi\colon \mathbf{Z}^n \to A, \quad (k_1, \ldots, k_n) \mapsto k_1 a_1 + \cdots + k_n a_n.$$

By the Homomorphism Theorem 4.100, $A \simeq \mathbf{Z}^n/N$, where $N = \operatorname{Ker}\varphi$. By Theorem 9.17, there exists a basis $\{e_1, \ldots, e_n\}$ of the group $\mathbf{Z}^n$ and natural numbers $u_1, \ldots, u_m$, $m \leq n$, such that $\{u_1 e_1, \ldots, u_m e_m\}$ is a basis of N and $u_i | u_{i+1}$ for $i = 1, \ldots, m-1$. Consider the homomorphism

$$\psi\colon \mathbf{Z}^n \to \mathbf{Z}_{u_1} \oplus \cdots \oplus \mathbf{Z}_{u_m} \oplus \underbrace{\mathbf{Z} \oplus \cdots \oplus \mathbf{Z}}_{n-m},$$

$$l_1 e_1 + \cdots + l_n e_n \mapsto ([l_1]_{u_1}, \ldots, [l_m]_{u_m}, l_{m+1}, \ldots, l_n).$$

Clearly, $\operatorname{Ker}\psi = N$. It follows that

$$A \simeq \mathbf{Z}_{u_1} \oplus \cdots \oplus \mathbf{Z}_{u_m} \oplus \underbrace{\mathbf{Z} \oplus \cdots \oplus \mathbf{Z}}_{n-m}. \tag{9.9}$$

(If $u_1 = \cdots = u_q = 1$, the first q factors have the form $\mathbb{Z}_1 = \mathbb{Z}/\mathbb{Z} = 0$ and should be dropped.)

Thus, A decomposes into a direct sum of cyclic subgroups. Each finite factor of this decomposition can, in turn, be decomposed into a direct sum of primary cyclic subgroups. Hence, we obtain the required decomposition of A.

It remains to show uniqueness. Let $\langle c\rangle_q$ denote the cyclic group of order q generated by c. Assume that A decomposes into a direct sum of primary cyclic and infinite cyclic subgroups:

$$A = \langle c_1\rangle_{p_1^{k_1}} \oplus \cdots \oplus \langle c_s\rangle_{p_s^{k_s}} \oplus \langle c_{s+1}\rangle_\infty \oplus \cdots \oplus \langle c_{s+t}\rangle_\infty \tag{9.10}$$

(prime numbers $p_1, \ldots, p_s$ are not necessarily all distinct). Consider the so-called *torsion subgroup*:

$$\operatorname{Tor} A := \{a \in A : ma = 0 \text{ for some } m \in \mathbb{Z},\ m \neq 0\}. \tag{9.11}$$

Clearly, $\operatorname{Tor} A$ is the sum of the first s factors in decomposition (9.10). Hence, $A/\operatorname{Tor} A \simeq \mathbb{Z}^t$. By definition, the subgroup $\operatorname{Tor} A$ does not depend on decomposition (9.10), thus we have shown that the value of t does not depend on this decomposition either.

Now, for every prime number p, we can consider the *p-torsion subgroup*

$$\operatorname{Tor}_p A := \{a \in A\colon p^k a = 0 \text{ for some } k \in \mathbb{Z}_+\}. \tag{9.12}$$

Clearly, $\operatorname{Tor}_p A$ is the sum of finite factors in decomposition (9.10) whose orders are powers of p. Hence, the sum of these factors does not depend on (9.10) either. Thus, we reduced the proof to the case of a primary group A.

Let $|A| = p^k$ and let A decompose into a direct sum of cyclic subgroups:

$$A = \langle c_1\rangle_{p^{k_1}} \oplus \cdots \oplus \langle c_r\rangle_{p^{k_r}}, \quad k_1 + \cdots + k_r = k. \tag{9.13}$$

We will prove by induction on k that the collection $\{k_1, \ldots, k_r\}$ does not depend on decomposition (9.13).

For $k = 1$, the statement is obvious. For $k > 1$, consider the subgroup

$$pA := \{pa\colon a \in A\} \subset A.$$

Obviously,

$$pA = \langle pc_1\rangle_{p^{k_1-1}} \oplus \cdots \oplus \langle pc_r\rangle_{p^{k_r-1}};$$

in particular, for $k_i = 1$, the corresponding factor simply disappears. Since the definition of the subgroup pA does not depend on decomposition (9.13), the induction hypothesis implies that the collection of k_i, $i = 1, \ldots, r$, such that $k_i \neq 1$ does not depend on this decomposition. As for the k_i's equal to 1, their number can be derived from the condition $k_1 + \cdots + k_r = k$ and does not depend on decomposition (9.13) either. □

Remark 9.29. In decomposition (9.13), the subgroups themselves are not defined uniquely for $r > 1$. For instance, for $k_1 = \cdots = k_r = 1$, the group A can be regarded as an r-dimensional vector space over the field $\mathbb{Z}_p$, so to decompose it into a direct sum of cyclic subgroups is the same as to decompose this vector space into a direct sum of one-dimensional subspaces. The latter decomposition is obviously not unique.

Remark 9.30. If the group A is finite, then there can be no infinite subgroups in its decomposition; therefore, it decomposes into a direct sum of primary cyclic subgroups.

We did not use the condition $u_i | u_{i+1}$ in the proof of Theorem 9.28. However, it allows us to reconstruct numbers $u_1, \ldots, u_m$, i.e., the invariant factors of the subgroup $N \subset \mathbb{Z}^n$, from the orders of factors in decomposition (9.10). Thus we can prove that invariant factors of a subgroup of a free abelian group L do not depend on the choice of basis of L satisfying the conditions of Theorem 9.17.

Namely, by following the proof of the theorem, we see that for a prime number p, its power in the factorization of u_m equals its maximum power among the numbers $p_1^{k_1}, \ldots, p_s^{k_s}$; the power of p in the factorization of u_{m-1} equals its maximum power among the remaining numbers $p_1^{k_1}, \ldots, p_s^{k_s}$; etc.

It also follows from the proof of the theorem that every finite abelian group A allows the decomposition

$$A = \langle a_1 \rangle_{u_1} \oplus \cdots \oplus \langle a_m \rangle_{u_m} \tag{9.14}$$

where $u_i | u_{i+1}$ for $i = 1, \ldots, m-1$. We can assume that $u_1 \neq 1$; otherwise a number of first factors in the decomposition can be dropped. Under these conditions, the numbers $u_1, \ldots, u_m$ are uniquely determined. They are called the *invariant factors* of A. Their product is equal to $|A|$.

The last invariant factor has a simple meaning.

Definition 9.31. The *exponent* of a finite group is the least common multiple of the orders of elements of this group.

Corollary 4.71 shows that the exponent of a finite group divides its order.

Proposition 9.32. *The exponent of a finite abelian group A equals its last invariant factor u_m.*

Proof. Clearly, $u_m a = 0$ for any $a \in A$. This means that the exponent of A divides u_m, but since A contains a cyclic subgroup of order u_m, the exponent equals u_m. □

Corollary 9.33. *A finite abelian group A is cyclic if and only if its exponent equals its order.*

Proof. A group A is cyclic if and only if it has only one factor in decomposition (9.14), but this means exactly that $u_m = |A|$. □

This criterion of cyclicity of a finite abelian group has an interesting application.

Theorem 9.34. *Every finite subgroup of the multiplicative group of a field (and, in particular, the multiplicative group of any finite field) is cyclic.*

Proof. Let G be a finite subgroup of the multiplicative group of a field K. Assume that its exponent equals m. Then $g^m = 1$ for every $g \in G$. Since the equation $x^m = 1$ has at most m solutions in K, $|G| \leq m$ and, hence, $|G| = m$. □

Exercise 9.35. Find a generator of each of the following groups: $\mathbb{Z}_7^*$; $\mathbb{Z}_{41}^*$.

Exercise 9.36. Prove that the group $\mathbb{Z}_{2^k}^*$ of invertible elements of the ring $\mathbb{Z}_{2^k}$ is not cyclic for $k > 2$; more precisely, $\mathbb{Z}_{2^k}^* = \langle 3 \rangle \times \langle -1 \rangle \cong \mathbb{Z}_{2^{k-2}} \oplus \mathbb{Z}_2$.

Remark 9.37. It can be shown that the group $\mathbb{Z}_n^*$ is cyclic if and only if $n = 2, 4, p^k$, or $2p^k$, where p is an odd prime number.

Example 9.38. For an odd prime p, $\mathbb{Z}_p^*$ is a cyclic group of an even order, hence the squares form in it a subgroup of index 2. Thus the map that sends a quadratic residue modulo p to 1 and a quadratic nonresidue to -1 is a homomorphism of the group $\mathbb{Z}_p^*$ to the (multiplicative) group $\{\pm 1\}$. The image of a residue class $[k]_p$ under this map is denoted $\left(\frac{k}{p}\right)$ and is called the *Legendre symbol*.

The residue class $[-1]_p$ is the only element of order 2 in the group $\mathbb{Z}_p^*$. It is quadratic if and only if this group contains an element of order 4, i.e., if $|\mathbb{Z}_p^*| = p - 1$ is divisible by 4. Thus,

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}.$$

Exercise 9.39. Prove that a polynomial $x^4 + 1$ is reducible over any finite field. (*Hint*: first prove that at least one of the elements $-1, 2, -2$ is a square in such a field.)

9.2. Ideals and Quotient Rings

Generalizing the construction of the ring of residue classes $\mathbb{Z}_n$ in Section 1.6, we can consider equivalence relations that agree with the operations in an arbitrary ring. Since a ring is, first of all, an additive abelian group, such a relation must be the congruence relation modulo an additive subgroup (see

Section 4.5 and, in particular, Exercise 4.84). Let us discuss what kind of subgroups defines a relation that agrees with multiplication as well.

Let A be a ring and $I \subset A$, an additive subgroup.

Proposition 9.40. *The equivalence relation modulo I agrees with multiplication if and only if I is invariant under left and right multiplication by any element of A.*

The above conditions say that for any $x \in I$ and $a \in A$, the inclusions $ax \in I$ and $xa \in I$ hold. An additive subgroup I satisfying these conditions is called a (*two-sided*) *ideal* of the ring A. A subgroup that satisfies the former (respectively, latter) condition is called a *left* (respectively, *right*) *ideal*. Clearly, in a commutative ring, there is no difference between left, right, and two-sided ideals.

Proof. Assume that the equivalence relation modulo I agrees with the operation of multiplication. Then for any $a \in A$,

$$x \equiv 0 \pmod{I} \quad \Longrightarrow \quad ax \equiv a \cdot 0 = 0 \pmod{I}.$$

This means that I is a left ideal. Similarly, one proves that I is a right ideal.

Conversely, let I be an ideal and let

$$a \equiv a' \pmod{I}, \qquad b \equiv b' \pmod{I},$$

i.e.,

$$a' = a + x, \qquad b' = b + y, \qquad x, y \in I.$$

Then

$$a'b' = ab + ay + xb + xy \equiv ab \pmod{I}.$$

□

So, if I is an ideal of A, we can define multiplication on the quotient group A/I by the following rule:

$$(a + I)(b + I) = ab + I.$$

It is easy to see that this operation satisfies the distributive law. The ring we have just constructed is called the *quotient ring* of the ring A by the ideal I and is denoted A/I. If A is commutative, associative, or has a unity, then the quotient ring has the respective property as well.

Example 9.41. A field has no nontrivial ideals (i.e., those different from the zero ideal and the field itself). Indeed, if x is a nonzero element of a field K, then every element of K can be presented in the form ax for $a \in K$, thus every ideal containing x coincides with K.

Example 9.42. Every additive subgroup of the ring $\mathbb{Z}$ is of the form $n\mathbb{Z}$, where $n \in \mathbb{Z}_+$ (see Example 4.52), and is an ideal. The quotient ring $\mathbb{Z}/n\mathbb{Z}$, $n \neq 0$, is just the ring of residue classes $\mathbb{Z}_n$.

We provide examples of ideals in other rings further down. Now, let us show how ideals and quotient rings relate to homomorphisms.

A map f from a ring A to a ring B is called a *homomorphism* if it agrees with the operations, i.e., if

$$f(x+y) = f(x) + f(y),$$
$$f(xy) = f(x)f(y)$$

for any $x, y \in A$. The image $\operatorname{Im} f$ of the homomorphism f is a subring of B and its kernel

$$\operatorname{Ker} f = \{x \in A \colon f(x) = 0\}$$

is an ideal of A.

According to the definition of the quotient ring A/I, the map

$$\pi \colon A \to A/I, \qquad a \mapsto a + I,$$

is a homomorphism. It is called the *canonical homomorphism* of the ring A onto the ring A/I. Its kernel is obviously the ideal I.

One has the following Homomorphism Theorem for rings which is an analogue of the Homomorphism Theorem 4.100 for groups.

Theorem 9.43. *Let $f \colon A \to B$ be a ring homomorphism. Then*

$$\operatorname{Im} f \simeq A/\operatorname{Ker} f.$$

More precisely, there exists an isomorphism

$$\varphi \colon \operatorname{Im} f \xrightarrow{\sim} A/\operatorname{Ker} f$$

that maps an element $b = f(a) \in \operatorname{Im} f$ to the coset $\pi(a) = a + \operatorname{Ker} f$.

Proof. We already know from Theorem 4.100 that the map φ is an additive group isomorphism. It remains to show that it preserves multiplication. Let $f(x) = u$ and $f(y) = v$. Then $f(xy) = uv$ and

$$\varphi(uv) = \pi(xy) = \pi(x)\pi(y) = \varphi(u)\varphi(v).$$

□

Example 9.44. Reduction modulo p that we considered in Section 3.6 is a homomorphism of the ring $\mathbb{Z}[t]$ to the ring $\mathbb{Z}_p[t]$. Its kernel is the ideal $p\mathbb{Z}[t]$ formed by polynomials with coefficients divisible by p. Therefore,

$$\mathbb{Z}[t]/p\mathbb{Z}[t] \simeq \mathbb{Z}_p[t].$$

Example 9.45. Let K be a field. Fix an element $c \in K$. As we practically proved in Section 3.1, the map

$$K[t] \to K, \qquad f \mapsto f(c),$$

is a ring homomorphism. By Bezout's theorem, its kernel consists of all polynomials divisible by $t - c$. Therefore,

$$K[t]/(t - c)K[t] \simeq K.$$

Example 9.46. Let $t^2 + pt + q \in \mathbb{R}[t]$ be a quadratic polynomial with a negative discriminant. Let $c \in \mathbb{C}$ be one of its imaginary roots. The map

$$\mathbb{R}[t] \to \mathbb{C}, \qquad f \mapsto f(c),$$

is a ring homomorphism. Its image coincides with $\mathbb{C}$ and its kernel consists of polynomials divisible by $(t - c)(t - \bar{c}) = t^2 + pt + q$. Therefore,

$$\mathbb{R}[t]/(t^2 + pt + q)\mathbb{R}[t] \simeq \mathbb{C}.$$

When A is an algebra over a field K, we add another condition to the definition of a *left, right*, or *two-sided* ideal: it should be preserved when multiplied by elements of K, i.e., also be a subspace of A. For a (two-sided) ideal I of A, we define multiplication of elements of A/I by elements of K as follows:

$$\lambda(a + I) = \lambda a + I.$$

This turns it into an algebra over K called the *quotient algebra* of A by the ideal I.

Remark 9.47. If A is an algebra with unity 1, then the ideals of the algebra A are the same as the ideals of the ring A. Indeed, let I be a left ideal of the ring A. Then for any $x \in I$ and $\lambda \in K$,

$$\lambda x = (\lambda 1)x \in I.$$

This means that I is a subspace, hence a left ideal of the algebra A. The same holds for the right ideals.

Example 9.48. A direct check shows that the matrices with all but the first column equal to zero form a left ideal in the algebra $\mathrm{L}_n(K)$ of matrices of order n. Similarly, the matrices with all but the first row equal to zero form a right ideal. However, there are no nontrivial two-sided ideals in the algebra $\mathrm{L}_n(K)$. Indeed, suppose $I \subset \mathrm{L}_n(K)$ is a nonzero two-sided ideal and $A = (a_{ij})$, a nonzero matrix in this ideal. Assume that $a_{kl} \neq 0$. For any i, j,

$$E_{ik}AE_{lj} = a_{kl}E_{ij} \in I,$$

thus $E_{ij} \in I$. Therefore, $I = \mathrm{L}_n(K)$.

Example 9.49. The niltriangular matrices form an ideal in the algebra of all triangular matrices.

Example 9.50. The functions whose value at a fixed point $x_0 \in X$ is 0, form an ideal in the algebra $F(X, K)$ of all K-valued functions on the set X.

A map f from an algebra A to an algebra B is called a *homomorphism* if it is linear and multiplication-preserving, i.e.,

$$f(xy) = f(x)f(y)$$

for any $x, y \in A$. The image $\operatorname{Im} f$ of the homomorphism f is a subalgebra of the algebra B and its kernel $\operatorname{Ker} f$ is an ideal of the algebra A.

For any ideal I of an algebra A, we define the *canonical homomorphism*

$$\pi\colon A \to A/I, \qquad a \mapsto a + I.$$

Its kernel is I.

One has the Homomorphism Theorem for algebras which is formulated just as the Homomorphism Theorem for rings.

Example 9.51. The map associating a triangular matrix to its diagonal part is a homomorphism of the algebra of triangular matrices onto the algebra of diagonal matrices. Its kernel is the ideal of niltriangular matrices. Thus, the quotient algebra of triangular matrices by the ideal of niltriangular matrices is isomorphic to the algebra of diagonal matrices.

Example 9.52. The map that sends a function $f \in F(X, K)$ to its value at a fixed point $x_0 \in X$ is a homomorphism of the algebra $F(X, K)$ to the field K regarded as a (one-dimensional) algebra over itself. Its kernel is the ideal $I(x_0)$ of functions whose value at the point x_0 is 0. Therefore,

$$F(X, K)/I(x_0) \simeq K.$$

Definition 9.53. A ring (respectively, an algebra) A decomposes into a direct sum of its subrings (respectively, subalgebras) $A_1, \ldots, A_k$ if

(i) it decomposes into the direct sum of $A_1, \ldots, A_k$ as an additive group (respectively, a vector space);

(ii) $A_i A_j = 0$ for $i \neq j$.

The latter condition is equivalent to saying that $A_1, \ldots, A_k$ are ideals (as long as condition (i) holds). Thus the following "componentwise" rule holds for multiplication:

$$(x_1 + \cdots + x_k)(y_1 + \cdots + y_k) = x_1 y_1 + \cdots + x_k y_k, \qquad x_i, y_i \in A_i.$$

Now let $A_1, \ldots, A_k$ be rings or algebras.

Definition 9.54. The *direct sum* of rings (respectively, algebras) $A_1, \ldots, A_k$ is their direct sum $A_1 \oplus \cdots \oplus A_k$ as additive groups (respectively, vector spaces) with componentwise multiplication:

$$(x_1, \ldots, x_k)(y_1, \ldots, y_k) = (x_1 y_1, \ldots, x_k y_k), \qquad x_i, y_i \in A_i.$$

Obviously, the operation of multiplication in $A_1 \oplus \cdots \oplus A_k$, which we have just defined, is distributive with respect to addition (respectively, bilinear), so that $A_1 \oplus \cdots \oplus A_k$ indeed becomes a ring (respectively, an algebra). If all the rings $A_1, \ldots, A_k$ are commutative, associative, or have a unity, then their direct sum has the respective property as well.

The direct sum of rings or algebras in the sense of Definition 9.53 is called *internal* and in the sense of Definition 9.54, *external*. These two notions are related as in the case of vector spaces.

Example 9.55. Let $n = kl$ for $(k, l) = 1$. The isomorphism of additive groups

$$\mathbb{Z}_n \xrightarrow{\sim} \mathbb{Z}_k \oplus \mathbb{Z}_l \tag{9.15}$$

that maps the unity $[1]_n$ of the ring $\mathbb{Z}_n$ to the unity $([1]_k, [1]_l)$ of the ring $\mathbb{Z}_k \oplus \mathbb{Z}_l$ (see Proposition 9.23) is actually a ring isomorphism. This follows from the fact that in a cyclic additive subgroup generated by the ring unity, multiplication can be expressed via addition by the following formula:

$$(s1)(t1) = (st)1, \quad s, t \in \mathbb{Z}.$$

Ring isomorphism (9.15) induces an isomorphism of multiplicative groups of invertible elements:

$$\mathbb{Z}_n^* \xrightarrow{\sim} \mathbb{Z}_k^* \times \mathbb{Z}_l^*. \tag{9.16}$$

Exercise 9.56. Using the conclusion of Example 9.55, obtain the following expression for the Euler function (see Example 4.72):

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_s}\right),$$

where $p_1, \ldots, p_s$ are all (distinct) prime divisors of n.

Example 9.57. The map sending a diagonal matrix to the sequence of its diagonal entries is an isomorphism of the algebra of diagonal matrices of order n over a field K to the direct sum of n copies of K.

From this point on, we always assume that A is a *commutative associative ring with unity*.

For any subset $S \subset A$, the collection of all "linear combinations"

$$a_1 x_1 + \cdots + a_m x_m, \quad x_1, \ldots, x_m \in S,\ a_1, \ldots, a_m \in A,$$

is the smallest ideal containing S. It is called the *ideal generated by the subset* S and is denoted (S). In particular, the ideal (u) generated by a single element u is called a *principal ideal*.

Definition 9.58. A *principal ideal domain* is an integral domain such that every ideal in it is principal.

In particular, every field is trivially a principal ideal domain.

Theorem 9.59. *Every Euclidean domain is a principal ideal domain.*

Proof. Obviously, the zero ideal is principal. Let I be a nonzero ideal of A and let u be an element of I with the least norm. Divide any element of I by u with a remainder: this remainder must belong to I, hence it is zero. This implies that $I = (u)$. □

Therefore, the rings $\mathbb{Z}$ and $K[t]$ (for a field K) are principal ideal domains.

Remark 9.60. It is easy to see that the property that every ideal is principal survives passing to the quotient ring. However, a quotient ring of a principal ideal domain need not be one, since it may cease being a domain, i.e., acquire zero divisors. For instance, all ideals in the ring $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ are principal, but it is a principal ideal domain only when n is prime (and then it is a field).

Remark 9.61. There exist principal ideal domains that are neither Euclidean domains nor fields. For example, such is the ring of numbers $a + b\sqrt{-19}$, where $a, b \in \mathbb{Z}$ or $a, b \in \mathbb{Z} + \frac{1}{2}$.

Properties of division that we proved for Euclidean domains in Section 3.5 generalize to arbitrary principal ideal domains.

Theorem 9.62. *In a principal ideal domain A, every pair of elements x, y has the greatest common divisor d, which can be presented in the form $d = ax + by$, $a, b \in A$.*

Proof. Consider the ideal

$$(x, y) = \{ax + by\colon a, b \in A\}$$

generated by elements x and y. There exists an element $d \in A$ such that $(x, y) = (d)$. This is the greatest common divisor of elements x and y. By construction, it can be presented in the form $d = ax + by$. □

Remark 9.63. The notation (x, y) for an ideal generated by elements x and y agrees quite well with the notation (x, y) for the greatest common divisor.

Factorization into primes also exists and is unique in principal ideal domains. Indeed, with the help of Theorem 9.62, the proof of uniqueness for a Euclidean domain in Section 3.5 carries verbatim to the case of a principal ideal domain. As for the existence, we will prove it later for a larger class of rings (see Theorem 9.153).

The following theorem generalizes Theorem 1.49.

Theorem 9.64. *Let u be a nonzero noninvertible element of a principal ideal domain A. The quotient ring $A/(u)$ is a field if and only if u is prime.*

Proof. For any $a \in A$, denote the coset $a + (u) \in A/(u)$ by $[a]$. If $u = vw$ for noninvertible v and w, then $[v][w] = 0$. However, $[v], [w] \neq 0$, thus the ring $A/(u)$ contains zero divisors and is not a field.

Conversely, if the element u is prime, then for any $x \notin (u)$, the elements x and u are relatively prime. Therefore, there exist a and b such that $ax + bu = 1$. Passing to cosets, we have $[a][x] = 1$ in $A/(u)$. Thus, every nonzero element of the ring $A/(u)$ is invertible, hence, $A/(u)$ is a field. □

Note that if a principal ideal domain is not a field, it contains noninvertible nonzero elements and hence prime elements.

Example 9.65. We will determine here when a prime number p is a prime element of the ring $\mathbb{Z}[i]$ of Gaussian integers (see Example 3.54). Since $\mathbb{Z}[i] \simeq \mathbb{Z}[t]/(t^2 + 1)$ (see Example 9.46),

$$\mathbb{Z}[i]/(p) \simeq \mathbb{Z}[t]/(t^2 + 1, p) \simeq \mathbb{Z}_p[t]/(t^2 + 1)$$

by Example 9.44. By Example 9.38, the polynomial $t^2 + 1$ is irreducible over $\mathbb{Z}_p$ if and only if $p \equiv -1 \pmod 4$. Theorem 9.64 (applied twice) implies that the latter condition is necessary and sufficient for p being prime in $\mathbb{Z}[i]$.

Let $p \equiv 1 \pmod 4$ and let $p = \pi_1 \cdots \pi_s$, $s \geq 2$, be the prime factorization of p in the ring $\mathbb{Z}[i]$. Passing to norms, we obtain

$$N(\pi_1) \cdots N(\pi_s) = N(p) = p^2,$$

implying that $s = 2$ and $N(\pi_1) = N(\pi_2) = p$. If $\pi_1 = a + bi$ for $a, b \in \mathbb{Z}$, then $a^2 + b^2 = p$ (and $\pi_2 = a - bi$). Therefore, every prime number of the form $4k + 1$ can be presented as a sum of squares of two integers.

Exercise 9.66. Prove that up to multiplication by invertible elements, prime elements of the ring $\mathbb{Z}[i]$ are prime natural numbers of the form $4k+3$, numbers of the form $a + bi$, for $a, b \in \mathbb{N}$ such that $a^2 + b^2$ is a prime natural number of the form $4k + 1$, and the number $1 + i$.

Exercise 9.67. Use the uniqueness of prime factorization in the ring $\mathbb{Z}[i]$ to prove that a natural number n can be presented as a sum of squares of two integers if and only if in its prime factorization (in $\mathbb{Z}$) every prime

factor of the form $4k+3$ occurs with even exponent. Find the number of presentations in this case.

The following theorem generalizes Example 9.55.

Theorem 9.68. *Let u and v be relatively prime elements of a principal ideal domain A. Then*

$$A/(uv) \simeq A/(u) \oplus A/(v). \tag{9.17}$$

Proof. The map

$$f\colon A \longrightarrow A/(u) \oplus A/(v), \qquad a \mapsto (a+(u), a+(v)),$$

is a ring homomorphism. Let a and b be elements of A such that $au+bv=1$. Then

$$f(bv) = (1+(u), 0+(v)), \qquad f(au) = (0+(u), 1+(v)),$$

implying that the homomorphism f is surjective. Obviously, $\operatorname{Ker} f = (uv)$. This establishes the isomorphism (9.17). □

Example 9.69. If $f \in K[t]$ is an irreducible polynomial over a field K, then the quotient ring $K[t]/(f)$ is a field. For instance, $\mathbb{R}[t]/(t^2+1) \simeq \mathbb{C}$ (see Example 9.46). On the other hand, if $f = (t-c_1)\cdots(t-c_n)$ for different $c_1, \ldots, c_n$, Theorem 9.68 implies that

$$K[t]/(f) \simeq K[t]/(t-c_1) \oplus \cdots \oplus K[t]/(t-c_n) \simeq \underbrace{K \oplus \cdots \oplus K}_{n}$$

(see Example 9.45).

9.3. Modules over Principal Ideal Domains

In view of properties (9.1)–(9.3), abelian groups can be regarded as "vector spaces over $\mathbb{Z}$." Likewise, one can define "vector spaces" over more general rings. They are called modules.

The notion of a module turns out to be very useful. In particular, the theory of modules over principal ideal domains, which we present in this section, contains the theory of finitely generated abelian groups discussed in Section 9.1 and the theorem about reduction of the matrix of a linear operator to the Jordan canonical form.

We start, however, with more general notions.

Let A be an associative ring with unity.

Definition 9.70. A (*left*) *A-module* (or a *module over A*) is an additive abelian group M, with the operation of (left) multiplication by elements of the ring A, that satisfies the following properties:

(i) $a(x+y) = ax + ay$ for any $a \in A$, $x, y \in M$;

(ii) $(a+b)x = ax + bx$ for any $a, b \in A$, $x \in M$;

(iii) $(ab)x = a(bx)$ for any $a, b \in A$, $x \in M$;

(iv) $1x = x$ for any $x \in M$.

In particular, a module over a field is a vector space and a module over $\mathbb{Z}$ is an additive abelian group. However, there exist other important examples of modules.

Example 9.71. A module over a polynomial ring $K[t]$ (K is a field) is a vector space over K with a linear operator that plays the role of multiplication by t.

Example 9.72. A ring A is a module over itself (the product of an element of the ring and an element of the module is defined as the product of these two elements in the ring).

Example 9.73. Any vector space V is naturally a module over the ring $\mathrm{L}(V)$ of all linear operators on V.

Remark 9.74. Similarly, one can define *right modules*. The difference is that in this case the elements of the ring A are put to the right of the elements of the module, so that when multiplying an element of this module by a product of elements of the ring, we first multiply it by the first element (and not the second as in the case of left modules). If the ring A is commutative, there is no difference between left and right modules (and elements of the ring can be put on either side of elements of the module).

A subset N of a module M is called a *submodule* if it is closed with respect to the operations of addition and multiplication by elements of the ring A. Every submodule is a module with respect to these operations.

Example 9.75. A submodule of an abelian group regarded as a $\mathbb{Z}$-module is just a subgroup.

Example 9.76. A submodule of a $K[t]$-module (see Example 9.71) is a subspace invariant under the operator of multiplication by t.

Example 9.77. A submodule of a ring A regarded as a (left) module over itself is a left ideal of this ring.

Just as we did this for vector spaces in Section 8.2 and for abelian groups in Section 9.1, we can define the (internal and external) *direct sum* of modules.

Now let us define a quotient module.

An equivalence relation R on an A-module M agrees with the operation of multiplication by elements of A if

$$x \overset{R}{\sim} x' \Rightarrow ax \overset{R}{\sim} ax'.$$

The relation of comparison modulo an additive subgroup $N \subset M$ agrees with multiplication by elements of A if and only if N is a submodule. In this case, we can define multiplication by elements of A on the quotient group M/N by the following rule:

$$a(x + N) = ax + N.$$

This makes this group into an A-module called the *quotient module* of M by N and denoted M/N.

In particular, this is how one defines the *quotient space* V/U of a vector space V by a subspace U. Quotient modules of $\mathbb{Z}$-modules are the same as quotient groups.

A map f of a module M into a module N (over the same ring) is called a *homomorphism* if

$$f(x + y) = f(x) + f(y),$$
$$f(ax) = af(x).$$

An invertible homomorphism is called an *isomorphism*.

If $f\colon M \longrightarrow N$ is a module homomorphism, then its image

$$\operatorname{Im} f = \{f(x)\colon x \in M\} \subset N$$

is a submodule of N and the kernel

$$\operatorname{Ker} f = \{x \in M\colon f(x) = 0\} \subset M$$

is a submodule of M.

For any submodule $N \subset M$, we define the *canonical homomorphism*

$$\pi\colon M \to M/N, \qquad x \mapsto x + N,$$

whose kernel is N.

Theorem 9.78 (Module Homomorphism Theorem). *Let $f\colon M \to N$ be a homomorphism of A-modules. Then*

$$\operatorname{Im} f \simeq M/\operatorname{Ker} f.$$

More precisely, there exists an isomorphism

$$\varphi\colon \operatorname{Im} f \xrightarrow{\sim} M/\operatorname{Ker} f$$

that maps an element $y = f(x) \in \operatorname{Im} f$ to the coset $\varphi(x) = x + \operatorname{Ker} f$.

Proof. We already know from Theorem 4.100 that the map φ is an isomorphism of additive groups. It remains to check that it commutes with multiplications by elements of the ring A. Let $f(x) = y$. Then $f(ax) = ay$ for $a \in A$ and

$$\varphi(ay) = \pi(ax) = a\pi(x) = a\varphi(y).$$

□

Consider an A-module M.

For any subset $S \subset M$, the collection of all linear combinations

$$a_1x_1 + \cdots + a_kx_k, \qquad x_i \in S,\ a_i \in A,$$

is the smallest submodule containing S. It is called the *submodule generated by the subset* S and is denoted $\langle S \rangle$. If $\langle S \rangle = M$, we say that M is *generated* by S or that S is the *generating set* of the module M. A module that allows a finite generating set is called *finitely generated.*

A module generated by a single element is called *cyclic.*

The ideal

$$\operatorname{Ann} M = \{a \in A\colon aM = 0\}$$

is called the *annihilator* of M. If $\operatorname{Ann} M \neq 0$, the module is called *periodic.*

Theorem 9.79. *Every cyclic A-module M is isomorphic to a module A/I, where I is a left ideal of the ring A. If A is commutative, I coincides with $\operatorname{Ann} M$, and thus is defined by M uniquely.*

Proof. Let $M = \langle x \rangle$ be a cyclic A-module. The map

$$f\colon A \longrightarrow M, \qquad a \mapsto ax,$$

is a module homomorphism and $\operatorname{Im} f = M$. By the Module Homomorphism Theorem, $M \simeq A/I$, where $I = \operatorname{Ker} f$. The second claim is obvious. □

A system $\{x_1, \ldots, x_n\}$ of elements of M is called *linearly independent* if $a_1x_1 + \cdots + a_nx_n = 0$ for $a_i \in A$ only if $a_1 = \cdots = a_n = 0$. A linearly independent generating system is called a *basis.*

A finitely generated module that has a basis is called *free.* A free cyclic module is isomorphic to A (as an A-module).

Just as for finitely generated abelian groups, one can build up a theory for finitely generated modules over principal ideal domains.

From this point on, we assume that A is a *principal ideal domain.*

Theorem 9.80. *All bases of a free A-module L contain the same number of elements.*

Proof. If A is a field, then the assertion of this theorem is known. If A is not a field, let p be a prime element of A. Then $A/(p)$ is a field and L/pL is a vector space over this field. If $\{e_1, \ldots, e_n\}$ is a basis of L, then $\{[e_1], \ldots, [e_n]\}$ is a basis of this vector space (here $[x]$ denotes the coset $x + pL$). Thus, $n = \dim L/pL$. □

The number of elements in a basis of a free module L is called its *rank* and is denoted $\operatorname{rk} L$.

Theorem 9.81. *Every submodule N of a free A-module L of rank n is a free A-module of rank $m \leq n$. Moreover, there exists a basis $\{e_1, \ldots, e_n\}$ of L and (nonzero) elements $u_1, \ldots, u_m \in A$ such that $\{u_1 e_1, \ldots, u_m e_m\}$ is a basis of the submodule N and $u_i | u_{i+1}$ for $i = 1, \ldots, m-1$.*

Proof. For $n > 1$, the first assertion of the theorem is the definition of the principal ideal domain. For $n > 1$, it is proved just as in the case $A = \mathbb{Z}$ (see Theorem 9.5).

Also, as in the case $A = \mathbb{Z}$, the proof of the second assertion is based on reducing the transition matrix C from a basis of L to a basis of N to the diagonal form using elementary transformations of these bases.

When A is a Euclidean domain, an elementary transformation of a basis of a free A-module is either of the following:

(i) adding a basis element multiplied by an element of A to another basis element;

(ii) interchanging two basis elements;

(iii) multiplying a basis element by an invertible element of A.

In this case, reduction of C to the diagonal form is performed just as in the proof of Proposition 9.13 with the only difference that we have to minimize not the entry c_{11} (which makes no sense) but its norm.

In the general case, the notion of an elementary transformation should be generalized. Let $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be an invertible matrix with entries from A. (A matrix is invertible if and only if its determinant is invertible.) Consider a system of elements $\{x_1, \ldots, x_p\}$ of an A-module M. Call a *quasi-elementary transformation* the replacement of two elements x_i and x_j by the linear combinations

$$ax_i + bx_j, \quad cx_i + dx_j.$$

Clearly, the inverse of a quasi-elementary transformation is also quasi-elementary. Also, elementary transformations are quasi-elementary.

Using a quasi-elementary transformation, every pair of elements $\{x, y\}$ of the ring A can be reduced to a pair of the form $\{d, 0\}$, where $d = (x, y)$. Indeed, there exist $a, b \in A$ such that $ax + by = d$. Consider the matrix

$\begin{pmatrix} a & b \\ -y/d & x/d \end{pmatrix}$. It is invertible since its determinant equals 1. The corresponding quasi-elementary transformation maps $\{x, y\}$ to $\{d, 0\}$.

Therefore, if a row or a column of the matrix C contains entries x, y, then using a quasi-elementary transformation of rows or columns, we can obtain entries $d, 0$. If we follow the general outline of the proof of Proposition 9.13, such transformations suffice to reduce C to diagonal form. □

Now we will study the structure of an arbitrary finitely generated A-module.

Every nontrivial cyclic A-module is isomorphic to either A or $A/(u)$ for a noninvertible nonzero element u.

If $(u, v) = 1$, it is easily seen that the ring isomorphism

$$A/(u, v) \xrightarrow{\sim} A/(u) \oplus A/(v)$$

from the proof of Theorem 9.68 is an isomorphism of A-modules. Therefore, if $u = p_1^{k_1} \cdots p_s^{k_s}$ is the prime factorization of u, we obtain the following isomorphism:

$$A/(u) \simeq A/(p_1^{k_1}) \oplus \cdots \oplus A/(p_s^{k_s}). \tag{9.18}$$

Definition 9.82. A finitely generated A-module M whose annihilator contains a power of a prime element $p \in A$ is *primary* (or, rather, *p-primary*).

Thus, every periodic cyclic A-module decomposes into a direct sum of primary cyclic submodules.

Theorem 9.83. *Every finitely generated A-module M decomposes into a direct sum of primary and free cyclic submodules. Moreover, the collection of annihilators of these submodules is defined uniquely.*

Proof. The proof of this theorem is analogous to that of Theorem 9.28. In particular, the existence of such a decomposition follows from Theorem 9.81 and isomorphism (9.18). To prove uniqueness (of the annihilators), one should consider the *torsion submodule*

$$\operatorname{Tor} M := \{x \in M : ax = 0 \text{ for some } a \in A,\ a \neq 0\}$$

and, for every prime $p \in A$, the *p-torsion submodule*

$$\operatorname{Tor}_p M := \{x \in M : p^k x = 0 \text{ for some } k \in \mathbb{Z}_+\}.$$

As in the case of abelian groups, uniqueness of the decomposition of a primary module into a direct sum of primary cyclic submodules is proved by induction. However, our original approach that used the group order does not work in the general case. Instead, we should use the following one: if a module M decomposes into a direct sum of p-primary cyclic submodules,

then the number of the summands equals the dimension of the submodule $\{x \in M : px = 0\}$ regarded as a vector space over the field $A/(p)$. □

As in the case of abelian groups, for any periodic A-module M, we obtain from the proof of Theorem 9.83 that

$$M \simeq A/(u_1) \oplus \cdots \oplus A/(u_m), \tag{9.19}$$

where $u_1, \ldots, u_m$ are noninvertible nonzero elements of the ring A such that $u_i | u_{i+1}$ for $i = 1, \ldots, m-1$. The elements $u_1, \ldots, u_m$ are determined uniquely up to multiplication by invertible elements. They are called *invariant factors* of the module M. Obviously,

$$\operatorname{Ann} M = (u_m). \tag{9.20}$$

In the case $A = K[t]$ (K a field), Theorem 9.83 describes the structure of linear operators on vector spaces over the field K (cf. Example 9.78). If the given vector space is finite-dimensional, we obviously deal with the finitely generated situation. Moreover, in this case, there are no free direct summands, since a free cyclic module over $K[t]$ is infinite-dimensional over K. The result looks especially simple if K is algebraically closed. Here primary cyclic modules have the form

$$K[t]/((t-\lambda)^m), \qquad \lambda \in K.$$

Such a module is an m-dimensional vector space over K with the basis

$$\{[(t-\lambda)^{m-1}], \ldots, [t-\lambda], [1]\},$$

where $[f(t)]$ denotes the class $f(t) + ((t-\lambda)^m)$. In this basis, the matrix of the operator of multiplication by t is the Jordan block

$$J(\lambda) = \underbrace{\begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & 0 \\ & & \ddots & \ddots \\ & 0 & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}}_{m}.$$

This implies

Theorem 9.84. *Every linear operator on a finite-dimensional vector space over an algebraically closed field has a Jordan canonical form in some basis. Moreover, this form is determined uniquely up to a permutation of the diagonal blocks.*

Recall that we have already proven the first claim of this theorem in Section 6.4.

It follows from (9.20) that the last invariant factor of a $K[t]$-module associated with a linear operator $\mathcal{A}$ is the minimal polynomial of $\mathcal{A}$ (see Theorem 6.65).

Exercise 9.85. Prove that in the basis $\{[t^{n-1}], [t^{n-2}], \dots, [t], [1]\}$, the matrix of the operator of multiplication by t on a $K[t]$-module $K[t]/(h(t))$, where $h(t) = t^n + a_1 t^{n-1} + \cdots + a_{n-1}t + a_n$, is

$$\begin{pmatrix} -a_1 & 1 & 0 & \dots & 0 & 0 \\ -a_2 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -a_{n-1} & 0 & 0 & \dots & 0 & 1 \\ -a_n & 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Also prove that its characteristic polynomial is $h(t)$. Conclude that the product of invariant factors of the $K[t]$-module associated to a linear operator $\mathcal{A}$ is equal to the characteristic polynomial of $\mathcal{A}$.

Exercise 9.86. Deduce the Cayley–Hamilton theorem (Corollary 6.69) from the above exercise.

Exercise 9.87. Find a canonical form for the matrix of a linear operator over the field of real numbers.

Exercise 9.88. Find a canonical form for the matrix of a linear operator on the four-dimensional vector space over the field $\mathbb{Z}_2$.

9.4. Noetherian Rings

For the rest of this chapter the word "ring" will stand for "*commutative associative ring with unity.*" Subrings are assumed to contain the unity, and homomorphisms, to map a unity to a unity.

A natural generalization of the class of principal ideal domains is the class of Noetherian rings.

Definition 9.89. A ring A is *Noetherian* if either of the following equivalent conditions holds:

(i) every ideal is generated by a finite number of elements;

(ii) every strictly ascending chain of ideals is finite, i.e., there exist no infinite chains of ideals $I_1 \subset I_2 \subset \cdots \subset I_n \subset \cdots$ $(I_n \neq I_{n+1})$.

Recall that we say that an ideal I is generated by elements $x_1, \dots, x_n \in A$ if $I = \{a_1x_1 + \cdots + a_nx_n : a_1, \dots, a_n \in A\}$. We denote this $I = (x_1, \dots, x_n)$.

The equivalence of conditions (i) and (ii) is proved as follows. Let $I_1 \subset I_2 \subset \cdots$ be an ascending chain of ideals. Then $I = \bigcup_{n=1}^{\infty} I_n$ is also an ideal. If it is generated by a finite number of elements, then they all belong to the ideal I_n for some sufficiently large n. Hence, $I = I_n$ and the chain is not strictly ascending.

Conversely, if an ideal I is not generated by a finite number of elements, then there exists a sequence of elements $x_1, x_2, \ldots \in I$ such that the sequence of ideals

$$(x_1) \subset (x_1, x_2) \subset \cdots$$

is strictly ascending.

Proposition 9.90. *Every quotient ring A/I of a Noetherian ring A is Noetherian.*

Proof. Let J be an ideal of the ring A/I. Then its full preimage under the canonical homomorphism $\pi : A \to A/I$ is an ideal of the ring A. If this ideal is generated by elements $x_1, \ldots, x_n \in A$, then J is generated by the elements $\pi(x_1), \ldots, \pi(x_n)$. □

The structure of finitely generated modules over Noetherian rings is not as simple as that of modules over principal ideal domains. However, the following theorem demonstrates that in some sense they still resemble finite-dimensional vector spaces.

Theorem 9.91. *Every submodule N of a finitely generated module M over a Noetherian ring A is finitely generated.*

Remark 9.92. In the case $M = A$ (i.e., when M is a free cyclic module), the assertion of the theorem is just the definition of a Noetherian ring.

Proof of Theorem 9.91. Let $M = \langle x_1, \ldots, x_n \rangle$. We will prove this theorem by induction on n.

For $n = 1$, we can assume that $M = A/I$, where I is an ideal of A (see Theorem 9.79). Then N is an ideal of the ring A/I and the assertion of the theorem follows from Proposition 9.90.

For $n > 1$, consider the submodule $M = \langle x_1, \ldots, x_{n-1} \rangle \subset M$ and set $N_1 = N \cap M_1$. By the induction hypothesis, N_1 is finitely generated. Let $N_1 = \langle y_1, \ldots, y_k \rangle$. The quotient module N/N_1 is a submodule of the cyclic module M/M_1, hence is finitely generated by the above. Let $N/N_1 = \langle z_1 + N_1, \ldots, z_l + N_1 \rangle$. Then $N = \langle y_1, \ldots, y_k, z_1, \ldots, z_l \rangle$. □

How does one show that a ring is Noetherian? One of the basic tools here is provided by the following theorem.

Theorem 9.93 (Hilbert's Basis Theorem). *The polynomial ring $A[x]$ over a Noetherian ring A is Noetherian itself.*

Proof. Consider an ideal I of the ring $A[x]$. Denote by $A[x]_n$ the set of polynomials of degree $\leq n$. This is a free A-module with a basis $\{1, x, \dots, x^n\}$. Put $I_n = I \cap A[x]_n$. By Theorem 9.91, I_n is a finitely generated A-module. Obviously, $I = \bigcup_{n=0}^{\infty} I_n$ and $xI_n \subset I_{n+1}$.

Denote by J_n the set of coefficients of x^n in all polynomials of I_n. It is clear that this is an ideal of A and that $J_n \subset J_{n+1}$. Since A is Noetherian, there exists m such that $J_n = J_m$ for all $n \geq m$. Therefore, for any polynomial $f \in I_n$, $n \geq m$, there exists $g \in I_m$ such that $f - x^{n-m}g \in I_{n-1}$. This shows that the ideal I of the ring $A[x]$ is generated by the subset I_m. Thus, if I_m is generated by polynomials $f_1, \dots, f_k$ as an A-module, then I is generated by the same polynomials as an $A[x]$-module. □

Corollary 9.94. *The ring of polynomials in any number of variables over a Noetherian ring is Noetherian.*

We say that a ring B is generated by elements $u_1, \dots, u_n$ over a subring A if every element of B can be presented as a polynomial in $u_1, \dots, u_n$ with coefficients in A. In this case there exists a homomorphism

$$f\colon A[x_1, \dots, x_n] \xrightarrow{\text{onto}} B, \qquad x_i \mapsto u_i$$

(where $A[x_1, \dots, x_n]$ stands for the polynomial ring in $x_1, \dots, x_n$ with coefficients in A) and hence,

$$B \simeq A[x_1, \dots, x_n]/\operatorname{Ker} f.$$

This is usually denoted $B = A[u_1, \dots, u_n]$, even though this does not mean that B is a polynomial ring in n independent variables ($u_1, \dots, u_n$ may be algebraically dependent).

Corollary 9.95. *Every ring that is finitely generated over a Noetherian ring is Noetherian itself.*

Remark 9.96. There also exists the following "absolute" version of Corollary 9.95 which does not depend on any particular subring.

A ring is said to be generated by elements $u_1, \dots, u_n$ if every element of this ring can be presented as a polynomial in $u_1, \dots, u_n$ with integer coefficients. In this case, it is isomorphic to a quotient ring of the ring $\mathbf{Z}[x_1, \dots, x_n]$, hence is Noetherian. Thus, every finitely generated ring is Noetherian.

When we work with rings, zero divisors (if they exist) often cause problems. There are methods to fight them. The "most awful" zero divisors are the nilpotent elements.

An element a of a ring A is called *nilpotent* if $a^m = 0$ for some natural number m. It is easy to see that the set of all nilpotent elements is an ideal of A. It is called the (*nilpotent*) *radical* of A and is denoted $\operatorname{rad} A$. The quotient ring $A/\operatorname{rad} A$ has no nilpotent elements (except for 0).

Example 9.97. Let A be a principal ideal domain. We will describe here $\operatorname{rad}(A/(u))$ for a nonzero noninvertible element $u \in A$. Let $u = p_1^{k_1} \cdots p_s^{k_s}$ be the factorization of u into primes. The element $a + (u) \in A/(u)$ is nilpotent if and only if $a^n \in (u)$ for some natural n. The uniqueness of prime factorization in A implies that this happens if and only if a is divisible by $p_1 \cdots p_s$. Therefore,

$$\operatorname{rad}(A/(u)) = (p_1 \cdots p_s)/(u).$$

Exercise 9.98. Prove that

$$\operatorname{rad}(A_1 \oplus \cdots \oplus A_k) = \operatorname{rad} A_1 \oplus \cdots \oplus \operatorname{rad} A_k.$$

An ideal of a ring A that is not equal to A is called *proper*.

Definition 9.99. A proper ideal I of a ring A is *prime* if the quotient ring A/I contains no zero divisors.

In other words, $ab \in I$ must imply that either $a \in I$ or $b \in I$.

For example, in a principal ideal domain A a nonzero ideal (p) is prime if and only if p is prime.

A proper ideal I of a ring A is called *maximal* if it is not contained in any larger proper ideal. The second definition of a Noetherian ring implies that a Noetherian ring has at least one maximal ideal.

Proposition 9.100. *An ideal I of a ring A is maximal if and only if the quotient ring A/I is a field.*

Proof. Obviously, the ideal I is maximal if and only if the quotient ring A/I has no nontrivial ideals. We know that a field does not contain such ideals (see Example 9.41). Conversely, let a ring K have no nontrivial ideals. Then for any nonzero element $a \in K$, the ideal (a) coincides with K. In particular, it contains 1, which means exactly that a is invertible. Thus, K is a field. □

Corollary 9.101. *Every maximal ideal is prime.*

Theorem 9.102. *The radical of a Noetherian ring coincides with the intersection of all its prime ideals.*

Proof. Clearly, the radical is contained in the intersection of all prime ideals. To prove the opposite inclusion, we have to check that if a is not nilpotent, then there exists a prime ideal that does not contain a.

If an element a is not nilpotent, then we can construct the ring $A' = A[a^{-1}]$ of fractions of the type b/a^n, $b \in A, n \in \mathbb{Z}_+$, just as we constructed the field of quotients of an integral domain in Section 3.10. By Corollary 9.95, the ring A' is Noetherian. Hence, it has a maximal ideal I'. Since a is invertible in A', $a \notin I'$. Set $I = I' \cap A$. The ring A/I embeds into the field A'/I', thus contains no zero divisors. Therefore, I is a prime ideal of A that does not contain a. □

Remark 9.103. Using transfinite tools (e.g., the Zorn lemma), it is easy to prove that every (not necessarily Noetherian) ring contains a maximal ideal. It follows that Theorem 9.102 holds, in fact, for all rings.

9.5. Algebraic Extensions

When a ring A is a subring of a ring B, we say that B is an *extension* of A. In this case, we deal not just with a ring but also with an algebra over A; this suggests how we may study B further. (The definition of an algebra over a ring is the same as that over a field.)

Let us introduce terminology that helps to describe this situation.

An element $u \in B$ is called *algebraic over* A if it satisfies a nontrivial algebraic equation with coefficients in A; otherwise we call u *transcendental*. In particular, every element $a \in A$ is algebraic over A since it satisfies the linear equation $x - a = 0$. The ring B is called an *algebraic extension* of A if every element of B is algebraic over A.

More generally, elements $u_1, \dots, u_n \in B$ are called *algebraically dependent* over A if they satisfy a nontrivial algebraic equation (in n indeterminates) with coefficients in A.

The set of elements of B that can be presented as $f(u_1, \dots, u_n)$ for a polynomial f with coefficients in A, is a subring (containing A). It is called the *subring generated over* A *by* $u_1, \dots, u_n$ and is denoted $A[u_1, \dots, u_n]$. If $u_1, \dots, u_n$ are algebraically independent, then this ring is isomorphic to the polynomial ring over A in n variables. In general, it is isomorphic to the quotient ring of the polynomial ring by the ideal of algebraic dependences of $u_1, \dots, u_n$. An extension B of a ring A is called *finitely generated* if there exist elements $u_1, \dots, u_n \in B$ such that $B = A[u_1, \dots, u_n]$.

If the ring B is an integral domain (and then, so is A), we can consider quotient fields $K = Q(A)$ and $L = Q(B)$ and assume that the events take place inside the "big" field L. The following diagram illustrates this approach:

$$\begin{array}{ccc} A & \subset & B \\ \cap & & \cap \\ K & \subset & L \end{array} \tag{9.21}$$

If elements $u_1, \ldots, u_n \in L$ are algebraically dependent over K, they are algebraically dependent over A, since the coefficients in the statement of dependence can be made "integral," i.e., belonging to A, by multiplying the statement of dependence by a common denominator.

Consider, first, algebraic extensions of fields.

The key to their understanding is the concept of a finite extension that we introduce below. The main idea in the proofs of next statements lies in the fact that a subspace of a finite-dimensional space is finite-dimensional.

If a field L is an extension of a field K, it can be regarded as a vector space over K. The dimension of this vector space is denoted $\dim_K L$.

Definition 9.104. An extension L of a field K is *finite* if $\dim_K L < \infty$. The value of $\dim_K L$ is called the *degree* of the extension L.

The following theorem suggests how one obtains finite extensions.

Theorem 9.105. *Let $h \in K[x]$ be an irreducible polynomial of degree n. Then $L = K[x]/(h)$ is a finite extension of K. Moreover, $\dim_K L = n$.*

Proof. That L is a field follows from the general Theorem 9.64. Since it is possible to uniquely divide with a remainder in $K[x]$, every element of L can be uniquely presented as

$$a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + (h), \qquad a_0, a_1, \ldots, a_{n-1} \in K.$$

This means that the cosets

$$1 + (h),\ x + (h),\ \ldots,\ x^{n-1} + (h)$$

form a basis of L over K. □

The element $\alpha = x + (h) \in L$ is clearly a root of h in L. Furthermore, $L = K[\alpha]$. For this reason, when we pass from K to L, we say that we *adjoin* to K a *root of the irreducible polynomial h.*

Extensions of this kind are called *simple*. We will demonstrate in Section 11.6 that every finite extension of a field of zero characteristic is simple (the Primitive Element Theorem). However, this fact does not help much for the discussion below—this is why we delay the proof (and the use) of this theorem.

Example 9.106. If $a \in K$ is an element which is not a square in K, then the field $K[\sqrt{a}]$ obtained by adjoining to K a root of the polynomial $x^2 - a$ is an extension of degree 2, also called a *quadratic extension* of K. In particular, $\mathbb{R}[\sqrt{-1}] = \mathbb{C}$.

Let L be an extension of K.

If an element $u \in L$ is algebraic over K, the set of all polynomials $f \in K[x]$ such that $f(u) = 0$ is a nonzero ideal of the ring $K[x]$. The generating element of this ideal is called the *minimal polynomial* of u and is denoted m_u (cf. the definition of the minimal polynomial of a linear operator in Section 6.5). Observe that the minimal polynomial is irreducible. Indeed, if $m_u = fg$, then either $f(u) = 0$ or $g(u) = 0$, thus the degree of either f or g equals the degree of m_u. The degree of m_u is called the *degree* of u over K.

Theorem 9.107. *An element $u \in L$ is algebraic over K if and only if $K[u]$ is a finite-dimensional vector space over K. With this condition, $K[u]$ is a field and its dimension (over K) equals the degree of u over K.*

Proof. If the space $K[u]$ is finite-dimensional over K, then it is generated by a finite number of powers of u. Hence, there exists n such that u^n can be expressed as a linear combination of lower powers. Thus, u is algebraic over K.

Conversely, assume that u is an algebraic element of degree n over K. Then u^n can be expressed as a linear combination of lower powers of u. By multiplying this expression successively by u and replacing u^n with its expression in terms of lower powers, we see that every power of u, hence every element of $K[u]$, can be expressed as a linear combination of $1, u, \ldots, u^{n-1}$. Therefore, $\dim_K K[u] \leq n$.

More precisely, consider the homomorphism

$$\varphi\colon K[x] \to L, \qquad f \mapsto f(u).$$

Its image is $K[u]$ and its kernel is the ideal generated by the minimal polynomial m_u of u. Thus,

$$K[u] \simeq K[x]/(m_u).$$

Since the polynomial m_u is irreducible, Theorem 9.105 implies that $K[u]$ is a field and that its dimension over K is $\deg m_u = n$. □

Corollary 9.108. *Every finite field extension is algebraic.*

Example 9.109. Let p be a prime number. Since the number $\varepsilon_p = \cos\frac{2\pi}{p} + i\sin\frac{2\pi}{p} \in \mathbb{C}$ is a root of the polynomial $x^{p-1} + \cdots + x + 1$, which is irreducible over $\mathbb{Q}$ (see Example 3.72), $\mathbb{Q}[\varepsilon_p]$ is an extension of $\mathbb{Q}$ of degree $p-1$. It contains all pth complex roots of unity. This field is called *cyclotomic*.

Theorem 9.110. *If L is a finite extension of a field K and M is a finite extension of L, then M is a finite extension of K. Moreover,*

$$\dim_K M = \dim_K L \cdot \dim_L M.$$

Proof. If $\{e_i\}$ is a basis of L over K and $\{f_j\}$ is a basis of M over L, then $\{e_i f_j\}$ is a basis of M over K. □

For any elements $u_1, \ldots, u_n \in L$, the set of elements of L that can be presented as ratios of elements of the ring $K[u_1, \ldots, u_n]$ is a subfield isomorphic to $Q(K[u_1, \ldots, u_n])$. It is called the *subfield generated over K by the elements $u_1, \ldots, u_n$* and is denoted $K(u_1, \ldots, u_n)$. In particular, if $u \in L$ is algebraic over K, then by Theorem 9.107, $K(u) = K[u]$ (this is the phenomenon of "cancellation of irrationality in the denominator").

If $K(u_1, \ldots, u_n) = L$, we say that the field L is generated over K by the elements $u_1, \ldots, u_n$.

Theorem 9.111. *If a field L is generated over K by a finite number of algebraic elements $u_1, \ldots, u_n$, then it is a finite extension of K.*

Proof. Consider the "tower of extensions"

$$K \subset K(u_1) \subset K(u_1, u_2) \subset \cdots \subset K(u_1, \ldots, u_n) = L.$$

Since $K(u_1, \ldots, u_m) = K(u_1, \ldots, u_{m-1})(u_m)$ and since u_m is algebraic over $K(u_1, \ldots, u_{m-1})$ (it is already algebraic over K), every "story" of this tower is a finite extension. By Theorem 9.110, L is a finite extension of K. □

Theorem 9.112. *Consider an extension L of a field K. The set $\bar{K}$ of all elements of L that are algebraic over K is a subfield, which is algebraically closed in L.*

(The latter means that every element of L which is algebraic over $\bar{K}$ belongs to $\bar{K}$, i.e., is already algebraic over K.)

Proof. If $u, v \in \bar{K}$, then by Theorem 9.111, $K(u, v) \subset \bar{K}$. In particular,

$$u + v,\ uv,\ u^{-1} \in \bar{K}.$$

This implies that $\bar{K}$ is a subfield of L.

Let $u \in L$ be algebraic over $\bar{K}$ and let $u_1, \ldots, u_n \in \bar{K}$ be the coefficients of an algebraic equation with a root u. By Theorem 9.111, $K' = K(u_1, \ldots, u_n)$ is a finite extension of K. Since u is algebraic over K', the field $K'(u)$ is a finite extension of K'. Hence, $K'(u)$ is a finite extension of K, thus $K'(u) \subset \bar{K}$. In particular, $u \in \bar{K}$. □

The field $\bar{K}$ is called the *algebraic closure of K in L*.

For instance, the *field of all algebraic numbers* is the algebraic closure $\overline{\mathbb{Q}}$ of the field $\mathbb{Q}$ in $\mathbb{C}$. Since the field $\mathbb{C}$ is algebraically closed, so is $\overline{\mathbb{Q}}$ (in the absolute sense and not just in $\mathbb{C}$). A finite extension of $\mathbb{Q}$ is called a *field of algebraic numbers* (there are many such fields). It is not difficult to

show that every field of algebraic numbers is isomorphic to a subfield of $\overline{\mathbf{Q}}$ (do this!).

In a simple extension of a field K obtained by adjoining a root of an irreducible polynomial f, this polynomial does not necessarily have more than one root (though it might). In general, if we want to obtain a field where f decomposes into linear factors, we need to construct further extensions.

Definition 9.113. A *splitting field* of a polynomial $f \in K[x]$ (not necessarily irreducible) is an extension L of K such that f decomposes into linear factors in $L[x]$ and L is generated over K by the roots of f.

A homomorphism (and, in particular, an isomorphism) of an extension of K acting trivially on K is called a *homomorphism* (an *isomorphism*) *over* K.

Theorem 9.114. *Any polynomial $f \in K[x]$ has a splitting field. Such a field is unique up to an isomorphism over K.*

To prove the second part of this theorem, we need the following

Lemma 9.115. *Let $P(\alpha)$ be an extension of a field P obtained by adjoining a root α of an irreducible polynomial $h \in P[x]$. Let φ be a homomorphism of P into a field F. The number of extensions of the homomorphism φ to a homomorphism $\psi\colon P(\alpha) \to F$ equals the number of roots of the polynomial h^φ in F (this polynomial is obtained from h by applying φ to each coefficient).*

Proof. Such an extension of ψ, if exists, is given by the following formula:

$$\psi(a_0 + a_1\alpha + \cdots + a_m\alpha^m) = \varphi(a_0) + \varphi(a_1)\beta + \cdots + \varphi(a_m)\beta^m, \quad a_0, a_1, \ldots, a_m \in P, \tag{9.22}$$

where $\beta = \psi(\alpha)$ is an element of F. Applying this formula to the equality $h(\alpha) = 0$, we obtain that $h^\varphi(\beta) = 0$. Conversely, if $\beta \in F$ is a root of the polynomial h^φ, then formula (9.22) defines a homomorphism $\psi\colon P(\alpha) \to F$. □

Proof of Theorem 9.114. Consider the following sequence of extensions:

$$K = K_0 \subset K_1 \subset K_2 \subset \cdots,$$

where K_i is obtained from K_{i-1} by adjoining a root of an irreducible factor f_i of f over K_{i-1} such that the degree of f_i is > 1. Since the number of irreducible factors of f increases at every step, this sequence cannot be infinite. Its last term $K_s = L$ is a splitting field of f.

Now let $\widetilde{L}$ be another splitting field. Let us construct a sequence of homomorphisms

$$\varphi_i\colon K_i \to \widetilde{L}, \qquad i = 0, 1, \ldots, s,$$

such that

$$\varphi_0 = \mathrm{id}, \qquad \varphi_i|_{K_{i-1}} = \varphi_{i-1}.$$

By the lemma, the ith step of this construction is possible if the polynomial $\widetilde{f}_i = f_i^{\varphi_{i-1}}$ has a root in $\widetilde{L}$. Since f_i divides f in $K_{i-1}[x]$, $\widetilde{f}_i$ divides f in $\widetilde{L}[x]$. Thus, homomorphisms φ_i exist. The last of them,

$$\varphi_s = \varphi\colon L \to \widetilde{L},$$

is an isomorphism since, according to the definition of a splitting field, the field $\widetilde{L}$ is a minimal extension of K where f decomposes into linear factors. □

Example 9.116. Let us find the degree of the splitting field L of the cubic polynomial

$$f = x^3 + a_1x^2 + a_2x + a_3 \in K[x], \qquad \mathrm{char}\, K \neq 2.$$

Consider three possible cases:

(i) f has three roots in K. Then $L = K$.

(ii) f has one root in K. Then L is a quadratic extension of K.

(iii) f has no roots in K, hence is irreducible over K. Now let $K_1 \supset K$ be the cubic extension obtained by adjoining a root α_1 of f to K. Two cases are possible:

(a) f has three roots in K_1; then $L = K_1$;

(b) f has only one root in K_1; then K is a quadratic extension of K_1, hence, $\dim_K L = 6$.

To distinguish cases (iii)(a) and (iii)(b) consider the discriminant of f. By definition, it equals

$$D = (\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2,$$

where $\alpha_1, \alpha_2, \alpha_3$ are the roots of f in L. (For the expression of D in terms of coefficients of f, see Section 3.9.) Let us prove that if f has no roots in K, then $\dim_K L = 3$ if and only if $D \in K^2$.

Observe that if $D \notin K^2$, then $D \notin K_1^2$, otherwise $K(\sqrt{D})$ would be a quadratic extension of K contained in K_1, which is impossible by the formula for the multiplication of dimensions in an extension tower (Theorem 9.110). Thus,

$$D \in K^2 \iff D \in K_1^2 \iff (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3) \in K_1.$$

Now, since $\alpha_1 \in K_1$ and α_2 and α_3 are the roots of a quadratic polynomial with coefficients from K_1, we have

$$(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3) \in K_1,$$

hence

$$D \in K^2 \iff \alpha_2 - \alpha_3 \in K_1 \iff \alpha_2, \alpha_3 \in K_1 \iff L = K_1$$

(here we used that $\operatorname{char} K \neq 2$).

Now we use Theorem 9.114 to describe all finite fields.

A finite field F has characteristic $p > 0$, a prime number. Its cyclic additive subgroup generated by the unity is a subfield isomorphic to the field of residue classes $\mathbb{Z}_p$. We identify this subfield with $\mathbb{Z}_p$. If $\dim_{\mathbb{Z}_p} F = n$, then

$$|F| = p^n.$$

Thus, the number of elements of any finite field is a power of a prime number.

Theorem 9.117. *For any prime p and any natural n, there exists a field with p^n elements. All fields with p^n elements are isomorphic.*

The proof of this theorem requires some preparation.

Let F be a (possibly infinite) field of characteristic $p > 0$. Consider the map

$$\varphi\colon F \to F, \qquad x \mapsto x^p.$$

Obviously, $\varphi(xy) = \varphi(x)\varphi(y)$. Moreover, as strange as it looks, $\varphi(x + y) = \varphi(x) + \varphi(y)$. Indeed, as we saw in Section 1.6,

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

Thus, φ is an endomorphism (a homomorphism into itself) of the field F. It is called the *Frobenius endomorphism*.

Since $\operatorname{Ker} \varphi = 0$, $\operatorname{Im} \varphi = F^p \simeq F$. Clearly, for a finite field, $F^p = F$, so here the Frobenius endomorphism is an automorphism.

Proof of Theorem 9.117. Let F be a finite field containing $q = p^n$ elements. Since the order of the multiplicative group F^* is $q - 1$, $a^{q-1} = 1$ for every $a \in F^*$. Therefore,

$$a^q = a \qquad \forall a \in F.$$

In other words, all elements of F are roots of the polynomial $x^q - x$. Therefore, F is a splitting field of this polynomial over $\mathbb{Z}_p$. By Theorem 9.114, this implies that all fields of q elements are isomorphic.

On the other hand, let F be the splitting field of the polynomial $f = x^q - x$ over $\mathbb{Z}_p$. Since $f' = -1$, this polynomial has no multiple roots. Its roots are the fixed points of the automorphism φ^n of F, where φ is the Frobenius automorphism. It is easy to see that the fixed points of any automorphism form a subfield. Therefore, the collection of roots of f is a

subfield of F containing q elements (hence it coincides with F). This proves that such a field exists. □

Corollary 9.118. *For any prime p and any natural n, there exists an irreducible polynomial of degree n over $\mathbb{Z}_p$.*

Proof. Let F be a field with $q = p^n$ elements and α, a generator of its multiplicative group (which, as we know, is cyclic). Then $F = \mathbb{Z}_p(\alpha)$, hence the minimal polynomial of α over $\mathbb{Z}_p$ has degree n. □

The field with q elements is denoted $\mathbb{F}_q$. (In particular, for a prime q, $\mathbb{F}_p = \mathbb{Z}_p$.)

Example 9.119. The only irreducible polynomial of the second degree over the field $\mathbb{Z}_2$ is the polynomial $x^2 + x + 1$. Adjoining to $\mathbb{Z}_2$ its root, we obtain the field $\mathbb{F}_4$.

Exercise 9.120. Write down the addition and the multiplication table for the field $\mathbb{F}_4$.

Let us regard a finite extension L of a field K as a vector space over K. There is a natural way to introduce an inner product on L.

Namely, for any $u \in L$, define the linear operator $T(u)$ on the space L:

$$T(u)x = ux, \qquad x \in L.$$

Call the trace of this operator the *trace* of u; denote it by $\operatorname{tr} u$. Clearly, the trace is a linear function on L. Define the inner product on L as

$$(u, v) = \operatorname{tr} uv. \tag{9.23}$$

This is a symmetric bilinear function on L. If $\operatorname{char} K = 0$, this function is nondegenerate because

$$(u, u^{-1}) = \operatorname{tr} 1 = \dim_K L \neq 0$$

for every nonzero element $u \in L$.

Example 9.121. We describe here the inner product on the cyclotomic field $\mathbb{Q}(\varepsilon_p) = \mathbb{Q}[\varepsilon_p]$ (see Example 9.109). As a vector space over $\mathbb{Q}$, the field $\mathbb{Q}(\varepsilon_p)$ is generated by the elements $1, \varepsilon_p, \varepsilon_p^2, \ldots, \varepsilon_p^{p-1}$ whose sum equals zero. As a basis of this space take, for instance, the elements $1, \varepsilon_p, \varepsilon_p^2, \ldots, \varepsilon_p^{p-2}$. By determining the matrices of operators $T(\varepsilon_p^k)$ in this basis, it is easy to see that

$$\operatorname{tr} 1 = p - 1, \quad \operatorname{tr} \varepsilon_p^k = -1, \quad k = 1, \ldots, p-1.$$

Therefore,

$$(\varepsilon_p^k, \varepsilon_p^l) = \begin{cases} p - 1, & \text{for } k + l \equiv 0 \pmod{p}, \\ -1, & \text{otherwise.} \end{cases}$$

The inner product of any two elements of the field $\mathbb{Q}(\varepsilon_p)$ can be easily determined if one of them is expressed as a rational linear combination of $1, \varepsilon_p, \varepsilon_p^2, \ldots, \varepsilon_p^{p-1}$ with coefficients adding up to zero (which is always possible). Namely, if $\sum_{k=0}^{p-1} x_k = 0$, then

$$\left(\sum_{k=0}^{p-1} x_k \varepsilon_p^k, \sum_{k=0}^{p-1} y_k \varepsilon_p^k\right) = p\left(x_0 y_0 + \sum_{k=1}^{p-1} x_k y_{p-k}\right).$$

A part of the above discussion on field extensions can be generalized to extensions of Noetherian rings as long as we appropriately modify the notions of an algebraic element and an algebraic extension.

Let a ring B be an extension of a ring A. An element $u \in B$ is called *integral algebraic* over A, or simply *integral*, if it satisfies a nontrivial algebraic equation with coefficients in A and the leading coefficient 1. In particular, elements of A itself are integral over A. An element $u \in B$ which is algebraic over A can be made integral if we multiply it by an appropriate nonzero element of A (namely, by the leading coefficient of the algebraic equation over A whose root is u).

A ring B is called an *integral extension* of A or simply *integral over* A if every element of B is integral over A.

When A is a field, these definitions are equivalent to the definitions of an algebraic element and an algebraic extension.

The following is the key definition.

Definition 9.122. An extension B of a ring A is *finite* if B is a finitely generated A-module.

Now we state partial analogues of Theorems 9.107, 9.110, 9.111, and 9.112 for ring extensions. Their proofs differ very little from the proofs of the corresponding theorems. One only needs to replace the word "basis" with the phrase "generating set" and use Theorem 9.91 instead of the statement that every subspace of a finite-dimensional vector space is finite-dimensional.

Theorem 9.123. *An element $u \in B$ is integral over A if and only if $A[u]$ is a finitely generated A-module.*

Corollary 9.124. *Every finite extension of a Noetherian ring is integral.*

Theorem 9.125. *If B is a finite extension of A, and C is a finite extension of B, then C is a finite extension of A.*

Theorem 9.126. *If a ring B is generated over A by a finite number of integral elements, then it is a finite extension of A.*

Recall that a finitely generated (and certainly finite) extension of a Noetherian ring is also Noetherian (Corollary 9.95).

Theorem 9.127. *Let B be an extension of a Noetherian ring A. The set $\overline{A}$ of all elements of B integral over A is a subring which is integrally closed in B.*

(The latter means that every element of B which is integral over $\overline{A}$ belongs to $\overline{A}$, i.e., that it is already integral over A.)

The ring $\overline{A}$ is called the *integral closure* of the ring A in B.

For example, all algebraic numbers integral over $\mathbb{Z}$—they are called *algebraic integers*—form a subring $\overline{\mathbb{Z}}$ in the field $\overline{\mathbb{Q}}$ of all algebraic numbers. The field of fractions of $\overline{\mathbb{Z}}$ coincides with $\overline{\mathbb{Q}}$.

Remark 9.128. In fact, Corollary 9.124 and, hence, Theorem 9.127 hold for arbitrary (not necessarily Noetherian) rings. This can be proved as follows. Let $B = Ae_1 + \cdots + Ae_n$. Then $e_i e_j = \sum_k c_{ijk} e_k$ for some $c_{ijk} \in A$. If A' is a subring of A containing all c_{ijk}'s, then $B' = A'e_1 + \cdots + A'e_n$ is a subring of B and, moreover, a finite extension of A'. For every $u = a_1 e_1 + \cdots + a_n e_n$, $a_i \in A$, take the subring generated by all c_{ijk}'s and a_i's as A'. It is Noetherian (see Remark 9.96), hence by Corollary 9.124, $u \in B'$ is integral over A' and, moreover, A.

The following theorem establishes a connection between finite field extensions and finite ring extensions.

An integral domain is called *normal* or *integrally closed* if it is integrally closed in its field of fractions. For example, the ring $\mathbb{Z}$ is normal by Theorem 3.67. On the other hand, the ring of polynomials without a linear term (i.e., of the form $a_0 + a_2 x^2 + \cdots$) or the ring of numbers of the form $a + b\sqrt{-3}$, $a, b \in \mathbb{Z}$, is not normal (prove this!).

Theorem 9.129. *Let A be a normal Noetherian domain with the field of fractions K. Let L be a finite extension of K and B, the integral closure of A in L. Assume that* $\operatorname{char} K = 0$. *Then B is a finite extension of A.*

(See diagram (9.21).)

Proof. We prove first that $\operatorname{tr} u \in A$ for every $u \in B$. Let $a_1, \ldots, a_m \in A$ be such that

$$u^m + a_1 u^{m-1} + \cdots + a_{m-1} u + a_m = 0.$$

Then

$$T(u)^m + a_1 T(u)^{m-1} + \cdots + a_{m-1} T(u) + a_m E = 0. \tag{9.24}$$

Let $P \supset K$ be the splitting field of the characteristic polynomial of the operator $T(u)$. It follows from (9.24) that all roots of this polynomial in P are integral over A. But the trace $\operatorname{tr} u = \operatorname{tr} T(u)$ is equal to the sum of

these roots, hence, it is also integral over A. On the other hand, $\operatorname{tr} u \in K$. It follows from normality of A that $\operatorname{tr} u \in A$.

Let $\{e_1, \dots, e_n\}$ be a basis of L over K. Multiplying $e_1, \dots, e_n$ by appropriate elements of A, we can obtain $e_1, \dots, e_n \in B$. Then $c_{ij} := (e_i, e_j) \in A$ for all i, j and

$$\Delta := \det(c_{ij}) \neq 0.$$

Let us determine when the element

$$u = x_1 e_1 + \dots + x_n e_n, \qquad x_1, \dots, x_n \in K,$$

is integral over A. This is obviously true if $x_1, \dots, x_n \in A$. Generally speaking, this condition is not necessary, but we will show now that the coefficients $x_1, \dots, x_n$ cannot be "too fractional."

Considering inner products of u with the basis vectors, we see that

$$\sum_j c_{ij} x_j = (e_i, u) \in A, \qquad i = 1, \dots, n. \tag{9.25}$$

If we regard equations (9.25) as a system of linear equations in indeterminates $x_1, \dots, x_n$, Cramer's rules imply that $x_1, \dots, x_n \in \Delta^{-1} A$.

Thus, the ring B is contained in the A-submodule generated by the elements $\Delta^{-1} e_1, \dots, \Delta^{-1} e_n$. Since A is Noetherian, it follows that B is a finitely generated A-module. □

Remark 9.130. Theorem 9.129 and its proof are also valid in the case of $\operatorname{char} K = p > 0$ whenever the inner product in L is nondegenerate. A finite extension L of K satisfying this condition is called *separable* (see Remark 11.58).

Let K be a field of algebraic numbers (i.e., a finite extension of the field $\mathbb{Q}$). The integral closure of the ring $\mathbb{Z}$ in K is called the *ring of integers* of K. It is denoted $\mathbb{Z}_K$. Theorem 9.129 implies that $\mathbb{Z}_K$ is a finitely generated (additive) abelian group. Since the group $\mathbb{Z}_K$ is torsion-free, it is free. Moreover, when multiplied by an appropriate (rational) integer, every element of K becomes integral. Thus, a basis of $\mathbb{Z}_K$ is also a basis of K regarded as a vector space over $\mathbb{Q}$, hence,

$$\operatorname{rk} \mathbb{Z}_K = \dim_{\mathbb{Q}} K.$$

Exercise 9.131. Prove that the integers of the field $\mathbb{Q}(\sqrt{d})$, where d is a square-free integer, are the numbers of the form $a + b\sqrt{d}$, where $a, b \in \mathbb{Z}$ or, whenever $d \equiv 1 \pmod 4$, $a, b \in \mathbb{Z} + \frac{1}{2}$.

Exercise 9.132. Prove that the integers of the cyclotomic field $\mathbb{Q}(\varepsilon_p)$ (see Examples 9.109 and 9.121) are the numbers

$$a_0 + a_1 \varepsilon_p + \dots + a_{p-2} \varepsilon_p^{p-2}, \qquad a_0, a_1, \dots, a_{p-2} \in \mathbb{Z}.$$

(*Hints*: (i) as in the proof of Theorem 9.129, prove at first that the denominators of the rational numbers $a_0, a_1, \ldots, a_{p-2}$ are powers of p;

(ii) instead of the decomposition in the powers of ε_p, consider the decomposition in the powers of $1 - \varepsilon_p$;

(iii) prove that in the ring of integers of the field $\mathbb{Q}(\varepsilon_p)$, the following numbers are associated:

$$1 - \varepsilon_p \sim 1 - \varepsilon_p^k, \quad k = 1, 2, \ldots, p-1, \qquad p \sim (1 - \varepsilon_p)^{p-1};$$

(iv) prove that if a rational integer is divisible by $1 - \varepsilon_p$, then it is divisible by p.)

9.6. Finitely Generated Algebras and Affine Algebraic Varieties

In this section we study algebras over a field K. Here, by an algebra, we always understand a commutative associative algebra with unity. An algebra A is called *finitely generated* if it is finitely generated over K. Observe that by Corollary 9.94, every finitely generated algebra is Noetherian.

The theory of finitely generated algebras lies at the foundation of the study of systems of algebraic equations; this is the subject of algebraic geometry.

Let A be an algebra without zero divisors. Elements $u_1, \ldots, u_n$ of A are called *algebraically dependent* if they are algebraically dependent over K.

Definition 9.133. A *transcendence basis* of A is a maximal algebraically independent system of elements or, equivalently, an algebraically independent system $\{u_1, \ldots, u_d\}$ such that A is an algebraic extension of the subalgebra $K[u_1, \ldots, u_d]$ generated by the elements $u_1, \ldots, u_d$. (Cf. the definition of a basis of a vector space.)

For example, $\{x_1, \ldots, x_n\}$ is a transcendence basis of the polynomial algebra $K[x_1, \ldots, x_n]$.

Proposition 9.134. *Every transcendence basis of an algebra A is a transcendence basis of its quotient field $Q(A)$ regarded as an algebra over K.*

Proof. Let $\{u_1, \ldots, u_d\}$ be a transcendence basis of A. Elements of $Q(A)$ that are algebraic over the subalgebra $K[u_1, \ldots, u_d]$ are the elements that are algebraic over the subfield

$$K(u_1, \ldots, u_d) = Q(K[u_1, \ldots, u_d]).$$

All such elements form a subfield in $Q(A)$ (Theorem 9.83). Since this subfield contains A, it coincides with $Q(A)$. □

Proposition 9.135. *Let $A = K[u_1, \ldots, u_n]$. Then every maximal algebraically independent subsystem of the system $\{u_1, \ldots, u_n\}$ is a transcendence basis of A.*

Proof. Let $\{u_1, \ldots, u_d\}$ be a maximal algebraically independent subsystem of the system $\{u_1, \ldots, u_n\}$. Consider the algebraic closure of the subfield $K(u_1, \ldots, u_d)$ in the field $Q(A)$. By assumption, it contains the elements $u_1, \ldots, u_n$, hence coincides with $Q(A)$. In particular, it contains A. □

Corollary 9.136. *Every finitely generated algebra without zero divisors has a transcendence basis.*

Proposition 9.137. *Let $\{u_1, u_2, \ldots, u_d\}$ be a transcendence basis of A and $v \in A$, an element that is transcendental over $K[u_2, \ldots, u_d]$. Then $\{v, u_2, \ldots, u_d\}$ is also a transcendence basis of A.*

Proof. Clearly, the elements $v, u_2, \ldots, u_d$ are algebraically independent. On the other hand, the elements $v, u_1, u_2, \ldots, u_d$ are algebraically dependent. Consider a nontrivial algebraic dependence of them. It should depend nontrivially on the element u_1. Thus, u_1 is algebraic over the subalgebra $K[v, u_2, \ldots, u_d]$. Therefore, the algebraic closure of the subfield $K(v_1, u_2, \ldots, u_d)$ in $Q(A)$ contains $K(u_1, u_2, \ldots, u_d)$, hence coincides with $Q(A)$. □

Theorem 9.138. *All transcendence bases of an algebra A (if such exist) have the same number of elements.*

This number is called the *transcendence degree* of A and is denoted $\operatorname{tr.deg} A$.

Proof. Let $\{u_1, \ldots, u_d\}$ and $\{v_1, \ldots, v_e\}$ be two transcendence bases. If all elements $v_1, \ldots, v_e$ are algebraic over $K[u_2, \ldots, u_d]$, then elements $u_2, \ldots, u_d$ form a transcendence basis of A by themselves, which is impossible. Hence, there exists an index i_1 such that the element v_{i_1} is transcendental over $K[u_2, \ldots, u_d]$. By Proposition 9.137, $\{v_{i_1}, u_2, \ldots, u_d\}$ is a transcendence basis of A. In the same manner, we can replace u_2 by an element v_{i_2}, etc. At the end, we obtain a transcendence basis of the form

$$\{v_{i_1}, v_{i_2}, \ldots, v_{i_d}\}.$$

Clearly, the indices $i_1, \ldots, i_d$ must be distinct and they should contain all indices $1, 2, \ldots, e$. Thus, $d = e$. □

Theorem 9.139 (Noether Normalization Lemma). *A finitely generated algebra $A = K[u_1, \ldots, u_n]$ without zero divisors has a transcendence basis $\{v_1, \ldots, v_d\}$ such that A is integral over $K[v_1, \ldots, v_d]$.*

Proof. We will prove this theorem under the assumption that the field K is infinite. In this case, we can construct the transcendence basis in question from linear combinations of $u_1, \ldots, u_n$.

We use induction on n. If the elements $u_1, \ldots, u_n$ are algebraically independent, they form such a transcendence basis. Otherwise, consider a nontrivial algebraic dependence between them:

$$f(u_1, \ldots, u_n) = 0, \qquad f \in K[x_1, \ldots, x_n].$$

Let $\deg f = m$. If x_n^m occurs in f with a nonzero coefficient, the element u_n is integral over the subalgebra $B = K[u_1, \ldots, u_{n-1}]$. By the induction hypothesis, there exists a transcendence basis $v_1, \ldots, v_d$ of B such that the algebra B is integral over $K[v_1, \ldots, v_d]$. This is the transcendence basis that we need.

The general case reduces to the above after a change of variables of the form

$$x_i = y_i + a_i y_n, \qquad i = 1, \ldots, n-1, \qquad x_n = y_n, \qquad a_1, \ldots, a_{n-1} \in k.$$

The polynomial

$$g(y_1, \ldots, y_{n-1}, y_n) = f(y_1 + a_1 y_n, \ldots, y_{n-1} + a_{n-1} y_n, y_n)$$

is also of degree m and y_n^m occurs in it with the coefficient

$$g_0(0, \ldots, 0, 1) = f_0(a_1, \ldots, a_{n-1}, 1),$$

where f_0 and g_0 are the leading homogeneous components of the polynomials f and g, respectively. Since f_0 is a nonzero homogeneous polynomial, it cannot be identically equal to zero on the hyperplane $x_n = 1$. Therefore, with a suitable choice of $a_1, \ldots, a_{n-1}$, y_n^m occurs in the polynomial g with a nonzero coefficient. Put

$$u_i = v_i + a_i v_n, \qquad i = 1, \ldots, n-1, \qquad u_n = v_n;$$

then

$$g(v_1, \ldots, v_n) = f(u_1, \ldots, u_n) = 0,$$

and the proof reduces to the previous case. □

Theorem 9.140. *If a finitely generated algebra A is a field, then it is a finite algebraic extension of the field K.*

Proof. By Theorem 9.139, there exists a transcendence basis $\{v_1, \ldots, v_d\}$ of A such that A is integral over the subalgebra

$$B = K[v_1, \ldots, v_d].$$

Let us prove that B is also a field. For any $u \in B$, there exists $u^{-1} \in A$. The element u^{-1} is integral over B, i.e.,

$$u^{-m} + b_1 u^{-m+1} + \cdots + b_{m-1} u^{-1} + b_m = 0$$

for some $b_1, \ldots, b_m \in B$. Thus,

$$u^{-1} = -b_1 - b_2 u - \cdots - b_m u^{m-1} \in B.$$

Since the algebra B is isomorphic to the algebra of polynomials in d variables, it cannot be a field when $d > 0$. Hence $d = 0$, i.e., A is an algebraic extension of K. □

Corollary 9.141. *If a finitely generated algebra A over an algebraically closed field K is itself a field, then $A = K$.*

Theorem 9.142. *Let A be a finitely generated algebra over an algebraically closed field K. Then for every nonnilpotent element $a \in A$, there exists a homomorphism $\varphi\colon A \to K$ such that $\varphi(a) \neq 0$.*

Proof. Following the proof of Theorem 9.102, consider a maximal ideal I' of the algebra $A' = A[a^{-1}]$. The field A'/I' is a finitely generated algebra over K, hence it coincides with K. Now we can take the restriction of the canonical homomorphism $A' \to A'/I' = K$ to A as φ. □

Let us apply this theorem to the study of systems of algebraic equations.

Let $M \subset K^n$ be the set of solutions of a system of algebraic equations

$$f_i(x_1, \ldots, x_n) = 0, \qquad i = 1, \ldots, m. \tag{9.26}$$

Consider the algebra

$$A = K[x_1, \ldots, x_n]/(f_1, \ldots, f_m) \tag{9.27}$$

and the canonical homomorphism $\pi\colon K[x_1, \ldots, x_n] \to A$. Set $\pi(x_i) = u_i$; then $A = K[u_1, \ldots, u_n]$

To each point $x \in K^n$, there corresponds the homomorphism

$$\psi_x\colon K[x_1, \ldots, x_n] \to K, \qquad f \mapsto f(x) \tag{9.28}$$

and, conversely, every homomorphism

$$\psi\colon K[x_1, \ldots, x_n] \to K$$

has the form ψ_x, where x is the point with coordinates $(\psi(x_1), \ldots, \psi(x_n))$.

If $x \in M$, the homomorphism ψ_x maps the ideal $(f_1, \ldots, f_m)$ to zero, hence, can be factored through the homomorphism π:

$$\begin{array}{ccc} K[x_1, \ldots, x_n] & \xrightarrow{\psi_x} & K \\ & {\scriptstyle\pi}\searrow \quad \nearrow{\scriptstyle\varphi_x} & \\ & A & \end{array} \tag{9.29}$$

The homomorphism $\varphi_x\colon A \to K$ thus obtained maps the generating elements $u_1, \ldots, u_n$ of A to the coordinates of the point x. Conversely, every homomorphism $\varphi\colon A \to K$ is a part of some commutative diagram of type (9.29), hence has the form φ_x for $x \in M$.

Therefore, the points of M are in one-to-one correspondence with homomorphisms of the algebra A to K. This quite trivial observation throws a bridge between commutative algebra and algebraic geometry and, in particular, allows Theorem 9.142 to take the following form:

Theorem 9.143 (Hilbert's Nullstellensatz). *Let M be the set of solutions of a system of algebraic equations* (9.26) *over an algebraically closed field K, and let a polynomial $f \in K[x_1, \dots, x_n]$ be identically zero on M. Then there exists a natural number k such that*

$$f^k \in (f_1, \dots, f_m). \tag{9.30}$$

Proof. Define the algebra A as above and put $a = \pi(f) \in A$. Condition (9.30) means that the element a is nilpotent. If this is not so, by Theorem 9.142, there exists a homomorphism $\varphi\colon A \to K$ such that $\varphi(a) \neq 0$. This homomorphism defines a point of M where the value of f is not zero. □

Observe that, on the other hand, every polynomial $f \in K[x_1, \dots, x_n]$ that satisfies condition (9.30) is identically zero on M.

Corollary 9.144. *The system of algebraic equations* (9.26) *over an algebraically closed field K is incompatible if and only if*

$$(f_1, \dots, f_m) \ni 1, \tag{9.31}$$

i.e., if there exist polynomials $g_1, \dots, g_m \in K[x_1, \dots, x_n]$ such that

$$f_1 g_1 + \dots + f_m g_m = 1. \tag{9.32}$$

Proof. Apply the Nullstellensatz to the polynomial $f = 1$. □

Definition 9.145. An *affine algebraic variety over K*, or an *algebraic variety in K^n*, is the set of solutions of a system of algebraic equations.

Let $M \subset K^n$ be an algebraic variety. K-valued functions on M that are restrictions of polynomials on the space K^n are called *polynomials on M*. They form an algebra called the *polynomial algebra on M*, which is denoted $K[M]$. The kernel of the restriction homomorphism

$$\rho\colon K[x_1, \dots, x_n] \to K[M]$$

is an ideal $I(M)$ that consists of all polynomials on K^n that are identically zero on M. We have

$$K[M] \simeq K[x_1, \dots, x_n]/I(M).$$

By Hilbert's Basis Theorem, the ideal $I(M)$ has a finite generating set:

$$I(M) = (f_1, \dots, f_m).$$

Clearly, the variety M can be defined by equations (9.26). The ideal $I(M)$ is called the *ideal of* M.

Every point $x \in M$ defines the homomorphism

$$\varphi_x \colon K[M] \to K, \qquad f \mapsto f(x). \tag{9.33}$$

The above discussion shows that we have a one-to-one correspondence between points of M and homomorphisms of the algebra $K[M]$ into K. Notice that as an algebra of functions on M, $K[M]$ has no nilpotent elements.

Conversely, let $A = K[u_1, \ldots, u_n]$ be a finitely generated algebra. Consider a homomorphism

$$\pi \colon K[x_1, \ldots, x_n] \to A, \qquad x_i \mapsto u_i.$$

Its kernel is an ideal I of the polynomial algebra $K[x_i, \ldots, x_n]$. Let

$$I = (f_1, \ldots, f_m),$$

and let $M \subset K^n$ be an algebraic variety defined by system (9.26). Then the points of M are in one-to-one correspondence with homomorphisms of A into K. However, the ideal $I(M)$ can be larger than I and so, the algebra $K[M]$ does not have to coincide with A.

In any case, $\operatorname{Ker} \rho \supset \operatorname{Ker} \pi$, hence, there exists a homomorphism

$$\sigma \colon A \to K[M].$$

Its kernel consists of elements of A that come (under π) from polynomials that are identically zero on M. By the Nullstellensatz, if K is algebraically closed,

$$\operatorname{Ker} \sigma = \operatorname{rad} A.$$

In particular, if A has no nilpotent elements (and K is algebraically closed), $A = K[M]$.

Thus, in the case of an algebraically closed field K, we have established a one-to-one correspondence between the algebraic varieties in K^n and the algebras on n generators that do not have nilpotent elements.

For any finitely generated algebra A over K, we call the set of all its homomorphisms into K the *spectrum* and denote it $\operatorname{Spec} A$. If we fix a set of n generators of A, then by the aforesaid, $\operatorname{Spec} A$ is identified with an algebraic variety in K^n.

In particular, for M an algebraic variety in K^n, $\operatorname{Spec} K[M]$ is identified with M if we choose the restrictions of coordinate functions of K^n as its generators. For other choices, we obtain other "models" of the spectrum that are regarded as algebraic varieties isomorphic to M. Thus, we regard as internal properties of M, the properties that can be expressed in terms of the algebra $K[M]$.

In order to study these properties, it is useful to employ the *Zariski topology*. In this topology, a set $N \subset M$ is closed if it is defined by equations of the form

$$f_i = 0, \qquad i = 1, \ldots, m,$$

where $f_1, \ldots, f_m \in K[M]$. For example, the closed subsets of K^n are exactly the algebraic varieties. The closed subsets of any algebraic variety $M \subset K^n$ are the algebraic varieties in K^n that are contained in M.

It is not difficult to check that this definition satisfies topology axioms, i.e., that the intersection of any number of closed sets and the union of a finite number of closed sets are closed. For instance, the union of subsets defined by equations $f_i = 0$, $i = 1, \ldots, m$, and $g_j = 0$, $j = 1, \ldots, p$, respectively, is defined by the equations $f_i g_j = 0$, $i = 1, \ldots, m$, $j = 1, \ldots, p$.

Except for some trivial cases, the Zariski topology is not Hausdorff. For example, the closed subsets of the line K^1 in the Zariski topology are the line itself and its finite subsets. Thus, if K is infinite, every two nonempty open subsets intersect.

Since it is so poor, the Zariski topology mostly plays an auxiliary role as a useful language for the study of algebraic varieties. But by itself, it can still express some rough properties of algebraic varieties.

Definition 9.146. A topological space is *Noetherian* if it does not contain an infinite strictly decreasing sequence of closed subsets.

To every closed subset N of an affine algebraic variety M, there corresponds the ideal $I_M(N)$ of the algebra $K[M]$ consisting of all polynomials that are identically zero on N. Furthermore, $N_1 \supset N_2$ if and only if $I_M(N_1) \subset I_M(N_2)$. Thus, since $K[M]$ is Noetherian, the variety M is a Noetherian topological space (in the Zariski topology).

Definition 9.147. A topological space M is *irreducible* if it is nonempty and satisfies either of the following equivalent properties:

(i) it cannot be presented as a union of two proper closed subsets;

(ii) any two of its nonempty open subsets have a nonempty intersection.

(Compare this definition with that of a connected topological space.)

Theorem 9.148. *An affine algebraic variety M is irreducible if and only if the algebra $K[M]$ has no zero divisors.*

Proof. Let $f_1, f_2 \in K[M]$ be nonzero polynomials such that $f_1 f_2 = 0$. Then $M = N_1 \cup N_2$, where N_i, $i = 1, 2$, is the closed subset defined by the equation $f_i = 0$.

Conversely, let $M = N_1 \cup N_2$, where N_1, N_2 are proper closed subsets. Choose nonzero polynomials $f_1 \in I_M(N_1), f_2 \in I_M(N_2)$. Then $f_1 f_2 = 0$. □

Proposition 9.149. *Every Noetherian topological space M can be uniquely presented in the form*

$$M = \bigcup_{i=0}^{s} M_i, \tag{9.34}$$

where $M_1, \dots, M_s$ are irreducible closed subsets such that none of them contains another.

Subsets M_i are called *irreducible components* of M.

Proof. Assume that there exist Noetherian topological spaces that cannot be presented as finite unions of their irreducible closed subsets. Call such spaces bad. Let M_0 be bad. Then M_0 is reducible, i.e., $M_0 = M_1 \cup N_1$, where M_1 and N_1 are proper closed subsets. Clearly, at least one of them is bad. Let it be M_1. Then $M_1 = M_2 \cup N_2$, where M_2 and N_2 are proper closed subsets and at least one of them is bad. Continuing this process, we obtain an infinite strictly decreasing sequence of closed subsets

$$M_0 \supset M_1 \supset M_2 \supset \cdots,$$

which is a contradiction since M_0 is Noetherian.

Thus, every Noetherian topological space M can be presented in the form (9.34), where $M_1, \dots, M_s$ are irreducible closed subsets. Removing those contained in others, we obtain a situation where none of these subsets contains another. Let us show that with this condition, decomposition (9.34) is unique.

Let $M = \bigcup_{j=1}^{t} N_j$ be another such decomposition. Then for any j,

$$N_j = \bigcup_{i=1}^{s} (M_i \cap N_j)$$

and irreducibility of N_j implies that there exists i such that $N_j \subset M_i$. Similarly, there exists k such that $M_i \subset N_k$, but then $N_j \subset N_k$. It follows that $j = k$ and $N_j = M_i$. Thus, $\{N_1, \dots, N_t\} \subset \{M_1, \dots, M_s\}$. The opposite inclusion is proved similarly. □

In particular, every affine algebraic variety uniquely decomposes into irreducible components.

Example 9.150. Let f be a polynomial of the second degree in n variables over an algebraically closed field K. The equation $f = 0$ defines an algebraic variety M in K^n. The following cases are possible:

(i) f does not split into linear factors; then M is an irreducible quadric;

(ii) f splits into two nonproportional linear factors; then M is a union of two hyperplanes, and these are its irreducible components;

(iii) f is a square of a linear polynomial; then M is a hyperplane (in this case, $I(M) \neq (f)$).

All this can be deduced from a more general Theorem 9.164, which is proved in the next section.

One of the most important characteristics of an irreducible algebraic variety is its dimension.

Definition 9.151. The *dimension* of an irreducible affine algebraic variety M is the number

$$\dim M = \text{tr. deg } K[M].$$

In particular,

$$\dim K^n = \text{tr. deg } K[x_1, \ldots, x_n] = n.$$

The dimension of an algebraic variety has the following property, which is similar to a property of the dimension of a vector space.

Theorem 9.152. *Let N be an irreducible closed subset of an irreducible affine algebraic variety M. Then* $\dim N \leq \dim M$ *and equality is attained only when $N = M$.*

Proof. Let

$$\rho\colon K[M] \to K[N]$$

be the restriction homomorphism. It is clear that if elements $\rho(f_1), \ldots, \rho(f_k)$ are algebraically independent in $K[N]$, the elements $f_1, \ldots, f_k$ are algebraically independent in $K[M]$. This implies the first assertion of the theorem.

Assume now that $N \neq M$. Let $\{\rho(f_1), \ldots, \rho(f_k)\}$ be a transcendence basis of $K[N]$ and $f \in I_M(N)$, $f \neq 0$. Let us prove that $f_1, \ldots, f_k, f$ are algebraically independent in $K[M]$; this will imply the second assertion.

Assume that $f_1, \ldots, f_k, f$ are algebraically dependent. This dependence can be written as

$$a_0(f_1, \ldots, f_k)f^m + a_1(f_1, \ldots, f_k)f^{m-1} + \cdots + a_m(f_1, \ldots, f_k) = 0,$$

where $a_0, a_1, \ldots, a_m$ are polynomials some of which are nonzero. We can assume that $a_m \neq 0$; otherwise, we can cancel f in the above equality. Applying ρ, we obtain

$$a_m(\rho(f_1), \ldots, \rho(f_k)) = 0,$$

contradicting the algebraic independence of $\rho(f_1), \ldots, \rho(f_k)$. □

9.7. Prime Factorization

One of the fundamental goals of arithmetics is to develop the theory of factorization into primes for the rings of algebraic integers. A similar problem over finitely generated algebras appears in algebraic geometry (e.g., in relation with the description of linear bundles over algebraic varieties). Despite the differences between these two types of rings, the problem of prime factorization can be treated somewhat similarly for both of them. This framework is presented in the end of this section; as for its beginning, here we prove existence and uniqueness of prime factorizations in certain types of rings.

Let A be an integral domain. Notice that for elements $a, b \in A$, the condition $b|a$ is equivalent to the condition $(a) \subset (b)$; thus, the condition $a \sim b$ is equivalent to the condition $(a) = (b)$.

Theorem 9.153. *In a Noetherian domain, every noninvertible nonzero element factorizes into a product of prime elements.*

(We assume that a product might consist of one factor only.)

Proof. Assume that there exist noninvertible nonzero elements which cannot be factorized into prime elements. We call such elements bad. Let a_0 be a bad element. Then it is certainly not prime, hence $a_0 = a_1 b_1$, where a_1 and b_1 are noninvertible elements. Clearly, at least one of the elements a_1 and b_1 is bad. Assume a_1 is bad. Then $a_1 = a_2 b_2$, where a_2 and b_2 are noninvertible elements and one of them is bad. Continuing this process, we obtain the following strictly increasing chain of ideals:

$$(a_0) \subset (a_1) \subset (a_2) \subset \cdots,$$

which cannot happen in a Noetherian domain. □

As for the uniqueness of prime factorization, it may hold only up to a permutation of factors and multiplication by invertible elements. Below we understand uniqueness in this sense only.

Inspecting the proof of uniqueness of prime factorization in Euclidean domains (see Section 3.5), we see that it relies on exactly one property of these rings: if a prime element p divides the product ab, then it divides either a or b. In other words, we use the fact that in a Euclidean domain, the ideal (p) generated by a prime element p is prime. This suggests the following theorem:

Theorem 9.154. *If in an integral domain A every principal ideal generated by a prime element is prime, then any element of A factorizes into primes in at most one way.*

Observe that a principal ideal generated by a nonprime noninvertible nonzero element is not prime in any integral domain.

Definition 9.155. An integral domain A is *factorial* (or a *unique factorization domain*) if every noninvertible nonzero element of A factorizes into prime elements and this factorization is unique in the above sense.

In particular, every principal ideal domain is factorial (see Section 9.2).

Obviously, in a factorial domain, a principal ideal generated by a prime element is prime.

In a factorial domain, every two elements a and b have the *greatest common divisor* $\mathrm{GCD}\{a, b\}$ which by definition is a common divisor divisible by all other common divisors. Namely, let

$$a = \prod_{i=1}^{s} p_i^{k_i}, \qquad b = \prod_{i=1}^{s} p_i^{l_i}, \qquad k_i, l_i \geq 0,$$

where $p_1, \ldots, p_s$ are prime. Then

$$\mathrm{GCD}\{a, b\} = \prod_{i=1}^{s} p_i^{\min\{k_i, l_i\}}.$$

The greatest common divisor is defined uniquely up to multiplication by an invertible element.

Elements a and b of a factorial domain are called *relatively prime* if $\mathrm{GCD}\{a, b\} = 1$, i.e., if the prime factorizations of a and b contain no common factors (up to the association relation).

The following theorem generalizes Theorem 3.67 (or rather, Corollary 3.68).

Theorem 9.156. *Every factorial domain is normal.*

The proof of this theorem is the same as the proof of Theorem 3.67.

Theorem 9.157. *The ring $A[x]$ of polynomials over a factorial domain A is also a factorial domain.*

Before attempting the proof, we need some preparation.

Call a polynomial $f \in A[x]$ *primitive* if its coefficients are relatively prime together.

Let K be the field of fractions of A. Obviously, every polynomial $h \in K[x]$ decomposes as $h = \lambda h_1$, where $h_1 \in A[x]$ is primitive and $\lambda \in K^*$.

Lemma 9.158 (Gauss Lemma). *If a polynomial $f \in A[x]$ factors into a product of two polynomials in the ring $K[x]$, then it factors into a product of two polynomials in $A[x]$ proportional to those in the first factorization.*

This lemma is proved just as Theorem 3.70. The only small difference is that the quotient ring of A by an ideal generated by a prime element is not a field in general. However, it is always an integral domain and this suffices for the proof.

Corollary 9.159. *If a polynomial $f \in A[x]$ factors in $K[x]$ into a product of polynomials of smaller degree, then it factors into a product of polynomials of smaller degree in $A[x]$.*

Proof of Theorem 9.157. Corollary 9.159 and obvious considerations imply that there are only two kinds of prime elements in the ring $A[x]$:

(i) prime elements of A;

(ii) primitive polynomials $h \in A[x]$ which are irreducible over K.

On the other hand, it is clear that all these elements are indeed prime and that every noninvertible nonzero element of the ring $A[x]$ factors into a product of such elements. If there exist two such factorizations of a polynomial $f \in A[x]$, then considering them in the ring $K[x]$ (which is known to be factorial), we conclude that the factors of the second type in these factorizations are associated in $K[x]$. Since they are primitive, they must also be associated in $A[x]$. So, after cancelling these factors, we obtain two prime factorizations in A and can use the factoriality of A. □

By induction, we conclude with

Corollary 9.160. *For any n, the polynomial ring $K[x_1, \ldots, x_n]$ in n variables over a field K is a factorial domain.*

Prime elements of the ring $K[x_1, \ldots, x_n]$ are called *irreducible polynomials*.

Obviously, every polynomial of first degree is irreducible.

Lemma 9.161. *If a polynomial $f \in K[x_1, \ldots, x_n]$ over an infinite field K is zero at all points of the hyperplane*

$$l := a_1 x_1 + \cdots + a_n x_n + b = 0,$$

then it is divisible by l.

Proof. By passing to another affine coordinate system, we can assume that $l = x_1$. Then the condition of the lemma means that all terms of f contain x_1, hence, f is divisible by l. □

In what follows, we provide two examples of factorization of a polynomial into linear factors, using the above lemma and the factoriality of the ring of polynomials.

Example 9.162. We calculate the Vandermonde determinant $V(x_1, \dots, x_n)$ (see Example 2.96) in a different way. Obviously, $V(x_1, \dots, x_n)$ is a polynomial in $x_1, \dots, x_n$. When $x_i = x_j$ (i, j different), it turns into zero because the corresponding matrix contains two equal rows in this case. By Lemma 9.161, we conclude that this polynomial is divisible by $x_i - x_j$ in the ring $K[x_1, \dots, x_n]$ whenever i, j are different. But then uniqueness of factorization in this domain implies that $V(x_1, \dots, x_n)$ is divisible by $\prod_{i>j}(x_i - x_j)$. It is easy to see that $V(x_1, \dots, x_n)$ is a homogeneous polynomial of degree $\frac{n(n-1)}{2}$. Thus,

$$V(x_1, \dots, x_n) = \mathrm{c} \prod_{i>j} (x_i - x_j), \qquad \mathrm{c} \in K.$$

Comparing the coefficients of $x_2 x_3^2 \cdots x_n^{n-1}$, we obtain $\mathrm{c} = 1$.

Exercise 9.163. In the same way, prove that

$$x^3 + y^3 + z^3 - 3xyz = (x + y + z)(x + \omega y + \overline{\omega} z)(x + \overline{\omega} y + \omega z),$$

where $\omega = -\frac{1}{2} + \imath \frac{\sqrt{3}}{2}$.

We now apply the result on factoriality of a polynomial ring to the description of $(n-1)$-dimensional algebraic varieties in K^n.

Let K be an algebraically closed field. For $f \in K[x_1, \dots, x_n]$, denote by $M(f)$ the algebraic variety in K^n defined by the equation $f = 0$.

Theorem 9.164. *The map $p \mapsto M(p)$ establishes a one-to-one correspondence between irreducible polynomials in n variables (considered up to the association relation) and $(n-1)$-dimensional irreducible algebraic varieties in K^n. Moreover, the ideal of $M(p)$ is generated by p.*

Proof. (i) Let $p \in K[x_1, \dots, x_n]$ be an irreducible polynomial. Then the ideal (p) is prime, hence, the variety $M(p)$ is irreducible and $I(M(p)) = (p)$. In particular, p is uniquely determined by $M(p)$ up to the association relation.

(ii) In the above notation, we have

$$K[M(p)] = K[x_1, \dots, x_n]/(p) = K[u_1, \dots, u_n],$$

where $u_1, \dots, u_n$ are the restrictions of coordinate functions $x_1, \dots, x_n$ of the space K^n to $M(p)$. Assume that the polynomial p is nontrivial in x_n. Then every polynomial in (p) is nontrivial in x_n. Thus, $u_1, \dots, u_{n-1}$ are algebraically independent and $\dim M(p) = n - 1$.

(iii) Conversely, let $M \subset K^n$ be an $(n-1)$-dimensional irreducible algebraic variety. Choose a nonzero polynomial $f \in I(M)$ and decompose it into irreducible factors. Since $I(M)$ is prime, at least one of these factors

lies in $I(M)$. Let it be some irreducible polynomial p. Then $M \subset M(p)$; however, since their dimensions coincide, we must have $M = M(p)$. □

Let $f \in K[x_1, \dots, x_n]$ be a noninvertible nonzero polynomial. Decompose it into irreducible factors:

$$f = p_1^{k_1} \cdots p_s^{k_s}.$$

Theorem 9.164 obviously implies that

$$M(f) = M(p_1) \cup \cdots \cup M(p_s)$$

is the decomposition of $M(f)$ into irreducible components.

Exercise 9.165. Determine $I(M(f))$.

We obtain similar results if, instead of K^n, we consider an irreducible affine variety M such that $K[M]$ is factorial. (The only place where this general case requires additional considerations is part (ii) of the proof of the theorem.)

However, if the algebra $K[M]$ is not factorial, it contains prime elements that generate nonprime principal ideals. At the same time, M contains $(n-1)$-dimensional irreducible subvarieties whose ideals are not principal.

Example 9.166. Let $Q \subset K^3$ be the quadratic cone defined by the equation $xy = z^2$. We have

$$K[Q] = K[x, y, z]/(xy - z^2) = K[u, v, w],$$

where u, v, w are related: $uv = w^2$. Obviously, u, v, w are prime elements of the algebra $K[Q]$ (as are all linear forms in them). Therefore, the relation $uv = w^2$ violates factoriality. This is related to the fact that the ideals (u), (v), and (w) are not prime (e.g., $uv \in (w)$ but $u \notin (w)$ and $v \notin (w)$) and also to the fact that the ideals of the (line) generators of Q are not principal (e.g., the ideal of the x-axis is (v, w) and the ideal of the y-axis is (u, w)).

Examples of this kind indicate that it might be more reasonable to consider the prime ideals of $K[M]$ corresponding to $(n-1)$-dimensional irreducible subvarieties instead of just the prime elements. And, indeed, this line of reasoning leads to a very beautiful theory—not just for finitely generated algebras but for Noetherian domains. We will now give a brief exposition of this theory.

Let A be a normal Noetherian domain. A *valuation* on A is a surjective map

$$\nu \colon A \setminus \{0\} \to \mathbf{Z}_+,$$

that satisfies the following conditions:

(i) $\nu(ab) = \nu(a) + \nu(b)$;

(ii) $\nu(a+b) \geq \min\{\nu(a), \nu(b)\}$.

The set of elements $a \in A$ such that $\nu(a) > 0$ is a prime ideal of A. It is called the *ideal of the valuation* ν and is denoted $\mathfrak{p}(\nu)$.

To every prime element $p \in A$ such that the ideal (p) is prime, there corresponds a valuation ν_p defined as follows: $\nu_p(a)$ is the largest power of p that divides a. It is obvious that $\mathfrak{p}(\nu_p) = (p)$. If the domain A is factorial, then for any pair of its noninvertible elements a and b,

$$b|a \iff \nu_p(a) \geq \nu_p(b) \quad \forall p.$$

However, in the general case the data provided by valuations of the form ν_p is not sufficient to determine if one element is divisible by another. The following two exercises suggest what reasonable generalization we should consider instead.

Exercise 9.167. Prove that a principal prime ideal that is not 0 or A is minimal among nonzero prime ideals of A.

Exercise 9.168. Prove that in a factorial domain, every minimal prime ideal is principal.

Minimal prime ideals of A are called its *prime divisors*. In the above case $A = K[M]$, prime divisors are the ideals of $(n-1)$-dimensional irreducible subvarieties of M.

One can show that every prime divisor $\mathfrak{p}$ is the ideal of a uniquely determined valuation $\nu_\mathfrak{p}$. The proof is based on the idea that the ideal $\mathfrak{p}$ becomes principal under a suitable embedding of A into a bigger ring $A\left[u^{-1}\right]$, where $u \in A \setminus \mathfrak{p}$. Clearly, if $\mathfrak{p} = (p)$, then $\nu_\mathfrak{p} = \nu_p$.

When $A = K[M]$ and $\mathfrak{p} = I(N)$ for an $(n-1)$-dimensional irreducible subvariety N of M, the value of $\nu_\mathfrak{p}(f)$ at $f \in K[M]$ has the meaning of the "order of zero" of f on N.

Example 9.169. In Example 9.166, the plane $x = 0$ touches the cone Q along the y-axis, the plane $y = 0$ touches it along the x-axis, and the plane $z = 0$ intersects it transversely in the x- and y-axes. Thus, if we denote by $\mathfrak{p}$ and $\mathfrak{q}$ the ideals of the axes x and y in the algebra $K[Q]$,

$$\begin{aligned} &\nu_\mathfrak{p}(u) = 0, \quad &&\nu_\mathfrak{p}(v) = 2, \quad &&\nu_\mathfrak{p}(w) = 1, \\ &\nu_\mathfrak{q}(u) = 2, \quad &&\nu_\mathfrak{q}(v) = 0, \quad &&\nu_\mathfrak{q}(w) = 1, \end{aligned}$$

which agrees with the relation $uv = w^2$.

The main properties of the valuations $\nu_\mathfrak{p}$ that justify their use are:

(i) for any $a \in A \setminus \{0\}$, the set of $\mathfrak{p}$ such that $\nu_\mathfrak{p}(a) > 0$ is finite;

(ii) for any $a, b \in A \setminus \{0\}$,

$$b|a \iff \forall \mathfrak{p} \quad \nu_{\mathfrak{p}}(a) \geq \nu_{\mathfrak{p}}(b).$$

Historically, this theory was first established for the rings of integers of cyclotomic fields in Kummer's work on Fermat's theorem. Unlike the general case, all nontrivial prime ideals are minimal in a ring of algebraic integers (as we will see below). Furthermore, in this case, this theory can be interpreted as a theorem on unique factorization of ideals into prime factors.

Namely, define multiplication on the set of ideals as

$$\mathfrak{ab} = \left\{ \sum_{i=1}^{k} a_i b_i \colon a_1, \ldots, a_k \in \mathfrak{a},\ b_1, \ldots, b_k \in \mathfrak{b} \right\}.$$

This multiplication is clearly commutative and associative; moreover, $(a)(b) = (ab)$. Thus, the semigroup of nonzero elements of A considered up to the association relation embeds into the semigroup of ideals.

It can be shown that if A is the ring of integers of a field of algebraic numbers, then every nonzero ideal of A factors into a product of prime ideals uniquely. The value of $\nu_{\mathfrak{p}}(a)$ is then interpreted as the exponent of $\mathfrak{p}$ in the factorization of the ideal (a).

Two ideals are called *equivalent* if they become equal after multiplication by suitable principal ideals. Classes of equivalent ideals of a ring A of algebraic integers form a group called the *class group* of A and denoted $\operatorname{Cl} A$. It measures how far A deviates from being factorial.

Let K be a field of algebraic numbers and $\mathbb{Z}_K$, the ring of its integers.

Theorem 9.170. *Every nonzero ideal $\mathfrak{a}$ of the ring $\mathbb{Z}_K$ is an* additive *subgroup of finite index.*

Proof. We know that there exists a basis $\{e_1, \ldots, e_n\}$ of K over $\mathbb{Q}$ such that

$$\mathbb{Z}_K = \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_n.$$

Let a be a nonzero element of $\mathfrak{a}$. The map $x \mapsto ax$ is a nonsingular linear transformation of the space K over $\mathbb{Q}$, hence $\{ae_1, \ldots, ae_n\}$ is also a basis of this space. It is contained in $\mathfrak{a}$, thus $\mathfrak{a}$ is a subgroup of finite index in $\mathbb{Z}_K$. □

Corollary 9.171. *Any nontrivial prime ideal $\mathfrak{p}$ of the ring $\mathbb{Z}_K$ is maximal.*

Proof. The quotient ring $\mathbb{Z}_K/\mathfrak{p}$ is finite and has no zero divisors. The rest follows from the next lemma. □

Lemma 9.172. *A finite integral domain A is a field.*

Proof. Let $a \in A$ be a nonzero element. Since A contains no zero divisors, the map

$$A \to A, \qquad x \mapsto ax$$

is injective, hence, surjective (as A is finite). In particular, there exists b such that $ab = 1$. □

Corollary 9.173. *A nontrivial prime ideal of the ring $\mathbb{Z}_K$ is a minimal prime ideal.*

Proof. If there existed a smaller prime ideal, it would not be maximal. □

Example 9.174. The ring of integers of the field $\mathbb{Q}(\sqrt{-5})$ is $\mathbb{Z}[\sqrt{-5}]$. Define the norm $N(c)$ of a number $c = a + b\sqrt{-5} \in \mathbb{Z}[\sqrt{-5}]$, $a, b \in \mathbb{Z}$, as follows:

$$N(c) = c\bar{c} = a^2 + 5b^2 \in \mathbb{Z}.$$

Clearly, this norm is multiplicative:

$$N(c_1c_2) = N(c_1)N(c_2).$$

Thus, if c is an invertible element of $\mathbb{Z}[\sqrt{-5}]$, $N(c) = \pm 1$. It follows that the only invertible elements of this ring are ± 1. If c is nonprime, noninvertible, and nonzero, then $N(c)$ can be presented as a product of two norms greater than 1. With this consideration, it is easy to show that all elements in the following equality:

$$2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}) \tag{9.35}$$

are simple. Therefore, the domain $\mathbb{Z}[\sqrt{-5}]$ is not factorial.

From the viewpoint of the theory of ideals, equality (9.35) can be explained as follows:

$$(2) = \mathfrak{p}^2, \qquad (3) = \mathfrak{q}_1\mathfrak{q}_2, \qquad (1 + \sqrt{-5}) = \mathfrak{p}\mathfrak{q}_1, \qquad (1 - \sqrt{5}) = \mathfrak{p}\mathfrak{q}_2,$$

where

$$\mathfrak{p} = (2, 1 + \sqrt{-5}) = (2, 1 - \sqrt{-5}), \quad \mathfrak{q}_1 = (3, 1 + \sqrt{-5}), \quad \mathfrak{q}_2 = (3, 1 - \sqrt{-5})$$

are prime ideals (prove this!). It can be shown that

$$\mathrm{Cl}\,\mathbb{Z}[\sqrt{-5}] \simeq \mathbb{Z}_2.$$

In general, the class group of the ring of integers of any field of algebraic numbers is finite. This means, in particular, that for every ideal, some power of it is a principal ideal.

Chapter 10

Groups

10.1. Direct and Semidirect Products

In Section 9.1, we considered direct sums of additive abelian groups. Of course, the name of the group operation does not matter; nothing stops us from doing the same for multiplicative groups, though in this case, it is natural to speak not about a direct sum but about a direct product. More importantly, we can jettison commutativity. Let us give appropriate rigorous definitions.

Definition 10.1. A group G decomposes into a *direct product* of subgroups $G_1, \ldots, G_k$ if

(i) every element $g \in G$ decomposes uniquely as $g = g_1 \cdots g_k$, $g_i \in G_i$;

(ii) $g_i g_j = g_j g_i$ for $g_i \in G_i$, $g_j \in G_j$, $i \neq j$.

If so, we write $G = G_1 \times \cdots \times G_k$. Obviously, if G is finite, then

$$|G| = |G_1| \cdots |G_k|.$$

Condition (i) implies that $G_i \cap G_j = \{e\}$ for $i \neq j$ but, as we have seen in the case of vector spaces, for $k > 2$ the latter condition is weaker than condition (i) (see Exercise 5.7).

The following rule for multiplying elements of G follows from condition (ii):

$$(g_1 \cdots g_k)(g_1' \cdots g_k') = (g_1 g_1') \cdots (g_k g_k'), \qquad g_i, g_i' \in G_i. \tag{10.1}$$

In particular, it is easy to see that every subgroup G_i is normal. The following lemma implies that condition (ii) can be replaced with the condition of normality of the subgroups $G_1, \ldots, G_k$.

Lemma 10.2. *Let G_1 and G_2 be normal subgroups of G such that $G_1 \cap G_2 = \{e\}$. Then $g_1 g_2 = g_2 g_1$ for any $g_1 \in G_1$, $g_2 \in G_2$.*

Proof. We have

$$g_1 g_2 g_1^{-1} g_2^{-1} = g_1 (g_2 g_1^{-1} g_2^{-1}) = (g_1 g_2 g_1^{-1}) g_2^{-1} \in G_1 \cap G_2 = \{e\},$$

implying $g_1 g_2 = g_2 g_1$. □

We now consider the case of two factors separately.

Proposition 10.3. *A group G decomposes into a direct product of its subgroups G_1 and G_2 of and only if*

(i) subgroups G_1 and G_2 are normal;

(ii) $G_1 \cap G_2 = \{e\}$;

(iii) $G = G_1 G_2$, i.e., every element $g \in G$ decomposes as $g = g_1 g_2$, $g_1 \in G_1$, $g_2 \in G_2$.

Proof. We already proved the "only if" part above. Conversely, assume that conditions (i)–(iii) hold. By Lemma 10.2, $g_1 g_2 = g_2 g_1$ for $g_1 \in G_1$, $g_2 \in G_2$. It remains to show that any $g \in G$ decomposes as $g = g_1 g_2$, where $g_1 \in G_1$, $g_2 \in G_2$, uniquely. Let

$$g_1 g_2 = g_1' g_2', \qquad g_1, g_1' \in G_1,\ g_2, g_2' \in G_2.$$

Then

$$g_1^{-1} g_1' = g_2 g_2'^{-1} \in G_1 \cap G_2 = \{e\},$$

hence,

$$g_1 = g_1', \qquad g_2 = g_2'.$$

□

Example 10.4. Let $G = \{e, a, b, c\}$ be a noncyclic group of order 4. It is easy to see that the square of any of the elements a, b, c is the identity and the product of either two of them (in any order) equals the third (the multiplication table of this group is given in Example 4.12). It follows that G is the direct product of any two of its cyclic subgroups of order 2; for instance,

$$G = \{e, a\} \times \{e, b\}.$$

Example 10.5. That any nonzero complex number can be uniquely presented in the trigonometric form means that

$$\mathbb{C}^* = \mathbb{R}_+^* \times \mathbb{T}$$

(for notations, see Examples 4.63 and 4.64).

Example 10.6. Let $G = \mathrm{GL}_n^+(\mathbb{R})$ be the group of matrices with positive determinant. Also, let G_1 be the group of scalar matrices λE, $\lambda > 0$, and $G_2 = \mathrm{SL}_n(\mathbb{R})$. Then $G = G_1 \times G_2$. Indeed, G_1 and G_2 are normal subgroups (and elements of G_1 commute with all elements of G), $G_1 \cap G_2 = \{e\}$, and $G = G_1 G_2$ since every matrix $A \in G$ decomposes as

$$A = \lambda A_1 = (\lambda E) A_1$$

for

$$\lambda = \sqrt[n]{\det A}, \qquad A_1 = \frac{1}{\lambda} A \in \mathrm{SL}_n(\mathbb{R}) = G_2.$$

Exercise 10.7. Find all n such that

$$\mathrm{GL}_n(\mathbb{R}) = \{\lambda E \colon \lambda \in \mathbb{R}^*\} \times \mathrm{SL}_n(\mathbb{R}).$$

We now define the external direct product of groups.

Definition 10.8. The *direct product* of groups $G_1, \ldots, G_k$ is the set of sequences $(g_1, \ldots, g_k)$, $g_i \in G_i$, with the componentwise operation of multiplication:

$$(g_1, \ldots, g_k)(g_1', \ldots, g_k') = (g_1 g_1', \ldots, g_k g_k').$$

Clearly, in this way we obtain a group. It is denoted $G_1 \times \cdots \times G_k$.

Identifying an element $g \in G_i$ with the sequence $(e, \ldots, g, \ldots, e) \in G_1 \times \cdots \times G_k$ (with g at the ith place), we embed G_i into $G_1 \times \cdots \times G_k$ as a subgroup. The group $G_1 \times \cdots \times G_k$ is the direct sum of these subgroups in the sense of Definition 10.1.

Conversely, if a group G decomposes into a direct product of its subgroups $G_1, \ldots, G_k$, the map

$$G_1 \times \cdots \times G_k \to G, \qquad (g_1, \ldots, g_k) \mapsto g_1 \cdots g_k$$

is an isomorphism according to rule (10.1).

Example 10.9. The group of (nonsingular) diagonal matrices of order n is isomorphic to the group

$$(K^*)^n = \underbrace{K^* \times \cdots \times K^*}_{n}.$$

Decomposition of a group into a direct product is not as common as decomposition into the so-called semidirect product. Before giving appropriate definitions, let us discuss group automorphisms.

Definition 10.10. An *automorphism* is an isomorphism from a group onto itself.

Example 10.11. The map $x \mapsto ax$, $a \neq 0$, is an automorphism of the additive group of a field.

Example 10.12. The map $X \mapsto (X^{\mathsf{T}})^{-1}$ is an automorphism of the group of nonsingular matrices.

Automorphisms of a group G form a group denoted $\operatorname{Aut} G$.

For every element $g \in G$, the map $a(g)\colon x \mapsto gxg^{-1}$, $x \in G$, is an automorphism:

$$a(g)(xy) = gxyg^{-1} = (gxg^{-1})(gyg^{-1}) = (a(g)x)(a(g)y).$$

Such an automorphism is called the *inner automorphism* defined by g.

The map $g \mapsto a(g)$ is a homomorphism from the group G to the group $\operatorname{Aut} G$:

$$a(gh)x = ghx(gh)^{-1} = g(hxh^{-1})g^{-1} = a(g)a(h)x.$$

Its kernel is the *center* Z of the group G:

$$Z = \{z \in G\colon zg = gz \ \forall g \in G\},$$

and its image is a subgroup of $\operatorname{Aut} G$ called the *group of inner automorphisms* of G. It is denoted $\operatorname{Inn} G$. By the homomorphism theorem,

$$\operatorname{Inn} G \simeq G/Z.$$

Example 10.13. It is easy to prove that the center of the group S_n is trivial whenever $n \geq 3$. Hence,

$$\operatorname{Inn} S_n \simeq S_n.$$

Example 10.14. The center of the group $\mathrm{GL}_n(K)$ (where K is a field) consists of scalar matrices. It is isomorphic to the group K^*. The quotient group $\mathrm{GL}_n(K)/\{\lambda E\colon \lambda \in K^*\}$ is just the projective group $\mathrm{PGL}_n(K)$ (the group of projective transformations of the $(n-1)$-dimensional projective space PK^n associated with the vector space K^n). Thus,

$$\operatorname{Inn} \mathrm{GL}_n(K) \simeq \mathrm{PGL}_n(K).$$

Consider $\varphi \in \operatorname{Aut} G$ and $g \in G$. A direct computation shows that

$$\varphi a(g)\varphi^{-1} = a(\varphi(g)).$$

Therefore, $\operatorname{Inn} G$ is a normal subgroup of the group $\operatorname{Aut} G$.

Of course, only nonabelian groups have nontrivial inner automorphisms.

Example 10.15. Here we will describe the group $\operatorname{Aut} S_3$. Since a group isomorphism preserves the order of an element, an automorphism φ of S_3 maps transpositions to transpositions. Moreover, since S_3 is generated by transpositions, an automorphism φ is uniquely determined by the way it permutes transpositions. There are three transpositions in S_3, hence

$$|\operatorname{Aut} S_3| \leq |S_3| = 6.$$

As we have seen before, the group $\operatorname{Inn} S_3$ contains exactly six elements. Therefore,

$$\operatorname{Aut} S_3 = \operatorname{Inn} S_3.$$

Example 10.16. Here we will describe the group $\operatorname{Aut} \mathbf{Z}_n$. Let $\varphi \in \operatorname{Aut} \mathbf{Z}_n$ and $\varphi([1]) = [k]$. Then

$$\varphi([l]) = \varphi(\underbrace{[1] + \cdots + [1]}_{l}) = \underbrace{[k] + \cdots + [k]}_{l} = [kl] = [k][l],$$

where multiplication in the last equality is understood in the sense of the ring $\mathbf{Z}_n$. Thus every automorphism of the group $\mathbf{Z}_n$ is of the form

$$\varphi_a \colon x \mapsto ax$$

for some $a \in \mathbf{Z}_n$. Conversely, for any $a \in \mathbf{Z}_n$, the map φ_a is a homomorphism of the group $\mathbf{Z}_n$ into itself and

$$\varphi_a \varphi_b = \varphi_{ab}.$$

Therefore, the homomorphism φ_a is invertible, i.e., is an automorphism, if and only if the element a is invertible in the ring $\mathbf{Z}_n$. Thus,

$$\operatorname{Aut} \mathbf{Z}_n \simeq \mathbf{Z}_n^*,$$

where $\mathbf{Z}_n^*$ is the group of invertible elements of the ring $\mathbf{Z}_n$.

Using the terminology of automorphisms, we can reformulate the definition of a normal subgroup as follows: a subgroup is normal if it is invariant with respect to all inner automorphisms of the group G.

Let N be a normal subgroup of a group G and H, any subgroup of G. Then

$$NH := \{nh \colon n \in N,\ h \in H\}$$

is a subgroup as the following equalities demonstrate:

$$\begin{aligned} (n_1h_1)(n_2h_2) &= n_1(h_1n_2h_1^{-1})h_1h_2, \\ (nh)^{-1} &= (h^{-1}n^{-1}h)h^{-1}. \end{aligned} \tag{10.2}$$

Moreover, $NH = HN$.

Definition 10.17. A group G decomposes as a *semidirect product* of subgroups N and H if

(i) N is a normal subgroup;

(ii) $N \cap H = \{e\}$;

(iii) $NH = G$.

The notation is $G = N \ltimes H$ (or $G = H \rtimes N$).

Properties (ii) and (iii) are equivalent to every element of G having a unique decomposition nh with $n \in N$, $h \in H$. In particular, if G is a finite group,

$$|G| = |N||H|.$$

Example 10.18. $S_n = A_n \leftthreetimes \langle(12)\rangle$.

Example 10.19. $S_4 = V_4 \leftthreetimes S_3$, where V_4 is the Klein 4-group (see Example 4.110) and S_3 is embedded into S_4 as the subgroup that fixes 4. Indeed, it is easy to see that for any $k \in \{1,2,3,4\}$, there exists a unique permutation in V_4 that moves 4 into k. It follows that every permutation $\sigma \in S_4$ decomposes uniquely as $\sigma = \tau\rho$, where $\tau \in V_4$, $\rho \in S_3$.

Example 10.20. $\mathrm{GL}_n(K) = \mathrm{SL}_n(K) \leftthreetimes \{\mathrm{diag}(\lambda, 1, \ldots, 1) \colon \lambda \in K^*\}$.

Example 10.21. The group $\mathrm{GA}(S)$ of affine transformations of an affine space S is the semidirect product of its (normal) subgroup $\mathrm{Tran}\, S$ of parallel translations and the group $\mathrm{GL}(V)$ of linear transformations of the associated vector space V that embeds into $\mathrm{GA}(S)$ as a subgroup fixing one point.

Example 10.22. The group $\mathrm{Isom}\, S$ of motions of a Euclidean affine space S is the semidirect product of the group of parallel translations and the group $\mathrm{O}(V)$ of orthogonal transformations of the associated Euclidean vector space.

If $G = N \leftthreetimes H$, then $G/N \simeq H$. However, one should not assume that for every normal subgroup N, it is possible to find a subgroup H (isomorphic to G/N) such that $G = N \leftthreetimes H$. For instance, for the (normal) subgroup $2\mathbb{Z}$ in the group $\mathbb{Z}$, there exists no such complementary subgroup.

Let $G = N \leftthreetimes H$. For any $h \in H$, denote by $\alpha(h)$ the restriction of the inner automorphism $a(h)$ of G to N. It is obvious that $\alpha(h) \in \mathrm{Aut}\, N$ and that the map $h \mapsto \alpha(h)$ is a homomorphism from H into the group $\mathrm{Aut}\, N$. The first of the formulas (10.2) can be rewritten as

$$(n_1h_1)(n_2h_2) = (n_1\alpha(h_1)n_2)(h_1h_2). \tag{10.3}$$

Now assume that for some groups N and H, there exists a homomorphism

$$\alpha \colon H \to \mathrm{Aut}\, N.$$

Define a multiplication on the Cartesian product $N \times H$ as follows:

$$(n_1, h_1)(n_2, h_2) = (n_1\alpha(h_1)n_2, h_1h_2). \tag{10.4}$$

This formula is suggested by formula (10.3). A direct check shows that operation (10.4) satisfies the axioms of a group operation. The resulting group G is called the (*external*) *semidirect product* of the groups N and H defined by the homomorphism α. It is denoted $N \overset{\alpha}{\leftthreetimes} H$ or simply $N \leftthreetimes H$. If

we identify the group N with the subgroup of G consisting of pairs of the form (n, e) and the group H, with the subgroup consisting of pairs of the form (e, h), then G becomes the semidirect product of these subgroups in the sense of Definition 10.17.

Conversely, if a group G splits into a semidirect product of its subgroups N and H and $\alpha\colon H \to \operatorname{Aut} N$ is the homomorphism defined as above, then the map

$$N \overset{\alpha}{\ltimes} H \to G, \qquad (n, h) \mapsto nh$$

is a group isomorphism.

The direct product is a particular case of the semidirect one: it occurs when α is the trivial homomorphism.

Below, we will denote the cyclic group of order n with a generator a by $\langle a \rangle_n$.

Example 10.23. Here we will describe the groups that are semidirect products of cyclic groups $\langle a \rangle_n$ and $\langle b \rangle_m$ of orders n and m, respectively.

A homomorphism

$$\alpha\colon \langle b \rangle_m \to \operatorname{Aut}\langle a \rangle_n \simeq \mathbb{Z}_n^*$$

is determined by the image of b, which acts by raising all elements of $\langle a \rangle_n$ to the power k (see Example 10.16). The number k (defined modulo n) must satisfy the following condition:

$$k^m \equiv 1 \pmod{n}.$$

In particular, if the number $|\mathbb{Z}_n^*| = \varphi(n)$ is relatively prime to m, then $k = 1$, which gives us direct product. So, for instance, any semidirect product of groups $\langle a \rangle_7$ and $\langle b \rangle_5$ is the direct product. Let us denote the semidirect product of the groups $\langle a \rangle_n$ and $\langle b \rangle_m$ corresponding to k by $\langle a \rangle_n \overset{k}{\ltimes} \langle b \rangle_m$. It is defined by the following relation:

$$bab^{-1} = a^k.$$

For example, $\langle a \rangle_n \overset{-1}{\ltimes} \langle b \rangle_2$ is the dihedral group D_n. Some of the semidirect products obtained in this way may turn out to be isomorphic. Namely, for $(r, m) = 1$, we can replace the element b with an element b^r that also generates the group $\langle b \rangle$; then k is replaced with k^r. This shows that k itself is not so essential as the cyclic subgroup generated in $\mathbb{Z}_n^*$ by the element $[k]$. For instance, there exist only two nonisomorphic groups that split into a semidirect product of the groups $\langle a \rangle_{11}$ and $\langle b \rangle_5$ (one of them is the direct product of these groups).

10.2. Commutator Subgroup

Let G be a group. A *commutator* of two elements $x, y \in G$ is an element

$$(x, y) = xyx^{-1}y^{-1}.$$

Its obvious properties are:

(i) $(x, y) = e \iff xy = yx$;

(ii) $(x, y)^{-1} = (y, x)$.

A subgroup generated by all commutators in the group G is called the *commutator subgroup* of G and is denoted (G, G) or G'. By property (ii), all elements of the commutator subgroup can be presented as products of commutators. The commutator subgroup is trivial if and only if G is abelian.

Clearly, if $\varphi\colon G \to H$ is a group homomorphism, $\varphi(G') \subset H'$. Moreover, if $\varphi(G) = H$, then $\varphi(G') = H'$. In particular, the commutator subgroup is invariant with respect to all inner automorphisms of the group, i.e., it is a normal subgroup.

Theorem 10.24. *The commutator subgroup G' of a group G is the smallest normal subgroup such that the resulting quotient group is abelian.*

Proof. (i) Denote G/G' by A and let $\pi : G \to A$ be the canonical homomorphism. Then $A' = \pi(G') = \{e\}$, hence A is abelian.

(ii) Let $N \subset G$ be a normal subgroup such that the quotient group $G/N = A$ is abelian. Let $\pi : G \to A$ be the canonical homomorphism. Then $\pi(G') = A' = \{e\}$, hence $G' \subset N$. □

To discuss further examples, we need the following propositions (which are also of independent interest).

Proposition 10.25. *The group A_n is generated by 3-cycles and, for $n \geq 5$, also by products of pairs of disjoint transpositions.*

Proof. Since S_n is generated by transpositions, A_n is generated by the products of pairs of transpositions. Both statements of the proposition follow from the relations

$$(ij)(jk) = (ijk),$$
$$(ij)(kl) = (ijk)(jkl),$$
$$(ij)(jk) = [(ij)(lm)][(jk)(lm)]$$

(here $i, j, k, \ldots$ are pairwise distinct). □

Proposition 10.26. *The group $\mathrm{SL}_n(K)$ is generated by elementary matrices of the first type, i.e., by the matrices $E + cE_{ij}$ $(i \neq j)$.*

Proof. We will use a slightly modified version of Gaussian elimination and show that it is possible to reduce any $A \in \mathrm{SL}_n(K)$ to the identity matrix by applying only elementary transformations of the first type. As in the proof of Theorem 4.59, this will imply the proposition.

Let $A = (a_{ij}) \in \mathrm{SL}_n(K)$, $n > 1$. First, by applying only elementary transformations of the first type, we will transform A so that $a_{11} = 1$. If $a_{i1} \neq 0$ for some $i > 1$, this can be achieved at once by adding the appropriately multiplied ith row to the first row. If $a_{i1} = 0$ for all $i > 1$, then $a_{11} \neq 0$. Here we add the first row to the second and proceed as in the previous case.

When $a_{11} = 1$, we multiply the first row by appropriate coefficients and subtract it from all other rows, obtaining $a_{i1} = 0$ for all $i > 1$. After this, we apply the same procedure to the matrix of order $n - 1$ obtained from A by deleting the first row and column. We continue as above, and finally arrive at a triangular matrix with all diagonal entries, except possibly the last one, equal to 1. However, the determinant of A is 1 by definition, and the above transformations do not change its value. Therefore, the last diagonal element of this triangular matrix is 1 as well.

Finally, we reduce this unitriangular matrix to the identity matrix via the standard reversed Gaussian elimination. □

Example 10.27. We will show here that $S_n' = A_n$. Since the group S_n/A_n is abelian, $S_n' \subset A_n$. As S_3 is not abelian and $|A_3| = 3$, $S_3' = A_3$. Therefore, for any n, S_n' contains all 3-cycles, hence coincides with A_n.

Example 10.28. Here we will show that $A_4' = V_4$ and $A_n' = A_n$ for $n \geq 5$. Since the group A_4/V_4 is abelian, $A_4' \subset V_4$. However, A_4 itself is not abelian, so $A_4' = V_4$. Therefore, for any n, A_n' contains all products of pairs of disjoint transpositions, thus, it coincides with A_n for $n \geq 5$.

Example 10.29. We will show that $\mathrm{SL}_n(K)' = \mathrm{GL}_n(K)' = \mathrm{SL}_n(K)$ whenever the field K contains more than 3 elements. (In fact, this is also true when $|K| \leq 3$ but only if $n \geq 3$.) Since the group $\mathrm{GL}_n(K)/\mathrm{SL}_n(K) \simeq K^*$ is abelian, $\mathrm{GL}_n(K)' \subset \mathrm{SL}_n(K)$. A direct computation shows that

$$\left(\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}, \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & (\lambda^2 - 1)c \\ 0 & 1 \end{pmatrix}.$$

Therefore, if K contains an element $\lambda \neq 0, \pm 1$, the group $\mathrm{SL}_n(K)'$ contains all elementary matrices of the first type, hence, coincides with $\mathrm{SL}_n(K)$.

Exercise 10.30. Find the commutator subgroup of $(a)_n \overset{k}{\ltimes} (b)_m$ (cf. Example 10.23).

Define *higher order commutator subgroups* $G^{(k)}$ of a group G as follows:

$$G^{(1)} = G', \qquad G^{(k+1)} = \left(G^{(k)}\right)'.$$

Definition 10.31. A group G is *solvable* if there exists a natural number m such that $G^{(m)} = \{e\}$.

Clearly, any subgroup and any quotient group of a solvable group is solvable. Conversely, if a normal subgroup N of G and the quotient group G/N are solvable, then G itself is solvable (prove this).

Example 10.32. It follows from Examples 10.27 and 10.28 that S_n is solvable for $n \leq 4$ and is not solvable for $n \geq 5$.

Example 10.33. We will prove here that the group $B_n(K)$ of (nonsingular) triangular matrices is solvable. The proof is by induction on n.

The group $\mathrm{B}_1(K) \simeq K^*$ is abelian. By deleting the last row and the last column from every triangular matrix of order n, we obtain a homomorphism

$$f\colon \mathrm{B}_n(K) \to \mathrm{B}_{n-1}(K).$$

By the induction hypothesis the group

$$\mathrm{B}_{n-1}(K) \simeq \mathrm{B}_n(K)/\mathrm{Ker}\, f$$

is solvable. The group $\mathrm{Ker}\, f$ consists of matrices of the type

$$\begin{pmatrix} 1 & & & c_1 \\ & \ddots & 0 & \vdots \\ & 0 & 1 & c_{n-1} \\ 0 & \dots & 0 & c_n \end{pmatrix}. \tag{10.5}$$

By assigning to each matrix like this the number c_n, we obtain a homomorphism $\mathrm{Ker}\, f \to K^*$ whose kernel consists of matrices of type (10.5) with $c_n = 1$. It is easy to see that this kernel is abelian. Therefore, the group $\mathrm{Ker}\, f$ is solvable, and so is $\mathrm{B}_n(K)$.

10.3. Group Actions

Recall that we denote the group of all bijective transformations of the set X (into itself) by $S(X)$. In particular, $S(\{1,\dots,n\}) = S_n$ is the symmetric or the permutation group.

Every subgroup of the group $S(X)$ is called a *transformation group* of the set X. We have already encountered many transformation groups in the course of this book.

The notion of a group action is closely related to that of a transformation group but it provides us with a more flexible language.

Definition 10.34. An *action* of a group G on a set X is a homomorphism

$$\alpha\colon G \to S(X).$$

In other words, to define an action of G on X means to indicate a transformation $\alpha(g) \in S(X)$ for each $g \in G$ so that

$$\alpha(gh) = \alpha(g)\alpha(h). \tag{10.6}$$

Properties of a homomorphism imply that the identity of G acts as the identity transformation id and that the inverse element acts as the inverse transformation.

The image of a "point" $x \in X$ under a transformation $\alpha(g)$ is denoted $\alpha(g)x$ or simply gx if it is clear what the action is. In this notation, property (10.6) becomes a kind of associativity:

$$(gh)x = g(hx).$$

When an action α of a group G on X is given, we write $G \overset{\alpha}{:} X$ or simply $G : X$.

Every transformation group $G \subset S(X)$ acts on X "tautologically" via the trivial homomorphism $G \to S(X)$.

Conversely, for any action $G \overset{\alpha}{:} X$, the subgroup $\operatorname{Im}\alpha \subset S(X)$ is a transformation group of X. By the homomorphism theorem,

$$\operatorname{Im}\alpha \simeq G/\operatorname{Ker}\alpha.$$

The normal subgroup $\operatorname{Ker}\alpha \subset G$ is called the *ineffectiveness kernel* of the action α. If $\operatorname{Ker}\alpha = \{e\}$, then the action α is called *effective*.

A particular case of an action is a *linear representation*, which is a homomorphism of G to the group $\mathrm{GL}(V)$ of (invertible) linear transformations of a vector space V.

Every action $G \overset{\alpha}{:} X$ naturally generates a number of other actions: on an invariant subset of X, on the set of all (or just some) subsets of X, on a quotient set of X by an invariant equivalence relation, etc. We notice, in particular, the linear representation α_* induced by α on the space of all (or just some) K-valued functions on X, which is defined as

$$(\alpha_*(g)f)(x) = f(\alpha(g)^{-1}x). \tag{10.7}$$

Usually we omit the symbol α; the formula thus becomes

$$(gf)(x) = f(g^{-1}x). \tag{10.8}$$

Every action $G : X$ restricted to a subgroup $H \subset G$ defines an action $H : X$.

For any group G we define three important actions of G on itself:

(i) the action l by left shifts:

$$l(g)x = gx;$$

(ii) the action r by right shifts:

$$r(g)x = xg^{-1};$$

(iii) the action a by conjugations (inner automorphisms):

$$a(g)x = gxg^{-1}.$$

The action l (as well as r) is effective, thus $G \simeq \operatorname{Im} l \subset S(G)$. In particular, this implies *Cayley's theorem*: every finite group of order n is isomorphic to a subgroup of S_n. As we saw in Section 10.1, the ineffectiveness kernel of the action a is the center Z of G.

An action $G : X$ defines an equivalence relation $\overset{G}{\sim}$ on X as follows:

$$x \overset{G}{\sim} y \qquad \text{if there exists } g \in G \text{ such that } y = gx$$

(check the axioms of an equivalence relation!). The equivalence classes here are called *orbits*. Thus, the orbit that contains a point x (called the orbit of x) is the subset

$$Gx = \{gx \colon g \in G\} \subset X.$$

If all elements of X are equivalent, i.e., if there is only one orbit, the action is called *transitive*.

Example 10.35. An orbit of the group of planar rotations about a point o is either a circle with the center at o or o itself.

Example 10.36. The group of all planar motions and even its subgroup of parallel translations act transitively.

Example 10.37. The Klein 4-group V_4 acts on the set $\{1, 2, 3, 4\}$ transitively.

Example 10.38. Restrict the action l (respectively, r) of G on itself to a subgroup $H \subset G$. The orbits of this action are the right (respectively, left) cosets of H in G.

Elements of G that are equivalent under the action a of G on itself are called *conjugate* and the orbits of this action are called *conjugacy classes*.

If a group G is defined as a transformation group of a set X, the description of its conjugacy classes can be obtained with the help of the following simple observation: if an element $g \in G$ maps a point x to a point y, then hgh^{-1} maps hx to hy.

Example 10.39. Assume that a permutation $\sigma \in S_n$ is a product of disjoint cycles:

$$\sigma = (i_1 i_2 \cdots i_p)(j_1 j_2 \cdots j_q) \cdots .$$

Then for any permutation $\tau \in S_n$, we have

$$\tau\sigma\tau^{-1} = (\tau(i_1)\tau(i_2)\cdots\tau(i_p))(\tau(j_1)\tau(j_2)\cdots\tau(j_q))\cdots .$$

This implies that two permutations are conjugate if and only if the sets of lengths of disjoint cycles in their decompositions coincide. Therefore, the number of conjugacy classes in the group S_n equals the number of (unordered) partitions of n into a sum of natural numbers.

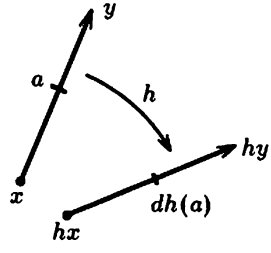


Figure 10.1

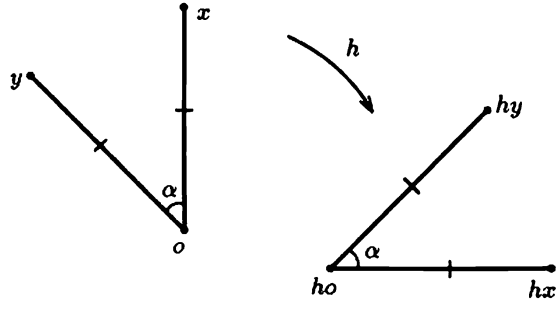


Figure 10.2

Example 10.40. The group $\mathrm{Isom}_+ E^2$ of proper motions of the Euclidean plane consists of parallel translations and rotations (Section 7.3). By the above-stated principle, the motion which is conjugate by a motion h to the parallel translation along a vector a is the parallel translation along $dh(a)$ (see Figure 10.1). (We already proved this in Section 4.2.) Similarly, the motion which is conjugate by h to the rotation about a point o through α is the rotation about ho through α (see Figure 10.2). Therefore, the conjugacy classes of the group $\mathrm{Isom}_+ E^2$ are of two types:

(i) the set of parallel translations along vectors of a given length $r \geq 0$;

(ii) the set of rotations through a given angle $\alpha \in (0, 2\pi)$.

Exercise 10.41. Describe conjugacy classes of the group of rotations of a cube.

Consider actions α and β of the same group G on sets X and Y, respectively. The map $f\colon X \to Y$ is called *equivariant* or, more precisely, G-equivariant if for any $g \in G$, the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle \alpha(g)}\downarrow & & \downarrow{\scriptstyle \beta(g)} \\ X & \xrightarrow[f]{} & Y \end{array}$$

is commutative. An equivariant bijection is called an *action isomorphism*. (A diagram of sets and maps is called *commutative* if the composition of maps along any path with the same beginning and end results in the same map.)

The subgroup

$$G_x = \{g \in G\colon gx = x\}$$

is called the *stabilizer* of a point x.

For a subgroup $H \subset G$, define an action of G on the set G/H of left cosets as follows:

$$g(uH) = (gu)H.$$

The following theorem shows that the action of a group G on an orbit Gx is determined by the stabilizer of x up to an isomorphism. This theorem generalizes Theorem 4.76 (and makes it more precise).

Theorem 10.42. *The map*

$$f\colon Gx \to G/G_x, \qquad y \mapsto G_x^y = \{g \in G\colon gx = y\}$$

is an action isomorphism.

Proof. (i) When $y = gx$, the set G_x^y coincides with the coset gG_x. Indeed,

$$g_1x = y \iff g^{-1}g_1x = x \iff g^{-1}g_1 \in G_x \iff g_1 \in gG_x.$$

(ii) It is clear from the definition that the map f is bijective.

(iii) The map f is equivariant. Indeed, let $y = ux$ for $u \in G$. Then for any $g \in G$,

$$f(gy) = f((gu)x) = (gu)G_x = g(uG_x) = gf(y).$$

□

Corollary 10.43. *Every transitive group action of G is isomorphic to its action on the set of left cosets of a subgroup of G.*

Corollary 10.44. *If G is finite,*

$$|Gx| = \frac{|G|}{|G_x|}. \tag{10.9}$$

(Here $|Gx|$ denotes the number of elements in the orbit Gx.)

It is easy to see that

$$G_{gx} = gG_xg^{-1}. \tag{10.10}$$

Since we can choose any point on the given orbit as the point x in the statement of Theorem 10.42, the actions of G on G/H and G/gHg^{-1} are isomorphic for any $H \subset G$ and $g \in G$.

The ineffectiveness kernel of the action $G : G/H$ is the intersection of stabilizers of all points, i.e., $\bigcap_{g\in G} gHg^{-1}$. This is the largest normal subgroup of G that is contained in H. In particular, H is normal if and only if it acts trivially on G/H.

Example 10.45. Consider a cube $K \subset E^3$. The isomorphism $S_4 \xrightarrow{\sim} \mathrm{Sym}_+ K$ (see Example 4.116) defines an action $S_4 : E^3$. This action, in turn, induces actions of the group S_4 on the set of vertices of the cube, on the set of its diagonals, etc. In the table below, we list several transitive actions $S_4 : X$ obtained in such a way. For each, we describe the stabilizer H of an element of X. In every case $|X||H| = |S_4| = 24$, as follows from Corollary 10.44.

Elements of X	$\lvert X\rvert$	$\lvert H\rvert$	H
cube edges	12	2	$\langle(12)\rangle$
diagonals of cube faces	12	2	$\langle(12)(34)\rangle$
cube vertices	8	3	$\langle(123)\rangle$
cube faces	6	4	$\langle(1234)\rangle$
pairs of opposite edges	6	4	$\langle(12),(34)\rangle$
pairs of opposite vertices (or diagonals)	4	6	S_3
pairs of opposite faces	3	8	$\langle V_4,(1234)\rangle$

Example 10.46. We will prove here that if $|G| = n < \infty$ and p is the least prime divisor of n, then every subgroup $H \subset G$ of index p is normal. Indeed, consider the action $H : G/H$. The number of elements in every orbit of this action divides $|H|$, thus it is either 1 or greater than or equal to p. Since $|G/H| = p$ and there exists at least one fixed point (the coset eH), we conclude that the action is trivial.

Exercise 10.47. Consider an action of a finite group G on a finite set X. Denote the set of orbits of this action by X/G. For every element $g \in G$, denote the set of points fixed by g in X by X^g. Prove *Burnside's Formula*

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

(*Hint*: compute in two ways the number of elements in the set $F = \{(g, x) \in G \times X : gx = x\}$.)

Exercise 10.48. Using Burnside's formula and the result of Exercise 10.41, determine the number of different ways to color the faces of a cube in three colors. (Two colorings are considered different if one cannot be obtained from the other by rotating the cube.)

For the action of G on itself by conjugations, the stabilizer of a point x is the subgroup

$$Z(x) = \{g \in G : gx = xg\}$$

called the *centralizer* of x. Denote by $C(x)$ the conjugacy class of x (it is an orbit of this action). For a finite group H, formula (10.9) implies

$$|C(x)| = \frac{|G|}{|Z(x)|}. \tag{10.11}$$

The action of G on itself by conjugations induces an action of G on the set of its subgroups. Subgroups that are equivalent with respect to this action are called *conjugate*. (For instance, the stabilizers of equivalent points under any action of G are conjugate subgroups according to (10.10).) For this action, the stabilizer of a subgroup $H \subset G$ is the subgroup

$$N(H) = \{g \in G : gHg^{-1} = H\}$$

called the *normalizer* of H. When the group G is finite, formula (10.9) shows that the number of subgroups that are conjugate to H equals $[G : N(H)]$ (the index of $N(H)$). Notice that $N(H) \supset H$, thus in the case of a finite G, $[G : N(H)]$ divides $[G : H]$.

10.4. Sylow Theorems

Let p be a prime number. Recall that a finite group G is called a p-group if $|G| = p^n$.

Theorem 10.49. *A nontrivial p-group has a nontrivial center.*

Proof. Let G be a nontrivial p-group with center Z. The set $G \setminus Z$ splits into nontrivial conjugacy classes; also by (10.11), the number of elements in each of these classes is divisible by p. Therefore, the number of elements in the center is also divisible by p. □

Corollary 10.50. *Every p-group is solvable.*

Proof. Let G be a nontrivial p-group with center Z. While proving the claim by induction on $n = \log_p |G|$, we can assume that the group G/Z is solvable. Since the group Z is solvable (and even abelian), we conclude that G is solvable too. □

Corollary 10.51. *Every group of order p^2 is abelian.*

Proof. Let G be a group of order p^2 with center Z. Assume that $Z \neq G$. Then $|Z| = p$ and $|G/Z| = p$, so that G/Z is cyclic. Let aZ be its generator. Then every element $g \in G$ has the form $g = a^k z$, $z \in Z$. But every two elements of this form commute, and this contradicts our assumption. □

Now let $|G| = p^n m$, where $(p, m) = 1$.

Definition 10.52. A *Sylow p-subgroup* of the group G is a subgroup of G of order p^n.

Theorem 10.53. *Sylow p-subgroups exist.*

Proof. If G is abelian, then its (only) Sylow p-subgroup is the p-torsion subgroup (see Section 9.1).

In the general case, we will prove this theorem by induction on $|G|$.

If $|G| = 1$, there is nothing to prove. Let $|G| > 1$. Consider the decomposition of G into conjugacy classes. Two cases are possible.

Case 1. There exists a nontrivial class $C(x)$ with the number of elements not divisible by p. Then p^n divides $|Z(x)|$ and by the induction hypothesis, there exists a subgroup of $Z(x)$ of order p^n. It is a Sylow p-subgroup of G.

Case 2. No such class exists. Then, as in the proof of Theorem 10.49, we obtain that $|Z|$ is divisible by p. Let $|Z| = p^{n_0} m_0$, where $(p, m_0) = 1$, and let $Z_1 \subset Z$ be the subgroup of order p^{n_0}. The order of the group G/Z_1 is $p^{n-n_0} m$, thus by the induction hypothesis, it contains a subgroup of order p^{n-n_0}. Its full preimage under the canonical homomorphism $G \longrightarrow G/Z_1$ is a Sylow p-subgroup of G. □

Theorem 10.54. *Any p-subgroup of the group G is contained in some Sylow p-subgroup. All Sylow p-subgroups are conjugate.*

Proof. Let $S \subset G$ be a Sylow p-subgroup and S_1 any p-subgroup. Consider the action of S_1 on G/S. Since the number of elements in any nontrivial orbit of S_1 is divisible by p and the number of elements of the set G/S is not divisible by p, S_1 has at least one fixed point in G/S. If gS is such a point, then $S_1 \subset gSg^{-1}$. This proves the first statement of the theorem.

Moreover, if S_1 is a Sylow p-subgroup, by comparing the orders, we see that $S_1 = gSg^{-1}$. □

Exercise 10.55. By a similar argument, prove that in a group of order p^n, every subgroup H of order p^k, $k < n$, has at least one fixed point in G/H that is different from eH. Conclude that $N(H) \neq H$ and that H is contained in a subgroup of order p^{k+1}.

Theorem 10.56. *The number of Sylow p-subgroups is congruent to 1 modulo p.*

Proof. Let S be a Sylow p-subgroup and $C(S)$, a class of subgroups conjugate to S. By Theorem 10.54, this is the set of all Sylow p-subgroups. When G acts on $C(S)$ by conjugations, the stabilizer of any subgroup $S' \in C(S)$ is its normalizer $N(S')$. Restrict this action to S. Then $C(S)$ splits into nontrivial S-orbits (the number of elements in each is divisible by p) and fixed points. We prove that the only fixed point is the subgroup S itself. This will imply that

$$|C(S)| \equiv 1 \pmod{p}.$$

Let $S' \in C(S)$ be a fixed point. This means that $S \subset N(S')$. Then S and S' are Sylow p-subgroups of $N(S')$, hence they are conjugate in this subgroup. Therefore, $S = S'$. □

Theorem 10.56 together with the fact that the number of Sylow p-subgroups divides the index of (any) Sylow p-subgroup, sometimes allows us to conclude that the Sylow p-subgroup is unique, and hence normal.

Example 10.57. We will prove here that every group G of order pq, where p and q are distinct prime numbers, is a semidirect product of cyclic subgroups of order p and q (see Example 10.23). Indeed, let $p > q$. Then it follows from Example 10.46 that the Sylow p-subgroup G_p is normal. If G_q is a Sylow q-subgroup, then $G_p \cap G_q = \{e\}$. Hence, $|G_pG_q| = pq = |G|$. Therefore,

$$G = G_p \ltimes G_q.$$

Example 10.58. We will prove here that every group G of order 45 is abelian. Indeed, let N_p, $p = 3, 5$, be the number of its Sylow p-subgroups. Then

$$\{N_3 \equiv 1 \pmod{3},\ N_3|5\} \implies N_3 = 1,$$
$$\{N_5 \equiv 1 \pmod{5},\ N_5|9\} \implies N_5 = 1.$$

Thus Sylow subgroups G_3 and G_5 are normal and

$$G = G_3 \times G_5.$$

But the group G_3 has order 9, hence is abelian by Corollary 10.51. Therefore, G is abelian.

Exercise 10.59. Prove that all groups of order < 60 are solvable. (*Hint:* prove that if $|G| = n < 60$, then for some prime p, $p|n$, the number N of Sylow p-subgroups of G does not exceed 4. If $N > 1$, consider the action of G by conjugation on the set of Sylow p-subgroups and obtain a nontrivial homomorphism $G \to S_N$.)

10.5. Simple Groups

Definition 10.60. A nontrivial group G is *simple* if it does not contain nontrivial normal subgroups (i.e., those different from $\{e\}$ and G).

A solvable simple group G is a cyclic group of prime order. Indeed, since $G' \neq G$, we must have $G' = \{e\}$, i.e., G is abelian. But all subgroups of an abelian group are normal. Hence, G is cyclic; moreover, its order is a prime.

Thus, there are two kinds of simple groups:

(i) abelian, which are only cyclic groups of prime order;

(ii) nonabelian (hence unsolvable).

An example of a nonabelian simple group is the group A_5 (for the proof of its simplicity, see below).

The following discussion explains the value of simple groups. Let us take a chain of subgroups:

$$G = G_0 \supset G_1 \supset \cdots \supset G_{m-1} \supset G_m = \{e\}, \tag{10.12}$$

where $G_{i+1} \lhd G_i$, $i = 0, 1, \ldots, m-1$. If the quotient group

$$F_i = G_i/G_{i+1}$$

contains a nontrivial normal subgroup N, then we can insert another group between G_i and G_{i+1}, namely the full preimage of N under the canonical homomorphism $G_i \to F_i$. Thus if G is finite, it is possible to construct a chain (10.12) where all quotient groups (factors) are simple. In either case, such a chain, if it exists, is called a *composition series* of G.

One can quite easily prove the *Jordan–Hölder theorem*: if a group G has a composition series, then the collection of its factors is determined uniquely up to a permutation. Therefore, to every group that has a composition series (e.g., a finite group), we can canonically associate a collection of simple groups. This is why the classification of simple groups is fundamental to understanding of the structure of all groups.

Exercise 10.61. Prove that a finite group is solvable if and only if all factors of its composition series are abelian.

The classification of nonabelian simple finite groups is tremendously difficult. It was obtained after a 30-year long effort by several hundred mathematicians worldwide and was completed by 1981.[1] We restrict ourselves to considering several examples only.

Proposition 10.62. *The group A_n is simple for $n \geq 5$.*

Lemma 10.63. *If the cycle decomposition of a permutation $\sigma \in A_n$ contains a cycle of even length or two cycles of the same odd length, then the conjugacy class of σ in A_n coincides with its conjugacy class in S_n.*

(Here any fixed element of $\{1, \ldots, n\}$ is regarded as a cycle of length 1.)

Proof. In either case, the centralizer of σ in S_n contains an odd permutation τ_0. Indeed, if the decomposition of σ contains a cycle of even length, then we can take it as τ_0, and if the decomposition contains cycles $(i_1 i_2 \cdots i_q)$, $(j_1 j_2 \cdots j_q)$ of equal odd length q, then we can take $\tau_0 = (i_1 j_1)(i_2 j_2) \cdots (i_q j_q)$. Now let τ be any odd permutation. Then

$$\tau \sigma \tau^{-1} = (\tau \tau_0) \sigma (\tau \tau_0)^{-1},$$

where $\tau \tau_0$ is already even. □

In particular, all products of pairs of disjoint transpositions and, for $n \geq 5$, all triple cycles are conjugate not only in S_n but also in A_n.

Proof of Proposition 10.62. Let $N \subset A_n$ be a normal subgroup containing a permutation $\sigma \neq e$. By taking an appropriate power of σ, we can assume that is has a prime order p. Then σ is a product of a number of disjoint cycles of length p.

Consider separately the following three possibilities:

(i) Let $p \geq 5$. Write σ as

$$\sigma = (i_1 i_2 i_3 i_4 \cdots i_p)\tau,$$

where τ is the product of remaining cycles (if there are none, put $\tau = e$), and conjugate it by the triple cycle $(i_1 i_2 i_3)$. We have

$$\sigma' = (i_1 i_2 i_3)\sigma(i_1 i_2 i_3)^{-1} = (i_2 i_3 i_1 i_4 \cdots i_p)\tau \in N$$

(cf. Example 10.39). Hence, $\sigma' \sigma^{-1} = (i_1 i_2 i_4) \in N$. Since all triple cycles are conjugate in A_n and all of them together generate A_n (see Proposition 10.25), we have $N = A_n$.

[1] The proof of this classification is distributed among many papers (total size about 10,000 pages) by different authors. Nobody can claim that he/she checked the entire proof. For this reason, some experts doubt that the proof is correct and complete.

(ii) Let $p = 3$. If σ is just a triple cycle, then, as above, we have $N = A_n$. Otherwise, write σ as

$$\sigma = (i_1 i_2 i_3)(j_1 j_2 j_3)\tau,$$

where τ is the product of remaining cycles, and conjugate it by $(i_1 j_1)(i_2 j_2)$. We have

$$\sigma' = (i_1 i_2 j_3)(j_1 j_2 i_3)\tau \in N,$$
$$\sigma'\sigma^{-1} = (i_1 j_1)(i_3 j_3) \in N.$$

Since all products of pairs of disjoint transpositions are conjugate in A_n and all of them together generate A_n (see Proposition 10.25), here we also have $N = A_n$.

(iii) Finally, let $p = 2$. Then the permutation σ is a product of an even number of disjoint transpositions. Write it as

$$\sigma = (i_1 i_2)(i_3 i_4)\tau,$$

where τ is the product of remaining transpositions, and conjugate it by $(i_1 i_2 i_3)$. We obtain

$$\sigma' = (i_2 i_3)(i_1 i_4)\tau \in N,$$
$$\sigma'\sigma^{-1} = (i_1 i_3)(i_2 i_4) \in N,$$

which implies, as above, that $N = A_n$. □

In particular, the group A_5 is a simple group of order 60. By Exercise 10.59, this is the least order that a nonabelian simple group might have. Observe that the above proof simplifies significantly for the case $n = 5$ as it comes down to considering the case where σ is a cycle of length 5 (modulo Lemma 10.63).

Exercise 10.64. Prove that the only nontrivial normal subgroup of A_4 is the Klein group V_4.

Exercise 10.65. Prove that every simple group G of order 60 is isomorphic to A_5. (*Hint:* by considering the action of G on the set of its Sylow 5-subgroups, obtain an inclusion $G \subset A_6$; then consider the action of G on A_6/G).

To present another example of a series of simple groups, we state (without proof) the following fact: for $n \geq 2$, the group

$$\mathrm{PSL}_n(K) = \mathrm{SL}_n(K)/\{\lambda E\colon \lambda \in K^*,\ \lambda^n = 1\}$$

is simple except when $n = 2$ and K is a finite field of two or three elements.

Exercise 10.66. Determine the order of the group $\mathrm{PSL}_2(\mathbb{F}_q)$, where $\mathbf{F}_q$ is the finite field of q elements. Prove that

$$\mathrm{PSL}_2(\mathbb{F}_2) \simeq S_3, \quad \mathrm{PSL}_2(\mathbb{F}_3) \simeq A_4, \quad \mathrm{PSL}_2(\mathbb{F}_4) \simeq \mathrm{PSL}_2(\mathbb{F}_5) \simeq A_5.$$

(*Hint:* consider the natural action of the group $G = \mathrm{PSL}_2(\mathbf{F}_q)$ on the projective plane $P\mathbb{F}_q^2$ over the field $\mathbb{F}_q$; for $q = 5$ argue as in Exercise 10.65.)

The group $\mathrm{PSL}_2(\mathbb{F}_7)$ is a simple group of order 168. Orderwise, this is the next nonabelian simple group after A_5. The group $\mathrm{PSL}_2(\mathbb{F}_9)$ is actually isomorphic to A_6.

Exercise 10.67. Prove that the group $\mathrm{PSL}_2(\mathbb{C})$ is simple.

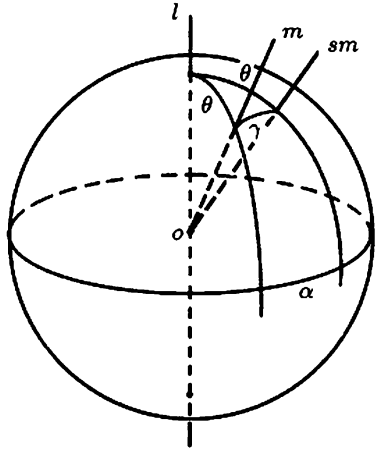


Figure 10.3

To demonstrate how geometric considerations can be applied to the proofs of simplicity of (infinite) groups, we show here that the group SO_3 is simple.

Every element of the group SO_3 is a rotation through an angle α about some axis. The conjugation by an element $g \in \mathrm{SO}_3$ of the rotation through α about an axis l results in the rotation through α about the axis gl (Example 10.40). Thus every normal subgroup of the group SO_3 that contains a rotation through α about some axis should contain a rotation through α about any axis.

Let m and m' be two lines; denote the angle between them by γ. It is easy to see (see Example 6.41) that the product of rotations through π

about m and m' is the rotation through 2γ about the axis perpendicular to the plane spanned by m and m'.

Assume now that $N \subset \mathrm{SO}_3$ is a normal subgroup that contains a rotation through an angle $\alpha \in (0, 2\pi)$ about an axis l. Let g be a rotation through π about an axis m such that it forms an angle $\theta \in [0, \frac{\pi}{2}]$ with l. Then

$$h = g(sgs^{-1}) = (gsg^{-1})s^{-1} \in N.$$

Since sgs^{-1} is the rotation through π about the axis sm, by the above remark, h is a rotation through 2γ, where γ is the angle between m and sm (see Figure 10.3). The angle γ equals 0 when $\theta = 0$, and equals α when $\theta = \frac{\pi}{2}$. Continuity implies that it attains all values between 0 and α. Therefore, N contains rotations through all angles between 0 and 2α. By taking powers of these rotations, we can obtain rotations through all angles. This shows that $N = \mathrm{SO}_3$.

It can be shown as well that the group SO_n is simple for any $n \geq 3$, except for $n = 4$. Nonsimplicity of the group SO_4 is a surprising fact discussed in Section 12.4.

10.6. Galois Extensions

Extensions of a field K obtained by adjoining roots of irreducible polynomials can turn out to be isomorphic. More generally, one of them can be isomorphic to a subfield of another, and it is not easy to figure out when this happens. The subject of this section, Galois theory, studies exactly homomorphisms (and, in particular, automorphisms) of algebraic field extensions.

We explained in Section 4.2 what role played by groups in geometry and physics. However, the theory of groups originated in Galois theory, where groups appear in a fundamentally different way. Ideas of Galois theory emerge in other mathematical fields. For instance, in topology the analogue of Galois theory is the theory of coverings (in particular, the analogue of the Galois group of a field is the fundamental group of a topological space) and in the theory of functions of a complex variable, such analogue is the theory of holomorphic maps of Riemann surfaces.

Let L be a field extension of K of finite degree n. The automorphisms of the field L over K form a group which we denote as $\mathrm{Aut}_K L$.

Proposition 10.68. $|\mathrm{Aut}_K L| \leq n$.

Proof. The field L can be obtained from K by taking successive simple extensions

$$K = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_s = L,$$

where K_i is obtained from K_{i-1} by adjoining a root α_i of an irreducible polynomial $f_i \in K_{i-1}[x]$. By Lemma 9.115, any homomorphism $\varphi_{i-1}\colon K_{i-1} \to L$

extends to a homomorphism $\varphi_i\colon K_i \to L$ in at most n_i ways, where

$$n_i = \deg f_i = \dim_{K_{i-1}} K_i.$$

Therefore, the identity automorphism of the field K extends to an automorphism of the field L in at most $n_1 n_2 \cdots n_s = n$ ways. □

Let $G \subset \operatorname{Aut}_K L$ be a (finite) subgroup of the group of automorphisms of the field L over K. Denote by L^G the subfield of G-invariant elements of L.

Theorem 10.69. *$L^G = K$ if and only if $|G| = n$. Moreover, if $L^G = K$, then for any two fields P, Q such that $K \subset P \subset Q \subset L$, there exist exactly $\dim_P Q$ extensions $\psi\colon Q \to L$ of any homomorphism $\varphi\colon P \to L$ over K.*

Proof. (i) By definition, $G \subset \operatorname{Aut}_{L^G} L$. Therefore,

$$|G| \le \dim_{L^G} L \le \dim_K L = n.$$

If $|G| = n$, then $\dim_{L^G} L = \dim_K L$, hence $L^G = K$.

(ii) Conversely, let $L^G = K$. For any element $\alpha \in L$, let $\{\alpha_1, \dots, \alpha_m\}$ be its G-orbit. Then

$$f = \prod_{i=1}^{m} (x - \alpha_i) \in L^G[x] = K[x] \tag{10.13}$$

is the minimal polynomial of α over K. By construction, it splits into different linear factors over $L[x]$.

Let us prove now the second statement of the theorem. Since every finite extension can be obtained by taking successive simple extensions, it suffices to consider the case when $Q = P(\alpha)$ is a simple extension of P. Let h be the minimal polynomial of α over P. Then in the ring $P[x]$, h divides the minimal polynomial f of α over K. Thus, h^φ divides f in the ring $\varphi(P)[x]$, hence it splits into different linear factors in $L[x]$. By Lemma 9.115, the homomorphism φ extends to a homomorphism $\psi\colon Q \to L$ in exactly $\deg h = \dim_P Q$ ways.

Applying the above discussion to the case of $P = K$, $Q = L$, we obtain $|\operatorname{Aut}_K L| = n$.

It remains to show that $G = \operatorname{Aut}_K L$. Let $\varphi \in \operatorname{Aut}_K L$. Then for any $\alpha \in L$, the element $\varphi(\alpha)$, as well as α, is a root of the polynomial (10.13), i.e., there exists an element $g \in G$ (it might depend on α) such that $\varphi(\alpha) = g\alpha$.

If the field L is finite, then we can take for α a generator of the group L^*. Then, of course, $\varphi = g \in G$. If L is infinite (and, then, so is K), for any $g \in G$, we define

$$L_g = \{\alpha \in L\colon \varphi(\alpha) = g\alpha\} \subset L.$$

Obviously, L_g is a subspace (even a subfield) over K. The above implies that

$$L = \bigcup_{g \in G} L_g.$$

We need to conclude from here that $L = L_g$ for some $g \in G$. This follows from the next lemma. □

Lemma 10.70. *A finite-dimensional vector space K over an infinite field cannot be covered by a finite number of proper subspaces.*

Proof. Let $V = \bigcup_{i=1}^{s} V_i$, where $V_1, \ldots, V_s$ are proper subspaces of V. For every i, consider a nonzero linear function $l_i \in V^*$ that becomes zero on V_i. Consider the polynomial $F = \prod_{i=1}^{s} l_i$. By our assumption, $F(v) = 0$ for each $v \in V$. Then F is the zero polynomial, and this is clearly not true. □

Definition 10.71. A finite extension L of a field K is a *Galois extension* if

$$|\operatorname{Aut}_K L| = \dim_K L.$$

In this case, the group $\operatorname{Aut}_K L$ is called the *Galois group* of the extension L and is denoted $\operatorname{Gal} L/K$.

Theorem 10.69 implies that if L is a Galois extension of K and $P \subset L$ is a subfield containing K, then L is a Galois extension of P.

A polynomial $f \in K[x]$ is called *separable* if it does not have multiple roots in any extension of K.

Denote by f' the formal derivative of a polynomial f.

Proposition 10.72. *A polynomial $f \in K[x]$ is separable if and only if $(f, f') = 1$. In particular, an irreducible polynomial f is separable if and only if $f' \neq 0$.*

Proof. First of all, observe that the greatest common divisor of any two polynomials $f, g \in K[x]$ can be found using the Euclidean algorithm, hence it is the same over any extension of K. On the other hand, if a polynomial f has a multiple irreducible factor h over an extension L of K, then $h | f'$ in $L[x]$, thus $(f, f') \neq 1$. In particular, this happens if f is not separable.

Conversely, if f is separable, then it splits into different linear factors over its splitting field, and it easily follows that $(f, f') = 1$. □

Corollary 10.73. *Every irreducible polynomial over a field of zero characteristic is separable.*

Corollary 10.74. *Every irreducible polynomial f over a field of characteristic $p \nmid \deg f$ is separable.*

Corollary 10.75. *Every irreducible polynomial over a finite field is separable.*

Proof. Let h be a nonseparable irreducible polynomial over a finite field K. Then $h' \neq 0$, hence

$$h = a_0 + a_1 x^p + a_2 x^{2p} + \cdots + a_m x^{mp}, \qquad a_0, a_1, \ldots, a_m \in K.$$

Since $K^p = K$ (see Section 9.5), there exist $b_0, b_1, \ldots, b_m \in K$ such that $b_k^p = a_k$. Thus, h has the form

$$h = (b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m)^p$$

and is reducible, a contradiction. □

An example of a nonseparable irreducible polynomial is the polynomial

$$x^p - t = (x - \sqrt[p]{t})^p$$

over the field $\mathbb{Z}_p(t)$.

Theorem 10.76. *Let $f \in K[x]$ be a polynomial such that all its irreducible factors are separable. Then its splitting field is a Galois extension of K.*

Proof. This actually follows from the proof of the second part of Theorem 9.114 for the case $\widetilde{L} = L$ if we take into account that under our assumptions, all polynomials $\widetilde{f}_i$ are separable. □

Observe that if L is the splitting field of a polynomial $f \in K[x]$, then every automorphism φ of the field L over K preserves the set $\{\alpha_1, \ldots, \alpha_n\}$ of roots of f and can only permute them. Since $L = K(\alpha_1, \ldots, \alpha_n)$, the automorphism φ is uniquely determined by the permutation that it performs on the set of roots. Thus, the group $\mathrm{Aut}_K L$ embeds into S_n.

Example 10.77. It follows from the formula for solutions of a quadratic equation that every quadratic extension of a field K of characteristic $\neq 2$ is of the form $K(\sqrt{d})$, where $d \in K \setminus K^2$. Every such extension is a Galois extension. Its Galois group is generated by the automorphism $a + b\sqrt{d} \mapsto a - b\sqrt{d}$, $a, b \in K$.

Example 10.78. The finite field $\mathbb{F}_q$, $q = p^n$, is a Galois extension of the field $\mathbb{Z}_p$. Its Galois group is the cyclic group of order n generated by the Frobenius automorphism.

Example 10.79. The cyclotomic field $K_n = \mathbb{Q}(e^{2\pi i/n})$ is the splitting field of the polynomial $x^n - 1$ over $\mathbb{Q}$ and is thus a Galois extension of the field $\mathbb{Q}$. Every automorphism of K_n induces an automorphism of the (cyclic) group C_n of roots of unity of order n contained in K_n. As we know, every automorphism of the group C_n is the raising to the kth power for some k

relatively prime to n. Thus, the group $\operatorname{Gal} K_n/\mathbb{Q}$ embeds into the group $\mathbb{Z}_n^*$, whose order is $\varphi(n)$ (where φ is Euler's function). In fact, this embedding is an isomorphism. To show this, it suffices to prove that for any prime p such that $p \nmid n$, there exists an automorphism of the field K_n that induces raising to the pth power in C_n. This means that if f is the minimal polynomial of $\varepsilon = e^{2\pi i/n}$ over $\mathbb{Q}$, then $f(\varepsilon^p) = 0$.

We may assume that f is a monic polynomial with integer coefficients. Then, by Gauss's lemma, $x^n - 1 = fg$, where $g \in \mathbb{Z}[x]$. It is easy to see that the polynomial $x^n - 1$ remains separable after the reduction modulo p. Therefore, the polynomials $[f]_p$ and $[g]_p$ are relatively prime.

Assume now that $f(\varepsilon^p) \neq 0$. Then $g(\varepsilon^p) = 0$, hence

$$g(x^p) = f(x)h(x), \qquad h \in \mathbb{Z}[x].$$

Reducing modulo p, we obtain

$$[g]_p^p = [f]_p[h]_p,$$

which contradicts the fact that $[f]_p$ and $[g]_p$ are relatively prime.

Therefore, $\dim_{\mathbb{Q}} K_n = \varphi(n)$.

Example 10.80. Let L be the splitting field of an irreducible cubic polynomial f with the discriminant D over a field K of characteristic $\neq 2, 3$ (see Example 9.116). Then L is a Galois extension of K. If $D \notin K^2$, then $\dim_K L = 6$ and $\operatorname{Gal} L/K \simeq S_3$. If $D \in K^2$, then $\dim_K L = 3$ and $\operatorname{Gal} L/K \simeq A_3$. The latter statement means that the Galois group performs only even permutations on the set of roots of f.

Example 10.81. Let

$$f = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$$

be a "generic" polynomial of degree n whose coefficients are regarded as elements of the field $K = k(a_1, \ldots, a_n)$ of rational functions in n independent variables over some field k. Let L be the splitting field of f over K and $x_1, \ldots, x_n \in L$, its roots.

By Viète's formula, $a_k = (-1)^k \sigma_k$, where $\sigma_1, \ldots, \sigma_n$ are elementary symmetric polynomials in $x_1, \ldots, x_n$. Therefore, $L = k(x_1, \ldots, x_n)$. Since

$$\operatorname{tr.deg} L = \operatorname{tr.deg} K = n,$$

$x_1, \ldots, x_n$ are algebraically independent over k. In particular, they are distinct, so that f is a separable polynomial and L is a Galois extension of the field K. Every permutation of the roots $x_1, \ldots, x_n$ defines an automorphism of the field L acting trivially on K. Therefore, $\operatorname{Gal} L/K \simeq S_n$. We have also proved that

$$k(x_1, \ldots, x_n)^{S_n} = k(\sigma_1, \ldots, \sigma_n).$$

10.7. Fundamental Theorem of Galois Theory

For any Galois extension L/K, Galois theory establishes a correspondence between the subfields of L containing K and the subgroups of the group $G = \operatorname{Gal} L/K$.

Namely, to every subfield $P \subset L$ containing K, there corresponds the subgroup

$$G_P = \{g \in G\colon g|_P = \mathrm{id}\} \subset G,$$

and to every subgroup $H \subset G$, there corresponds the subfield

$$L^H = \{\alpha \in L\colon h\alpha = \alpha\ \forall h \in H\} \subset L.$$

Theorem 10.69 shows that

$$|G_P| = \dim_P L, \tag{10.14}$$

$$\dim_{L^H} L = |H|. \tag{10.15}$$

Theorem 10.82 (Fundamental Theorem of Galois Theory). *The maps $P \mapsto G_P$ and $H \mapsto L^H$ described above, are inverse to each other and thus establish a one-to-one correspondence between the set of subfields of L containing K and the set of subgroups of G. Moreover, a Galois extension of K contained in L corresponds to a normal subgroup of G and vice versa.*

Proof. Obviously,

$$L^{G_P} \supset P.$$

Also, (10.14) and (10.15) imply that

$$\dim_{L^{G_P}} L = |G_P| = \dim_P L.$$

Therefore,

$$L^{G_P} = P.$$

Likewise, we can prove that

$$G_{L^H} = H.$$

Now, since by Theorem 10.69, every automorphism of a subfield P extends to an automorphism of L, the field P is a Galois extension of K if and only if the transformations in G that leave it invariant induce $\dim_K P$ distinct automorphisms on it. But formula (10.14) implies that

$$\dim_K P = |G : G_P|.$$

Thus, P is a Galois extension if and only if every transformation in G leaves it invariant.

Since $P = L^H$ for $H = G_P$,

$$gP = L^{gHg^{-1}}.$$

Therefore, the subfield P is invariant under every transformation in G if and only if the subgroup H is normal. □

Example 10.83. Let f be an irreducible cubic polynomial over a field K of characteristic $\neq 2$ and $D \notin K^2$ (see Example 9.116). Let L be the splitting field of f. Then $\operatorname{Gal} L/K \simeq S_3$. The subgroup $A_3 \subset S_3$ corresponds to the quadratic extension $K(\sqrt{D})$ contained in L.

Example 10.84. Let φ be the Frobenius automorphism of the finite field $\mathbf{F}_{p^n}$, p prime. As we know, $\operatorname{Gal} \mathbf{F}_{p^n}/\mathbf{Z}_p = \langle\varphi\rangle_n$. Any subgroup of the group $\langle\varphi\rangle_n$ has the form $\langle\varphi^m\rangle_{n/m}$, where $m|n$. The corresponding subfield is the field of fixed points of the automorphism φ^m. It has dimension m over $\mathbf{Z}_p$, i.e., it is isomorphic to $\mathbf{F}_{p^m}$.

Example 10.85. Let p be an odd prime number. The Galois group G of the cyclotomic field $K_p = \mathbb{Q}(\varepsilon_p)$ (see Examples 9.109 and 10.79) is a cyclic group of order $p-1$. Let $H \subset G$ be the (unique) subgroup of index 2. Then $P = K_p^H$ is a quadratic extension of $\mathbb{Q}$. Let us prove that

$$P = \begin{cases} \mathbb{Q}(\sqrt{p}) & \text{for } p \equiv 1 \pmod 4, \\ \mathbb{Q}(\sqrt{-p}) & \text{for } p \equiv -1 \pmod 4. \end{cases}$$

A generator of G acts on the group C_p of pth roots of unity by raising to the rth power (here r is such that $[r]_p$ generates $\mathbf{Z}_p^*$). Consider the following number:

$$\alpha = \varepsilon_p - \varepsilon_p^r + \varepsilon_p^{r^2} - \cdots - \varepsilon_p^{r^{p-2}} = \sum_{k=1}^{p-1} \left(\frac{k}{p}\right) \varepsilon_p^k,$$

where $\left(\frac{k}{p}\right)$ is the Legendre symbol (see Example 9.38). Clearly,

$$g(\alpha) = \begin{cases} \alpha & \text{for } g \in H, \\ -\alpha & \text{for } g \in G \setminus H. \end{cases}$$

Therefore, $\alpha \in P$ and $\alpha^2 \in \mathbb{Q}$.

By Examples 9.121 and 9.38, we have

$$\begin{aligned} \alpha^2 &= \frac{1}{p-1} \operatorname{tr} \alpha^2 = \frac{1}{p-1}(\alpha, \alpha) \\ &= \frac{p}{p-1} \sum_{k=1}^{p-1} \left(\frac{k}{p}\right)\left(\frac{-k}{p}\right) = p\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} p, \end{aligned}$$

as required.

Galois theory was created in relation with the problem of solvability of algebraic equations by radicals.

We say that an element α of a field extension of K is *radical* over K if it can be expressed in terms of elements of K using arithmetic operations and root extraction (of any degree). In other words, this means that α belongs to the last field in a chain of extensions

$$K = K_0 \subset K_1 \subset \cdots \subset K_s,$$

where $K_i = K_{i-1}(\alpha_i)$ for α_i such that $\alpha_i^{n_i} \in K_{i-1}$, $n_i \in \mathbb{N}$.

Proposition 10.86. *If a polynomial $f \in K[x]$ is irreducible and at least one of its roots is radical over K, then all of its roots are radical over K.*

Proof. Let α_1 and α_2 be roots of K in some extensions of the field K. Then there exists an isomorphism of the field $K(\alpha_1)$ onto the field $K(\alpha_2)$ mapping α_1 to α_2. Thus, if α_1 is radical over K, then so is α_2. □

An algebraic equation $f(x) = 0$, where $f \in K[x]$, is said to be *solvable by radicals* over K if all its roots are radical over K. This is equivalent to saying that the splitting field L of f over K is contained in a field obtained from K by successively adjoining roots of certain elements.

The main achievement of E. Galois (1830) regarding solvability of algebraic equations by radicals is the following

Theorem 10.87. *Let f be an irreducible polynomial over a field K of zero characteristic and let L be its splitting field over K. The equation $f(x) = 0$ is solvable by radicals over K if and only if the group* $\operatorname{Gal} L/K$ *is solvable.*

The proof of this theorem is based on the fact that if P is a field containing n distinct roots of unity of degree n, then its extension of the form $P(\alpha)$, where $\alpha^n = a \in P$, is a Galois extension with a cyclic Galois group whose order divides n. Below, we will provide a complete proof of a simpler version of this theorem that speaks of solvability by quadratic radicals.

The fact that the group S_n is solvable only for $n \leq 4$ together with Example 10.81 implies that a generic algebraic equation of degree n over an (arbitrary) field K of zero characteristic is solvable by radicals only for $n \leq 4$. Solvability of a generic equation of degree n over K by radicals means that it is possible to describe uniformly, i.e., by a general formula, the roots of any equation of degree n in terms of its coefficients and some fixed elements of K using arithmetic operations and root extraction. Absence of such a formula does not mean that a particular equation cannot be solved by radicals. For instance, every algebraic equation over $\mathbb{C}$ is solvable by radicals since all its roots lie in $\mathbb{C}$.

Traditionally, solvability of algebraic equations by radicals over $\mathbb{Q}$ attracted the biggest interest. One can deduce from Theorem 10.87 that for

any $n \geq 5$, there exists an irreducible polynomial $f \in \mathbb{Q}[x]$ of degree n such that the equation $f(x) = 0$ is not solvable by radicals.

Solvability by *quadratic radicals* is defined just as solvability by radicals, except that now only extraction of square roots is allowed.

Theorem 10.88. *Let f be an irreducible polynomial over a field K of characteristic $\neq 2$ and L, the splitting field of f over K. The equation $f(x) = 0$ is solvable by quadratic radicals over K if and only if*

$$\dim_K L = 2^n, \qquad n \in \mathbb{N}. \tag{10.16}$$

Proof. (i) Assume the equation $f(x) = 0$ is solvable by quadratic radicals. Then there exists a chain of quadratic extensions

$$K = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_s$$

such that $L \subset K_s$. We have

$$\dim_K L \,|\, \dim_K K_s = 2^s,$$

implying (10.16).

(ii) Conversely, let $\dim_K L = 2^n$. Then the group $G = \operatorname{Gal} L/K$ is a 2-group, hence it is solvable. Consider its composition series

$$G = G_0 \supset G_1 \supset G_2 \supset \cdots \supset G_s = \{e\}.$$

Obviously, $|G_{i-1}|/|G_i| = 2$ for each i. Denote $K_i = L^{G_i}$. We obtain the following chain of quadratic extensions:

$$K = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_s = L,$$

which proves that the equation $f(x) = 0$ is solvable by quadratic radicals. □

Remark 10.89. Since

$$\deg f = \dim_K K(\alpha),$$

where $\alpha \in L$ is a root of the polynomial f, equality (10.16) implies that $\deg f$ is a power of 2. The converse is false.

Remark 10.90. In the second part of the proof we used that L is a Galois extension of the field K. This is certainly true if $\operatorname{char} K = 0$. If $\operatorname{char} K = p > 2$, this follows from the fact that f is separable since its degree, a power of 2, is not divisible by p.

Solvability of equations by quadratic radicals was of interest in connection with compass and straightedge constructions.

Every problem on a compass and straightedge construction can be stated as follows: given a length unit and intervals of lengths $a_1, \ldots, a_k$, construct an interval of length α. Studying possible elementary steps of constructions,

one can prove that the above problem can be solved if and only if the number α is a quadratic radical over the field $K = \mathbb{Q}(a_1, \ldots, a_k)$.

Remark 10.91. When we speak about a real number being a quadratic radical over a field $K \subset \mathbb{R}$, the original definition does not preclude us from extracting square roots from negative numbers, which pushes us into the complex region. However, to construct a complex number means to construct its real and imaginary parts, and arithmetic operations over and extractions of square roots from complex numbers reduce to arithmetic operations over real numbers and extractions of square roots from positive numbers. All these operations can be performed by a straightedge and compass.

In particular, if α is transcendental over K, the problem of constructing it is unsolvable. This is how one proves that it is impossible to square the circle (if the radius of the circle is chosen as the length unit, this problem is equivalent to constructing an interval of length π).

If α is algebraic over K with the minimal polynomial $f \in K[x]$, Theorem 10.88 implies that the problem of constructing α is solvable if and only if the degree of the splitting field of f is a power of 2. In particular, it is necessary that the degree of f itself be a power of 2.

Example 10.92. The problem of doubling the cube reduces to that of constructing an interval of length $\sqrt[3]{2}$. Since the polynomial $x^3 - 2$ is irreducible over $\mathbb{Q}$ and its degree is not a power of 2, this problem is unsolvable.

Example 10.93. The problem of trisecting the angle equal to φ reduces to that of constructing an interval of length $\cos(\varphi/3)$, given an interval of length $\cos\varphi$. By the well-known formula,

$$\cos\varphi = 4\cos^3\frac{\varphi}{3} - 3\cos\frac{\varphi}{3},$$

hence $\alpha = \cos(\varphi/3)$ is a root of the polynomial

$$f = 4x^3 - 3x - \cos\varphi \in K[x],$$

where $K = \mathbb{Q}(\cos\varphi)$. If we are looking for a universal method of trisecting an angle that does not depend on the angular measure φ, we have to regard $\cos\varphi$ as an independent variable. Then the polynomial f is irreducible over K (check this!) and the problem is unsolvable, just as in the previous example. For particular angles (e.g., the right angle) the problem may be solvable, but one can produce values of φ for which it is not. The criterion for solvability is the presence of a root of f in K. For instance, if $\varphi = \frac{\pi}{3}$, then $K = \mathbb{Q}$ and $f = 4x^3 - 3x - \frac{1}{2}$ has no roots in K; therefore, in this case the problem is unsolvable.

Example 10.94. The problem of dividing a circle into n equal parts (cyclotomy) reduces to the problem of constructing an interval of length $\cos\frac{2\pi}{n}$ or, equivalently, of constructing the number $e^{2\pi\imath/n} = \cos\frac{2\pi}{n} + \imath\sin\frac{2\pi}{n}$. Therefore, cyclotomy is possible if and only if the degree of the cyclotomic field $K_n = \mathbb{Q}(e^{2\pi\imath/n})$ is a power of 2. It is known that the degree of K_n equals $\varphi(n)$ (see Example 10.79). If n is a prime number, $\varphi(n) = n - 1$, thus we must have $n = 2^m + 1$. It is easy to see that the number $2^m + 1$ is prime only if m is a power of 2. Thus, n must be of the form

$$n = 2^{2^k} + 1.$$

Such numbers are called *Fermat numbers*. For $k = 0, 1, 2, 3, 4$, we obtain the prime numbers

$$3,\ 5,\ 17,\ 257,\ 65537,$$

but for $k = 5$, the number is already not prime. These are the only prime Fermat numbers known at present.

Galois theory allows us to give a conceptual proof of the fundamental theorem of algebra of complex numbers; it uses only the following two properties of the fields $\mathbb{R}$ and $\mathbb{C}$:

(i) every polynomial of odd degree over $\mathbb{R}$ has a root in $\mathbb{R}$;

(ii) it is possible to extract a square root from any number in $\mathbb{C}$.

Both these properties can be easily proven without resorting to the fundamental theorem (see Sections 3.4 and 1.5, respectively).

Property (i) implies that over $\mathbb{R}$, there exist no irreducible polynomials of odd degree greater than 1. Hence, there exists no nontrivial finite extension of odd degree (because the minimal polynomial of every element of such an extension must be of odd degree).

Let $f \in \mathbb{C}[x]$. Denote by $\overline{f}$ the polynomial obtained from f by replacing all its coefficients with their conjugates. Then $\overline{f\overline{f}} = \overline{f}f = f\overline{f}$, hence $f\overline{f} \in \mathbb{R}[x]$. On the other hand, if $c \in \mathbb{C}$ is a root of the polynomial $f\overline{f}$, then either c or $\overline{c}$ is a root of f. Therefore, it suffices to show that every polynomial of positive degree with real coefficients has a root in $\mathbb{C}$.

Let $f \in \mathbb{R}[x]$ be a polynomial of positive degree with a splitting field $L \supset \mathbb{R}$ over $\mathbb{R}$. Let $G = \operatorname{Gal} L/\mathbb{R}$. Consider a Sylow 2-subgroup H of G and the field $L^H = K$. Since $\dim_{\mathbb{R}} K = |G : H|$ is an odd number, $G = H$ by the above, i.e., G is a 2-group. But then Theorem 10.88 implies that the field L is contained in a field obtained from $\mathbb{R}$ by successively adjoining square roots. It follows from property (ii) above that $L = \mathbb{R}$ or $\mathbb{C}$. Hence, f has a root in $\mathbb{C}$.

Chapter 11

Linear Representations and Associative Algebras

In applications of group theory, the most important role is played by their linear representations. There are two sources for linear representations of groups:

(i) for a group G of differentiable transformations with a common fixed point, taking the differential at this point is a linear representation of G;

(ii) a group action on a set X defines, according to formula (10.7), a linear representation of this group in the space of functions on X.

On the other hand, the matrix algebra is a rich object where calculations are extremely effective. It thus becomes a benchmark for the study of many algebraic structures. A comparison with the matrix algebra is achieved via a linear representation.

11.1. Invariant Subspaces

For a vector space V over a field K, we denote the (associative) algebra of all linear operators on V by $\mathrm{L}(V)$. If the space V is finite-dimensional, the linear operators can be represented as matrices in some basis, and this establishes an isomorphism between the algebra $\mathrm{L}(V)$ and the matrix algebra $\mathrm{L}_n(K)$, $n = \dim V$.

Definition 11.1. A *linear representation* of a set X in a vector space V is a map

$$R\colon X \to \mathrm{L}(V). \tag{11.1}$$

The space V is called the *representation space* and its dimension, the *dimension of the representation*. The operators $R(x)$, $x \in X$, are called the *representation operators*.

If the set X is endowed with some operations, it is natural to require for the representation to agree with them. Thus, a *linear representation of a group* is defined by the conditions

$$R(xy) = R(x)R(y), \qquad R(e) = \mathcal{E}$$

(so, it can be defined as a homomorphism into $\mathrm{GL}(V)$), while a *linear representation of an associative algebra* is defined by the conditions

$$R(x+y) = R(x) + R(y), \qquad R(xy) = R(x)R(y),$$
$$R(\lambda x) = \lambda R(x), \qquad \lambda \in K.$$

Nonetheless, at first we will study properties of linear representations that are independent of any operations on the set X.

Definition 11.2. Let $R\colon X \to \mathrm{L}(V)$ and $S\colon X \to \mathrm{L}(U)$ be two linear representations of the same set X over the same field. A *morphism* of the representation R to the representation S is a linear map $\varphi\colon V \to U$ satisfying the following property: for any $x \in X$, the diagram

$$\begin{array}{ccc} V & \xrightarrow{R(x)} & V \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ U & \xrightarrow[S(x)]{} & U \end{array}$$

is commutative. An invertible morphism is called an *isomorphism* of representations.

Linear representations R and S are called *isomorphic* if there exists an isomorphism of R to S. In this case, we write $R \simeq S$. In respective bases of V and U, isomorphic representations have the same matrices.

Example 11.3. A linear representation of a one-point set is just a linear operator on a vector space. Two linear representations of a one-point set over an algebraically closed field are isomorphic if and only if the matrices of the corresponding linear operators have the same Jordan canonical form. Thus, in this case, the classes of isomorphic representations are parameterized by Jordan matrices.

Remark 11.4. The problem of describing linear representations of a two-point set (to say nothing of bigger sets) is considered "wild." This means that from the current viewpoint, it cannot be solved in any reasonable way. However, the only interesting linear representations are those of sets with operations (first of all, groups) and in this case, the difficulty in describing the representations depends on other reasons rather than on the number of elements.

Every linear representation $R\colon X \to \mathrm{L}(V)$ over a field K can be viewed as a linear representation over an extension L of K, once the representation operators are extended to linear operators on $V(L)$ (see Section 8.1). In the basis of $V(L)$ composed of vectors of V, such an extension of the representation is given by the same matrices as the original representation.

Proposition 11.5. *Let $R\colon X \to \mathrm{L}(V)$ and $S\colon X \to \mathrm{L}(W)$ be linear representations of a set X over an infinite field K and let L be an extension of K. Then, if R and S are isomorphic over L, they are already isomorphic over K.*

Proof. Write down the matrices of representations R and S in some bases of V and W. That R and S are isomorphic over K means that there exists a nonsingular matrix C with entries in K such that

$$CR(x) = S(x)C \qquad \forall x \in X. \tag{11.2}$$

Relations (11.2) represent a system of homogeneous linear equations in entries of C with coefficients from K. Let $\{C_1, \dots, C_m\}$ be a fundamental system of its solutions. If the representations R and S are not isomorphic over K, then $\det(\lambda_1 C_1 + \dots + \lambda_m C_m) = 0$ for every $\lambda_1, \dots, \lambda_m \in K$, hence $\det(t_1 C_1 + \dots + t_m C_m)$ is the zero polynomial in $t_1, \dots, t_m$. But then $\det(\lambda_1 C_1 + \dots + \lambda_m C_m) = 0$ for any $\lambda_1, \dots, \lambda_m \in L$, and this means that R and S are not isomorphic over L. □

In the understanding of the structure of linear representations, an important role is played by invariant subspaces.

Consider a representation $R\colon X \to \mathrm{L}(V)$. A subspace $U \subset V$ is called *invariant* with respect to R if it is invariant under all representation operators $R(x), x \in X$. It is obvious that a sum or an intersection of invariant subspaces is also an invariant subspace.

An invariant subspace $U \subset V$ gives rise to two new representations of X: the *subrepresentation*

$$R_U\colon X \to \mathrm{L}(U), \qquad R_U(x) = R(x)|_U,$$

and the *quotient representation*

$$R_{V/U}\colon X \to \mathrm{L}(V/U), \qquad R_{V/U}(x)(v + U) = R(x)v + U.$$

The representation R_U (respectively, $R_{V/U}$) is uniquely determined by the property that the embedding $U \to V$ (respectively, the canonical map $V \to V/U$) is a morphism of representations.

In the matrix form, this looks as follows: if a basis of the space V is chosen so that its first vectors form a basis of U, then

$$R(x) = \begin{pmatrix} R_U(x) & * \\ 0 & R_{V/U}(x) \end{pmatrix}.$$

Definition 11.6. A linear representation (11.1) is *irreducible* if $V \neq 0$ and there exist no nontrivial subspaces $U \subset V$ that are invariant under R.

Obviously, every one-dimensional representation is irreducible.

Example 11.7. Consider the representation Π of the additive group $\mathbb{R}$ by rotations of the Euclidean space E^2 which is defined in the orthonormal basis $\{e_1, e_2\}$ by the formula

$$\Pi(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

It is irreducible because no one-dimensional subspace is mapped into itself by all rotations. However, if one considers this representation over the complex numbers, it becomes reducible. More precisely, the one-dimensional subspaces spanned by the vectors $e_1 - ie_2$ and $e_1 + ie_2$, respectively, are invariant, and in the basis formed by these vectors, this representation has the form

$$\Pi(t) = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$$

(see Example 6.19).

Example 11.8. The isomorphism $S_4 \xrightarrow{\sim} \mathrm{Sym}_+ K$ (see Example 4.116) defines a linear representation of the group S_4 in the space E^3. Let us prove that it is irreducible. Since the orthogonal complement of an invariant subspace is invariant as well (see Proposition 6.34), it suffices to prove that there exist no one-dimensional invariant subspaces, and this is obvious. In fact, this representation is irreducible not just over $\mathbb{R}$ but over $\mathbb{C}$ as well. This is a consequence of the following general proposition.

Proposition 11.9. *Let $R\colon X \to \mathrm{L}(V)$ be an irreducible real linear representation of odd dimension. Then the complexification of R is also irreducible.*

Proof. Assume that $W \subset V(\mathbb{C})$ is a nontrivial invariant subspace. Observe that if a subspace of $V(\mathbb{C})$ is invariant under complex conjugation, then with every vector it contains its real and imaginary parts as well, hence it is a

complexification of a subspace of the space V. It follows that the invariant subspaces $W \cap \overline{W}$ and $W + \overline{W}$ are complexifications of some invariant subspaces of V. By irreducibility of R, we must have

$$W \cap \overline{W} = 0, \quad W + \overline{W} = V(\mathbb{C}),$$

i.e., $V(\mathbb{C}) = W \oplus \overline{W}$. But then $\dim V(\mathbb{C}) = 2 \dim W$, which contradicts the condition that the dimension of V is odd. □

Example 11.10. Let V be an n-dimensional vector space with a basis $\{e_1, \ldots, e_n\}$. The linear representation M of the group S_n defined by the formulas

$$M(\sigma)e_i = e_{\sigma(i)}, \qquad \sigma \in S_n,$$

is called the *monomial representation*. This representation is reducible; one can present at least two nontrivial invariant subspaces: the one-dimensional subspace $\langle e_1 + \cdots + e_n \rangle$ and the $(n-1)$-dimensional subspace

$$V_0 = \left\{ \sum_i x_i e_i : \sum_i x_i = 0 \right\}.$$

Let us prove that whenever $\operatorname{char} K = 0$, the representation $M_0 = M_{V_0}$ is irreducible. Indeed, let $U \subset V_0$ be an invariant subspace with a nonzero vector $x = \sum_i x_i e_i \in U$. Since $\sum_i x_i = 0$, not all numbers $x_1, \ldots, x_n$ are equal. Without loss of generality, assume that $x_1 \neq x_2$. Then

$$x - M((12))x = (x_1 - x_2)(e_1 - e_2) \in U,$$

hence $e_1 - e_2 \in U$. Applying representation operators to $e_1 - e_2$, we obtain that $e_i - e_j \in U$ for all i, j, but then $U = V_0$.

Example 11.11. Let A be an associative algebra. Then the formula

$$T(a)x = ax, \qquad a, x \in A,$$

determines a linear representation T of the algebra A on itself called the (*left*) *regular representation*. We emphasize that this is an algebra representation, i.e., that

$$T(a+b) = T(a) + T(b), \qquad T(ab) = T(a)T(b), \qquad T(\lambda a) = \lambda T(a).$$

For instance, the second of these properties is equivalent to associativity of multiplication in A. Invariant subspaces of this representations are nothing but left ideals of A.

If $\varphi\colon V \to U$ is a morphism of the representation $R\colon X \to \mathrm{L}(V)$ to the representation $S\colon X \to \mathrm{L}(U)$, then $\operatorname{Im}\varphi$ is an invariant subspace of U and $\operatorname{Ker}\varphi$ is an invariant subspace of V. Hence, the following theorem holds:

Theorem 11.12. *Every morphism of an irreducible representation is either an isomorphism or the zero map.*

Unless specified otherwise, in the sequel we will always assume that only finite-dimensional linear representations are considered.

Theorem 11.13 (Schur's Lemma). *Every endomorphism (i.e., a morphism to itself) of an irreducible representation over an algebraically closed field is a multiple of the identity operator.*

Proof. Let $R\colon X \to \mathrm{L}(V)$ be the given representation. A linear operator $\varphi \in \mathrm{L}(V)$ is an endomorphism of R if it commutes with all representation operators. Therefore, if φ is an endomorphism of R, then so is $\varphi - \lambda\mathcal{E}$ for every $\lambda \in K$. Choose as λ an eigenvalue of φ. By Theorem 11.12, we obtain that $\varphi - \lambda\mathcal{E} = 0$. □

Corollary 11.14. *Let $R\colon X \to \mathrm{L}(V)$ and $S\colon X \to \mathrm{L}(U)$ be two irreducible representations of a set X over an algebraically closed field. Then every two morphisms of R to S are proportional.*

Proof. If one of the morphisms is zero, there is nothing to prove. Thus, we have only to prove that the two isomorphisms are proportional. But if $\varphi\colon V \to U$ and $\psi\colon V \to U$ are two isomorphisms of R to S, then $\psi^{-1}\varphi\colon V \to V$ is an automorphism of R. By Schur's lemma, $\psi^{-1}\varphi = \lambda\mathcal{E}$, thus $\varphi = \lambda\psi$. □

Corollary 11.15. *Every irreducible representation of an abelian group over an algebraically closed field is one-dimensional.*

Proof. In the case of an abelian group, all representation operators commute with one another, hence, each of them is an endomorphism of the representation. By Schur's lemma, they are all scalar. Therefore, every subspace is invariant and a representation is irreducible only if it is one-dimensional. □

Definition 11.16. A linear representation $R\colon X \to \mathrm{L}(V)$ is *completely reducible* if every invariant subspace $U \subset V$ has a complementary invariant subspace, i.e., an invariant subspace W such that $V = U \oplus W$.

Observe that, as strange as it sounds, every irreducible representation is completely reducible.

Example 11.17. The formula

$$R(t) = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}, \qquad t \in \mathbb{R},$$

defines a two-dimensional real linear representation of the additive group $\mathbb{R}$. Its nontrivial invariant subspaces are the one-dimensional subspaces spanned by the basis vectors (i.e., coordinate axes). Since these subspaces are complements of each other, the representation R is completely reducible.

Example 11.18. The formula

$$S(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

defines another linear representation of this group. In this case, the only nontrivial invariant subspace is the one-dimensional subspace spanned by the first basis vector. Thus, the representation S is not completely reducible.

Proposition 11.19. *Every subrepresentation and every quotient representation of a completely reducible representation is completely reducible.*

Proof. Let $R\colon X \to \mathrm{L}(V)$ be a completely reducible representation and $U \subset V$, an invariant subspace. For every invariant subspace $U_1 \subset U$, there exists an invariant complement in V. Denote it V_2. The subspace $U_2 = U \cap V_2$ is then an invariant subspace which is complementary to U_1 in V.

Now let $\pi\colon V \to V/U$ be the canonical map and $W_1 \subset V/U$, an invariant subspace. Then $V_1 = \pi^{-1}(W_1)$ is an invariant subspace of V (that contains U). If $V_2 \subset V$ is its complementary invariant subspace, $W_2 = \pi(V_2)$ is the invariant subspace that is complementary to W_1 in V/U. □

Below we present a different characterization of completely reducible representations.

Theorem 11.20. *(i) If a representation $R\colon X \to \mathrm{L}(V)$ is completely reducible, then V decomposes into a direct sum of minimal invariant subspaces.*

(ii) Conversely, if the space V decomposes into a sum (not necessarily direct) of minimal invariant subspaces $V_1, \dots, V_m$, then the representation R is completely reducible. Moreover, for any invariant subspace $U \subset V$, we can take a sum of certain V_i's as its invariant complement.

(Here a minimal invariant subspace is a subspace that is minimal among nonzero invariant subspaces.)

Proof. (i) Consider a minimal invariant subspace V_1, find a complementary invariant subspace of it, consider a minimal invariant subspace V_2 of it, etc.

(ii) Let $U \subset V$ be an invariant subspace. For any subset $I \subset \{1, \dots, n\}$, put $V_I = \sum_{i \in I} V_i$. Let I be a maximal subset (possibly empty) such that $U \cap V_I = 0$. Then for any $j \notin I$, we should have $U \cap V_{I \cup \{j\}} \neq 0$, implying

$$(U \oplus V_I) \cap V_j \neq 0.$$

Since V_j is a minimal invariant subspace, $V_j \subset U \oplus V_I$. Therefore,

$$V = U \oplus V_I.$$

□

Example 11.21. If $\operatorname{char} K = 0$, the monomial representation of the group S_n defined in Example 11.10 is completely reducible, since in this case, the space V decomposes into a direct sum of minimal invariant subspaces:

$$V = \langle e_1 + \cdots + e_n \rangle \oplus V_0.$$

Exercise 11.22. Let Ad be the linear representation of the group $\mathrm{GL}_n(K)$ in the space $\mathrm{L}_n(K)$ defined as

$$\mathrm{Ad}(A)X = AXA^{-1}.$$

Prove that if $\operatorname{char} K = 0$, then $\langle E \rangle$ and $\langle X \in \mathrm{L}_n(K) \colon \operatorname{tr} X = 0 \rangle$ are minimal invariant subspaces. Conclude that the representation Ad is completely reducible.

Definition 11.23. The *sum* of linear representations $R_i \colon X \to \mathrm{L}(V_i)$, $i = 1, \ldots, m$, is the linear representation

$$R = R_1 + \cdots + R_m \colon X \to \mathrm{L}(V_1 \oplus \cdots \oplus V_m),$$

defined by the formula

$$R(x)(v_1, \ldots, v_m) = (R_1(x)v_1, \ldots, R_m(x)v_m).$$

In matrix form,

$$R(x) = \begin{pmatrix} R_1(x) & 0 & \ldots & 0 \\ 0 & R_2(x) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & R_m(x) \end{pmatrix}.$$

If $R \colon X \to \mathrm{L}(V)$ is a linear representation and the space V decomposes into a direct sum of invariant subspaces $V_1, \ldots, V_m$, then $R \simeq R_1 + \cdots + R_m$, where $R_i = R_{V_i}$.

Theorem 11.20 implies the following description of completely reducible representations:

Corollary 11.24. *A linear representation is completely reducible if and only if it is isomorphic to a sum of irreducible representations.*

Corollary 11.25. *Let $R \colon X \to \mathrm{L}(V)$ be a completely reducible representation which is isomorphic to a sum of irreducible representations $R_1, \ldots, R_m$. Then every subrepresentation and every quotient representation of R are isomorphic to a sum of some of the representations R_i.*

Proof. Let

$$V = V_1 \oplus \cdots \oplus V_m$$

be a decomposition into a direct sum of invariant subspaces such that $R_{V_i} \simeq R_i$ and let $U \subset V$ be an invariant subspace. By Theorem 11.20, there exists a subset $I \subset \{1, \ldots, m\}$ such that $V = U \oplus V_I$. Clearly,

$$R_{V/U} \simeq R_{V_I} \simeq \sum_{i \in I} R_i.$$

Now, let $J = \{1, \ldots, m\} \setminus I$. Then $V = V_J \oplus V_I$, hence

$$R_U \simeq R_{V/V_I} \simeq R_{V_J} \simeq \sum_{j \in J} R_j.$$

□

Example 11.26. Every irreducible linear representation of a one-point set X over an algebraically closed field K is one-dimensional. Thus, every completely reducible representation of X over K is a linear operator whose matrix is diagonal in some basis.

Let $R\colon X \to \mathrm{L}(V)$ be a completely reducible representation and let

$$V = V_1 \oplus \cdots \oplus V_m \tag{11.3}$$

be a decomposition of V into a direct sum of minimal invariant subspaces.

Definition 11.27. The *isotypic component* of a representation R corresponding to an irreducible representation S of X is the sum $V_{(S)}$ of all summands V_i in the decomposition (11.3) such that $R_{V_i} \simeq S$, as well as the restriction $R_{(S)}$ of R to this sum.

It is clear from this definition that the space V decomposes into a direct sum of the isotypic components corresponding to different irreducible representations of X.

Example 11.28. For a completely reducible representation of a one-point set over an algebraically closed field, its isotypic components are the eigenspaces of the corresponding linear operator.

A representation R is called *isotypic* or, rather, *S-isotypic* if $R = R_{(S)}$.

Isotypic representations can be described as follows. Let $S\colon X \to \mathrm{L}(U)$ be an irreducible representation and Z a vector space. Define a representation

$$R\colon X \to \mathrm{L}(U \otimes Z) \tag{11.4}$$

by the formula

$$R(x)(u \otimes z) = (S(x)u) \otimes z.$$

If $\{z_1, \ldots, z_m\}$ is a basis of Z, the decomposition

$$U \otimes Z = (U \otimes z_1) \oplus \cdots \oplus (U \otimes z_m) \tag{11.5}$$

is a decomposition of the space $U \otimes Z$ into a direct sum of invariant subspaces such that the restriction of R to each of them is isomorphic to S.

Theorem 11.29. *If the base field K is algebraically closed, then every invariant subspace of $U \otimes Z$ has the form $U \otimes Z_0$, where Z_0 is a subspace of Z.*

Proof. Since any sum of subspaces of the form $U \otimes Z_0$ is a subspace of the same kind, it suffices to prove the theorem for minimal invariant subspaces.

Let $W \subset U \otimes Z$ be a minimal invariant subspace. By the decomposition (11.5), for any $w \in W$, we have

$$w = \varphi_1(w) \otimes z_1 + \cdots + \varphi_m(w) \otimes z_m,$$

where $\varphi_1, \ldots, \varphi_m$ are morphisms of the representation R_W to S. By Corollary 11.15, $\varphi_i = \lambda_i \varphi$, where $\lambda_i \in K$ and φ is a fixed isomorphism of R_W to S. Thus,

$$w = \varphi(w) \otimes (\lambda_1 z_1 + \cdots + \lambda_m z_m),$$

so that

$$W = U \otimes (\lambda_1 z_1 + \cdots + \lambda_m z_m\rangle.$$

□

Exercise 11.30. Prove that if the field K is algebraically closed, every endomorphism of the representation (11.4) is of the form

$$u \otimes z \mapsto u \otimes Cz,$$

where C is a linear operator on Z.

Theorem 11.31 (Burnside's Theorem). *Let $R\colon X \to \mathrm{L}(V)$ be an irreducible representation of a set X over an algebraically closed field. Then the subalgebra of the algebra $\mathrm{L}(V)$ generated by the set $R(X)$ coincides with $\mathrm{L}(V)$, except when $\dim V = 1$ and $R(X) = 0$.*

Notice that the subalgebra generated by $R(X)$ consists of all linear combinations of products of operators $R(x)$, $x \in X$. Thus, this theorem claims that except for the trivial case, every linear operator is a linear combination of products of operators $R(x)$, $x \in X$.

Proof. As we know, the space $\mathrm{L}(V)$ can be identified with $V \otimes V^*$ so that each decomposable element $u \otimes \alpha \in V \otimes V^*$ correspond to the operator

$$u \otimes \alpha\colon v \mapsto \alpha(v)u.$$

Under this identification, the products of an operator $u \otimes \alpha$ with any linear operator $\mathcal{A} \in \mathrm{L}(V)$ look like

$$\mathcal{A}(u \otimes \alpha) = \mathcal{A}u \otimes \alpha, \tag{11.6}$$

$$(u \otimes \alpha)\mathcal{A} = u \otimes \mathcal{A}^*\alpha, \tag{11.7}$$

where $\mathcal{A}^* \in \mathrm{L}(V^*)$ is the adjoint operator determined by the formula

$$(\mathcal{A}^*\alpha)(v) = \alpha(\mathcal{A}v).$$

Observe that, since the canonical one-to-one correspondence between the subspaces of V and the subspaces of V^* that associates to each subspace its annihilator, the representation

$$R^*\colon X \to \mathrm{L}(V^*)$$

defined as

$$R^*(x) = R(x)^*$$

is irreducible.

Let us define representations T_l and T_r of X in the space $\mathrm{L}(V)$ by the formulas

$$T_l(x)\mathcal{A} = R(x)\mathcal{A}, \qquad T_r(x)\mathcal{A} = \mathcal{A}R(x).$$

Formulas (11.6) and (11.7) imply that these representations are isotypic.

Denote the subalgebra of $\mathrm{L}(V)$ generated by the set $R(X)$ by A. Clearly, it is a subspace of $\mathrm{L}(V)$ that is invariant under the representation T_l as well as with respect to the representation T_r. By Theorem 11.29, it can be presented in the form $A = V \otimes W_0$ for a subspace W_0 of V^* and, at the same time, in the form $A = V_0 \otimes V^*$ for a subspace V_0 of V. This is possible only if A is $\mathrm{L}(V)$ or 0. In the latter case, $\dim V = 1$; otherwise R would be reducible. □

Exercise 11.32. The *tensor product* of group representations $R\colon G \to \mathrm{GL}(V)$ and $S\colon H \to \mathrm{GL}(W)$ is the representation

$$R \otimes S\colon G \times H \to \mathrm{GL}(V \otimes W)$$

defined as

$$(R \otimes S)(g, h) = R(g) \otimes S(h).$$

(See the definition of the tensor product of linear operators in Section 8.1.) Prove that the tensor product of irreducible representations of groups G and H is irreducible over an algebraically closed field.

Consider now the class of completely reducible linear representations that, in some sense, are the opposite of isotypic representations. Namely, we say that a completely reducible representation *has a simple spectrum* if it is a sum of pairwise nonisomorphic irreducible representations or, in other words, if all its (nonzero) isotypic components are irreducible.

Example 11.33. A completely reducible representation of a one-point set over an algebraically closed field has a simple spectrum if and only if all roots of the characteristic polynomial of the corresponding linear operator are simple.

For representations with a simple spectrum, the invariant subspaces and endomorphisms are especially easy to describe.

Proposition 11.34. *Let $R\colon X \to \mathrm{L}(V)$ be a completely reducible representation with a simple spectrum. Consider a decomposition* (11.3) *of V into a direct sum of minimal invariant subspaces. Then*

(i) every invariant subspace of V is a sum of terms in the decomposition (11.3);

(ii) if the base field K is algebraically closed, then every endomorphism φ of R has the following form:

$$\varphi(x) = \lambda_i x \quad \text{for} \quad x \in V_i, \quad \lambda_1, \dots, \lambda_m \in K. \tag{11.8}$$

Proof. (i) Every invariant subspace is a sum of minimal invariant subspaces. By the definition of a representation with a simple spectrum, every minimal invariant subspace is an isotypic component, hence it coincides with a summand of the decomposition (11.3).

(ii) Every summand of the decomposition (11.3) is invariant under φ and, by Schur's lemma, the action of φ on this summand is scalar. □

Corollary 11.35. *For a completely reducible representation with a simple spectrum, the decomposition of the representation space into a direct sum of minimal invariant subspaces is unique.*

A linear representation $R\colon G \to \mathrm{GL}(V)$ of a group G (over a field K of characteristic $\neq 2$) is called *orthogonal* (respectively, *symplectic*) if there exists a nondegenerate symmetric (respectively, skew-symmetric) bilinear function on V that is invariant under every representation operator.

Exercise 11.36. Prove that if $R\colon G \to \mathrm{GL}(V)$ is an irreducible representation of G over an algebraically closed field, then

(a) every nonzero invariant bilinear function on V is nondegenerate;

(b) any two such functions are proportional;

(c) every such function is either symmetric or skew-symmetric.

11.2. Complete Reducibility of Linear Representations of Finite and Compact Groups

For some classes of groups, it is possible to prove the complete reducibility of all their linear representations.

We begin with the finite groups. For them, the proof is purely algebraic and is based on the simple idea of the lemma below.

Let S be a finite-dimensional affine space over a field K.

Lemma 11.37 (Fixed Point Lemma). *Let G be a finite group of affine transformations of S whose order is not divisible by* $\operatorname{char} K$. *Then G has a fixed point in S.*

Proof. The center of mass of the orbit of any point $p \in S$ is one such fixed point:

$$\operatorname{cent} Gp = \frac{1}{|G|} \sum_{g \in G} gp.$$

□

Now let V be a finite-dimensional vector space over a field K, and $G \subset \mathrm{GL}(V)$ a group of linear transformations.

Theorem 11.38. *Let G be a finite group whose order is not divisible by* $\operatorname{char} K$. *Then for every G-invariant subspace $U \subset V$, there exists a G-invariant complementary subspace W.*

Proof. To specify a subspace W that is a complement of U is the same as to specify the projection $\mathcal{P}$ onto U along W. The subspace W is invariant if and only if $\mathcal{P}$ commutes with all transformations in G.

The set of all projections onto U is described by the following linear equations:

$$\mathcal{P}v \in U \quad \forall v \in V, \qquad \mathcal{P}u = u \quad \forall u,$$

hence, it is a plane in the space $\mathrm{L}(V)$ of all linear operators on V. Denote this plane by S.

The group G acts on $\mathrm{L}(V)$ by conjugations. This action leaves S invariant, thus inducing affine transformations on S. Hence, we obtain a finite group of affine transformations of the plane S. Take its fixed point as the required projection $\mathcal{P}$. □

Corollary 11.39. *Every linear representation of a finite group G over a field K whose characteristic does not divide $|G|$, is completely reducible.*

Proof. The asserton follows by applying the theorem to the image of the group G under the given linear representation. □

Example 11.40. In Example 11.8, we constructed a three-dimensional representation of the group S_4. We will prove here in another way that it is irreducible not only over $\mathbb{R}$ but also over $\mathbb{C}$. Since it is completely reducible in either case, it suffices to show that it has no one-dimensional invariant subspaces, i.e., that the representation operators have no common eigenvectors. We know (see Proposition 6.18) that to any eigenvector of a real linear operator corresponding to an imaginary eigenvalue, there corresponds

a two-dimensional invariant real subspace. However, this representation has no two-dimensional invariant real subspaces.

For $K = \mathbb{R}$ or $\mathbb{C}$, a reasonable generalization of finite groups is compact topological groups.

Definition 11.41. A *topological group* is a group G with a Hausdorff topology such that the group operations

$$\mu\colon G \times G \to G, \qquad (x, y) \mapsto xy,$$
$$\iota\colon G \to G, \qquad x \mapsto x^{-1}$$

are continuous maps. A *topological group homomorphism* is a group homomorphism which is also a continuous map.

Examples of topological groups are the additive and the multiplicative group of the fields $\mathbb{R}$ and $\mathbb{C}$ and also the groups of nonsingular matrices over these fields. Every group (e.g., a finite one) can be regarded as a topological group with discrete topology.

A subgroup of a topological group with the induced topology is a topological group. A direct product of topological groups is also a topological group.

A topological group is called *compact* if it is also a compact topological space. In particular, all finite groups are compact. Examples of infinite compact topological groups are the "circle"

$$\mathbb{T} = \{z \in \mathbb{C}^* : |z| = 1\},$$

the orthogonal group O_n, and the unitary group U_n. Let us prove that O_n is compact. This group is determined by the equations

$$\sum_k x_{ik}x_{jk} = \delta_{ij}$$

in the n^2-dimensional space $\mathrm{L}_n(\mathbb{R})$ of all real matrices $X = (x_{ij})$ of order n, hence, it is closed in $\mathrm{L}_n(\mathbb{R})$. These equations also imply that $|x_{ij}| \leq 1$; therefore, O_n is also bounded in $\mathrm{L}_n(\mathbb{R})$. Thus, it is compact. Compactness of U_n is proved in a similar way.

Every closed subgroup of a compact group is compact. A direct product of compact groups is compact. For instance, the direct product of n copies of the circle $\mathbb{T}$ is a compact group called the *n-dimensional torus* and denoted $\mathbb{T}^n$. The image of a compact group under a (continuous) homomorphism (in particular, a linear representation) is a compact group.

There are analogues of the fixed point lemma and Theorem 11.38 for compact groups. Their proofs use the notion of the *center of mass* of a convex set.

Let M be a nonempty bounded convex set in a real affine space S. If $\operatorname{aff} M = S$, define the center of mass $\operatorname{cent} M$ of the set M by the formula

$$\operatorname{cent} M = \mu(M)^{-1} \int_M x\mu(dx),$$

where μ is the standard measure on S invariant under parallel translations. The measure μ is defined up to a constant multiple, but the formula implies that the freedom in the choice of μ does not affect the result. The integral on the right-hand side can be defined coordinatewise or directly, as the limit of integral sums which are (up to the factor before the integral) barycentric linear combinations of points of S and, hence, are well defined. The first definition shows that the integral exists, and the second, that it does not depend on the choice of coordinates. In general, we define $\operatorname{cent} M$ as above but replace S with the space $\operatorname{aff} M$.

Since the definition of the center of mass is stated in terms of affine geometry, for any affine transformation α,

$$\operatorname{cent} \alpha(M) = \alpha(\operatorname{cent} M).$$

In particular, if the set M is invariant under an affine transformation, its center of mass is a fixed point of this transformation.

The definition of the center of mass implies that $\operatorname{cent} M \in \overline{M}$. Actually,

$$\operatorname{cent} M \in M^\circ,$$

where M° is the interior of the set M with respect to the space $\operatorname{aff} M$. Indeed, for any affine-linear function f which is nonnegative on M and nonzero in $\operatorname{aff} M$, we have

$$f(\operatorname{cent} M) = \mu(M)^{-1} \int_M f(x)\mu(dx) > 0.$$

Lemma 11.42 (Fixed Point Lemma). *Let G be a compact group of affine transformations of a real affine space S. Let $M \subset S$ be a nonempty convex set which is invariant under G. Then G has a fixed point in M.*

Note that we can regard the whole space S as M.

Proof. The center of mass of the convex hull of the orbit of any point $p \in M$ is a required fixed point. □

Theorem 11.43. *Let G be a compact group of linear transformations of a vector space V over the field $K = \mathbb{R}$ or $\mathbb{C}$. Then for any G-invariant subspace $U \subset V$, there exists a G-invariant complementary subspace W.*

Proof. This proof repeats that of Theorem 11.38 verbatim. One should only notice that when $K = \mathbb{C}$, the plane S of projections onto U must be regarded as a real affine space. □

Corollary 11.44. *Every real or complex linear representation of a compact topological group is completely reducible.*

Example 11.45. By Theorem 11.43 and Corollary 11.14, every (continuous) complex linear representation of a compact abelian group is a sum of one-dimensional representations, i.e., its matrices are diagonal in some basis. In particular, this can be applied to finite abelian groups and the group $\mathbb{T}$.

There is another way to prove the complete reducibility of linear representations of compact groups. It is also of interest.

Theorem 11.46. *Let G be a compact group of linear transformations of a real (respectively, complex) vector space V. Then there exists a G-invariant positive definite quadratic (respectively, Hermitian) function on V.*

Proof. The set of all positive definite quadratic (respectively, Hermitian) functions is a G-invariant convex set in the space of all quadratic (respectively, Hermitian) functions. A fixed point of G in this set is a required function. □

Corollary 11.47. *Every compact (and, in particular, finite) subgroup of the group $\mathrm{GL}_n(\mathbb{R})$ (respectively, $\mathrm{GL}_n(\mathbb{C})$) is conjugate to a subgroup of O_n (respectively, U_n).*

Theorem 11.46 allows us to give another proof of Theorem 11.43: for an invariant complement of U, one can take the orthogonal complement with respect to the inner product defined by the invariant quadratic (respectively, Hermitian) function.

11.3. Finite-Dimensional Associative Algebras

The linear representation approach leads, first of all, to a rather good description of the structure of finite-dimensional associative algebras.

Let A be a finite-dimensional associative (but not necessarily commutative) algebra over a field K.

An element $a \in A$ is called *nilpotent* if $a^n = 0$ for some $n \in \mathbb{N}$. The algebra A is called *nilpotent* if all its elements are nilpotent. Every subalgebra and every quotient algebra of a nilpotent algebra are nilpotent. On the other hand, if an ideal I and the quotient algebra A/I are nilpotent, then A is nilpotent.

Example 11.48. The algebra of niltriangular matrices (triangular with a zero diagonal) of order n is nilpotent. Moreover, the product of any n elements of this algebra is zero. As we will see now, every nilpotent algebra has this property.

Theorem 11.49. *For every nilpotent algebra A, there exists $n \in \mathbb{N}$ such that the product of any n elements of A is zero.*

For any subspaces $B, C \subset A$, we denote by BC the linear span of all products of the form bc, $b \in B$, $c \in C$. In this notation, the statement of this theorem can be written as $A^n = 0$ for some $n \in \mathbb{N}$.

Proof. Let $B \subset A$ be a maximal subspace for which there exists $n \in \mathbb{N}$ such that $B^n = 0$. Assume that $B \neq A$ and let $a \in A \setminus B$. Since $aB^n = 0$, there exists $k \geq 0$ such that $aB^k \not\subset B$ but $aB^{k+1} \subset B$. Replacing a with a suitable element from aB^k, we obtain

$$aB \subset B. \tag{11.9}$$

For some $m \in \mathbb{N}$, we have

$$a^m = 0. \tag{11.10}$$

Let $C = B \oplus \langle a \rangle$. Conditions (11.9) and (11.10) imply that $C^{nm} = 0$, and this contradicts the definition of B. □

Remark 11.50. Let A be a nilpotent (finite-dimensional) algebra. Then there exists n such that $a^n = 0$ for all $a \in A$ (show this). An infinite-dimensional algebra A is called nilpotent if there exists $n \in \mathbb{N}$ such that the product of any n elements of A is zero. There are infinite-dimensional algebras which are not nilpotent, while all their elements are. Such algebras are called *nilalgebras*.

Unlike the commutative case, all nilpotent elements of an associative algebra A do not have to form an ideal (or even a subspace). However, if I and J are two nilpotent ideals, their sum

$$I + J := \{x + y \colon x \in I,\, y \in J\}$$

is also a nilpotent ideal since it contains the nilpotent ideal I, and the quotient algebra

$$(I + J)/I \simeq J/(I \cap J)$$

by I is nilpotent too. Thus, there exists the largest nilpotent ideal. It is called the *radical* of A and is denoted $\operatorname{rad} A$.

In the commutative case, it coincides with the radical of A in the sense of Section 9.4.

An algebra A is called *semisimple* if $\operatorname{rad} A = 0$.

Example 11.51. By Example 9.97, the algebra $K[t]/(h)$ is semisimple if and only if the polynomial h has no multiple irreducible factors.

When $\operatorname{char} K = 0$, semisimple algebras can be characterized in another way.

Let $T\colon A \to \mathrm{L}(A)$ be a regular representation of A (see Example 11.11).

Define the "inner product" on A by the formula

$$(a, b) = \operatorname{tr} T(ab) = \operatorname{tr} T(a)T(b). \tag{11.11}$$

This is a symmetric bilinear function (possibly degenerate). Moreover, it has the following property:

$$(ab, c) = (a, bc),$$

which follows from the associativity of multiplication in A. (When A is a field, this inner product coincides with the one introduced in Section 9.5.)

Proposition 11.52. *The orthogonal complement $I^{\perp}$ of an ideal I of an algebra A is also an ideal.*

Proof. Let $x \in I^{\perp}$, $a \in A$, $y \in I$. Then

$$(xa, y) = (x, ay) = 0, \qquad (ax, y) = (y, ax) = (ya, x) = 0.$$

□

Proposition 11.53. *If $\operatorname{char} K = 0$, every element $a \in A$ that is orthogonal to all of its own powers is nilpotent.*

Proof. Let

$$(a^n, a) = \operatorname{tr} T(a)^{n+1} = 0 \qquad \forall n \in \mathbb{N}.$$

Consider an extension of K where the characteristic polynomial f of the operator $T(a)$ splits into linear factors:

$$f = t^{k_0} \prod_{i=1}^{s} (t - \lambda_i)^{k_i}, \qquad \lambda_1, \ldots, \lambda_s \text{ are different and nonzero.}$$

Then

$$\operatorname{tr} T(a)^{n+1} = \sum_{i=1}^{s} k_i \lambda_i^{n+1} = 0.$$

Let n assume values from 1 to s. Consider the above equalities as a square system of homogeneous linear equations with respect to $k_1, \ldots, k_s$. Its determinant differs from the Vandermonde determinant only by the factor $\lambda_1 \cdots \lambda_s$, hence it is nonzero. Therefore, $k_1 = \cdots = k_s = 0$ in K, which is impossible if $\operatorname{char} K = 0$. Thus, $s = 0$ and this means that the operator

$T(a)$ is nilpotent, i.e., $T(a)^m = 0$ for some $m \in \mathbb{N}$. But then

$$a^{m+1} = T(a)^m a = 0.$$

□

Theorem 11.54. *(i) If the inner product* (11.11) *is nondegenerate, then the algebra A is semisimple.*

(ii) Conversely, if the algebra A is semisimple and $\operatorname{char} K = 0$, *the inner product* (11.11) *is nondegenerate.*

Proof. (i) Let I be a nilpotent ideal of A. Then for every $x \in I$ and $a \in A$, the element ax is nilpotent (it belongs to I). Thus

$$(a, x) = \operatorname{tr} T(ax) = 0.$$

Therefore, $I \subset A^\perp = 0$.

(ii) Conversely, if $\operatorname{char} K = 0$, by Propositions 11.52 and 11.53, $A^\perp$ is a nilpotent ideal. □

Remark 11.55. In fact, we proved a stronger result: if $\operatorname{char} K = 0$, then $\operatorname{rad} A$ is the kernel of the inner product (11.11).

Example 11.56. It follows from formula (11.6) that the regular representation of the algebra $\mathrm{L}(V)$ is isotypic. More precisely, it is isomorphic to nR, where $n = \dim V$ and R is the *tautological representation* of $\mathrm{L}(V)$ in the space V (i.e., the identity map $\mathrm{L}(V) \to \mathrm{L}(V)$). Thus, the inner product (11.11) on $\mathrm{L}(V)$ has the form

$$(\mathcal{A}, \mathcal{B}) = n \operatorname{tr} \mathcal{A}\mathcal{B}. \tag{11.12}$$

If $\mathcal{E}_{ij}$ is the operator defined (in a fixed basis) by the matrix unit E_{ij}, then

$$\operatorname{tr} \mathcal{E}_{ij}\mathcal{E}_{kl} = \begin{cases} 1, & \text{for } k = j, l = i, \\ 0, & \text{otherwise.} \end{cases}$$

It follows that if $\operatorname{char} K$ does not divide n, the inner product (11.12) is nondegenerate, hence, the algebra $\mathrm{L}(V)$ is semisimple. (We will see in Example 11.62 that it is semisimple also when $\operatorname{char} K$ divides n.)

Example 11.57. Let $A = K[t]/(h)$. We can view A as a $K[t]$-module. Then h is the characteristic polynomial of the operator of multiplication by t (see Exercise 9.85). Let $c_1, \ldots, c_n$ be its roots (with multiplicities) in the splitting field. Then for every polynomial $f \in K[t]$, the characteristic polynomial of the operator of multiplication by $f(t)$ has roots $f(c_1), \ldots, f(c_n)$. However, the operator of multiplication by $f(t)$ on the $K[t]$-module A is the same as

the operator of multiplication by $[f] = f + (h)$ on the algebra A, i.e., the operator $T([f])$. Thus,

$$\operatorname{tr} T([f]) = \sum_i f(c_i).$$

Therefore, in the basis $\{[1], [t], [t^2], \ldots, [t^{n-1}]\}$ of A, the matrix of the inner product (11.11) is

$$\begin{pmatrix} s_0 & s_1 & s_2 & \ldots & s_{n-1} \\ s_1 & s_2 & s_3 & \ldots & s_n \\ s_2 & s_3 & s_4 & \ldots & s_{n+1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ s_{n-1} & s_n & s_{n+1} & \ldots & s_{2n-2} \end{pmatrix}, \tag{11.13}$$

where $s_k = c_1^k + \cdots + c_n^k$. Observe that we can express the power sums s_k in terms of the coefficients of the polynomial h without determining its roots.

By Theorem 11.54, if $\operatorname{char} K = 0$, the algebra A is semisimple if and only if the matrix (11.13) is nonsingular. On the other hand (see Example 11.51), it is semisimple if and only if the polynomial h has no multiple irreducible factors, which, in the case of $\operatorname{char} K = 0$, is equivalent to $c_1, \ldots, c_n$ being distinct. Therefore, we conclude that the polynomial h over a field of zero characteristic is separable (i.e., has no multiple roots in any extension of K) if and only if the matrix (11.13) is nonsingular.

Remark 11.58. The latter statement is valid in any characteristic and can be proven directly. Namely, matrix (11.13) can be presented as the product

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ c_1 & c_2 & \ldots & c_n \\ c_1^2 & c_2^2 & \ldots & c_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ c_1^{n-1} & c_2^{n-1} & \ldots & c_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & c_1 & c_1^2 & \ldots & c_1^{n-1} \\ 1 & c_2 & c_2^2 & \ldots & c_2^{n-1} \\ 1 & c_3 & c_3^2 & \ldots & c_3^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & c_n & c_n^2 & \ldots & c_n^{n-1} \end{pmatrix}.$$

Therefore, its determinant equals

$$\prod_{i>j} (c_i - c_j)^2,$$

i.e., the discriminant of h (see Section 3.9); this implies the above statement.

In particular, we obtain that for an irreducible polynomial h over a field K of positive characteristic, the field $L = K[x]/(h)$ is a separable extension of K in the sense of Remark 9.130 if and only if h is separable.

Exercise 11.59. Prove that when $\operatorname{char} K = 0$, the number of distinct roots of h equals the rank of the matrix (11.13).

An algebra A is called *simple* if $A \neq 0$ and A has no nontrivial ideals (i.e., ideals different from 0 or A).

Example 11.60. Every extension L of a field K is a simple (commutative) algebra over K.

Exercise 11.61. Prove the converse: every simple commutative algebra over K is either a field containing K (i.e., an extension of K) or the one-dimensional algebra with the zero multiplication.

For every algebra A, the subspace A^2 is an ideal (as is any product of two ideals). If A is nilpotent, $A^2 \neq A$. Thus, a simple algebra A cannot be nilpotent except for the trivial case when $A^2 = 0$ and $\dim A = 1$, i.e., when A is the one-dimensional algebra with the zero multiplication. Except for this case, every simple algebra is semisimple.

Example 11.62. The algebra $\mathrm{L}(V)$ is simple (see Example 9.48), thus semisimple.

Theorem 11.63. *Every semisimple associative algebra A decomposes into a direct sum of (nontrivial) simple algebras:*

$$A = A_1 \oplus \cdots \oplus A_s, \tag{11.14}$$

and every ideal of A is a sum of some terms in this decomposition.

Proof. We will prove this theorem under the assumption that $\operatorname{char} K = 0$. If the algebra A is simple, there is nothing to prove (here $s = 1$). Let it be nonsimple and let $A_1 \subset A$ be a minimal ideal. Then either

$$A = A_1 \oplus A_1^{\perp} \tag{11.15}$$

or $A_1 \subset A_1^{\perp}$. In the latter case, the ideal A_1 is nilpotent by Proposition 11.53, hence this case is impossible. In the former case, the decomposition (11.15) implies that every ideal of the algebra A_1 and every ideal of the algebra $A_1^{\perp}$ are ideals of A. Thus, the algebra A_1 is simple and the algebra $A_1^{\perp}$ is semisimple. If the algebra $A_1^{\perp}$ is not simple, apply the same procedure to it, and so on.

Now let I be an ideal of A. Denote by π_k the projection onto the kth term in the decomposition (11.14). Obviously, $I_k = \pi_k(I)$ is an ideal of the algebra A_k. If $I_k \neq 0$, then $I_k = A_k$, hence

$$A_k = A_k^2 = A_k I_k = A_k I \subset I.$$

This implies that I is a sum of terms in the decomposition (11.14). □

In particular, every (finite-dimensional) semisimple commutative associative algebra A is a direct sum of several finite extensions of K (see Exercise 11.61). If K is algebraically closed, then A is simply a sum of several copies of K.

Exercise 11.64. Prove the latter statement (about the algebraically closed case) by methods of commutative algebra. (*Hint*: consider Spec A and use Hilbert's Nullstellensatz.)

Example 11.65. Consider a polynomial $h \in K[x]$ without multiple irreducible factors. Let $h = p_1 \cdots p_s$ be its decomposition into irreducible factors over K. By Theorem 9.107, the following algebras are isomorphic:

$$K[t]/(h) \simeq K[t]/(p_1) \oplus \cdots \oplus K[t]/(p_s). \tag{11.16}$$

This provides a decomposition of the semisimple algebra $K[t]/(h)$ into a direct sum of simple algebras (finite extensions of K). In particular, when $K = \mathbb{R}$, in the decomposition (11.16) every real root of h corresponds to a one-dimensional summand isomorphic to $\mathbb{R}$, and every pair of conjugate imaginary roots corresponds to a two-dimensional summand isomorphic to $\mathbb{C}$ (see Examples 9.45 and 9.46).

Exercise 11.66. Compute by two methods the inner product (11.11) in the algebra $\mathbb{R}[x]/(h)$ for a polynomial $h \in \mathbb{R}[x]$ without multiple complex roots. Prove that the number of pairs of conjugate imaginary roots of h equals the negative index of inertia of the symmetric matrix (11.13). In particular, all roots of the polynomial h are real if and only if the matrix (11.13) is positive definite.

As for simple algebras, the following theorem describes their structure in the case of an algebraically closed field. The general case is the subject of Section 11.6.

Theorem 11.67. *Every nontrivial simple associative algebra A over an algebraically closed field K is isomorphic to an algebra of the form $\mathrm{L}(V)$, where V is a vector space over K. Moreover, every nontrivial irreducible representation of A is isomorphic to the tautological representation of $\mathrm{L}(V)$.*

(By a trivial irreducible representation, we understand the one-dimensional representation R with $R(A) = 0$.)

Proof. Consider the restriction of the regular representation of A to a minimal invariant subspace V (left ideal) of A. We obtain an irreducible representation; denote it by R. Its kernel is an ideal of A. By Burnside's theorem, either

$$A \simeq R(A) = \mathrm{L}(V)$$

or $\dim V = 1$ and $R(A) = 0$. In the latter case, $AV = 0$, hence

$$A_0 := \{x \in A \colon Ax = 0\} \neq 0;$$

however, it is easy to see that A_0 is an ideal of A. Thus, $A_0 = A$, contradicting nontriviality of A.

Now let $A = \mathrm{L}(V)$ for a vector space V, and let R be the tautological representation of A in V. Then the regular representation T of A is isomorphic to nR, where $n = \dim V$. Let $S\colon A \to \mathrm{L}(U)$ be an irreducible representation of A. Choose a nonzero vector $u_0 \in U$ and consider the map

$$\varphi\colon A \to U, \qquad a \mapsto S(a)u_0.$$

Since

$$\varphi(T(a)x) = \varphi(ax) = S(ax)u_0 = S(a)S(x)u_0 = S(a)\varphi(x),$$

φ is a morphism of T into S. If S is nontrivial, $\operatorname{Im}\varphi = U$, so that the representation S is isomorphic to a quotient representation of T. Since S is irreducible, $S \simeq R$ (by Corollary 11.25). □

Theorems 11.63 and 11.67 imply that every semisimple associative algebra over an algebraically closed field is isomorphic to an algebra of the form

$$A = \mathrm{L}(V_1) \oplus \cdots \oplus \mathrm{L}(V_s), \tag{11.17}$$

where $V_1, \ldots, V_s$ are vector spaces. If $\dim V_i = n_i$, $i = 1, \ldots, s$,

$$\dim A = n_1^2 + \cdots + n_s^2. \tag{11.18}$$

We can determine the number s if we know the center of A. In general, the *center* of an associative algebra A is a (commutative) subalgebra

$$Z(A) = \{z \in A\colon az = za \ \forall a \in A\}.$$

It is known that the center of the algebra $\mathrm{L}(V)$ is one-dimensional and consists of scalar transformations (see Exercise 1.80). Thus, for a semisimple algebra A presented as the decomposition (11.17),

$$\dim Z(A) = s. \tag{11.19}$$

Let R_i, $i = 1, \ldots, s$, denote the irreducible representation of A in the space V_i defined by the projection onto $\mathrm{L}(V_i)$ in the decomposition (11.17). Notice that the representations $R_1, \ldots, R_s$ are pairwise nonisomorphic because they have different kernels.

Theorem 11.68. *Every nontrivial irreducible representation of algebra* (11.17) *is isomorphic to one of the representations* $R_1, \ldots, R_s$.

Proof. Let $S\colon A \to \mathrm{L}(U)$ be a nontrivial irreducible representation of A. Since $S(A) = \mathrm{L}(U)$ is a simple algebra, $\operatorname{Ker} S$ is the sum of all terms of the decomposition (11.17) except for one, say, $\mathrm{L}(V_i)$. But then S is actually defined by a representation of the algebra $\mathrm{L}(V_i)$, hence, it is isomorphic to R_i by Theorem 11.67. □

Also, another theorem is true: *every linear representation of a semisimple associative algebra (over any field) is completely reducible.*

11.4. Linear Representations of Finite Groups

The theory of associative algebras provides important information about linear representations of finite groups.

Let G be a finite group of order n and K a field.

Definition 11.69. The *group algebra* of the group G over K is the algebra KG whose basis elements are indexed by elements of the group G so that the product of basis elements with indices $g, h \in G$ is the basis element with the index gh.

Usually basis elements of the algebra KG are identified with the corresponding elements of the group G. With this identification, every element of the algebra KG is written as

$$a = \sum_{g \in G} a_g g, \qquad a_g \in K. \tag{11.20}$$

Associativity of multiplication in G implies associativity of multiplication in KG.

Every linear representation R of the group G in a vector space V over a field K uniquely extends to a linear representation of the algebra KG in the same space by the following formula:

$$R\left(\sum_{g \in G} a_g g\right) = \sum_{g \in G} a_g R(g).$$

Conversely, the restriction of every linear representation of the algebra KG to G is a linear representation of the group G. This establishes a one-to-one correspondence between representations of a group and its group algebra.

Obviously, corresponding representations of G and KG have the same collection of invariant subspaces. In particular, irreducible representations of the group correspond to irreducible representations of the group algebra and vice versa.

Theorem 11.70. *If* $\operatorname{char} K$ *does not divide* n, *the algebra* KG *is semisimple.*

Proof. We use Theorem 11.54. Calculate the inner product (11.11) on the algebra KG. It is easy to see that for any $g \in G$,

$$\operatorname{tr} T(g) = \begin{cases} n, & g = e, \\ 0, & g \neq e. \end{cases}$$

Thus, for every $g, h \in G$,

$$(g,h) = \begin{cases} n, & gh = e, \\ 0, & gh \neq e. \end{cases} \tag{11.21}$$

When $\operatorname{char} K$ does not divide n, this inner product is nondegenerate, hence KG is semisimple. □

For the rest of this section we assume that $K = \mathbb{C}$. Theorem 11.70 and results from Section 11.3 imply that the algebra $\mathbb{C}G$ is a direct sum of matrix algebras.

Theorem 11.71. *The group G has only a finite number of irreducible complex representations up to isomorphism. Their dimensions $n_1, \dots, n_s$ satisfy the relation*

$$n_1^2 + \cdots + n_s^2 = n, \tag{11.22}$$

while their number s equals the number of conjugacy classes of G.

Proof. The first statement of the theorem and relation (11.22) follow from Theorem 11.68 and formula (11.18). By formula (11.19), the number s equals the dimension of the center of the algebra $\mathbb{C}G$. Let us determine this center.

Element (11.20) lies in the center of $\mathbb{C}G$ if and only if it commutes with all elements of G, i.e., if

$$hah^{-1} = \sum_{g \in G} a_g (hgh^{-1}) = \sum_{g \in G} a_{h^{-1}gh} g = a$$

for every $h \in G$. This means that in the expression of a, the coefficients of conjugate elements of G are equal. Therefore, the center of $\mathbb{C}G$ is the linear span of elements of the form $\sum_{g \in C} g$, where C is a conjugacy class. Hence, the dimension of the center equals the number of conjugacy classes. □

Example 11.72. For an abelian group, every irreducible representation is one-dimensional (Corollary 11.15). Their number equals n because in this case, every conjugacy class consists of one element. This agrees with formula (11.22).

Example 11.73. Since under every homomorphism of a group G into an abelian group, its commutant is mapped into the identity, one-dimensional representations of every group G reduce to representations of the quotient group G/G'. In particular, for every n, the group S_n has exactly two one-dimensional representations: the trivial one and the nontrivial one that maps a permutation to its sign.

Example 11.74. For the group S_3, there definitely exist three pairwise nonisomorphic irreducible representations:

R_1, the trivial one-dimensional representation;

R_1', the sign of a permutation;

R_2, the two-dimensional representation under which S_3 is isomorphic to the symmetry group of an equilateral triangle (check that this representation is irreducible not only over $\mathbb{R}$ but over $\mathbb{C}$ as well!).

Since there exist exactly three conjugacy classes in S_3 (or, also, since $1^2 + 1^2 + 2^2 = 6$), this is the complete list of irreducible complex representations of S_3.

Example 11.75. Similarly, one obtains the following complete list of irreducible complex representations of the group S_4:

R_1, the trivial one-dimensional representation;

R_1', the sign of a permutation;

R_2, the composition of the homomorphism $S_4 \to S_4/V_4 \simeq S_3$ and the two-dimensional irreducible representation of S_3;

R_3, the isomorphism onto the group of rotations of a cube;

R_3', the isomorphism onto the symmetry group of a regular tetrahedron.

Remark 11.76. It follows from Examples 11.74 and 11.75 that all irreducible complex representations (and, thus, all complex representations) of the groups S_3 and S_4 are complexifications of real representations. It can be shown that the same is true for the group S_n for any n. Thus, together with Proposition 11.5, this implies that one can work with real representations of S_n just as with the complex ones. In particular, all theory presented in this section holds for real representations of S_n.

Exercise 11.77. Describe all irreducible representations of the dihedral group D_n.

Exercise 11.78. Prove that every irreducible representation of a group $G \times H$ is a tensor product of irreducible representations of groups G and H (see the definition in Exercise 11.32).

Let

$$R_i \colon G \to \mathrm{GL}(V_i), \qquad i = 1, \ldots, s,$$

be all irreducible complex representations of the group G. Then, after the appropriate identification, we can assume that

$$\mathbb{C}G = \mathrm{L}(V_1) \oplus \cdots \oplus \mathrm{L}(V_s), \tag{11.23}$$

and R_i is just the projection onto the ith summand in this decomposition.

Subspaces $\mathrm{L}(V_i)$ are the isotypic components of the regular representation T of the algebra $\mathbb{C}G$, and the restriction of T to $\mathrm{L}(V_i)$ is isomorphic to $n_i R_i$, $n_i = \dim V_i$. Thus, for every $a, b \in \mathbb{C}G$,

$$(a,b) = \sum_{i=1}^{s} n_i \operatorname{tr} R_i(a) R_i(b) \tag{11.24}$$

(see formula (11.12)).

Consider now the space $\mathbb{C}[G]$ of all complex-valued functions on G. Since every function φ on G extends uniquely to a linear function on $\mathbb{C}G$ according to the formula

$$\varphi\left(\sum_{g\in G} a_g g\right) = \sum_{g\in G} a_g \varphi(g),$$

the space $\mathbb{C}[G]$ is naturally identified with the dual space of $\mathbb{C}G$.

On the other hand, the inner product on the space $\mathbb{C}G$ determines its isomorphism with the dual space. In particular, under this isomorphism, an element $g \in G$ corresponds to the function φ_g defined as

$$\varphi_g(h) = (g,h) = \begin{cases} n, & gh = e, \\ 0, & gh \neq e, \end{cases}$$

i.e., to the δ-function $\delta_{g^{-1}}$ (at the point g^{-1}) multiplied by n.

With the help of the above isomorphism, we can transfer the inner product from the space $\mathbb{C}G$ to the space $\mathbb{C}[G]$. Then, for the δ-functions, we obtain

$$(\delta_g, \delta_h) = \frac{1}{n^2}(g^{-1}, h^{-1}) = \begin{cases} \frac{1}{n}, & gh = e, \\ 0, & gh \neq e, \end{cases}$$

and for any two functions φ and ψ,

$$(\varphi, \psi) = \frac{1}{n} \sum_{g\in G} \varphi(g)\psi(g^{-1}). \tag{11.25}$$

Let us calculate now the inner products of the matrix entries of irreducible representations of G.

In each space V_i, $i = 1, \ldots, s$, choose a basis and denote by φ_{ijk}, $j, k = 1, \ldots, n_i$, the (j,k)th matrix entry of the operator $R_i(g)$ in this basis. The function $\varphi_{ijk} \in \mathbb{C}[G]$ so defined is called the (j,k)th *matrix entry* of the representation R_i.

On the other hand, denote by $\mathcal{E}_{ijk}$ the linear operator on V_i whose matrix in the chosen basis is the matrix unit E_{jk}. The decomposition (11.23)

implies that the elements $\mathcal{E}_{ijk}$ form a basis of the space $\mathbb{C}G$. It follows from formula (11.24) that

$$(\mathcal{E}_{ijk}, \mathcal{E}_{ikj}) = n_i \tag{11.26}$$

and that other inner products of the elements $\mathcal{E}_{ijk}$ are zero.

Under the isomorphism of the spaces $\mathbb{C}G$ and $\mathbb{C}[G]$, the element $\mathcal{E}_{ijk}$ corresponds to the matrix entry φ_{ikj} with the coefficient n_i, in view of formula (11.26). Therefore,

$$(\varphi_{ijk}, \varphi_{ikj}) = \frac{1}{n_i} \tag{11.27}$$

and the other scalar products of matrix entries are zero.

Of particular interest are the sums of diagonal matrix entries; these are called the characters of representations R_i.

Generally speaking, let $R\colon G \to \mathrm{GL}(V)$ be a representation of a group G.

Definition 11.79. The *character* of the representation R is the function $\chi \in \mathbb{C}[G]$ defined as

$$\chi(g) = \operatorname{tr} R(g).$$

Obviously, the character of the sum of two representations equals the sum of their characters.

Since the traces of conjugate operators are equal,

$$\chi(hgh^{-1}) = \chi(g) \qquad \forall g, h \in G.$$

Functions $\chi \in \mathbb{C}[G]$ with this property are called *central*. They form a subspace in $\mathbb{C}[G]$ that we denote $Z\mathbb{C}[G]$. Clearly, for a finite group G, the dimension of this subspace equals the number of conjugacy classes, i.e., $\dim Z\mathbb{C}[G] = s$.

In particular, let χ_i be the character of the representation R_i, $i = 1, \ldots, s$. Formula (11.27) implies

Theorem 11.80. *The characters $\chi_1, \ldots, \chi_s$ form an orthonormal basis of the space $Z\mathbb{C}[G]$, i.e.,*

$$(\chi_i, \chi_j) = \delta_{ij}. \tag{11.28}$$

Let $R\colon G \to \mathrm{GL}(V)$ be a linear representation with character χ.

Corollary 11.81. *The multiplicity of an irreducible representation R_i in the decomposition of R equals (χ, χ_i).*

Proof. If $R \simeq \sum_{i=1}^{s} k_i R_i$, then $\chi = \sum_{i=1}^{s} k_i \chi_i$, hence, $(\chi, \chi_i) = k_i$. □

Corollary 11.82. *A representation R is irreducible if and only if $(\chi, \chi) = 1$.*

Proof. If $R \simeq \sum_{i=1}^{s} k_i R_i$, then $(\chi, \chi) = \sum_{i=1}^{s} k_i^2 = 1$ if and only if one of the multiplicities k_i equals 1 while the others are 0. □

Instead of the bilinear inner product (11.25) on the space $\mathbb{C}[G]$, we can consider the Hermitian inner product

$$(\varphi|\psi) = \frac{1}{n} \sum_{g \in G} \varphi(g)\overline{\psi(g)}, \tag{11.29}$$

which is more useful for actual calculations. If in every space V_i we choose a basis which is orthonormal with respect to an invariant Hermitian inner product (see Theorem 11.54), the representation operators are written as unitary matrices, i.e., the following relations hold:

$$\varphi_{ikj}(g^{-1}) = \overline{\varphi_{ijk}(g)}.$$

In terms of the Hermitian metric (11.29), relations (11.27) and (11.28) mean that *the matrix entries φ_{ijk} form an orthogonal basis of the space $\mathbb{C}[G]$ and*

$$(\varphi_{ijk}|\varphi_{ijk}) = \frac{1}{n_i},$$

while the characters χ_i form an orthonormal basis of the space $Z\mathbb{C}[G]$.

Example 11.83. The character of a one-dimensional representation coincides with the only matrix entry or, to put it differently, with the representation itself. The cyclic group $\langle a \rangle_n$ has n one-dimensional complex representations $R_0, R_1, \ldots, R_{n-1}$ defined by the conditions

$$R_k(a) = \omega^k, \qquad \omega = e^{2\pi i/n}.$$

Thus, the characters of this group are given by the following table:

	χ_0	χ_1	...	χ_{n-1}
e	1	1	...	1
a	1	ω	...	ω^{n-1}
a^2	1	ω^2	...	$\omega^{2(n-1)}$
...	...	...	...	...
a^{n-1}	1	ω^{n-1}	...	$\omega^{(n-1)^2}$

The orthogonality relations for the characters mean in this case that if the character table is divided by $\sqrt{n}$, we get a unitary matrix.

Example 11.84. Using the description of the irreducible representations of the group S_4 provided in Example 11.75, it is not difficult to obtain the following character table for this group:

	χ_1	χ_1'	χ_2	χ_3	χ_3'	
e	1	1	2	3	3	1
(12)	1	−1	0	−1	1	6
(12)(34)	1	1	2	−1	−1	3
(123)	1	1	−1	0	0	8
(1234)	1	−1	0	1	−1	6

The leftmost column of this table lists the representatives of conjugacy classes of S_4, and the rightmost column lists the number of elements in each class: these are needed for the calculations of inner products. For instance,

$$(\chi_2, \chi_3) = (\chi_2|\chi_3) = \frac{1}{24}(1\cdot 2\cdot 3+6\cdot 0\cdot(-1)+3\cdot 2\cdot(-1)+8\cdot(-1)\cdot 0+6\cdot 0\cdot 1) = 0.$$

Example 11.85. Let V be the (6-dimensional) space of functions on the set of faces of a cube. The isomorphism between S_4 and the symmetry group of the cube defines a linear representation of S_4 in the space V. Denote this representation by R and its character by χ. Every element $g \in S_4$ permutes the faces of the cube and thus $R(g)$ permutes the δ-functions of the faces. Therefore, $\chi(g) = \operatorname{tr} R(g)$ is the number of faces preserved by g. Thus, we obtain the following table for the values of χ:

	e	(12)	(12)(34)	(123)	(1234)
χ	6	0	2	0	2

Calculating the inner products of this character and the characters of irreducible representations of S_4 (Example 11.84), we obtain

$$(\chi|\chi_1) = 1, \quad (\chi|\chi_1') = 0, \quad (\chi|\chi_2) = 1, \quad (\chi|\chi_3) = 1, \quad (\chi|\chi_3') = 0.$$

Thus,

$$R \simeq R_1 + R_2 + R_3.$$

Exercise 11.86. Describe explicitly the minimal invariant subspaces of the representation R in Example 11.85.

Exercise 11.87. Let $G \subset S_n$ be a doubly transitive permutation group. (This means that for every two ordered pairs of different symbols, there exists a permutation of G that maps the first pair to the second.) Prove that the representation of G in the space of functions on the set $\{1, \dots, n\}$ decomposes into a sum of exactly two irreducible representations, one of which is the trivial one-dimensional representation. (*Hint*: use the expression that Burnside's formula (Exercise 10.47) gives for the number of orbits of G on the set $\{1, \dots, n\} \times \{1, \dots, n\}$.)

Exercise 11.88. Determine the character table of the group A_5.

Exercise 11.89. Let $R\colon G \to \mathrm{GL}(V)$ be a linear representation of a finite group G. Prove that the projection $\mathcal{P}_i$ of the space V onto its isotypic component that corresponds to the irreducible representation R_i of G can be given by the formula

$$\mathcal{P}_i = \frac{n_i}{n} \sum_{g \in G} \chi_i(g^{-1}) R(g),$$

where $n = |G|$, $n_i = \dim R_i$, and χ_i is the character of R_i. (*Hint*: prove that the element

$$\frac{n_i}{n} \sum_{g \in G} \chi_i(g^{-1}) g \in \mathbb{C}G$$

is the unity of the ith summand in the decomposition (11.23); for this, calculate its inner products with elements of G using (11.21) and (11.24).)

Apart from the operation of addition of representations that we considered above, there exist other important operations on linear representations of (arbitrary) groups.

For every linear representation $R\colon G \to \mathrm{GL}(V)$, it is possible to define the *dual representation* $R^*\colon G \to \mathrm{GL}(V^*)$ by the rule

$$(R^*(g)\alpha)(x) = \alpha(R(g)^{-1}x), \qquad \alpha \in V^*, x \in V, \tag{11.30}$$

i.e., by the standard rule describing how a transformation acts on functions. In the matrix language, this looks as follows:

$$R^*(g) = (R(g)^{\mathsf{T}})^{-1}. \tag{11.31}$$

Therefore, the character of the dual representation is determined by the formula

$$\chi_{R^*}(g) = \chi_R(g^{-1}). \tag{11.32}$$

The definition of the dual representation can be rewritten in the following symmetric form:

$$(R^*(g)\alpha)(R(g)x) = \alpha(x).$$

It follows that $R^{**} = R$ (under the canonical identification of V^{**} and V). It might happen that $R^* \simeq R$; in this case, the representation R is called *self-dual*.

For a complex linear representation of a finite group, in a basis that is orthonormal with respect to an invariant Hermitian inner product, formulas (11.31) and (11.32) become

$$R^*(g) = \overline{R(g)}, \quad \chi_{R^*}(g) = \overline{\chi_R(g)}. \tag{11.33}$$

Example 11.90. For irreducible (one-dimensional) representations of a cyclic group, we have, in the notation of Example 11.83,

$$R_0^* \simeq R_0, \quad R_k^* \simeq R_{n-k}, \quad k = 1, \ldots, n-1.$$

Example 11.91. It follows from Example 10.39 that in the group S_n, every element is conjugate to its inverse. Thus, every linear representation of S_n is self-dual.

Exercise 11.92. Prove that if R is an irreducible group representation, the representation R^* is also irreducible.

Exercise 11.93. Prove that all representations of a finite group G are self-dual if and only if every element of G is conjugate to its inverse.

Now we define the multiplication of linear representations of a group G.

The *product* of linear representations $R\colon G \to \mathrm{GL}(V)$ and $S\colon G \to \mathrm{GL}(W)$ is the linear representation

$$RS\colon G \to \mathrm{GL}(V \otimes W), \quad g \mapsto R(g) \otimes S(g).$$

(See the definition of the tensor product of linear operators in Section 8.1.)

Remark 11.94. Sometimes the representation RS is called the tensor product of representations R and S, but we reserve this term for the representation of the direct product of two groups, as defined in Exercise 11.32.

Exercise 11.95. Choose bases in the spaces V and W and write R and S in these bases. We write an element of the space $V \otimes W$ as the matrix Z of its coordinates (see formula (8.10)). Prove that in this notation, the representation RS is described by the formula

$$(RS)(g)Z = R(g)ZS(g)^{\mathsf{T}}. \tag{11.34}$$

Formula (8.28) implies that

$$\chi_{RS} = \chi_R \chi_S. \tag{11.35}$$

As a rule, the product of irreducible representations is not irreducible. The decomposition of the product of irreducible representations into irreducible components is one of the fundamental problems of representation theory. Due to formula (11.35), for representations of finite groups this problem can be solved with the help of characters.

Example 11.96. Let us decompose the square of the representation R_3 of the group S_4 (Example 11.75) into a sum of irreducible representations. Using the character table of S_4 given in Example 11.84, we obtain

$$(\chi_3^2|\chi_1) = 1, \quad (\chi_3^2|\chi_1') = 0, \quad (\chi_3^2|\chi_2) = 1,$$
$$(\chi_3^2|\chi_3) = 1, \quad (\chi_3^2|\chi_3') = 1.$$

Therefore,

$$R_3^2 \simeq R_1 + R_2 + R_3 + R_3'. \tag{11.36}$$

Similarly, one defines the product of several representations as well as the symmetric and the exterior power of a representation. For instance, the symmetric square of a representation $R\colon G \to \mathrm{GL}(V)$ is the representation

$$S^2R\colon G \to \mathrm{GL}(S^2(V)), \quad g \mapsto S^2(R(g)).$$

(See the definition of the symmetric square of a linear operator in Section 8.3.)

Formula (8.53) implies that

$$\chi_{S^2R}(g) = \frac{1}{2}(\chi_R(g)^2 + \chi_R(g^2)). \tag{11.37}$$

If we identify the space $S^2(V)$ with the space $ST^2(V)$ of symmetric tensors, the representation S^2R becomes nothing but the restriction of the representation R^2 to the invariant subspace $ST^2(V)$. A similar statement holds for the exterior square Λ^2R of a representation R. Since $T^2(V) = ST^2(V) \oplus \Lambda T^2(V)$,

$$R^2 \simeq S^2R + \Lambda^2R. \tag{11.38}$$

Example 11.97. Hooke's law in the theory of elasticity describes the relation between the deformation tensor σ and the stress tensor τ of a convex solid at a given point. Both these tensors are symmetric operators on the space E^3. (For the definition of the deformation tensor, see Example 6.45.) By lifting indices we can regard them as elements of the space $S^2(E^3)$. Hooke's law says that $\sigma = \mathcal{H}\tau$, where $\mathcal{H}$ is a symmetric operator on the space $S^2(E^3)$ called the elasticity tensor. It describes the elasticity properties of a given convex solid at a given point (under given temperature and pressure). Since $\dim S^2(E^3) = 6$, the dimension of the space of symmetric operators on $S^2(E^3)$ equals $\frac{6\cdot 7}{2} = 21$. Thus, in the general case, the elasticity tensor depends on 21 parameters which should be determined from experiment.

The picture simplifies if the solid has a crystal structure. Namely, let $G = d\Gamma$, where Γ is the symmetry group of this crystal structure (see Example 9.11). Then the operator $\mathcal{H}$ must commute with all operators $S^2\mathcal{A}$ for $\mathcal{A} \in G$. The general form of such an operator can be determined using representation theory. The larger the group G, the smaller the number of parameters on which it depends.

Consider, for example, the crystal of table salt (see Figure 4.2 in Section 4.2). Here G is the symmetry group of a cube, i.e., in the notation of Example 11.75, $G = R_3(S_4) \times \{\pm\mathcal{E}\}$. The second factor acts trivially on

$S^2(E^3)$ and can thus be discounted. Thus, the operator $\mathcal{H}$ must be an endomorphism of the representation S^2R_3 of S_4. From formula (11.37), we obtain the following table for the values of the character χ of this representation:

	e	(12)	(12)(34)	(123)	(1234)
χ	6	2	2	0	0

Calculating its inner products with the characters of irreducible representations, we conclude that

$$S^2R_3 \simeq R_1 + R_2 + R_3'. \tag{11.39}$$

In particular, the representation S^2R_3 has a simple spectrum. By Proposition 11.34 (see also Remark 11.76), the general form of this endomorphism depends on three parameters. Thus, in order to find the elasticity tensor of the crystal of table salt, one has to determine experimentally just 3 parameters (instead of 21!).

Remark 11.98. Due to isomorphism (11.38), decomposition (11.39) could be found by subtracting from decomposition (11.36) the representation Λ^2R_3, which can be easily shown to be isomorphic to R_3.

Exercise 11.99. Prove that an irreducible representation of a group is self-dual if and only if it is orthogonal or symplectic (for definitions, see the end of Section 11.1).

Exercise 11.100. Prove that an irreducible complex representation of a finite group is a complexification of a real representation if and only if it is orthogonal.

Exercise 11.101. Prove that the sum of dimensions of orthogonal irreducible representations of a finite group G minus the sum of dimensions of its symplectic irreducible representations equals the number of solutions of the equation $x^2 = e$ in G. (*Hint*: calculate the trace of the antiautomorphism of the group algebra $\mathbb{C}G$ induced by the inversion in G in the basis consisting of elements of G and in a basis that agrees with decomposition (11.23).)

11.5. Invariants

Every action of a group G on a set X defines a linear representation of this group in the space $F(X, K)$ of K-valued functions on X (see formula (10.8)).

Definition 11.102. A function $f \in F(X, K)$ is an *invariant* of (this action of) the group G if $gf = f$ for every $g \in G$.

In other words, an invariant is a function that is constant on the orbits of G. Knowing invariants helps in describing orbits, which is an important problem. Namely, if an invariant f assumes different values at two points, then these points must belong to different orbits. An ideal solution to this problem is to list invariants $f_1, \dots, f_m$ such that for any two points in different orbits, at least one of the invariants assumes different values at these points. In this case, we say that the invariants $f_1, \dots, f_m$ *separate the orbits*.

Example 11.103. For the group $\mathrm{GL}_n(\mathbb{C})$ consider the linear representation Ad in the space $\mathrm{L}_n(\mathbb{C})$ defined as

$$\mathrm{Ad}(A)X = AXA^{-1}.$$

Let $f_X(t) = \det(tE - X)$ be a characteristic polynomial of X. Write it as

$$f_X(t) = t^n - f_1(X)t^{n-1} + \cdots + (-1)^n f_n(X).$$

Then $f_k(X)$ is the sum of the principal minors of X of order k (Exercise 6.15). Since characteristic polynomials of similar matrices are equal, $f_1, \dots, f_n$ are invariants of this action of $\mathrm{GL}_n(\mathbb{C})$. However, they do not separate orbits. Indeed, two matrices lie in the same orbit if and only if they have the same Jordan canonical form, whereas the values of invariants $f_1, \dots, f_n$ determine only the eigenvalues of the matrix.

Example 11.104. Invariants of the symmetric group S_n acting on K^n by permuting the coordinates are functions in n variables that do not change under any permutation of variables. In particular, invariant polynomials are the symmetric polynomials.

The space $F(X, K)$ is an algebra with respect to ordinary function multiplication, and transformations in G are automorphisms of this algebra. It follows that the invariants form a subalgebra in $F(X, K)$.

Usually, one searches for invariants not among all functions but only among those that are, in some sense, "good." The most common is the situation where $X = V$ is a vector space over a field K and the action of G is defined by its linear representation in V. In this case, one usually confines the search for invariants to the algebra $K[V]$ of polynomials on V. (This is what we did in Example 11.103.) The subalgebra of invariants in $K[V]$ is denoted $K[V]^G$.

We say that the orbits of a linear group $G \subset \mathrm{GL}(V)$ are *separated by invariants* if for every two orbits, there exists an invariant $f \in K[V]^G$ that assumes different values on them.

Theorem 11.105. *If $G \subset \mathrm{GL}(V)$ is a finite group and its order is not divisible by* $\operatorname{char} K$, *then its orbits are separated by invariants.*

Proof. Let O_1 and O_2 be two different orbits. There exists a polynomial $f \in K[V]$ such that the value of f is 1 at every point of O_1 and 0 at every point of O_2. The set of all polynomials of degree $\leq \deg f$ with this property is a plane S in the space of all polynomials of degree $\leq \deg f$. The group G preserves this plane and acts on it by affine transformations. By Lemma 11.37, there exists a fixed point of G on S. This is an invariant we need. □

Exercise 11.106. In the above proof, we used the fact that for any finite number of points in V, there exists a polynomial assuming prearranged values at these points. Prove this.

Example 11.107. Defining a vector $(x_1, x_2, \ldots, x_n) \in K^n$ up to a permutation of its coordinates is equivalent to giving the polynomial

$$(x - x_1)(x - x_2) \cdots (x - x_n) \in K[x].$$

Coefficients of this polynomial are, up to a sign, elementary symmetric polynomials in $x_1, x_2, \ldots, x_n$. Therefore, the orbits of the group S_n in the space K^n (see Example 11.104) are separated by elementary symmetric polynomials (which are invariants of S_n) and, certainly, by all invariants.

If the algebra $K[V]^G$ is generated by invariants $f_1, \ldots, f_m$ and these invariants assume the same values at two points, then all invariants assume the same values at these points. Thus, if the orbits of G are separated by (all) invariants, they are also separated by invariants $f_1, \ldots, f_m$. For instance, in the previous example, we could say beforehand (whenever $\operatorname{char} K$ does not divide $|G|$) that the orbits of S_n must be separated by elementary symmetric polynomials, since these polynomials generate the algebra of all symmetric functions.

The following theorem is a particular case of the modern version of *Hilbert's finiteness theorem*. Hilbert himself proved this theorem in 1891 for linear representations of $\mathrm{SL}_n(K)$, but proofs based on his ideas can be applied to a more general setting and, in particular, to the case of simple groups.

Theorem 11.108. *If G is a finite group and its order is not divisible by* $\operatorname{char} K$, *then the algebra $K[V]^G$ is finitely generated.*

The statement of this theorem means that there are invariants $f_1, \ldots, f_m$ such that every invariant can be represented as a polynomial in $f_1, \ldots, f_m$ (but perhaps not uniquely).

Proof. Define the so-called *Reynolds operator* $\natural$ (this symbol is called "natural") as

$$f^{\natural} = \operatorname{cent} Gf = \frac{1}{|G|} \sum_{g \in G} gf. \tag{11.40}$$

This linear operator has the following properties:

(i) $f^{\natural} \in K[V]^G$ for every $f \in K[V]$;

(ii) $f^{\natural} = f$ for every $f \in K[V]^G$;

(iii) $(fh)^{\natural} = fh^{\natural}$ for every $f \in K[V]^G$, $h \in K[V]$.

In other words, this is a projection onto the algebra of invariants that commutes with multiplications by invariants.

Observe that a polynomial is invariant if and only if all its homogeneous components are invariant and that the Reynolds operator maps a homogeneous polynomial to a homogeneous polynomial of the same degree.

Now let $I \subset K[V]$ be an ideal generated by all homogeneous invariants of positive degree. By Hilbert's basis theorem, I is generated by a finite number of polynomials. Clearly, these can be chosen among homogeneous invariants. Let these be invariants $f_1, \ldots, f_m$, and let

$$K[f_1, \ldots, f_m] \subset K[V]^G$$

be the subalgebra that they generate. We will show that it coincides with the algebra of invariants. For this, we prove by induction on n that every homogeneous invariant of degree n lies in the algebra $K[f_1, \ldots, f_m]$.

For $n = 0$, there is nothing to prove (the algebra $K[f_1, \ldots, f_m]$ contains the constants by definition). Let f be a homogeneous invariant of positive degree. Since $f \in I$, there exist polynomials $h_1, \ldots, h_m \in K[V]$ such that

$$f = \sum_{i=1}^{m} f_i h_i.$$

Without loss of generality, we can assume that h_i is a homogeneous polynomial of degree

$$\deg h_i = \deg f - \deg f_i < \deg f.$$

Applying the operator $\natural$ to the previous equality, we obtain

$$f = \sum_{i=1}^{m} f_i h_i^{\natural}.$$

By induction hypothesis, $h_i^{\natural} \in K[f_1, \ldots, f_m]$. Hence, $f \in K[f_1, \ldots, f_m]$. $\square$

In particular cases, finding a finite generating set of the algebra of invariants explicitly can be a difficult problem.

Example 11.109. Here we give another proof that the algebra of invariants of the symmetric group S_n, i.e., the algebra of symmetric polynomials (see Example 11.104), is generated by elementary symmetric polynomials $\sigma_1, \ldots, \sigma_n$. We showed in Example 10.81 that

$$K(x_1, \ldots, x_n)^{S_n} = K(\sigma_1, \ldots, \sigma_n),$$

and $\sigma_1, \ldots, \sigma_n$ are algebraically independent. Since $x_1, \ldots, x_n$ are the roots of the polynomial

$$(x - x_1) \cdots (x - x_n) = x^n - \sigma_1 x^{n-1} + \cdots + (-1)^n \sigma_n$$

with coefficients in $K[\sigma_1, \ldots, \sigma_n]$, the algebra $K[x_1, \ldots, x_n]$ of all polynomials and, a fortiori, its subalgebra $K[x_1, \ldots, x_n]^{S_n}$ are integral extensions of the algebra $K[\sigma_1, \ldots, \sigma_n]$. At the same time,

$$K[x_1, \ldots, x_n]^{S_n} \subset K(x_1, \ldots, x_n)^{S_n} = K(\sigma_1, \ldots, \sigma_n).$$

Since the algebra $K[\sigma_1, \ldots, \sigma_n]$ is isomorphic to the algebra of polynomials in n variables and, hence, factorial, it is normal (integrally closed in its quotient field). Therefore,

$$K[x_1, \ldots, x_n]^{S_n} = K[\sigma_1, \ldots, \sigma_n].$$

Example 11.110. One should not think that, as in the previous example, the algebra of invariants is always generated by algebraically independent elements. Such a situation is rather an exception than a rule. For example, consider the group

$$G = \{\pm E\} \subset \mathrm{GL}(V), \qquad \operatorname{char} K \neq 2.$$

A homogeneous polynomial is an invariant of this group if and only if it has an even degree. Thus, here the minimal set of generators of the algebra of invariants consists of polynomials $f_{ij} = x_i x_j$ and these are related:

$$f_{ij} f_{kl} = f_{ik} f_{jl}.$$

Remark 11.111. Theorems 11.105 and 11.108 still hold for finite groups whose order is divisible by $\operatorname{char} K$ but the proofs provided above are no longer valid.

When $K = \mathbf{R}$, the above theorems can be generalized to arbitrary compact groups.

Theorem 11.112. *The orbits of a compact group G of linear transformations of a real vector space V are separated by invariants.*

Proof. Following the proof of Theorem 11.105, we cannot expect now that a polynomial assuming 1 on O_1 and 0 on O_2 exists. However, by the Weierstrass approximation theorem (on uniform approximation by polynomials of a continuous function on a compact set), there exists a polynomial f that is

positive on O_1 and negative on O_2. The collection of polynomials of degree $\leq \deg f$ with this property is a G-invariant convex set M in the space of all polynomials of degree $\leq \deg f$. A fixed point of G in this set is an invariant we need. □

Remark 11.113. For complex vector spaces, an analogous theorem fails as the example of the circle $\mathbb{T} \subset \mathbb{C}^* = \mathrm{GL}_1(\mathbb{C})$ demonstrates.

Theorem 11.114. *Let G be a compact group of linear transformations of a vector space V over the field $K = \mathbb{R}$ or $\mathbb{C}$. Then the algebra $K[V]^G$ is finitely generated.*

As Theorem 11.108, this theorem is a particular case of Hilbert's finiteness theorem.

Proof. The proof can follow that of Theorem 11.108 once we define the Reynolds operator with properties (i)–(iii). This can be done by replacing summation over a finite group in formula (11.40) with a properly defined integration over a compact group. (For instance, in the case of $G = \mathbb{T}$, this is the standard integration over a circle.) However, we will use a different construction.

By complete reducibility of linear representations of a compact group (see Corollary 11.44), the space $K[V]_n$ of homogeneous polynomials of degree n on V decomposes into the direct sum of the subspace $K[V]_n^G$ of G-invariant polynomials and a G-invariant subspace $(K[V]_n)_G$. Let

$$K[V]_G = \bigoplus_{n=0}^{\infty} (K[V]_n)_G.$$

Clearly, the subspace $K[V]_G$ is invariant under G and

$$K[V] = K[V]^G \oplus K[V]_G. \tag{11.41}$$

Now, let us define the operator $\natural$ as the projection onto $K[V]^G$ with respect to decomposition (11.41). By construction, this projection commutes with the action of G. It remains only to check that it commutes with multiplication by invariants. For this, it suffices to prove that

$$K[V]^G K[V]_G \subset K[V]_G.$$

The multiplication by an invariant $f \in K[V]^G$ commutes with the action of G, i.e., it is an endomorphism of the representation of G in the space $K[V]$. Since the subspace $K[V]_G$ is, by construction, a complement of $K[V]^G$, the representation of G on $K[V]_G$ decomposes into a sum of nontrivial irreducible representations. The same can be said about the representation of

G in $fK[V]_G$. Hence, the projection of this subspace onto $K[V]^G$ is zero, i.e.,

$$fK[V]_G \subset K[V]_G,$$

and this completes the proof. □

Example 11.115. Consider the linear representation R of the group O_n in the space L_n^+ of real symmetric matrices of order n determined by the formula

$$R(A)X = AXA^{-1}(= AXA^{\top}).$$

Consider the characteristic polynomial of X,

$$\det(tE - X) = t^n - f_1(X)t^{n-1} + f_2(X)t^{n-2} - \cdots + (-1)^n f_n(X).$$

Let us prove that

$$\mathbb{R}[\mathrm{L}_n^+]^{R(\mathrm{O}_n)} = \mathbb{R}[f_1, \ldots, f_n]$$

and that $f_1, \ldots, f_n$ are algebraically independent. For this, recall that every symmetric matrix is orthogonally similar to a diagonal matrix. Thus, every invariant f of the group $R(\mathrm{O}_n)$ is determined uniquely by its restriction to the space D of diagonal matrices. Since diagonal matrices that differ only in the order of their diagonal elements are orthogonally similar, $f|_D$ is a symmetric polynomial in diagonal elements $x_1, \ldots, x_n$. A direct check shows that the restrictions of the invariants $f_1, \ldots, f_n$ to D are the elementary symmetric polynomials in $x_1, \ldots, x_n$. Then the statements we are proving follow from the theorem about symmetric polynomials. Notice that in this example, the orbits are separated by invariants, as they should be, according to Theorem 11.112.

11.6. Division Algebras

Since the field of complex numbers is algebraically closed, it follows that the only finite-dimensional algebras over $\mathbb{R}$ that are also fields are $\mathbb{R}$ and $\mathbb{C}$. However, if we drop the condition of commutativity for multiplication, we can construct one more such algebra, namely the algebra of quaternions. This algebra also plays an important role in mathematics and its applications. The theory turns out to be even more substantial if, as a base field, we consider an arbitrary field instead of $\mathbb{R}$ (e.g., $\mathbb{Q}$).

Definition 11.116. A *division ring* is an associative ring with unity where every nonzero element has an inverse. An algebra that is also a division ring is called a *division algebra*.

Remark 11.117. A ring that consists of zero only is not regarded as a division ring.

In other words, a division ring is a "noncommutative field." Similarly to a field, a division ring has no zero divisors and its nonzero elements form a multiplicative group (though, not necessarily abelian). The multiplicative group of nonzero elements of a division ring D is denoted D^*.

Every division ring D can be viewed as a division algebra over its *center*

$$Z(D) = \{z \in D\colon za = az\ \forall a \in D\},$$

which is obviously a field.

In a division algebra D with the unity 1 over a field K, the elements of the form $\lambda 1$, $\lambda \in K$, form a subring isomorphic to K. It is contained in the center $Z(D)$ of D. Usually these elements are identified with the corresponding elements of K. With this identification, $Z(D) \supset K$. An algebra D is called *central* if $Z(D) = K$.

Exercise 11.118. Prove that a finite-dimensional associative algebra is a division algebra if and only if it contains no zero divisors.

The simplest and most important examples of noncommutative division algebras are the quaternion algebras.

A (*generalized*) *quaternion algebra* over a field K of characteristic $\neq 2$ is an algebra $D = D(\alpha, \beta)$, $\alpha, \beta \in K^*$, generated by elements $\imath, \jmath$ satisfying the following relations:

$$\imath^2 = \alpha, \qquad \jmath^2 = \beta, \qquad \imath\jmath = -\jmath\imath.$$

It is easy to see that the elements $1, \imath, \jmath$, and $k = \imath\jmath$ constitute a basis of the algebra D over K with elements $\imath, \jmath, k$ pairwise anticommutative and

$$k^2 = -\alpha\beta.$$

In particular, for $K = \mathbb{R}$, the algebra $D(-1,-1)$ is the ordinary algebra of quaternions $\mathbb{H}$ discovered by Hamilton in 1843.

The algebra $D(1,1)$ is isomorphic to the matrix algebra $\mathrm{L}_2(K)$. This isomorphism is established as follows:

$$1 \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \imath \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \jmath \leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad k \leftrightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

In order to determine if it is possible to divide in a quaternion algebra, define for every quaternion

$$q = x + y\imath + z\jmath + uk, \qquad x, y, z, u \in K,$$

the *conjugate* quaternion $\overline{q}$ by the formula

$$\overline{q} = x - y\imath - z\jmath - uk.$$

It is easy to see that the linear map $q \mapsto \overline{q}$, called the *standard involution*, is an *antiautomorphism* of the algebra D, i.e.,

$$\overline{q_1 q_2} = \overline{q_2}\,\overline{q_1}.$$

(By linearity, it suffices to check this for the basis elements.) The element

$$N(q) = q\overline{q} = x^2 - \alpha y^2 - \beta z^2 + \alpha\beta u^2 \in K \tag{11.42}$$

is called the *norm* of the quaternion q. Clearly, q is invertible if and only if $N(q) \neq 0$ (and in this case, $q^{-1} = N(q)^{-1}\overline{q}$).

The algebra $D = D(\alpha, \beta)$ is a division algebra if and only if the equation

$$x^2 - \alpha y^2 - \beta z^2 + \alpha\beta u^2 = 0$$

has no nonzero solutions in the field K or, as this is sometimes said, if the quadratic function (11.42) does not represent zero over K. In particular, this condition holds for $K = \mathbb{R}$ and $\alpha = \beta = -1$ since in this case the quadratic form (11.42) is positive definite.

Finally, observe that for any $a, b \in K^*$, we have

$$(a\imath)^2 = a^2\alpha, \qquad (b\jmath)^2 = b^2\beta, \qquad (a\imath)(b\jmath) = -(b\jmath)(a\imath).$$

This shows that

$$D(a^2\alpha, b^2\beta) \simeq D(\alpha, \beta).$$

Exercise 11.119. Prove that in the algebra $D(1,1) \simeq \mathrm{L}_2(K)$, the norm of a matrix is its determinant. Interpret in matrix terms the standard involution of this algebra.

Exercise 11.120. Prove that $D(\alpha, 1) \simeq \mathrm{L}_2(K)$ for every $\alpha \in K^*$.

If D is a finite-dimensional division algebra over a field K, the subalgebra $K[x]$ is commutative for every $x \in D$, hence, it is a field. Thus, every finite-dimensional division algebra over an algebraically closed field K coincides with K.

When dealing with division algebras over a field that is not algebraically closed, it is always useful to study what happens over algebraic extensions of this field. For instance, in the study of real algebras, it is useful to investigate their complexifications. Allowing algebraic extensions, we put ourselves in the situation equivalent to that over an algebraically closed field. On the other hand, many properties of algebras are preserved in such extensions.

Let A be an algebra over a field K and P an extension of K. The vector space $A(P) = P \otimes_K A$ can be turned into an algebra over P if we define the product of its elements by the rule

$$(\lambda \otimes u)(\mu \otimes v) = \lambda\mu \otimes uv.$$

Identifying every element $a \in A$ with the element $1 \otimes a \in A(P)$, we obtain an inclusion of the algebra A into $A(P)$. If $\{e_1, \dots, e_n\}$ is a basis of A over K, then the multiplication in K is defined by the formulas

$$e_i e_j = \sum_k c_{ijk} e_k.$$

Elements $c_{ijk} \in K$ are called the *structure constants* of the algebra A in the basis $\{e_1, \dots, e_n\}$. The same formulas define the multiplication in the algebra $A(P)$ in the basis $\{e_1, \dots, e_n\}$. However, it makes sense to consider extensions because there exist other bases of $A(P)$ where the structure constants might look simpler.

To gain something from this method, we need, of course, to prove beforehand that some properties of an algebra are invariant when the base field is extended.

Proposition 11.121. *A semisimple finite-dimensional associative algebra A over a field K of zero characteristic remains semisimple if we pass to any extension P of K.*

Proof. We use here the criterion for semisimplicity of a finite-dimensional associative algebra which relates to the inner product (see Theorem 11.54). Obviously, in a basis that consists of elements of A, the matrix of the inner product in $A(P)$ is the same as in A. Thus, it is nonsingular, implying that $A(P)$ is semisimple. □

For example, let L be a finite extension of a field K. Regard it as an algebra over K. This algebra is semisimple (even simple); therefore, by the above, for every extension P of K, the algebra $L(P)$ is also semisimple, hence, it is a direct sum of finite extensions of P.

Let $\alpha \in L \setminus K$ be an element with the minimal polynomial h over K. Then

$$L \supset K[\alpha] \simeq K[x]/(h),$$

hence

$$L(P) \supset P[\alpha] \simeq P[x]/(h).$$

If the polynomial h is reducible over P, in particular, if it has a root in P, then $P[\alpha]$ and, a fortiori, $L(P)$ are not fields. Thus, by taking successive simple algebraic extensions of K, we can obtain a finite extension P such that

$$L(P) \simeq \underbrace{P \oplus \cdots \oplus P}_{n}, \qquad n = \dim_K L. \tag{11.43}$$

If $P \supset K$ is such an extension, we say that L *splits* over P.

As an example of an application of the above approach, let us prove the *Primitive Element Theorem*:

Theorem 11.122. *Every finite extension L of a field K of zero characteristic is simple, i.e., generated over K by a single element.*

Proof. Let $\dim_K L = n$. If L is not generated as an algebra over K by a single element, then for every $\alpha \in L$, the elements $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ are linearly dependent. This can be expressed by setting the determinant formed by columns of coordinates of $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ in some basis of L over K identically equal to zero. As a function of coordinates of α, this determinant is a polynomial with coefficients in K. If it is zero for all values assumed by the variables in K, it is a zero polynomial, hence it equals zero in every extension P over K. This, in turn, means that the algebra $L(P)$ is not generated over P by a single element.

However, if L splits over P, it is easy to prove that $L(P)$ is generated by a single element. Indeed, consider an element

$$\alpha = (\alpha_1, \ldots, \alpha_n) \in \underbrace{P \oplus \cdots \oplus P}_{n}$$

with different coordinates $\alpha_1, \ldots, \alpha_n \in P$. Then the determinant formed by the coordinates of the elements $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ is the Vandermonde determinant for $\alpha_1, \ldots, \alpha_n$, thus nonzero. □

We now turn to the study of noncommutative finite-dimensional division algebras. Clearly, every division algebra is simple.

Let D be a finite-dimensional central division algebra over a field K.

Proposition 11.123. *The algebra D remains simple over every extension P of K.*

Proof. Let $I \subset D(P)$ be a nonzero ideal. Consider a shortest nonzero linear combination of the form

$$a = \sum_{i=1}^{s} \lambda_i a_i, \qquad \lambda_i \in P,\ a_i \in D,$$

belonging to I. Clearly, $a_1, \ldots, a_s$ are linearly independent over K; otherwise we could have reduced the number of summands. Likewise, $\lambda_1, \ldots, \lambda_s$ are linearly independent over K.

By multiplying the element a by a_1^{-1} (on either side), we obtain $a_1 = 1$ and still remain inside I. If $s > 1$, then $a_2 \notin K = Z(D)$, thus, there exists an element $c \in D^*$ such that $ca_2 \neq a_2 c$. We have

$$a - cac^{-1} = \sum_{i=2}^{s} \lambda_i (a_i - ca_i c^{-1}) \in I,$$

while $a - cac^{-1} \neq 0$ since $\lambda_2, \ldots, \lambda_s$ are linearly independent over K and $a_2 - ca_2c^{-1} \neq 0$. This contradicts the definition of a. Therefore, $s = 1$. But then, $I \ni 1$ and $I = D(P)$. □

Theorem 11.124. *There exists a finite-dimensional extension P of K such that $D(P) \simeq \mathrm{L}_n(P)$ for some $n \in \mathbb{N}$.*

Corollary 11.125. $\dim D = n^2$.

The number n is called the *degree* of the algebra D and is denoted $\deg D$. For instance, the degree of a quaternion algebra is 2.

Proof of Theorem 11.124. Let $\bar{K}$ be a maximal algebraic extension of K. (Its existence is proved with the use of Zorn's lemma.) This algebraically closed field is called the *algebraic closure* of K. By Theorem 11.67,

$$D(\bar{K}) \simeq \mathrm{L}_n(\bar{K})$$

for some $n \in \mathbb{N}$. Let $e_{ij} \in D(\bar{K})$ be the elements that are identified by this isomorphism with matrix units, and let $P \subset \bar{K}$ be the subfield generated over K by the coordinates of all these elements in a basis of $D(\bar{K})$ composed of elements of D. Obviously, P is a finite extension of K and $D(P) \simeq \mathrm{L}_n(P)$. □

If $P \supset K$ is an extension such that $D(P) \simeq \mathrm{L}_n(P)$, we say that D *splits* over P.

Example 11.126. A quaternion algebra $D(\alpha, \beta)$ splits over the field $P = K(\sqrt{\alpha}, \sqrt{\beta})$.

We gain important information about the algebra D from the study of its maximal commutative subalgebras or, to put it differently, its maximal subfields.

Theorem 11.127. *Every maximal subfield F of the algebra D has dimension n over K. Every isomorphism of maximal subfields extends to an inner automorphism of D.*

(We do not claim, however, that all maximal subfields are isomorphic.)

Proof. First of all, observe that if F is a maximal commutative subalgebra of D and P is an extension of K, then $F(P)$ is a maximal commutative subalgebra of $D(P)$. Indeed, the maximality condition means that F coincides with its own centralizer

$$Z_D(F) = \{x \in D \colon ax = xa \ \forall a \in F\},$$

and this is equivalent to

$$\dim Z_D(F) = \dim F.$$

However, the coordinate definition of the centralizer is written as a system of homogeneous linear equations with coefficients in K, and the dimension of the space of solutions of any homogeneous system of linear equations does not change if we pass to an extension field.

Now, let $\bar{K}$ be the algebraic extension of K. Then $D(\bar{K}) \simeq \mathrm{L}_n(\bar{K})$ and

$$F(\bar{K}) \simeq \underbrace{\bar{K} \oplus \cdots \oplus \bar{K}}_{m},$$

where $m = \dim F$. The latter means that $F(\bar{K})$ has a basis $\{e_1, \ldots, e_m\}$ such that

$$e_i^2 = e_i, \qquad e_i e_j = 0 \quad \text{for} \quad i \neq j.$$

If we view the algebra $\mathrm{L}_n(\bar{K})$ as an algebra of linear operators, the elements $e_1, \ldots, e_m$ correspond to pairwise commuting projections. In a suitable basis, these projections have diagonal matrices.

Identify the algebra $D(\bar{K})$ with $\mathrm{L}_n(\bar{K})$ via a fixed isomorphism. Then the above discussion implies that there exists an element $c \in D(\bar{K})^*$ such that $cF(\bar{K})c^{-1}$ consists of diagonal matrices. But since $cF(\bar{K})c^{-1}$ is a maximal commutative subalgebra, it coincides with the subalgebra of all diagonal matrices. Hence, $m = n$, which proves the first assertion of the theorem.

Now let $F_1, F_2 \subset D$ be two maximal commutative subalgebras and $\varphi\colon F_1 \xrightarrow{\sim} F_2$, an isomorphism. Then φ extends to an isomorphism

$$\overline{\varphi}\colon F_1(\bar{K}) \xrightarrow{\sim} F_2(\bar{K}).$$

The above discussion shows that there exists an element $c \in D(\bar{K})^*$ such that $c_1 F(\bar{K})c^{-1} = F_2(\bar{K})$. More precisely, since every automorphism of the algebra of diagonal matrices is simply a permutation of diagonal entries, hence induced by an inner automorphism of the algebra $\mathrm{L}_n(\bar{K}) = D(\bar{K})$, we can assume that

$$cac^{-1} = \overline{\varphi}(a) \qquad \text{for } a \in F_1(\bar{K}). \tag{11.44}$$

We need to prove that there exists a nonzero element $x \in D$ such that $xax^{-1} = \varphi(a)$ for $a \in F_1$ or, equivalently, that

$$xa = \varphi(a)x \qquad \forall a \in F_1.$$

In the coordinate form, this conditions are written as a system of homogeneous linear equations with coefficients in K. It follows from (11.44) that this system has a nonzero solution in $\bar{K}$; but then it has a nonzero solution in K as well. □

Let us apply the theory that we have just developed to the description of division algebras over $\mathbb{R}$ and over finite fields.

Theorem 11.128 (Frobenius Theorem). *Every finite-dimensional division algebra D over $\mathbb{R}$ is isomorphic to either $\mathbb{R}$, or $\mathbb{C}$, or $\mathbb{H}$.*

Proof. The center $Z(D)$ of D is a finite extension of $\mathbb{R}$, hence it is isomorphic to either $\mathbb{R}$ or $\mathbb{C}$. In the latter case, D can be viewed as an algebra over $\mathbb{C}$; thus, since $\mathbb{C}$ is an algebraically closed field, $D \simeq \mathbb{C}$.

Now let $Z(D) = \mathbb{R}$, i.e., let D be a central algebra. Since every maximal subfield of D is isomorphic to either $\mathbb{R}$ or $\mathbb{C}$, $\deg D = 1$ or 2. In the former case, $D = \mathbb{R}$. Now consider the latter case.

Choose a maximal subfield of D and identify it with $\mathbb{C}$. By Theorem 11.127, the complex conjugation in $\mathbb{C}$ extends to an inner automorphism of D, i.e., there exists an element $\jmath \in D$ such that $\jmath z \jmath^{-1} = \overline{z}$ for every $z \in \mathbb{C}$. Clearly, $\jmath \notin \mathbb{C}$. Therefore, $D = \mathbb{C} \oplus \mathbb{C}\jmath$. Now, since $\jmath^2$ commutes with $\jmath$ and with all elements of $\mathbb{C}$, $\jmath^2 \in Z(D) = \mathbb{R}$. Multiplying $\jmath$ by a suitable real number, we obtain $\jmath^2 = \pm 1$. However, the case $\jmath^2 = 1$ is impossible because then $(\jmath + 1)(\jmath - 1) = 0$. Therefore,

$$\imath^2 = \jmath^2 = -1, \qquad \imath\jmath = -\jmath\imath,$$

implying $D \simeq \mathbb{H}$. □

Theorem 11.129 (Wedderburn's Theorem). *Every finite division ring is commutative, i.e., is a field.*

Proof. Let D be a finite division ring with the center K. Then D is a finite-dimensional central division algebra over K. The first statement of Theorem 11.127 implies that all maximal subfields of D contain the same number of elements, hence are isomorphic. The second statement implies that they can be obtained from one another by inner automorphisms of D. Furthermore, every element $\alpha \in D$ is contained in the subfield $K[\alpha]$, hence, in a maximal subfield.

Let F be a maximal subfield of D. The above implies that the group D^* is covered by subgroups conjugate to F^*. The number of these subgroups equals $[D^* : N(F^*)]$ and, at any rate, does not exceed $[D^* : F^*]$. Therefore,

$$|D^*| \leq |F^*|[D^* : F^*] = |D^*|.$$

However, the equality is impossible here for the simple reason that all these subgroups contain the identity. The only exception is the trivial case, when $D^* = F^*$, thus, $D = F(= K)$. □

Unlike the case of $\mathbb{R}$ and the finite fields, there exist central division algebras of any degree over many other fields, for instance, over $\mathbb{Q}$.

Example 11.130. Let $F = \mathbb{Q}(\theta)$, where θ is a root of the irreducible polynomial

$$f = t^3 - 3t + 1.$$

The discriminant of f equals $81 = 9^2$; therefore, F is a Galois extension of $\mathbb{Q}$ of degree 3 (see Example 10.80). Let σ be a generator of the group $\operatorname{Gal} F/\mathbb{Q}$. Consider formal expressions

$$a_0 + a_1 s + a_2 s^2, \qquad a_0, a_1, a_2 \in F. \tag{11.45}$$

Define the multiplication of such formal expressions using the distributive laws, the associativity law, and the relations

$$s^3 = 2, \qquad sa = \sigma(a)s, \qquad a \in F.$$

We obtain a 9-dimensional noncommutative algebra D over $\mathbb{Q}$ containing the field F as a subalgebra. Let us prove that D is a division algebra.

The algebra D can be presented by matrices of order 3 over F. Namely, we map every element $a \in F$ to the matrix

$$T(a) = \begin{pmatrix} a & 0 & 0 \\ 0 & \sigma(a) & 0 \\ 0 & 0 & \sigma^2(a) \end{pmatrix}$$

and thus obtain an embedding of the field F into $\mathrm{L}_3(F)$ (as a $\mathbb{Q}$-subalgebra). Now, consider the matrix

$$S = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \end{pmatrix}.$$

It is easy to check that

$$S^3 = 2E, \qquad ST(a) = T(\sigma(a))S.$$

Therefore, matrices of the form

$$T(a_0) + T(a_1)S + T(a_2)S^2 = \begin{pmatrix} a_0 & a_1 & a_2 \\ 2\sigma(a_2) & \sigma(a_0) & \sigma(a_1) \\ 2\sigma^2(a_1) & 2\sigma^2(a_2) & \sigma^2(a_0) \end{pmatrix}, \tag{11.46}$$

where $a_0, a_1, a_2 \in F$, form a $\mathbb{Q}$-subalgebra of the algebra $\mathrm{L}_3(F)$ isomorphic to D. It consists of all matrices $A \in \mathrm{L}_3(F)$ satisfying the condition

$$SAS^{-1} = \sigma(A), \tag{11.47}$$

where $\sigma(A)$ denotes the matrix obtained from A by applying σ to all its entries.

Obviously, if a matrix A satisfying condition (11.47) is nonsingular, the matrix A^{-1} also satisfies condition (11.47). Thus, to prove that D is a division algebra, it suffices to check that every nonzero matrix of the form (11.46) is nonsingular.

To prove the latter statement, we apply reduction modulo 2. Let O be the ring of integers of the field F. Obviously, O is invariant with respect to the group $\operatorname{Gal} F/\mathbb{Q}$ so that if $a \in O$, then $T(a) \in \mathrm{L}_3(O)$. Furthermore, O contains the subring $\mathbb{Z}[\theta] = \{u_0 + u_1\theta + u_2\theta^2 \colon u_0, u_1, u_2 \in \mathbb{Z}\}$. Since the polynomial

$$[f]_2 = t^3 + t + 1 \in \mathbb{Z}_2[t]$$

is irreducible over $\mathbb{Z}_2$, the quotient ring

$$\mathbb{Z}[\theta]/2\mathbb{Z}[\theta] \simeq \mathbb{Z}_2[t]/[f]_2\mathbb{Z}_2[t]$$

is the field with eight elements. There exists a natural homomorphism

$$\mathbb{Z}[\theta]/2\mathbb{Z}[\theta] \to O/2O. \tag{11.48}$$

Since the additive group of O is isomorphic to $\mathbb{Z}^3$, $|O/2O| = 8$. It follows that the homomorphism (11.48) is, in fact, an isomorphism; thus the ring $O/2O$ is also a field.

By multiplying the element (11.45) in D by a suitable rational number, we can obtain the situation where the numbers a_0, a_1, a_2 belong to O and at least one of them does not belong to $2O$. If $a_0 \in 2O$ but $a_1 \notin 2O$, then, multiplying by $s^{-1} = s^2/2$, we obtain $a_0 \notin 2O$. If $a_0, a_1 \in 2O$ but $a_2 \notin 2O$, we obtain the same result when multiplying by $s^{-2} = s/2$. Therefore, it suffices to prove that elements (11.45) in D such that

$$a_0, a_1, a_2 \in O, \qquad a_0 \notin 2O$$

are invertible.

With these conditions, all elements of matrix (11.46) belong to O and the reduction of this matrix modulo 2 is a strictly triangular matrix over the field $O/2O$. The determinant of the latter matrix is nonzero. It follows that the determinant of matrix (11.46) is nonzero; this completes the proof.

Since $\dim D = 9$, D is a central division algebra of degree 3.

Division algebras appear naturally in the study of irreducible linear representations over fields that are not algebraically closed. Namely, let $R : X \to \mathrm{L}(V)$ be a nontrivial irreducible representation of a set X over a field K. Consider the set D of endomorphisms of R. Obviously, this is a subalgebra in $\mathrm{L}(V)$. By Theorem 11.12, every nonzero endomorphism of R is invertible. Therefore, D is a division algebra.

The space V can be viewed as a D-module or, as it is sometimes called, a vector space over D. (It is easy to see that every finitely generated module over a division algebra has a basis—just as a vector space over a field.) The set $R(X)$ is contained in the algebra $\mathrm{L}_D(V)$ of linear transformations of this vector space (which is isomorphic to an algebra of matrices over D). It is

not difficult to obtain the following generalization of Theorem 11.31: *the subalgebra of* $\mathrm{L}(V)$ *generated by the set* $R(X)$ *coincides with* $\mathrm{L}_D(V)$.

This implies, in turn, the following generalization of Theorem 11.67: *every nontrivial simple finite-dimensional associative algebra over a field* K *is isomorphic to the algebra of all linear representations of a vector space over a division algebra over* K.

In particular, by the Frobenius theorem, irreducible real linear representations separate into three types with $D = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, respectively.

Exercise 11.131. Prove that the representations of these three types are described by whether their complexifications remain irreducible, split into a sum of two nonisomorphic irreducible representations, or split into a sum of two isomorphic irreducible representations, respectively.

If we do not require associativity, the definition of a division algebra should be changed. An algebra D (not necessarily associative) is called a *division algebra* if for every $a, b \in D$, $a \neq 0$, each of the equations $ax = b$ and $ya = b$ has a solution.

Exercise 11.132. Prove that for associative algebras, this definition is equivalent to the previous one.

Exercise 11.133. Prove that the statement of Exercise 11.118 also holds for nonassociative algebras.

Dropping the associativity condition, we can produce new interesting examples of division algebras, even over $\mathbb{R}$.

Example 11.134. We present here a construction of the octonion algebra $\mathbb{O}$, which is the most interesting example of a nonassociative division algebra over $\mathbb{R}$.

Let V be a three-dimensional vector space over the field $\mathbb{Z}_2$. Consider the 8-dimensional algebra $\mathbb{O}$ over $\mathbb{R}$ with the basis $\{e_a \colon a \in V\}$ and the multiplication table

$$e_a e_b = \varepsilon(a, b) e_{a+b},$$

where the coefficients $\varepsilon(a, b)$ are equal to ± 1 and are determined according to the following rules:

(i) $\varepsilon(0, b) = \varepsilon(a, 0) = 1$, so that $e_0 = 1$ is the unity of the algebra $\mathbb{O}$;

(ii) $\varepsilon(a, a) = -1$ for $a \neq 0$, so that the square of every "imaginary unit" e_a, $a \neq 0$, equals -1;

(iii) $\varepsilon(a, b) = -\varepsilon(b, a)$ for $a, b \neq 0$, $a \neq b$, so that the imaginary units anticommute;

(iv) $\varepsilon(a,b) = \varepsilon(b,c) = \varepsilon(c,a)$ for $a,b,c \neq 0$, $a+b+c=0$, so that any two imaginary units generate a subalgebra isomorphic to the algebra of quaternions;

(v) $\varepsilon(a,b)\varepsilon(b,c)\varepsilon(c,d)\varepsilon(d,a) = -1$ for distinct $a,b,c,d \neq 0$, $a+b+c+d=0$.

Nonzero vectors of V can be viewed as points of the projective plane PV over the field $\mathbb{Z}_2$. With this interpretation, condition (iv) pertains to triples of points on one line and condition (v), to quadruples of points neither three of which lie on one line.

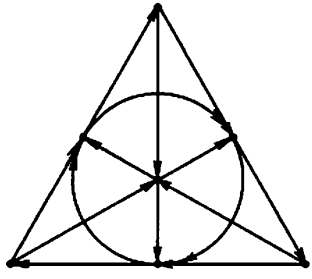


Figure 11.1

An example of a choice of coefficients $\varepsilon(a,b)$, $a,b \neq 0$, $a \neq b$, satisfying conditions (iii)-(v) is given in Figure 11.1, where the lines of PV are denoted by six lines and one circle, and the arrow from a point a to a point b stands for $\varepsilon(a,b) = 1$. (Other coefficients can be reconstructed via rules (iii) and (iv).) It is not difficult to show that any other choice of coefficients $\varepsilon(a,b)$ reduces to this one by multiplication of some imaginary units by -1.

The algebra $\mathbb{O}$ constructed above is called the *octonion algebra* or the *Cayley algebra*.

As in the case of quaternions, the linear map $u \mapsto \overline{u}$ that preserves the unity and multiplies all imaginary units by -1 is an antiautomorphism of the algebra $\mathbb{O}$. The element $N(u) = u\overline{u}$, called the *norm* of the octonion u, lies in $\mathbb{R}$ and equals the sum of squares of its coordinates. If $u \neq 0$, then $N(u) \neq 0$ and

$$u^{-1} = N(u)^{-1}\overline{u} \tag{11.49}$$

is the inverse of u. Conditions (i)-(iii) suffice to establish all these properties; however, since $\mathbb{O}$ is not associative, they do not insure that it is a division algebra.

When associativity is not valid, the weaker property of alternativity is sometimes sufficient. An algebra is called *alternative* if the *associator*

$$[uvw] = (uv)w - u(vw)$$

of any three of its elements is skew-symmetric in u, v, w. In particular, it follows that when either two of the elements are the same, associativity holds.

Exercise 11.135. Prove that in an alternative algebra, a subalgebra generated by any two elements is associative.

Let us show that the algebra $\mathbb{O}$ is alternative. By linearity, it suffices to check that the associator of any three basis vectors e_a, e_b, e_c, $a, b, c \in V$, is skew-symmetric. If a, b, c are linearly dependent, e_a, e_b, e_c lie in an associative algebra and there is nothing to check. Assume that a, b, c are linearly independent. Let us prove that not just the associator but also the products $(e_a e_b)e_c$, $e_a(e_b e_c)$ are skew-symmetric in a, b, c. For instance, consider the first one. Clearly, it is skew-symmetric in a and b. Thus, it suffices to check that it is skew-symmetric in b and c. Conditions (iii) and (iv) imply that

$$(e_a e_b)e_c = \varepsilon(a, b)\varepsilon(a + b, c)e_{a+b+c} = -\varepsilon(a, b)\varepsilon(a + b + c, c)e_{a+b+c}$$

and, similarly,

$$(e_a e_c)e_b = \varepsilon(a, c)\varepsilon(a + c, b)e_{a+b+c} = -\varepsilon(c, a)\varepsilon(b, a + b + c)e_{a+b+c}.$$

But condition (v) implies that

$$\varepsilon(a, b)\varepsilon(a + b + c, c) = -\varepsilon(c, a)\varepsilon(b, a + b + c).$$

It follows that

$$(e_a e_b)e_c = -(e_a e_c)e_b.$$

The skew-symmetricity of the second product is shown similarly.

Formula (11.47) for the inverse element in $\mathbb{O}$ implies that

$$u^{-1} \in \langle 1, u \rangle;$$

thus if two elements in the product of three are inverses of each other, then, just as when they coincide, we have associativity. Therefore, for $u \neq 0$, the element $u^{-1}v$ is a solution of the equation $ux = v$ and the element vu^{-1} is a solution of the equation $yu = v$. Hence, $\mathbb{O}$ is a division algebra.

There exists the following theorem: *every alternative finite-dimensional division algebra over* $\mathbb{R}$ *is isomorphic to either* $\mathbb{R}$, *or* $\mathbb{C}$, *or* $\mathbb{H}$, *or* $\mathbb{O}$.

Lie Groups

The definition of a Lie group is similar to that of a topological group. Namely, a Lie group is a group G that possesses a structure of a differentiable manifold such that the group operations

$$\mu\colon G \times G \to G, \qquad (x, y) \mapsto xy,$$
$$\iota\colon G \to G, \qquad x \mapsto x^{-1}$$

are differentiable. In other words, (local) coordinates of a product are differentiable functions of (local) coordinates of the factors, and the coordinates of the inverse of an element are differentiable functions of the coordinates of this element. A Lie group can be viewed as a topological one but its structure is richer.

In the above definition, we can consider either real or complex manifolds. Respectively, we thus come to the definitions of a real and a complex Lie group. To cover both cases at once, we denote the base field (respectively, $\mathbb{R}$ or $\mathbb{C}$) by K.

Examples of Lie groups are the additive and the multiplicative group of K and the group of nonsingular matrices $\mathrm{GL}_n(K)$ (or, in geometric terms, the group $\mathrm{GL}(V)$ of invertible linear transformations of an n-dimensional vector space V over the field K). In the latter case, the coordinates are the matrix entries.

The theory of Lie groups encompasses algebra, analysis, and geometry. For this reason, its methods and concepts play an important role in the majority of areas of mathematics and theoretical physics.

12.1. Definition and Simple Properties of Lie Groups

We will not use the general definition of a Lie group that we supplied above. To simplify the presentation, we restrict ourselves to linear Lie groups: the ones that are subgroups of $\mathrm{GL}_n(K)$. Actually, almost every Lie group can be presented as a linear Lie group.

By a differentiable function, we mean a function that has continuous partial derivatives of the first order in all of its domain. However, in all examples below, the functions are analytic and, in the case $K = \mathbb{C}$, it is well known that every differentiable function is analytic.

Recall that a subset $M \subset K^n$ is called a d-dimensional *differentiable manifold* if for every point p, it can be defined in some neighborhood of p by a system of equations

$$f_i(x_1, \ldots, x_n) = 0, \qquad i = 1, \ldots, m, \tag{12.1}$$

where $m = n - d$ and $f_1, \ldots, f_m$ are differentiable functions such that the rank of their Jacobian matrix at p equals m.

Remark 12.1. Any open subset of K^n is determined locally by an empty system of equations and so, by definition, it is an n-dimensional differentiable manifold. On the other hand, every discrete subset is defined locally by the system of equations of the form $x_i = c_i$, $i = 1, \ldots, n$, hence, it is a zero-dimensional differentiable manifold.

The restriction on the rank of the Jacobian matrix of the functions $f_1, \ldots, f_m$ means that it has a minor of order m that is nonzero at the point p. Without loss of generality, assume that

$$\begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_m} \end{vmatrix} (p) \neq 0. \tag{12.2}$$

Then, the implicit function theorem implies that, as in the case of a system of linear equations, we can regard "free" variables $x_{m+1}, \ldots, x_n$ as parameters of a point of M in some neighborhood of p; the "principal" variables $x_1, \ldots, x_m$ can be expressed in terms of the free ones via differentiable functions:

$$\begin{cases} x_1 = \varphi_1(x_{m+1}, \ldots, x_n), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_m = \varphi_m(x_{m+1}, \ldots, x_n). \end{cases} \tag{12.3}$$

More precisely, let $p_1, \ldots, p_n$ be coordinates of p. Then there exists a neighborhood U of the point $(p_1, \ldots, p_m)$ in the space K^m of principal variables and a neighborhood V of the point $(p_{m+1}, \ldots, p_n)$ in the space

K^d of free variables such that the intersection $M \cap (U \times V)$ is the graph of the differentiable map $\varphi\colon V \to U$ defined by equations (12.3), i.e., a point $(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) \in U \times V$ belongs to M if and only if conditions (12.3) hold.

The tangent space $T_p(M)$ at a point p of the manifold M defined by equations (12.1) consists of vectors $(dx_1, \ldots, dx_n) \in K^n$ that satisfy the system of homogeneous linear equations obtained by differentiating equations (12.1) at p:

$$df_i(p) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(p) dx_j = 0, \qquad i = 1, \ldots, m. \tag{12.4}$$

Observe that the restriction on the rank of the Jacobian matrix of the functions $f_1, \ldots, f_m$ is equivalent to the condition that the dimension of the space of solutions of the system (12.4) be equal to $n - m$. Sometimes checking the latter is easier than calculating the rank of the Jacobian matrix. If condition (12.2) holds, we can take $dx_{m+1}, \ldots, dx_n$ as free variables in system (12.4).

The space $T_p(M)$ can be described as the set of all tangent vectors to the curves on M passing through p. This implies, in particular, that the tangent space does not depend on the choice of the system of equations that defines M in a neighborhood of p.

Let us emphasize that here we view the tangent space $T_p(M)$ as a subspace of the vector space K^n and not as the parallel plane passing through the point p.

Definition 12.2. A *linear Lie group* is a subgroup G of $\mathrm{GL}_n(K)$ that is a differentiable manifold in the space $\mathrm{L}_n(K)$ of all matrices.

Since every subgroup $G \subset \mathrm{GL}_n(K)$ is invariant under multiplications by its elements which are linear transformations of the space $\mathrm{L}_n(K)$, it suffices to check that G satisfies the definition of a differentiable manifold at one point of G, say, at the identity e. (The identity of a linear group is the identity matrix E.)

Exercise 12.3. If G is a linear Lie group, then

$$T_g(G) = T_e(G)g$$

for every matrix $g \in G$.

In what follows, by a Lie group we understand a linear Lie group and denote the tangent space of a Lie group G at the identity simply by $T(G)$.

Example 12.4. The group $\mathrm{GL}_n(K)$ is an open subset of $\mathrm{L}_n(K)$, thus, it is an n^2-dimensional Lie group.

Example 12.5. The group $\mathrm{SL}_n(K)$ is an (n^2-1)-dimensional Lie group. Indeed, it is defined by the equation $\det g = 1$ and, differentiating it at the identity, we obtain the linear equation

$$d\det g = \operatorname{tr} dg = 0,$$

which defines an (n^2-1)-dimensional space of matrices with zero trace.

In Section 6.5, we defined the derivative of a matrix function of one variable. Likewise, we can define partial derivatives of a matrix function of several variables. Define the differential of a matrix function $\Phi = \Phi(x_1, \ldots, x_n)$ by the formula

$$d\Phi = \sum_{i=1}^{n} \frac{\partial \Phi}{\partial x_i} dx_i.$$

Passing to the matrix entries, it is easy to prove that

$$d(\Phi + \Psi) = d\Phi + d\Psi, \tag{12.5}$$

$$d(\Phi\Psi) = (d\Phi)\Psi + \Phi(d\Psi). \tag{12.6}$$

Using the latter formula in the calculation of the differential $d(\Phi\Phi^{-1})$ (which is zero), we obtain

$$d(\Phi^{-1}) = -\Phi^{-1}(d\Phi)\Phi^{-1}.$$

Example 12.6. The orthogonal group $\mathrm{O}_n(K)$ is defined by the matrix equation $gg^{\top} = E$. Due to its obvious symmetry, we regard this equation as a system of $\frac{n(n+1)}{2}$ equations in matrix entries. Differentiating it at the identity, we obtain the linear equation

$$d(gg^{\top}) = dg + (dg)^{\top} = 0,$$

which defines the $\frac{n(n-1)}{2}$-dimensional space of skew-symmetric matrices. Since

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

$\mathrm{O}_n(K)$ is an $\frac{n(n-1)}{2}$-dimensional Lie group over K. Note that the group $\mathrm{O}_n(\mathbb{R})$ is usually denoted simply by O_n.

Exercise 12.7. Prove that the pseudo-orthogonal group $\mathrm{O}_{k,l}$, $k + l = n$, (see Section 7.3) is an $\frac{n(n-1)}{2}$-dimensional real Lie group.

Exercise 12.8. For an even $n = 2m$, consider the group of linear transformations of K^n that preserve the nonsingular skew-symmetric bilinear function

$$\alpha(x, y) = \sum_{i=1}^{m} (x_i y_{m+i} - x_{m+i} y_i).$$

This group is called *symplectic* and is denoted $\mathrm{Sp}_n(K)$. Prove that $\mathrm{Sp}_n(K)$ is an $\frac{n(n+1)}{2}$-dimensional Lie group.

Example 12.9. The group $\mathrm{B}_n(K)$ of nonsingular triangular matrices is an open subset of the space of all triangular matrices, hence, it is an $\frac{n(n+1)}{2}$-dimensional Lie group.

Example 12.10. Every discrete (and, in particular, finite) subgroup of $\mathrm{GL}_n(K)$ is a zero-dimensional Lie group.

We can also consider real Lie groups that consist of complex matrices: these are understood as subgroups of the group $\mathrm{GL}_n(\mathbb{C})$ that are differentiable manifolds in the space $\mathrm{L}_n(\mathbb{C})$ viewed as a $2n^2$-dimensional real vector space.

Example 12.11. The group U_n is determined in $\mathrm{L}_n(\mathbb{C})$ by the matrix equation

$$gg^* = E \tag{12.7}$$

(here $g^* = \overline{g}^{\mathsf{T}}$). This equation can be regarded as a system of n^2 real equations in real and imaginary parts of the entries x_{ij} of the matrix g:

$$\sum_k |x_{ik}|^2 = 1, \qquad i = 1, \ldots, n,$$

$$\Re \sum_k x_{ik}\overline{x}_{jk} = \Im \sum_k x_{ik}\overline{x}_{jk} = 0, \qquad i, j = 1, \ldots, n;\ i < j.$$

Differentiating equation (12.7) at the identity, we obtain the linear equation

$$dg + (dg)^* = 0,$$

which defines the n^2-dimensional subspace of skew-Hermitian matrices. The group U_n is an n^2-dimensional real Lie group, since $2n^2 - n^2 = n^2$.

Example 12.12. The group $\mathrm{SU}_n = \mathrm{U}_n \cap \mathrm{SL}_n(\mathbb{C})$ is an (n^2-1)-dimensional real Lie group (prove this). Its tangent space at the identity consists of the skew-Hermitian matrices with zero trace.

Proposition 12.13. *Every Lie group $G \subset \mathrm{GL}_n(K)$ is closed in $\mathrm{GL}_n(K)$.*

Proof. Let $\overline{G}$ be the closure of G in $\mathrm{GL}_n(K)$. By continuity, $\overline{G}$ is a subgroup as well, while the definition of a differentiable manifold implies that G is open in $\overline{G}$. Now consider $g \in \overline{G}$. The coset gG is open in $\overline{G}$, hence, it intersects G. But then $gG = G$ and, in particular, $g \in G$. □

The main approach in the theory of Lie groups is to pass from considering a group G to considering its tangent space $T_e(G) = T(G)$ (which, as we will see later, has an algebra structure). However, if, for instance, the group G is discrete, then its tangent space is zero, thus carries no information about the structure of G. In the general case, a Lie group G is well controlled by its tangent space at the identity only if G is connected.

Recall that a topological space is called *connected* if it cannot be decomposed into a union of two disjoint proper closed subsets. A union of two intersecting connected subsets of a topological space M is connected. It follows that the relation "$x \sim y$ if x and y lie in the same connected subset" is an equivalence relation on M. Classes of this equivalence are called *connected components* of the space M.

If M is a differentiable manifold, then every point of M has a connected neighborhood (say, homeomorphic to a ball). This implies that the connected components of M are open in M. At the same time, they are closed in M as each of them is the complement of the union of the others.

The connected component of a Lie group G that contains the identity is denoted G°.

Proposition 12.14. *G° is a normal subgroup of G, and all other connected components are the cosets of G°.*

Proof. The left or right multiplication by an element $g \in G$ is a homeomorphism of the topological space G onto itself and, thus, can only permute its connected components. Therefore, $gG^\circ = G^\circ g$ is the connected component that contains g. In particular, if $g \in G^\circ$, then $gG^\circ = G^\circ$. This implies that G° is closed with respect to multiplication.

Similarly, passing to the inverse element is a homeomorphism of the topological space G into itself and can only permute its connected components. Since $(G^\circ)^{-1}$ contains the identity, $(G^\circ)^{-1} = G^\circ$. Therefore, G° is a subgroup. The rest has already been shown before. □

Example 12.15. We will prove here that the group $\mathrm{SL}_n(K)$ is connected. For fixed different i and j, the matrices of the form $E + cE_{ij}$, $c \in K$, make up a connected subset that contains the identity. Therefore, all elementary matrices of the first type lie in $\mathrm{SL}_n(K)^\circ$. However, we know that they generate $\mathrm{SL}_n(K)$ (see Section 10.2). Thus, $\mathrm{SL}_n(K)^\circ = \mathrm{SL}_n(K)$.

Example 12.16. One can prove similarly (do it!) that the group $\mathrm{GL}_n(\mathbb{C})$ is connected, while the group $\mathrm{GL}_n(\mathbb{R})$ contains two connected components, one of which is the group of matrices with positive determinant.

Example 12.17. Let us show that the group O_n consists of two connected components, one of which is SO_n (the other consists of orthogonal matrices with determinant -1). Let $n = 2m$ or $2m+1$. Consider orthogonal matrices

of the form

$$(12.8)\qquad \begin{pmatrix} \Pi(\varphi_1) & & 0 \\ & \ddots & \\ & & \Pi(\varphi_m) \\ 0 & & \begin{matrix} \\ (1) \end{matrix} \end{pmatrix}, \qquad \varphi_1, \ldots, \varphi_m \in \mathbb{R},$$

where the block in parentheses (of order 1) appears if $n = 2m + 1$. These matrices form a connected subset—it is homeomorphic to a direct product of m circles, i.e., to an m-dimensional torus. Since this subset contains the identity, it lies in O_n°. However, we know that every matrix in SO_n is conjugate in O_n to a matrix of the form (12.8) (see Section 6.3). Therefore, $\mathrm{O}_n^\circ \supset \mathrm{SO}_n$. On the other hand, since O_n is a union of two cosets of SO_n, each obviously being closed, we obtain $\mathrm{O}_n^\circ = \mathrm{SO}_n$.

Example 12.18. Similarly, one shows that the groups U_n and SU_n are connected.

Exercise 12.19. Prove that the group $\mathrm{SO}_{n,1}$ consists of two connected components and that $\mathrm{SO}_{n,1}^\circ$ is the subgroup of transformations preserving both connected components of the hyperboloid

$$x_1^2 + \cdots + x_n^2 - x_{n+1}^2 = -1.$$

(*Hint*: prove that every transformation in the group $\mathrm{SO}_{n,1}$ that preserves both connected components of the above hyperboloid is a product of a hyperbolic rotation (Lorentz transformation) in a two-dimensional subspace that contains the basis vector e_{n+1} and a transformation from SO_n which leaves e_{n+1} invariant.)

Proposition 12.20. *A connected Lie group is generated by any neighborhood of the identity.*

Proof. Let U be a neighborhood of the identity in the Lie group G. Denote by $\widetilde{G}$ the subgroup generated by U. For every $g \in G$, the subset gU, which is a neighborhood of g in G, is contained in the coset $g\widetilde{G}$. This shows that all cosets of $\widetilde{G}$ are open in G. On the other hand, they are closed in G since each of them is the complement of the union of the others. Therefore, if G is connected, there exists only one such coset, i.e., $G = \widetilde{G}$. □

Remark 12.21. In this section, we chose the language of matrices. But it is clear that instead of matrices, we could speak about linear transformations defined by them. For instance, one can speak about the Lie group $\mathrm{GL}(V)$ of nonsingular linear transformations of an n-dimensional vector space V

over K, about the Lie group $\mathrm{O}(V)$ of orthogonal transformations of an n-dimensional Euclidean space V, and so on. In the sequel, when useful, we will sometimes switch to this language of linear transformations.

Remark 12.22. By Lemma 7.105, the group $\mathrm{GA}(S)$ of affine transformations of an n-dimensional affine space S embeds naturally into the group $\mathrm{GL}(V)$ of linear transformation of an $(n+1)$-dimensional vector space V. In a suitable basis, the image of this embedding consists of matrices of the form

$$\left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \vdots \\ a_{n1} & \dots & a_{nn} & b_n \\ \hline 0 & \dots & 0 & 1 \end{array}\right),$$

where the matrix $A = (a_{ij})$ is nonsingular. Thus, the group $GA(S)$, as well as some of its subgroups such as the group of motions of an n-dimensional Euclidean space, can be regarded as a linear Lie group.

12.2. The Exponential Map

A Lie group G and its tangent space $T(G)$ are related via the exponential map.

We defined the exponential of a matrix in Section 6.5. It gives rise to a map

$$\exp\colon \mathrm{L}_n(K) \to \mathrm{GL}_n(K), \tag{12.9}$$

called the *exponential map*.

The definition of the exponential of a matrix implies that $\exp 0 = E$ and

$$\exp X = E + X + o(\|X\|). \tag{12.10}$$

This shows that the differential of the exponential map at 0 is the identity map. In particular, the Jacobian matrix of the exponential map at 0 is nonsingular, hence by the implicit function theorem, the map exp yields a diffeomorphism from a neighborhood of 0 in the space $\mathrm{L}_n(K)$ to a neighborhood of the identity in the group $\mathrm{GL}_n(K)$. The inverse map (defined in a neighborhood of the identity of $\mathrm{GL}_n(K)$) is denoted log.

Remark 12.23. The map log is determined by the series

$$\log(E+X) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{X^n}{n},$$

which is absolutely convergent for $\|X\| < 1$.

Remark 12.24. It can be shown that the map exp yields a diffeomorphism from the open subset of $\mathrm{L}_n(K)$ that consists of matrices with complex eigenvalues λ such that $|\Im\lambda| < \pi$ to the open subset of $\mathrm{GL}_n(K)$ that consists of matrices without negative eigenvalues. The map (12.9) is not a diffeomorphism on the entire $\mathrm{L}_n(K)$.

The following proposition generalizes the well-known formula

$$e^a = \lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n.$$

Proposition 12.25. *Let $g(t)$, $|t| < c$, be a differentiable curve on the group $\mathrm{GL}_n(K)$ such that*

$$g(0) = E, \qquad g'(0) = A. \tag{12.11}$$

Then

$$\exp A = \lim_{n\to\infty} g\left(\frac{1}{n}\right)^n. \tag{12.12}$$

Proof. At $t = 0$, the curve $\log g(t)$ has the same tangent vector as $g(t)$, i.e., A. This means that

$$\log g(t) = tA + o(t),$$

hence

$$g(t) = \exp(tA + o(t)).$$

In particular,

$$g\left(\frac{1}{n}\right) = \exp\left(\frac{A}{n} + o\left(\frac{1}{n}\right)\right).$$

Taking the nth power of the latter identity, we obtain

$$g\left(\frac{1}{n}\right)^n = \exp(A + o(1)).$$

This implies (12.12). □

Theorem 12.26. *Let $G \subset \mathrm{GL}_n(K)$ be a Lie group. Then*

$$\exp T(G) \subset G. \tag{12.13}$$

Moreover, the map exp *is a diffeomorphism from a neighborhood of* 0 *in the space $T(G)$ to a neighborhood of the identity of G.*

Proof. For any $A \in T(G)$, there exists a curve $g(t)$ on G satisfying condition (12.11). Since the group G is closed in $\mathrm{GL}_n(K)$ (see Proposition 12.13), (12.12) implies that $\exp A \in G$.

To prove the second assertion of the theorem, rewrite the map

$$\exp\colon T(G) \to G \tag{12.14}$$

in local coordinates of neighborhoods of the identity in G and zero in $T(G)$. As such, we can choose the free variables of the system of equations defining G near the identity and the corresponding free variables of the system of homogeneous linear equations defining $T(G)$. (Recall that here the variables are matrix entries.) We obtain a differentiable map f from a neighborhood of zero of the space K^d (where $d = \dim G$) to this space. Formula (12.10) implies that the differential of f at zero is the identity map. Hence, by the implicit function theorem, the map f is a diffeomorphism from a (possibly smaller) neighborhood of zero in K^d to a region in this space. Passing back to the map exp, we obtain the second assertion of this theorem. □

Example 12.27. When $G = \mathrm{SL}_n(K)$, condition (12.13) says that if $\operatorname{tr} A = 0$, then $\det \exp A = 1$ (see Example 12.5).

Example 12.28. When $G = \mathrm{O}_n$ (respectively, U_n), condition (12.13) says that the exponent of a skew-symmetric (respectively, skew-Hermitian) matrix is an orthogonal (respectively, unitary) matrix (see Examples 12.6 and 12.11). This can also be checked directly (try it).

Theorem 12.29. *A connected Lie group is uniquely determined by its tangent space at the identity.*

Proof. Theorem 12.26 and Proposition 12.20 imply that a connected Lie group $G \subset \mathrm{GL}_n(K)$ coincides with the subgroup generated by the set $\exp T(G)$. □

Notice that the above theorem does not claim the existence of a Lie group with a given tangent space. In fact, very few subspaces of the space of matrices are tangent spaces to Lie groups. A necessary condition for this is provided in the next section.

Remark 12.30. Generally speaking, $G \neq \exp T(G)$, i.e., the map (12.14) does not have to be surjective (even for a connected Lie group G). For example, the tangent space $T(G)$ of the group $G = \mathrm{SL}_2(\mathbb{R})$ consists of matrices with zero trace. Such a matrix has complex eigenvalues λ, $-\lambda$, where either $\lambda \in \mathbb{R}$ or $\lambda \in i\mathbb{R}$. In either case,

$$\operatorname{tr} \exp A = e^{\lambda} + e^{-\lambda} \geq -2.$$

Therefore, matrices $g \in G$ with $\operatorname{tr} g < -2$ do not belong to $\exp T(G)$ (e.g., the matrix $\begin{pmatrix} -2 & 0 \\ 0 & -1/2 \end{pmatrix}$).

Let $G \subset \mathrm{GL}_n(K)$, $H \subset \mathrm{GL}_m(K)$ be Lie groups. A map $f\colon G \to H$ is called a *Lie group homomorphism* if it is a group homomorphism and is differentiable, i.e., the entries of the matrix $f(g)$ are differentiable functions of the entries of $g \in G$. The differential of a homomorphism f at the

identity is a linear map from the space $T(G)$ to the space $T(H)$. We will simply denote it df without stating explicitly at which point the differential is taken.

Theorem 12.31. *Let $f\colon G \to H$ be a Lie group homomorphism. Then*

$$f(\exp A) = \exp df(A) \tag{12.15}$$

for any $A \in T(G)$.

Proof. We will use Proposition 12.25. Let $g(t)$ be a curve on the group G satisfying conditions (12.11). Then the curve $h(t) = f(g(t))$ on the group H satisfies the following conditions:

$$h(0) = E, \qquad h'(0) = df(A).$$

Therefore,

$$f(\exp A) = f\left(\lim_{n\to\infty} g\left(\frac{1}{n}\right)^n\right) = \lim_{n\to\infty} h\left(\frac{1}{n}\right)^n = \exp df(A).$$

□

Example 12.32. Applied to the homomorphism

$$\det\colon \mathrm{GL}_n(\mathbb{C}) \to \mathbb{C}^* \ (= \mathrm{GL}_1(\mathbb{C})),$$

formula (12.15) implies that

$$\det \exp A = e^{\operatorname{tr} A}$$

for any matrix $A \in \mathrm{L}_n(\mathbb{C})$ (cf. Example 12.5).

Theorem 12.33. *A homomorphism of a connected Lie group into another Lie group is uniquely determined by its differential at the identity.*

Proof. Let $f\colon G \to H$ be a Lie group homomorphism. Theorems 12.26 and 12.31 imply that if we know df, we can determine $f(g)$ for an element g in some neighborhood U of the identity in G. But if G is connected, then by Proposition 12.20, it is generated by U. Thus, we can determine $f(g)$ for all $g \in G$. □

The above theorem does not claim the existence of a Lie group homomorphism with a given differential. In fact, very few linear maps of tangent spaces are differentials of Lie group homomorphisms. A necessary condition for this is provided in the next section.

Exercise 12.34. Prove that the kernel $\operatorname{Ker} f$ of a Lie group homomorphism $f\colon G \to H$ is a Lie group whose tangent space coincides with $\operatorname{Ker} df$.

Exercise 12.35. In the same notation, prove that if $\operatorname{Im} df = T(H)$ and H is connected, then $\operatorname{Im} f = H$.

12.3. Tangent Lie Algebra and the Adjoint Representation

The matrix

$$[A, B] = AB - BA$$

is called the *commutator* of matrices $A, B \in \mathrm{L}_n(K)$. This should not be confused with the group commutator, $(A, B) = ABA^{-1}B^{-1}$, which is defined for nonsingular matrices; however, these two notions are closely related.

Proposition 12.36. *For any matrices* $A, B \in \mathrm{L}_n(K)$,

$$[A, B] = \frac{\partial^2}{\partial t\, \partial s}(\exp tA, \exp sB)\Big|_{t=s=0}. \tag{12.16}$$

Proof. By differentiating the group commutator

$$(\exp tA, \exp sB) = (\exp tA)(\exp sB)(\exp tA)^{-1}(\exp sB)^{-1}$$

with respect to s at $s = 0$, we obtain

$$\frac{\partial}{\partial s}(\exp tA, \exp sB)\Big|_{s=0} = (\exp tA)B(\exp tA)^{-1} - B.$$

Differentiating this expression with respect to t at $t = 0$, we obtain

$$\frac{\partial^2}{\partial t\, \partial s}(\exp tA, \exp sB)\Big|_{t=s=0} = AB - BA = [A, B].$$

□

Theorem 12.37. *The tangent space* $T(G)$ *of a Lie group* $G \subset \mathrm{GL}_n(K)$ *is closed with respect to taking the commutator, i.e.,*

$$A, B \in T(G) \implies [A, B] \in T(G).$$

Proof. Fix t and consider the curve

$$g(s) = (\exp tA, \exp sB) \in G.$$

Since $g(0) = E$,

$$g'(0) = \frac{\partial}{\partial s}(\exp tA, \exp sB)\Big|_{s=0} \in T(G).$$

Therefore,

$$\frac{\partial^2}{\partial t\, \partial s}(\exp tA, \exp sB)\Big|_{t=s=0} = [A, B] \in T(G).$$

□

A subspace of the space of matrices that is closed with respect to taking the commutator is called a *linear Lie algebra*. Thus, the tangent space $T(G)$ of a (linear) Lie group G is a linear Lie algebra. It is called the *tangent algebra* of the group G.

Example 12.38. For the group $\mathrm{SL}_n(K)$, the above theorem says that if $\operatorname{tr} A = \operatorname{tr} B = 0$, then $\operatorname{tr}[A, B] = 0$. In fact, the latter equality is always true. In other words,

$$\operatorname{tr} AB = \operatorname{tr} BA$$

for any two matrices A, B (see Section 5.3).

Example 12.39. For the group $\mathrm{O}_n(K)$, the above theorem says that the commutator of two skew-symmetric matrices A, B is also a skew-symmetric matrix. This can be easily checked directly:

$$[A, B]^\top = (AB - BA)^\top = B^\top A^\top - A^\top B^\top = BA - AB = -[A, B].$$

The commutator is *anticommutative*, i.e.,

$$[A, B] + [B, A] = 0, \tag{12.17}$$

and satisfies the *Jacobi identity*

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0. \tag{12.18}$$

This identity can be easily checked by direct calculation. It is a consequence of the associativity of matrix multiplication.

Any algebra satisfying identities (12.17) and (12.18) is called a *Lie algebra*. For instance, the space E^3 with the operation of cross product is a Lie algebra (see Example 1.75). The space $\mathrm{L}_n(K)$ is a Lie algebra with respect to taking the commutator. Theorem 12.37 means that the tangent space of any Lie group $G \subset \mathrm{GL}_n(K)$ is a subalgebra of this algebra.

Theorem 12.40. *The differential of a Lie group homomorphism is a homomorphism of their tangent algebras.*

Proof. Let $f\colon G \to H$ be a Lie group homomorphism and let $A, B \in T(G)$. By Theorem 12.31,

$$f((\exp tA, \exp sB)) = (f(\exp tA), f(\exp sB)) = (\exp t\, df(A), \exp s\, df(B)).$$

As in the proof of Theorem 12.37, regard the commutator $(\exp tA, \exp sB)$ with a fixed t as a curve on G parameterized by s. When $s = 0$, this curve passes through the identity. The map df sends the tangent vector of this curve at $s = 0$ to the tangent vector of its image in H. Thus,

$$df\left(\frac{\partial}{\partial s}(\exp tA, \exp sB)\Big|_{s=0}\right) = \frac{\partial}{\partial s}(\exp t\, df(A), \exp s\, df(B))\Big|_{s=0}.$$

By differentiating with respect to t at $t = 0$, we obtain

$$df([A, B]) = [df(A), df(B)]$$

because of (12.16) and the fact that the linear map df and the operator of differentiation commute. □

A (differentiable) homomorphism of a Lie group G into the Lie group $\mathrm{GL}(V)$ is called a *linear representation* of the group G (as a Lie group) on the space V.

For every Lie group G, there exists a remarkable linear representation on the space $T(G)$; it plays an important role in the theory of Lie groups. Here is how we construct it.

Every element $g \in G$ defines an inner automorphism $a(g)$ of the group G as follows:

$$a(g)x = gxg^{-1}, \qquad x \in G. \tag{12.19}$$

This automorphism is the restriction to G of a linear transformation $X \mapsto gXg^{-1}$ on the space $\mathrm{L}_n(K)$. In particular, it is differentiable. Its differential at the identity is denoted $\mathrm{Ad}(g)$ and is called the *adjoint operator* of $g \in G$. The operator $\mathrm{Ad}(g)$ is determined by the same formula as $a(g)$:

$$\mathrm{Ad}(g)X = gXg^{-1}, \qquad X \in T(G).$$

Since $a(xy) = a(x)a(y)$,

$$\mathrm{Ad}(xy) = \mathrm{Ad}(x)\,\mathrm{Ad}(y).$$

(However, this can be easily checked directly.) Furthermore, if $g = (g_{ij})$, $g^{-1} = (\tilde{g}_{ij})$, and $X = (x_{ij})$, then $\mathrm{Ad}(g)X = (y_{ij})$, where

$$y_{ij} = \sum_{k,l} g_k x_{kl} \tilde{g}_{lj}. \tag{12.20}$$

As coordinates in the space $T(G)$, we can take matrix entries such that other matrix entries can be expressed as their linear combinations. Formula (12.20) shows that the matrix entries of $\mathrm{Ad}(g)$ are rational (hence, differentiable) functions of the entries of g. Therefore, the map

$$\mathrm{Ad}\colon G \to \mathrm{GL}(T(G))$$

is a linear representation of the Lie group G on the space $T(G)$. It is called the *adjoint representation* of G.

Exercise 12.41. Prove that if $f\colon G \to H$ is a Lie group homomorphism, then

$$f(\mathrm{Ad}(g)X) = \mathrm{Ad}(f(g))df(X), \qquad g \in G,\ X \in T(G).$$

Exercise 12.42. Prove that the kernel of the adjoint representation of a connected Lie group is its center.

A homomorphism of a Lie algebra L into the Lie algebra $\mathrm{L}(V)$ of linear transformations of a vector space V (with the operation of taking the commutator) is called a *linear representation* of L (as a Lie algebra) on V. By

Theorem 12.40, the differential of a linear representation of a Lie group G is a linear representation of its tangent Lie algebra $T(G)$.

The differential of the adjoint representation Ad of a Lie group G is called the *adjoint representation* of the Lie algebra $T(G)$ and is denoted ad.

Theorem 12.43.

$$\mathrm{ad}(A)X = [A, X], \quad A, X \in T(G).$$

Proof. Let $g(t)$ be a curve on G satisfying conditions (12.11). Then

$$\mathrm{ad}(A) = \frac{\partial}{\partial t}\,\mathrm{Ad}(g(t))\Big|_{t=0}.$$

Therefore, for any $X \in T(G)$,

$$\mathrm{ad}(A)X = \frac{\partial}{\partial t}\,\mathrm{Ad}(g(t))X\Big|_{t=0} = \frac{\partial}{\partial t} g(t)Xg(t)^{-1}\Big|_{t=0} = AX - XA = [A, X].$$

□

The fact that ad is a linear representation of the algebra $T(G)$ means that

$$\mathrm{ad}([A, B]) = [\mathrm{ad}(A), \mathrm{ad}(B)]$$

or, taking the above theorem into account, that

$$[[A, B], C] = [A, [B, C]] - [B, [A, C]]$$

for any $A, B, C \in T(G)$. The last identity is equivalent to the Jacobi identity. This result can be viewed as a conceptual proof of the Jacobi identity for matrix commutators.

On the other hand, this discussion suggests the way to define the adjoint representation ad for any Lie algebra L (not necessarily associated with any Lie group) as

$$\mathrm{ad}(a)x = [a, x], \qquad a, x \in L.$$

Example 12.44. The adjoint representation of the Lie algebra $(E^3, \times)$ is defined as

$$\mathrm{ad}(a)x = a \times x.$$

Since the *triple product* $(a, b, c) = (a \times b, c)$ is skew-symmetric, the operator $\mathrm{ad}(a)$ is also skew-symmetric. This defines a homomorphism of the Lie algebra $(E^3, \times)$ into the Lie algebra $T(\mathrm{SO}_3)$ of skew-symmetric matrices of order 3. It is easy to see that this homomorphism has a trivial kernel. Since both algebras are three-dimensional, we, in fact, constructed an isomorphism.

Exercise 12.45. Determine explicitly the matrices of the adjoint operators of the vectors of an orthonormal basis of E^3 in the same basis.

Exercise 12.46. Prove that for any matrices $A, X \in \mathrm{L}_n(K)$, the following equality holds:

$$(\exp A)X(\exp A)^{-1} = \sum_{k=0}^{\infty} \frac{1}{k!}\underbrace{[A,[A,\ldots,[A}_{k},X]\ldots]].$$

Exercise 12.47. The *center* of a Lie algebra L is a subalgebra

$$Z(L) = \{z \in L\colon [z,u] = 0\ \forall u \in L\}.$$

Prove that the center of a connected Lie group G is a Lie group whose tangent algebra coincides with the center of the Lie algebra $T(G)$. (*Hint*: use Exercises 12.42 and 12.34.)

Example 12.48. Consider the adjoint representation of the Lie group SU_2. The Lie algebra $T(\mathrm{SU}_2)$ consists of skew-symmetric matrices with zero trace, i.e., matrices of the form

$$X = \begin{pmatrix} ix_1 & x_2 + ix_3 \\ -x_2 + ix_3 & -ix_1 \end{pmatrix}. \tag{12.21}$$

Observe that

$$\det X = x_1^2 + x_2^2 + x_3^2.$$

Thus, $\det X$ is a positive definite quadratic function on the space $T(\mathrm{SU}_2)$. Take it as the norm; this turns $T(\mathrm{SU}_2)$ into a (three-dimensional) Euclidean space. Since

$$\det \mathrm{Ad}(g)X = \det gXg^{-1} = \det X,$$

we see that the adjoint operators of the elements of the group SU_2 are orthogonal, i.e., $\mathrm{Ad}(\mathrm{SU}_2) \subset \mathrm{O}_3$. Since the group SU_2 is connected, so is its image. Therefore, $\mathrm{Ad}(\mathrm{SU}_2) \subset \mathrm{SO}_3$.

Furthermore, $\operatorname{Ker}\mathrm{Ad}$ consists of matrices that commute with all matrices of the form (12.21). Using the fact that every matrix commuting with the matrix $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$ is diagonal, it is easy to show that

$$\operatorname{Ker}\mathrm{Ad} = \{\pm E\}.$$

Similarly, we can prove that $\operatorname{Ker}\mathrm{ad} = 0$. Since $\dim \mathrm{SU}_2 = \dim \mathrm{SO}_3 = 3$, we see that $\operatorname{Im}\mathrm{ad} = T(\mathrm{SO}_3)$. It follows from Theorems 12.31 and 12.26 and Proposition 12.20 that

$$\mathrm{Ad}(\mathrm{SU}_2) = \mathrm{SO}_3.$$

Therefore, the adjoint representation is a homomorphism of the group SU_2 onto the group SO_3 with the kernel $\{\pm E\}$.

Note that the group SU_2 consists of the matrices of the form

$$\begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}, \qquad a, b \in \mathbb{C}, \quad |a|^2 + |b|^2 = 1$$

and coincides with the group of quaternions of norm 1 in the matrix model of the algebra $\mathbb{H}$ (see Exercise 1.82). Thus, the group SU_2 is the (three-dimensional) sphere in the four-dimensional Euclidean space $\mathbb{H}$. The group SO_3 is obtained from this sphere by identifying the antipodal points, hence it is the three-dimensional real projective space. The space $T(SU_2)$ coincides with the space of purely imaginary quaternions.

Exercise 12.49. In a similar way, prove that the adjoint representation maps the group $SL_2(\mathbb{R})$ onto the connected component of the group $SO_{2,1}$ (see Exercise 12.19) and that the kernel of this map is $\{\pm E\}$.

Exercise 12.50. Prove that the adjoint representation is a homomorphism of the group $SL_2(\mathbb{C})$ onto the group $SO_3(\mathbb{C})$ with the kernel $\{\pm E\}$.

12.4. Linear Representations of Lie Groups

Linear representations of Lie groups have been studied extensively; here we can only discuss the starting points of this theory. The principal idea is to replace the study of linear representations of a Lie group with that of linear representations of its tangent Lie algebra.

Let G be a Lie group and

$$R\colon G \to \mathrm{GL}(V),$$

a (finite-dimensional) representation of G. Then

$$dR\colon T(G) \to \mathrm{L}(V)$$

is a linear representation of the Lie algebra $T(G)$.

Theorem 12.51. *Every subspace $U \subset V$ that is invariant with respect to G is invariant with respect to $T(G)$. If G is connected, then the converse is also true: every subspace invariant with respect to $T(G)$ is also invariant with respect to G.*

Proof. (i) Let U be an invariant subspace with respect to G and let $A \in T(G)$. Consider a curve $g(t)$ in G satisfying conditions (12.11). Then

$$dR(A) = \frac{\partial}{\partial t} R(g(t))\bigg|_{t=0},$$

hence, for every vector $u \in U$,

$$dR(A)u = \frac{\partial}{\partial t} R(g(t))u\bigg|_{t=0} \in U.$$

(ii) Conversely, let U be an invariant subspace with respect to $T(G)$. By Theorem 12.31, for every $A \in T(G)$,

$$R(\exp A) = \exp dR(A),$$

hence, for every $u \in U$,

$$R(\exp A)u = \sum_{k=0}^{\infty} \frac{1}{k!} dR(A)^k u \in U.$$

Therefore, the subspace U is invariant with respect to $\exp T(G)$. If G is connected, it follows that U is invariant with respect to all of G. □

Thus, if G is connected, the collections of invariant subspaces of a representation R of G and the representation dR of $T(G)$ are the same.

Corollary 12.52. *A linear representation R of a connected Lie group G is irreducible (respectively, completely reducible) if and only if the representation dR of the Lie algebra $T(G)$ is irreducible (respectively, completely reducible).*

Exercise 12.53. Let G be a connected Lie group and H a connected Lie subgroup of G. Prove that the following conditions are equivalent:

(i) H is a normal subgroup of G;

(ii) the subspace $T(H)$ is invariant under the adjoint representation of G;

(iii) $T(H)$ is an ideal of $T(G)$.

A connected Lie group is called *simple* if it does not contain nontrivial connected normal Lie subgroups. A Lie algebra is called *simple* if it does not contain nontrivial ideals.

Exercise 12.54. Prove that if the tangent algebra of a connected Lie group G is simple, then G is simple (as a Lie group). (The converse is also true.)

Exercise 12.55. Prove that the Lie group SO_3 is simple. (In fact, SO_3 has no nontrivial normal subgroups; see Section 10.5.)

The classification of simple Lie groups is as important for the theory of Lie groups as the classification of simple finite groups is for the theory of finite groups. It was obtained in late 19th–early 20th century by W. Killing and E. Cartan (first for complex and then for real Lie groups). This is one of the most amazing mathematical achievements.

It can be shown that the Lie group SO_n is simple whenever $n \geq 5$. However, the Lie group SO_4 is not simple, as the following example demonstrates.

Example 12.56. As we saw in Example 12.48, the group SU_2 can be regarded as the group of quaternions of norm 1. Consider the linear representation R of the Lie group $G = SU_2 \times SU_2$ on the space $\mathbb{H}$ defined as

$$R(p, q)x = pxq^{-1}, \qquad x \in \mathbb{H}.$$

Since

$$N(pxq^{-1}) = N(p)N(x)N(q)^{-1} = N(x)$$

for $p, q \in \mathrm{SU}_2$, we see that $R(G) \subset \mathrm{O}_4$. The connectedness implies that $R(G) \subset \mathrm{SO}_4$. Thus, we have defined a homomorphism

$$R : \mathrm{SU}_2 \times \mathrm{SU}_2 \to \mathrm{SO}_4.$$

If $(p, q) \in \operatorname{Ker} R$, then, in particular, $R(p, q)1 = pq^{-1} = 1$, implying $p = q$. Then, as in Example 12.48, we obtain $p = q = \pm 1$. Since $\dim G = \dim \mathrm{SO}_4 = 6$, it follows that $R(G) = \mathrm{SO}_4$. Therefore,

$$\mathrm{SO}_4 \simeq (\mathrm{SU}_2 \times \mathrm{SU}_2)/\{(E, E), (E, -E)\}.$$

In particular, under the homomorphism R, each factor in the product $\mathrm{SU}_2 \times \mathrm{SU}_2$ is mapped into a connected normal Lie subgroup of SO_4. Hence, SO_4 is not a simple Lie group.

Exercise 12.57. Prove that the Lie group $\mathrm{SL}_n(K)$ is simple for $n \geq 2$.

Complex and real Lie groups are closely related.

Let G be a connected complex Lie group.

Definition 12.58. A connected real Lie subgroup $H \subset G$ is a *real form* of G if

$$T(G) = T(H) \oplus \imath T(H).$$

Remark 12.59. This definition can be extended to nonconnected Lie groups once we require every connected component of G to intersect H.

Example 12.60. The group $\mathrm{SL}_n(\mathbb{R})$ is a real form of the group $\mathrm{SL}_n(\mathbb{C})$.

Example 12.61. The group SU_n is also a real form of the group $\mathrm{SL}_n(\mathbb{C})$, while the group U_n is a real from of the group $\mathrm{GL}_n(\mathbb{C})$. This follows from the fact that every complex matrix can be presented uniquely as a sum of an Hermitian and a skew-Hermitian matrix and that the space of Hermitian matrices can be obtained from the space of skew-Hermitian matrices by the multiplication by $\imath$.

Example 12.62. The group SO_n is a real form of the group $\mathrm{SO}_n(\mathbb{C})$ (which can be proven to be connected).

Theorem 12.63. *Let R: $G \to \mathrm{GL}(V)$ be a complex linear representation of a connected complex Lie group G and let H be a real form of G. Then the collections of invariant subspaces of $R(G)$ and $R(H)$ are the same.*

Proof. By Theorem 12.51, a subspace $U \subset V$ is invariant under G (respectively, H) if and only if it is invariant under $T(G)$ (respectively, $T(H)$). But since

$$dR(T(G)) = dR(T(H)) + \imath dR(T(H)),$$

the invariance of U with respect to $T(G)$ is equivalent to its invariance with respect to $T(H)$. □

This theorem can be used to show complete reducibility of linear representations of certain complex Lie groups.

Definition 12.64. A connected complex Lie group is *reductive* if it has a compact real form.

So, in view of the above examples, the groups $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{C})$, and $\mathrm{SO}_n(\mathbb{C})$ are reductive. It can be shown (though this is not easy) that every noncommutative simple complex Lie group is reductive.

Remark 12.65. It is more natural to exclude connectedness from the definition of a reductive group: then the class of reductive groups includes all finite groups.

Theorem 12.63 and the complete reducibility of linear representations of compact groups proven in Section 11.2 immediately imply

Theorem 12.66. *Every linear representation of a reductive complex Lie group is completely reducible.*

This method of proof is due to H. Weyl and is called the *unitary trick*. It can be used in the proofs of other theorems as well. For instance, it allows us to extend Hilbert's finiteness theorem, which we proved in Section 11.5 for compact groups, to reductive groups.

Using the above theory, let us find all irreducible linear representations of the Lie group $\mathrm{SL}_2(\mathbb{C})$. This example plays a key role in the theory of linear representations of arbitrary simple Lie groups.

Denote the tangent Lie algebra of $\mathrm{SL}_2(\mathbb{C})$ by $\mathrm{sl}_2(\mathbb{C})$. Fix the following basis of this algebra:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{12.22}$$

A direct calculation shows that

$$[H, E_+] = 2E_+, \qquad [H, E_-] = -2E_-, \qquad [E_+, E_-] = H. \tag{12.23}$$

Let $R\colon \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}(V)$ be a linear representation. Let

$$dR(H) = \mathcal{H}, \qquad dR(E_+) = \mathcal{E}_+, \qquad dR(E_-) = \mathcal{E}_-.$$

The operators $\mathcal{H}, \mathcal{E}_+, \mathcal{E}_-$ satisfy the relations

$$[\mathcal{H}, \mathcal{E}_+] = 2\mathcal{E}_+, \qquad [\mathcal{H}, \mathcal{E}_-] = -2\mathcal{E}_-, \qquad [\mathcal{E}_+, \mathcal{E}_-] = \mathcal{H},$$

which follow from relations (12.23).

Lemma 12.67. *Let v be an eigenvector of the operator $\mathcal{H}$ with the eigenvalue λ. Then, if nonzero, the vector $\mathcal{E}_+v$ (respectively, $\mathcal{E}_-v$) is an eigenvector of $\mathcal{H}$ with the eigenvalue $\lambda+2$ (respectively, $\lambda-2$).*

Proof. We have

$$\mathcal{H}\mathcal{E}_+v = \mathcal{E}_+\mathcal{H}v + [\mathcal{H},\mathcal{E}_+]v = \lambda\mathcal{E}_+v + 2\mathcal{E}_+v = (\lambda+2)\mathcal{E}_+v.$$

□

Lemma 12.68. *There exists an eigenvector v_0 of the operator $\mathcal{H}$ such that $\mathcal{E}_+v_0 = 0$.*

Proof. Since the operator $\mathcal{H}$ has only a finite number of eigenvalues, there exists an eigenvalue λ_0 of $\mathcal{H}$ such that λ_0+2 is not an eigenvalue. The corresponding eigenvector v_0 is the one we need. □

Every such vector v_0 is called a *highest weight vector* of the representation R.

Let v_0 be a highest weight vector and λ_0 the corresponding eigenvalue of $\mathcal{H}$. Consider vectors

$$v_k = (\mathcal{E}_-)^k v_0, \qquad k = 0, 1, 2, \ldots .$$

By Lemma 12.67,

$$\mathcal{H}v_k = (\lambda_0 - 2k)v_k.$$

Lemma 12.69. $\mathcal{E}_+v_k = c_kv_{k-1}$, *where* $c_k = k(\lambda_0 - k + 1)$.

Proof. We will prove this formula by induction on k. The formula is valid for $k=0$ assuming that $v_{-1}=0$. Assume that it is valid for k. Then

$$\begin{aligned}\mathcal{E}_+v_{k+1} &= \mathcal{E}_+\mathcal{E}_-v_k = \mathcal{E}_-\mathcal{E}_+v_k + [\mathcal{E}_+,\mathcal{E}_-]v_k \\ &= c_k\mathcal{E}_-v_{k-1} + \mathcal{H}v_k = c_kv_k + (\lambda_0 - 2k)v_k = c_{k+1}v_k,\end{aligned}$$

where

$$c_{k+1} = c_k + \lambda_0 - 2k = (k+1)\lambda_0 - k^2 - k = (k+1)(\lambda_0 - k).$$

□

Since eigenvectors of the operator $\mathcal{H}$ that correspond to different eigenvalues are linearly independent, there exists a number n such that the vectors $v_0, v_1, v_2, \ldots, v_n$ are nonzero and linearly independent, whereas $v_{n+1} = 0$. Then it follows from Lemma 12.69 that $c_{n+1} = 0$, i.e., that $\lambda_0 = n$.

Now, it follows from the previous formulas that the linear span of the vectors $v_0, v_1, v_2, \ldots, v_n$ is invariant with respect to the operators $\mathcal{H}, \mathcal{E}_+, \mathcal{E}_-$, hence with respect to all of the algebra $\mathrm{sl}_2(\mathbb{C})$.

If the representation R is irreducible, then

$$V = \langle v_0, v_1, \ldots, v_n \rangle \tag{12.24}$$

and the operators $\mathcal{H}, \mathcal{E}_+, \mathcal{E}_-$ have the following form in the basis $\{v_0, v_1, \ldots, v_n\}$:

$$\mathcal{H}v_k = (n-2k)v_k, \tag{12.25}$$

$$\mathcal{E}_+ v_k = k(n-k+1)v_{k-1}, \tag{12.26}$$

$$\mathcal{E}_- v_k = v_{k+1}, \tag{12.27}$$

assuming $v_{-1} = v_{n+1} = 0$. The number n is called the *highest weight* of the representation R. Formulas (12.25)–(12.27) show that an irreducible representation of the group $\mathrm{SL}_2(\mathbb{C})$ is completely determined by its highest weight.

Conversely, if condition (12.24) holds, the representation R is irreducible. Indeed, any nonzero invariant subspace is, in particular, invariant under the operator $\mathcal{H}$; hence it is a linear span of some of its eigenvectors $v_0, v_1, \ldots, v_n$. But, since it is invariant under the operators $\mathcal{E}_+$ and $\mathcal{E}_-$, it must contain all these vectors, i.e., coincide with V.

The remaining question is whether a representation with a given highest weight exists. One can check directly that the operators $\mathcal{H}, \mathcal{E}_+, \mathcal{E}_-$ given by formulas (12.25)–(12.27) define a linear representation of the Lie algebra $\mathrm{sl}_2(\mathbb{C})$ and then, using a general theorem omitted in this course, prove that there exists a linear representation of the group $\mathrm{SL}_2(\mathbb{C})$ whose differential is the above representation. We take a different approach and construct the required representation explicitly.

We assume that the group $\mathrm{SL}_2(\mathbb{C})$ acts on the space $\mathbb{C}^2$ with the basis (x, y) tautologically, i.e., that the element

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C})$$

acts by the formulas

$$\begin{cases} gx = ax + cy, \\ gy = bx + dy. \end{cases} \tag{12.28}$$

This action induces a linear action of $\mathrm{SL}_2(\mathbb{C})$ on the space $S^n(\mathbb{C}^2)$. Regard x and y as coordinates in the dual space $(\mathbb{C}^2)^*$ and identify the symmetric algebra of $\mathbb{C}^2$ with the polynomial algebra on $(\mathbb{C}^2)^*$ (see Section 8.3). Then the elements of the space $S^n(\mathbb{C}^2)$ are homogeneous polynomials of degree n in x and y, also known as *binary forms* of degree n. The action of $\mathrm{SL}_2(\mathbb{C})$ is then interpreted as a linear change of variables by formulas (12.28). We

thus obtain a linear representation; denote it by R_n. By definition,

$$(R_n(g)f)(x,y) = f(ax+cy, bx+dy)$$

for every binary form f of degree n.

Let us calculate the differential of the representation R_n. Put

$$dR_n(H) = \mathcal{H}, \qquad dR(E_+) = \mathcal{E}_+, \qquad dR(E_-) = \mathcal{E}_-.$$

Taking into account that

$$\exp tH = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}, \qquad \exp tE_+ = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad \exp tE_- = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix},$$

we obtain

$$(\mathcal{H}f)(x,y) = \frac{\partial}{\partial t} f(e^t x, e^{-t} y)\Big|_{t=0} = x\frac{\partial f(x,y)}{\partial x} - y\frac{\partial f(x,y)}{\partial y},$$

$$(\mathcal{E}_+ f)(x,y) = \frac{\partial}{\partial t} f(x, y+tx)\Big|_{t=0} = x\frac{\partial f(x,y)}{\partial y},$$

$$(\mathcal{E}_- f)(x,y) = \frac{\partial}{\partial t} f(x+ty, y)\Big|_{t=0} = y\frac{\partial f(x,y)}{\partial x}$$

for a binary form $f \in S^n(\mathbb{C}^2)$.

In particular, for $f_0 = x^n$,

$$\mathcal{H}f_0 = nf_0, \qquad \mathcal{E}_+ f_0 = 0,$$

i.e., f_0 is a highest weight vector of the representation R_n with the eigenvalue n. Moreover,

$$f_k = \mathcal{E}_-^k f_0 = n(n-1)\cdots(n-k+1)x^{n-k}y^k,$$

so that the forms $f_0, f_1, \ldots, f_n$ comprise a basis of the space $S^n(\mathbb{C}^2)$. Therefore, R_n is an irreducible representation with the highest weight n.

Simultaneously, we have just proved that every irreducible linear representation of the algebra $\mathrm{sl}_2(\mathbb{C})$ is a differential of an (irreducible) linear representation of the group $\mathrm{SL}_2(\mathbb{C})$.

The results obtained also provide a description of irreducible complex linear representations of the Lie groups $\mathrm{SL}_2(\mathbb{R})$ and SU_2, which are real forms of the group $\mathrm{SL}_2(\mathbb{C})$. Indeed, if H is a real form of $\mathrm{SL}_2(\mathbb{C})$ and $S\colon H \to \mathrm{GL}(V)$ is its irreducible complex linear representation, then the representation dS of the Lie algebra $T(H)$ extends uniquely to a linear representation of the Lie algebra $\mathrm{sl}_2(\mathbb{C})$. By the aforesaid, this representation is the differential of a linear representation of the group $\mathrm{SL}_2(\mathbb{C})$. It follows that the irreducible complex linear representations of the group H are exactly the restrictions of the irreducible linear representations of $\mathrm{SL}_2(\mathbb{C})$ to H.

Exercise 12.70. For $n > 0$, prove that

$$\operatorname{Ker} R_n = \begin{cases} \{E\}, & n \text{ odd}, \\ \{\pm E\}, & n \text{ even}. \end{cases}$$

Exercise 12.71. Describe irreducible linear representation of the Lie group $\mathrm{SO}_3(\mathbb{C})$.

Exercise 12.72. Describe irreducible complex linear representations of the Lie groups SO_3 and $\mathrm{SO}^{\circ}_{2,1}$.

Answers to Selected Exercises

1.79. $E_{ij}E_{kl} = \delta_{jk}E_{il}$ (here δ_{ij} is the Kronecker delta).

2.35. q^n.

2.39. $(q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{n-1})$.

2.41. $\frac{(q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{k-1})}{(q^k-1)(q^k-q)(q^k-q^2)\cdots(q^k-q^{k-1})}$.

2.69. 3.

3.65. $2 = (1+i)(1-i) \sim (1+i)^2$, 3 is prime, $5 = (2+i)(2-i)$.

3.66. x, x^2+x+1, x^3+x^2+1, x^3+x+1, x^4+x^3+1, x^4+x+1, $x^4+x^3+x^2+x+1$.

4.112. $|\mathrm{GL}_2(\mathbb{Z}_p)| = p(p+1)(p-1)^2$, $|\mathrm{SL}_2(\mathbb{Z}_p)| = p(p-1)^2$.

5.5. Three distinct one-dimensional subspaces of a two-dimensional vector space.

5.7. Same as in Exercise 5.5.

5.52. $\left(\frac{n(n+1)}{2}, \frac{n(n-1)}{2}\right)$.

5.68. $\begin{vmatrix} 1 & -\cos\alpha_{12} & -\cos\alpha_{13} & -\cos\alpha_{14} \\ -\cos\alpha_{12} & 1 & -\cos\alpha_{23} & -\cos\alpha_{24} \\ -\cos\alpha_{13} & -\cos\alpha_{23} & 1 & -\cos\alpha_{34} \\ -\cos\alpha_{14} & -\cos\alpha_{24} & -\cos\alpha_{34} & 1 \end{vmatrix}$, where α_{ij} is the angle between the ith and the jth faces; the dihedral angle in the regular tetrahedron equals $\arccos 1/3$.

5.85. $\sum_k \overline{c}_{ki}c_{kj} = \delta_{ij}$; $\sum_k \overline{c}_{ik}c_{jk} = \delta_{ij}$.

6.10. $\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix}$.

6.46. The matrix D is determined up to a permutation of the diagonal entries. Given D, the matrices O_1 and O_2 are determined up to a transformation $O_1 \to O_1O$, $O_2 \to O^{-1}O_2$, where O is an orthogonal matrix that commutes with D.

6.63. $t, t-1$.

6.85. (i) $\max_i \sum_j |a_{ij}|$, where $(a_{ij}) = A$ is the matrix of the operator $\mathcal{A}$; (ii) $\sqrt{\max_i \lambda_i}$, where $\lambda_1, \ldots, \lambda_n$ are the (nonnegative) eigenvalues of the selfadjoint operator $\mathcal{A}^*\mathcal{A}$; (iii) $\max_j \sum_i |a_{ij}|$.

7.12. 5.

7.13. $\dim(U_1 + U_2)$ if $P_1 \cap P_2 \neq \varnothing$; $\dim(U_1 + U_2) + 1$ if $P_1 \cap P_2 = \varnothing$.

7.44. The faces of positive dimension are determined by the following condition: $k < n/2$ coordinates x_i are 0 and $l < n/2$ coordinates are 1, given that $k + l < n - 1$ (this is automatic for an even n). A vertex is a point having $[n/2]$ coordinates equal to 0 and $[n/2]$ coordinates equal to 1; for an odd n, the remaining coordinate equals $1/2$.

7.45. Convex hulls of all subsets of the set of the vertices of the simplex.

7.64. When p_1, p_2, p_3 are permuted, their ratio assumes the values c, $-c-1$, $\frac{1}{c}$, $-\frac{1}{c}-1$, $-\frac{1}{c+1}$, $-\frac{c}{c+1}$. The smallest number of different values equals 2 if the equation $x^2+x+1=0$ has a solution in the field K, and 3 otherwise.

7.71. The rotation about the third point through the sum of the angles if it is not 2π, and a parallel translation otherwise.

7.73. See the answer to Exercise 10.41 with the list of all elements of the group $\mathrm{Sym}_+ K$. Additionally, $\mathrm{Sym}\, K$ contains six reflections through planes passing through the edges, three reflections through planes parallel to the faces, eight mirror rotations through $\pi/3$ about the axes passing through the vertices, six mirror rotations through $\pi/2$ about axes passing through the centers of the faces, and the central symmetry.

7.74. $(x_1, x_2) \mapsto (tx_1, t^{-1}x_2)$; $(x_1, x_2) \mapsto (tx_2, t^{-1}x_1)$ for $t \in \mathbb{R}^*$.

7.75. If the inner squares of the sides of one triangle (regarded as vectors) equal the inner squares of the respective sides of the other triangle, these triangles are equal.

7.100. $\frac{1}{y_1}, \frac{y_2}{y_1}, \ldots, \frac{y_n}{y_1}$.

7.109. Permutations from V_4 do not change the cross-ratio. Permutations preserving the point p_4 change the cross-ratio just as the simple ratio of points p_1, p_2, p_3 (see the answer to Exercise 7.64) with the minus sign, i.e., the values of the cross-ratio are: $\delta, 1-\delta, \frac{1}{\delta}, 1-\frac{1}{\delta}, \frac{1}{1-\delta}, \frac{\delta}{\delta-1}$.

9.24. $[10]_{15}, [6]_{15}$.

9.35. $[3]_7, [6]_{41}$.

9.67. If $n = 2^m p_1^{k_1} \cdots p_s^{k_s} q_1^{l_1} \cdots q_t^{l_t}$, where $p_1, \ldots, p_s$ are distinct prime numbers of the form $4k+1$ and $q_1, \ldots, q_t$ are distinct prime numbers of the form $4k+3$ and $l_1, \ldots, l_t$ are even, this number equals $4(k_1+1)\cdots(k_s+1)$.

9.87. For example, a block-diagonal matrix that consists of Jordan blocks and blocks of the following form:

$$\begin{pmatrix} a & 1 & & & & & \\ -b & a & & & & & \\ & & a & 1 & & & \\ & & -b & a & & & \\ & & & & a & 1 & \\ & & & & -b & a & \end{pmatrix} \quad (b > 0).$$

9.88. For example, a block-diagonal matrix of order 4 that consists of Jordan blocks and blocks of the form

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

9.120.

$+$	0	1	α	$\alpha+1$
0	0	1	α	$\alpha+1$
1	1	0	$\alpha+1$	α
α	α	$\alpha+1$	0	1
$\alpha+1$	$\alpha+1$	α	1	0

$\times$	0	1	α	$\alpha+1$
0	0	0	0	0
1	0	1	α	$\alpha+1$
α	0	α	$\alpha+1$	1
$\alpha+1$	0	$\alpha+1$	1	α

9.165. $(p_1 \cdots p_s)$.

10.7. For odd n.

10.30. $\langle a^{k-1} \rangle_{n/(n,k-1)}$.

10.41. Five classes: the identity motion (1 element); rotations through $\frac{2\pi}{3}$ about axes passing through the vertices (8 elements); rotations through π about axes passing through the centers of the edges (6 elements); rotations through $\frac{\pi}{2}$ about axes passing through the centers of the faces (6 elements);

rotations through π about axes passing through the centers of the faces (3 elements).

10.48. 57.

10.66. $\frac{1}{2}q(q^2-1)$ for q odd and $q(q^2-1)$ for q a power of 2.

11.77. In the notation of Example 10.23, $D_n = (a)_n \ltimes (b)_2$. For an odd n, the group D_n has two one-dimensional representations

$$a \mapsto 1, \quad b \mapsto \pm 1$$

and $\frac{n-1}{2}$ two-dimensional irreducible representations

$$a \mapsto \begin{pmatrix} \omega^k & 0 \\ 0 & \omega^{-k} \end{pmatrix}, \quad b \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad 1 \le k < \frac{n}{2},$$

where $\omega = e^{2\pi i/n}$. For an even n, D_n has four one-dimensional representations

$$a \mapsto \pm 1, \quad b \mapsto \pm 1$$

and $\frac{n}{2} - 1$ two-dimensional irreducible representations described as in the case of odd n.

11.86. The one-dimensional subspace of constants; the two-dimensional subspace of "even" functions assuming the same value on opposite faces of the cube, with the sum of all values equal zero; the three-dimensional subspace of "odd" functions assuming opposite values on the opposite faces of the cube.

11.88.

	χ_1	χ_2	χ_3'	χ_4	χ_5	
e	1	3	3	4	5	1
(12)(34)	1	−1	−1	0	1	15
(123)	1	0	0	1	−1	20
(12345)	1	$\frac{3-\sqrt{5}}{2}$	$\frac{3+\sqrt{5}}{2}$	−1	0	12
(12354)	1	$\frac{3+\sqrt{5}}{2}$	$\frac{3-\sqrt{5}}{2}$	−1	0	12

11.119. $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$.

12.45. $\left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$.

12.71. For every $n \in \mathbb{Z}_+$, there exists a unique $(2n+1)$-dimensional irreducible representation S_n of the group $\mathrm{SO}_3(\mathbb{C})$ related to the representation R_{2n} of the group $\mathrm{SL}_2(\mathbb{C})$ via the following commutative diagram:

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbb{C}) & \xrightarrow{R_{2n}} & \mathrm{GL}_{2n+1}(\mathbb{C}) \\ {\scriptstyle \mathrm{Ad}}\searrow & & \nearrow{\scriptstyle S_n} \\ & \mathrm{SO}_3(\mathbb{C}) & \end{array}$$

(see Exercise 12.50).

12.72. For every $n \in \mathbb{Z}_+$, there exists a unique $(2n+1)$-dimensional irreducible complex representation of the group SO_3 (respectively, $SO^\circ_{2,1}$) obtained by restriction of the representation S_n of the group $SO_3(\mathbb{C})$ (see the answer to Exercise 12.71).

Bibliography

[1] Sh. Axler, *Linear algebra done right.* Springer-Verlag, New York, 1997.

[2] M. E. Atiyah and I. G. Macdonald, *Introduction to commutative algebra.* Addison-Wesley, Reading, MA, 1969.

[3] R. Bellman, *Introduction to matrix analysis.* SIAM, Philadelphia, PA, 1997.

[4] M. Berger, *Geometry.* I, II. Springer-Verlag, Berlin, 1987.

[5] G. Birkhoff and S. Mac Lane, *A survey of modern algebra.* Third edition. Macmillan, New York, 1965.

[6] D. K. Faddeev, *Lectures in algebra.* "Nauka", Moscow, 1984. (Russian)

[7] D. K. Faddeev and I. S. Sominskii, *Problems in higher algebra.* Freeman, San Francisco–London, 1965.

[8] F. R. Gantmacher, *The theory of matrices.* Vols. 1, 2. Chelsea Publishing Co., New York, 1959.

[9] I. M. Gelfand, *Lectures on linear algebra.* Dover, New York, 1989.

[10] I. M. Glazman and Yu. I. Lyubich, *Finite-dimensional linear analysis: a systematic presentation in problem form.* The M.I.T. Press, Cambridge, MA, 1974.

[11] P. Halmos, *Finite-dimensional vector spaces.* Springer-Verlag, New York–Heidelberg, 1974.

[12] I. N. Herstein, *Noncommutative rings.* MAA, Washington, DC; Wiley, New York, 1968.

[13] ———, *Topics in algebra.* Second edition. Xerox College Publishing, Lexington, MA, 1975.

[14] ———, *Abstract algebra.* Third edition. Prentice Hall, Upper Saddle River, NJ, 1996.

[15] T. Hungerford, *Algebra.* Springer-Verlag, New York–Berlin, 1980.

[16] N. Jacobson, *Basic algebra.* I, II. Freeman, New York, 1985, 1989.

[17] A. I. Kostrikin, *Introduction to algebra.* Parts I–III. Fizmatlit, Moscow, 2000; English transl. of the first edition, Springer-Verlag, New York–Berlin, 1982.

[18] A. I. Kostrikin (ed.), *Problems in algebra.* "Faktorial", Moscow, 1995. (Russian)

[19] A. I. Kostrikin and Yu. I. Manin, *Linear algebra and geometry.* Gordon and Breach, Amsterdam, 1997.

[20] P. Lancaster, *Theory of matrices*. Academic Press, New York–London, 1969.

[21] S. Lang, *Algebraic number theory*. Second edition. Springer-Verlag, New York, 1994.

[22] ———, *Algebra*. Revised third edition, Springer-Verlag, New York, 2002.

[23] H. Matsumura, *Commutative algebra*. Benjamin/Cummings, Reading, MA, 1980.

[24] A. P. Mishina and I. V. Proskuryakov, *Higher algebra. Linear algebra, polynomials, general algebra*. Pergamon Press, Oxford–New York–Paris, 1965.

[25] R. S. Pierce, *Associative algebras*. Springer-Verlag, New York–Berlin, 1982.

[26] I. V. Proskuryakov, *Problems in linear algebra*. "Nauka", Moscow, 1984; English transl. of the 1974 Russian edition, "Mir", Moscow, 1978.

[27] I. R. Shafarevich, *Basic notions of algebra*. Springer-Verlag, Berlin, 1997.

[28] ———, *Basic algebraic geometry. 1. Varieties in projective space*. Springer-Verlag, Berlin, 1994.

[29] ———, *Basic algebraic geometry. 2. Schemes and complex manifolds*. Springer-Verlag, Berlin, 1994.

[30] J.-P. Serre, *Linear representations of finite groups*. Springer-Verlag, New York–Heidelberg, 1977.

[31] G. E. Shilov, *Linear algebra*. Prentice-Hall, Englewood Cliffs, NJ, 1971.

[32] B. L. van der Waerden, *Algebra*. Vol. I, II. Springer-Verlag, New York, 1991.

[33] E. B. Vinberg, *The algebra of polynomials*. "Prosveshchenie", Moscow, 1980. (Russian)

[34] ———, *Linear representations of groups*. Birkhäuser, Basel, 1989.

Index

Titles in This Series

56 **E. B. Vinberg,** A course in algebra, 2003

55 **C. Herbert Clemens,** A scrapbook of complex curve theory, second edition, 2003

54 **Alexander Barvinok,** A course in convexity, 2002

53 **Henryk Iwaniec,** Spectral methods of automorphic forms, 2002

52 **Ilka Agricola and Thomas Friedrich,** Global analysis: Differential forms in analysis, geometry and physics, 2002

51 **Y. A. Abramovich and C. D. Aliprantis,** Problems in operator theory, 2002

50 **Y. A. Abramovich and C. D. Aliprantis,** An invitation to operator theory, 2002

49 **John R. Harper,** Secondary cohomology operations, 2002

48 **Y. Eliashberg and N. Mishachev,** Introduction to the h-principle, 2002

47 **A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi,** Classical and quantum computation, 2002

46 **Joseph L. Taylor,** Several complex variables with connections to algebraic geometry and Lie groups, 2002

45 **Inder K. Rana,** An introduction to measure and integration, second edition, 2002

44 **Jim Agler and John E. M^{c}Carthy,** Pick interpolation and Hilbert function spaces, 2002

43 **N. V. Krylov,** Introduction to the theory of random processes, 2002

42 **Jin Hong and Seok-Jin Kang,** Introduction to quantum groups and crystal bases, 2002

41 **Georgi V. Smirnov,** Introduction to the theory of differential inclusions, 2002

40 **Robert E. Greene and Steven G. Krantz,** Function theory of one complex variable, 2002

39 **Larry C. Grove,** Classical groups and geometric algebra, 2002

38 **Elton P. Hsu,** Stochastic analysis on manifolds, 2002

37 **Hershel M. Farkas and Irwin Kra,** Theta constants, Riemann surfaces and the modular group, 2001

36 **Martin Schechter,** Principles of functional analysis, second edition, 2002

35 **James F. Davis and Paul Kirk,** Lecture notes in algebraic topology, 2001

34 **Sigurdur Helgason,** Differential geometry, Lie groups, and symmetric spaces, 2001

33 **Dmitri Burago, Yuri Burago, and Sergei Ivanov,** A course in metric geometry, 2001

32 **Robert G. Bartle,** A modern theory of integration, 2001

31 **Ralf Korn and Elke Korn,** Option pricing and portfolio optimization: Modern methods of financial mathematics, 2001

30 **J. C. McConnell and J. C. Robson,** Noncommutative Noetherian rings, 2001

29 **Javier Duoandikoetxea,** Fourier analysis, 2001

28 **Liviu I. Nicolaescu,** Notes on Seiberg-Witten theory, 2000

27 **Thierry Aubin,** A course in differential geometry, 2001

26 **Rolf Berndt,** An introduction to symplectic geometry, 2001

25 **Thomas Friedrich,** Dirac operators in Riemannian geometry, 2000

24 **Helmut Koch,** Number theory: Algebraic numbers and functions, 2000

23 **Alberto Candel and Lawrence Conlon,** Foliations I, 2000

22 **Günter R. Krause and Thomas H. Lenagan,** Growth of algebras and Gelfand-Kirillov dimension, 2000

21 **John B. Conway,** A course in operator theory, 2000

20 **Robert E. Gompf and András I. Stipsicz,** 4-manifolds and Kirby calculus, 1999

19 **Lawrence C. Evans,** Partial differential equations, 1998

18 **Winfried Just and Martin Weese,** Discovering modern set theory. II: Set-theoretic tools for every mathematician, 1997

17 **Henryk Iwaniec,** Topics in classical automorphic forms, 1997

16 **Richard V. Kadison and John R. Ringrose,** Fundamentals of the theory of operator algebras. Volume II: Advanced theory, 1997

15 **Richard V. Kadison and John R. Ringrose,** Fundamentals of the theory of operator algebras. Volume I: Elementary theory, 1997

14 **Elliott H. Lieb and Michael Loss,** Analysis, 1997

13 **Paul C. Shields,** The ergodic theory of discrete sample paths, 1996

12 **N. V. Krylov,** Lectures on elliptic and parabolic equations in Hölder spaces, 1996

11 **Jacques Dixmier,** Enveloping algebras, 1996 Printing

10 **Barry Simon,** Representations of finite and compact groups, 1996

9 **Dino Lorenzini,** An invitation to arithmetic geometry, 1996

8 **Winfried Just and Martin Weese,** Discovering modern set theory. I: The basics, 1996

7 **Gerald J. Janusz,** Algebraic number fields, second edition, 1996

6 **Jens Carsten Jantzen,** Lectures on quantum groups, 1996

5 **Rick Miranda,** Algebraic curves and Riemann surfaces, 1995

4 **Russell A. Gordon,** The integrals of Lebesgue, Denjoy, Perron, and Henstock, 1994

3 **William W. Adams and Philippe Loustaunau,** An introduction to Gröbner bases, 1994

2 **Jack Graver, Brigitte Servatius, and Herman Servatius,** Combinatorial rigidity, 1993

1 **Ethan Akin,** The general topology of dynamical systems, 1993

Great book! The author's teaching experience shows in every chapter.

—Efim Zelmanov, University of California, San Diego

Vinberg has written an algebra book that is excellent, both as a classroom text or for self-study. It is plain that years of teaching abstract algebra have enabled him to say the right thing at the right time.

—Irving Kaplansky, MSRI

This is a comprehensive text on modern algebra written for advanced undergraduate and basic graduate algebra classes. The book is based on courses taught by the author at the Mechanics and Mathematics Department of Moscow State University and at the Mathematical College of the Independent University of Moscow.

The unique feature of the book is that it contains almost no technically difficult proofs. Following his point of view on mathematics, the author tried, whenever possible, to replace calculations and difficult deductions with conceptual proofs and to associate geometric images to algebraic objects. Another important feature is that the book presents most of the topics on several levels, allowing the student to move smoothly from initial acquaintance to thorough study and deeper understanding of the subject.

Presented are basic topics in algebra such as algebraic structures, linear algebra, polynomials, groups, as well as more advanced topics like affine and projective spaces, tensor algebra, Galois theory, Lie groups, associative algebras and their representations. Some applications of linear algebra and group theory to physics are discussed.

Written with extreme care and supplied with more than 200 exercises and 70 figures, the book is also an excellent text for independent study.

ISBN 0-8218-3318-9

9 780821 833186

GSM/56