DIE GRUNDLEHREN DER

MATHEMATISCHEN WISSENSCHAFTEN

IN EINZELDARSTELLUNGEN MIT BESONDERER BERÜCKSICHTIGUNG DER ANWENDUNGSGEBIETE

BAND 97

Springer-Verlag Berlin Heidelberg GmbH

LINEAR ALGEBRA

BY

WERNER H. GREUB

MATHEMATICS DEPARTMENT
UNIVERSITY OF TORONTO

SECOND EDITION

WITH 6 FIGURES

1963

Springer-Verlag Berlin Heidelberg GmbH

Library of Congress Catalog Card Number 63—17509

ISBN 978-3-662-01547-6 ISBN 978-3-662-01545-2 (eBook)
DOI 10.1007/978-3-662-01545-2

Originally published by Springer-Verlag OHG. Berlin · Göttingen · Heidelberg in 1963.

Softcover reprint of the hardcover 2nd edition 1963

TO

ROLF NEVANLINNA

Preface to the second edition

Besides the very obvious change from German to English, the second edition of this book contains many additions as well as a great many other changes. It might even be called a new book altogether were it not for the fact that the essential character of the book has remained the same; in other words, the entire presentation continues to be based on an axiomatic treatment of linear spaces.

In this second edition, the thorough-going restriction to linear spaces of finite dimension has been removed. Another complete change is the restriction to linear spaces with real or complex coefficients, thereby removing a number of relatively involved discussions which did not really contribute substantially to the subject. On p. 6 there is a list of those chapters in which the presentation can be transferred directly to spaces over an arbitrary coefficient field.

Chapter I deals with the general properties of a linear space. Those concepts which are only valid for finitely many dimensions are discussed in a special paragraph.

Chapter II now covers only linear transformations while the treatment of matrices has been delegated to a new chapter, chapter III. The discussion of dual spaces has been changed; dual spaces are now introduced abstractly and the connection with the space of linear functions is not established until later.

Chapters IV and V, dealing with determinants and orientation respectively, do not contain substantial changes. Brief reference should be made here to the new paragraph in chapter IV on the trace of an endomorphism — a concept which is used quite consistently throughout the book from that time on.

Special emphasize is given to tensors. The original chapter on Multilinear Algebra is now spread over four chapters: Multilinear Mappings (Ch. VI), Tensor Algebra (Ch. VII), Exterior Algebra (Ch. VIII) and Duality in Exterior Algebra (Ch. IX). The chapter on multilinear mappings consists now primarily of an introduction to the theory of the tensor-product. In chapter VII the notion of vector-valued tensors has been introduced and used to define the contraction. Furthermore, a treatment of the transformation of tensors under linear mappings has been added. In Chapter VIII the antisymmetry-operator is studied in greater detail and the concept of the skew-symmetric power is introduced. The dual product (Ch. IX) is generalized to mixed tensors. A special paragraph in this chapter covers the skew-symmetric powers of the unit tensor and shows their significance in the characteristic polynomial. The paragraph "Adjoint Tensors" provides a number of applications of the duality theory to certain tensors arising from an endomorphism of the underlying space.

There are no essential changes in Chapter X (Inner product spaces) except for the addition of a short new paragraph on normed linear spaces. In the next chapter, on linear mappings of inner product spaces, the orthogonal projections (§ 3) and the skew mappings (§ 4) are discussed in greater detail. Furthermore, a paragraph on differentiable families of automorphisms has been added here.

Chapter XII (Symmetric Bilinear Functions) contains a new paragraph dealing with Lorentz-transformations.

Whereas the discussion of quadrics in the first edition was limited to quadrics with centers, the second edition covers this topic in full.

The chapter on unitary spaces has been changed to include a more thorough-going presentation of unitary transformations of the complex plane and their relation to the algebra of quaternions.

The restriction to linear spaces with complex or real coefficients has of course greatly simplified the construction of irreducible subspaces in chapter XV. Another essential simplification of this construction was achieved by the simultaneous consideration of the dual mapping. A final paragraph with applications to Lorentz-transformation has been added to this concluding chapter.

Many other minor changes have been incorporated — not least of which are the many additional problems now accompanying each paragraph.

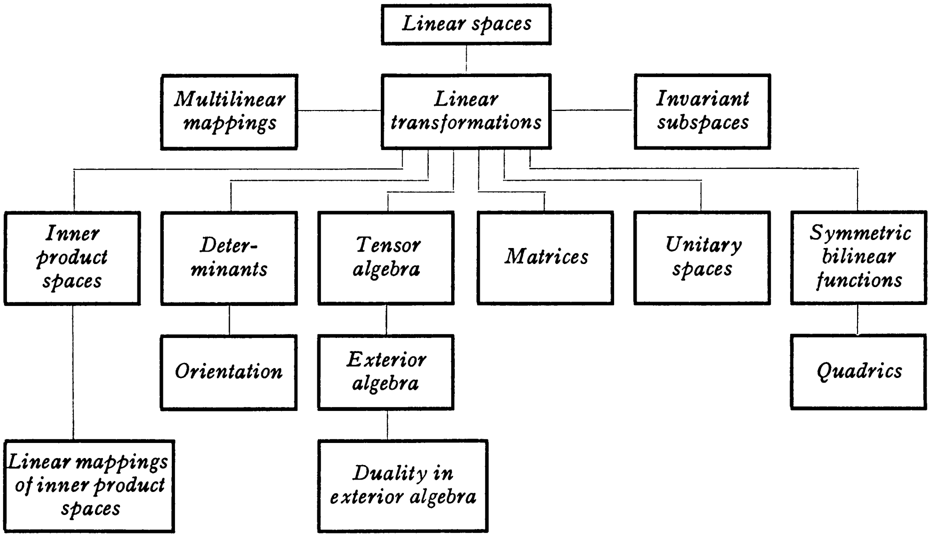
Last, but certainly not least, I have to express my sincerest thanks to everyone who has helped me in the preparation of this second edition. First of all, I am particularly indebted to CORNELIE J. RHEINBOLDT who assisted in the entire translating and editing work and to Dr. WERNER C. RHEINBOLDT who cooperated in this task and who also made a number of valuable suggestions for improvements, especially in the chapters on linear transformations and matrices. My warm thanks also go to Dr. H. BOLDER of the Royal Dutch/Shell Laboratory at Amsterdam for his criticism on the chapter on tensor-products and to Dr. H. H. KELLER who read the entire manuscript and offered many important suggestions. Furthermore, I am grateful to Mr. GIORGIO PEDERZOLI who helped to read the proofs of the entire work and who collected a number of new problems and to Mr. KHADJA NESAMUDDIN KHAN for his assistance in preparing the manuscript.

Finally I would like to express my thanks to the publishers for their patience and cooperation during the preparation of this edition.

Toronto, April 1963 WERNER H. GREUB

Contents

Interdependence of Chapters



Chapter I

Linear spaces

§ 1. The axioms of a linear space

1.1. Additive groups. A set E: $(x, y \ldots)$ is called an *additive group* if to every pair x and y there is assigned a third element of E, called the *sum* of x and y and written as $x + y$, such that the following axioms hold:

I.1. $x + y = y + x$ (commutative law)

I.2. $(x + y) + z = x + (y + z)$ (associative law)

I.3. There exists a zero-element 0 such that $x + 0 = x$ for every $x \in E$.

I.4. To every element x there exists an *inverse element* $-x$ such that $x + (-x) = 0$.

The zero-element is uniquely determined. In fact, assume there are two such elements 0 and $0'$. Then for every $x \in E$

$$x + 0 = x \quad \text{and} \quad x + 0' = x .$$

Substituting $x = 0'$ in the first and $x = 0$ in the second equation we obtain

$$0' + 0 = 0' \quad \text{and} \quad 0 + 0' = 0$$

and hence by the commutative law $0 = 0'$.

For every $x \in E$, there is only one inverse element $-x$. We prove more generally that to any two elements a and b there is exactly one element x such that

$$x + a = b . \tag{1.1}$$

To show first the uniqueness, let x_1 and x_2 be two solutions of (1.1). Then

$$x_1 + a = b \quad \text{and} \quad x_2 + a = b$$

and consequently

$$x_1 + a = x_2 + a .$$

Now let $-a$ be a negative element of a. Adding $-a$ to the above equation we obtain by the associative law

$$x_1 + (a + (-a)) = x_2 + (a + (-a))$$

and hence $x_1 = x_2$. This result applied to the vector $b = 0$ yields the uniqueness of $-a$.

To show that (1.1) always has a solution, consider the element

$$x = b + (-a)\,. \tag{1.2}$$

Then

$$x + a = b + (-a) + a = b + 0 = b\,.$$

The element x defined by (1.2) is called the *difference* of b and a and is denoted by $b - a$.

1.2. Real linear spaces. A *real linear space* or *real vector space* is an additive group with the following additional structure: There is defined a multiplication between the real numbers $\lambda, \mu \ldots$ and the elements of E; in other words, to every pair (λ, x) an element λx of E is assigned, subject to the following axioms:

II.1. $(\lambda\,\mu)\,x = \lambda(\mu\,x)$ (associative law)

II.2. $\left.\begin{aligned}(\lambda + \mu)\,x &= \lambda x + \mu x\\ \lambda(x + y) &= \lambda x + \lambda y\end{aligned}\right\}$ (distributive laws)

II.3. $1 \cdot x = x$.

The elements of a linear space are called *vectors* and the coefficients *scalars*. From the first distributive law II.2 we obtain by inserting $\mu = 0$

$$\lambda x = \lambda x + 0 \cdot x$$

and adding the vector $-(\lambda x)$

$$0 \cdot x = 0\,^{*)}.$$

Similarly, the second law II.2. yields for $y = 0$

$$\lambda \cdot 0 = 0\,.$$

These two equations state that $\lambda x = 0$ if $\lambda = 0$ or $x = 0$. Conversely, the equation $\lambda x = 0$ implies that $\lambda = 0$ or $x = 0$. In fact, assume that $\lambda \neq 0$. Then it follows from the axioms II.3 and II.1. that

$$x = 1 \cdot x = \left(\frac{1}{\lambda} \cdot \lambda\right) x = \frac{1}{\lambda}\,(\lambda x) = \frac{1}{\lambda} \cdot 0 = 0\,.$$

Altogether we have shown that $\lambda x = 0$ if and only if $\lambda = 0$ or $x = 0$.

Substituting $\mu = -\lambda$ in the first distributive law we obtain

$$\lambda x + (-\lambda x) = 0\,,$$

whence

$$(-\lambda)\,x = -\lambda x\,.$$

Similarly, the second distributive law yields

$$\lambda(-x) = -\lambda x\,.$$

---

*) It should be observed that the symbol 0 on the left-hand side denotes the *scalar* zero and on the right-hand side the *vector* zero.

Finally we observe that the two distributive laws hold for any finite number of terms,

$$\left(\sum_\nu \lambda^\nu\right) x = \sum_\nu \lambda^\nu x$$

$$\lambda \sum_\nu x_\nu = \sum_\nu \lambda x_\nu$$

as can be shown by induction.

1.3. Examples: 1. Consider the set of all ordered n-tuples of real numbers

$$x = (\xi^1 \ldots \xi^n)$$

where n is a fixed integer. Addition of two n-tuples

$$x = (\xi^1 \ldots \xi^n) \quad \text{and} \quad y = (\eta^1 \ldots \eta^n)$$

is defined by

$$x + y = (\xi^1 + \eta^1 \ldots \xi^n + \eta^n)$$

and multiplication by a real number by

$$\lambda x = (\lambda \xi^1, \ldots \lambda \xi^n) .$$

The linear space thus obtained is called the real *n-dimensional number-space* and is denoted by R^n. Its zero-vector is the n-tuple

$$0 = (0 \ldots 0)$$

and the inverse of a vector x is given by the n-tuple

$$-x = (-\xi^1 \ldots -\xi^n) .$$

2. Denote by C the set of all real valued continuous functions f in the interval $0 \leqq t \leqq 1$. Defining addition and multiplication by a real number as

$$(f + g)(t) = f(t) + g(t)$$

and

$$(\lambda f)(t) = \lambda f(t)$$

we obtain a linear space. The zero-vector of this linear space is the identically vanishing function.

Instead of all continuous functions we could also consider the set of all differentiable functions or the set of all continuously differentiable functions.

3. Let S be an arbitrary set. Consider all real valued functions in S which assume the value zero except for finitely many points of S. If addition and multiplication is defined as in example 2 this set becomes a linear space $C(S)$. For every element $a \in S$ denote by f_a the function defined by

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

Then every function $f \in C(S)$ can be written as a finite linear combination

$$f = \sum_{a \in A} f(a) \cdot f_a$$

where A is the set of all points x for which $f(x) \neq 0$. Identifying every point $a \in S$ with the corresponding function f_a we can write the above equation as

$$f = \sum_{a \in A} f(a) \cdot a\,.$$

In this notation a function $f \in C(S)$ appears as a "formal linear combination" of the elements of S.

The space $C(S)$ is called the *linear space generated by* S.

1.4. Linear dependence. A system of p ($p \geqq 1$) vectors $(x_1 \ldots x_p)$ of a linear space E is called *linearly dependent* if there exist coefficients $(\lambda^1 \ldots \lambda^p)$, not all zero, such that

$$\sum_\nu \lambda^\nu x_\nu = 0\,.$$

Otherwise the vectors $(x_1 \ldots x_p)$ are called *linearly independent*. One single vector x is obviously linearly dependent if and only if $x = 0$.

If the system $(x_1 \ldots x_p)$ is linearly dependent, so is every other system $(x_1 \ldots x_p \ldots x_q)$ containing the vectors $x_1 \ldots x_p$. In fact, assume that

$$\lambda^1 x_1 + \cdots + \lambda^p x_p = 0$$

with at least one $\lambda^\nu \neq 0$. Then the relation

$$\lambda^1 x_1 + \cdots + \lambda^p x_p + 0 \cdot x_{p+1} + \cdots 0 \cdot x_q = 0$$

shows that the vectors $(x_1 \ldots x_q)$ are again linearly dependent. In particular, every system containing the zero-vector is linearly dependent.

From this result it follows that a system of linearly independent vectors remains linearly independent if some vectors are omitted.

1.5. Cartesian Product. Consider two linear spaces E and F. Form the product set $E \times F$ defined as the set of all pairs (x, y) with $x \in E$ and $y \in F$. In $E \times F$ introduce addition and multiplication by real numbers as follows:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$
$$\lambda(x, y) = (\lambda x, \lambda y)\,.$$

It is easy to see that these two operations satisfy the axioms of a linear space. The space $E \times F$ thus obtained is called the *Cartesian product* of E and F. In the same way the Cartesian product of any finite number of linear spaces can be defined.

1.6. Complex linear spaces. Instead of using real numbers as coefficients in a linear space, one can also take complex numbers. In this way one obtains a *complex linear space*. More precisely, let E be an additive group.

Assume that to every complex number λ and every vector $x \in E$ a vector $\lambda x \in E$ is assigned such that the three axioms II of sec. 1.2 are satisfied. Then E is called a *complex linear space.*

As an example, consider the set of all ordered n-tuples of complex numbers

$$z = (\zeta^1 \ldots \zeta^n)$$

with operations defined as in sec. 1.3 for the real number-space. The complex linear space C^n thus obtained is called the *n-dimensional complex number-space.*

1.7. Linear spaces over an arbitrary coefficient-field. The only properties of the real or the complex numbers used in the axioms of a linear space are those based upon the additive and multiplicative structure of these numbers. This fact suggests the generalization of the concept of a linear space by using as scalars the elements of an arbitrary commutative coefficient-field.

A *commutative field* Λ is a set of elements α, $\beta \ldots$ with two operations, addition and multiplication, subject to the following conditions:

I. Laws of addition:

1. $\alpha + \beta = \beta + \alpha$ (commutative law).
2. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ (associative law).
3. There exists an element 0 such that $\alpha + 0 = \alpha$ for every element $\alpha \in \Lambda$.
4. To every element α there exists an element $-\alpha$ such that $\alpha + (-\alpha) = 0$.

The above axioms assert that Λ is an abelian group.

II. Laws of multiplication:

1. $\alpha\beta = \beta\alpha$ (commutative law).
2. $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ (associative law).
3. There exists an element $\varepsilon \in \Lambda$ such that $\varepsilon\alpha = \alpha$ for every $\alpha \in \Lambda$.
4. To every element $\alpha \neq 0$ there exists an element α^{-1} such that $\alpha\alpha^{-1} = \varepsilon$.

III. The distributive law:

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma .$$

1.8. Let Λ be a given commutative field. A *linear space over the coefficient-field* Λ is an additive group in which a multiplication

$$(\lambda, x) \to \lambda x, \quad \lambda \in \Lambda, \quad x \in E$$

is defined such that the axioms II of sec. 1.2 are satisfied*). All properties derived from these axioms remain true for a linear space over an arbitrary commutative coefficient-field.

---

*) The number 1 in axiom II.3 has to be replaced by ε.

For the sake of simplicity we shall concern ourselves in the following chapters only with real and complex linear spaces. In other words, whenever we speak about linear spaces without further specification, a real or a complex space is understood. However, all the developments of the chapters I, II, III, VI*) and VII apply word for word to a linear space over an arbitrary commutative coefficient-field. The results of the chapters IV, VIII and IX can be carried over to linear spaces over a commutative coefficient-field, whose characteristic**) is different from 2.

Problems: 1. Show:

a) The set of all real numbers of the form $a + b\sqrt{5}$ with a and b integers forms an additive group,

b) the set of all real numbers of the form $\alpha + \beta\sqrt{3}$ with α and β rational forms a field,

c) the set of all complex numbers of the forms $\gamma + i\delta$ where γ and δ are real and $i = \sqrt{-1}$ forms a field.

2. Show that axiom II.3 can be replaced by the following one: The equation $\lambda x = 0$ holds if and only if $\lambda = 0$ or $x = 0$.

3. Given a system of linearly independent vectors $(x_1, \dots, x_p)$, prove that the system $x_1 \dots x_i + \lambda x_j, \dots, x_p (i \neq j)$ with arbitrary λ is again linearly independent.

4. Show that the set of all solutions of the homogeneous linear differential equation

$$\frac{d^2 y}{dt^2} + p\frac{dy}{dt} + qy = 0,$$

where p and q are functions of t, is a vector space.

5. Which of the following sets of functions are linearly dependent?

a) $f_1 = 3t$; $f_2 = t + 5$; $f_3 = 2t^2$; $f_4 = (t+1)^2$

b) $f_1 = (t+1)^2$; $f_2 = t^2 - 1$; $f_3 = 2t^2 + 2t - 3$

c) $f_1 = 1$; $f_2 = e^{it}$; $f_3 = e^{-it}$

d) $f_1 = t^2$; $f_2 = t$; $f_3 = 1$

e) $f_1 = 1 - t$; $f_2 = t(1-t)$; $f_3 = 1 - t^2$.

6. Let E be a real linear space. Consider the set $E \times E$ of ordered pairs (x, y) with $x \in E$ and $y \in E$. Show that the set $E \times E$ becomes a complex linear space C by the operations:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

and

$$(\alpha + i\beta)(x, y) = (\alpha x - \beta y, \alpha y + \beta x) \quad (\alpha, \beta \text{ real numbers}).$$

---

*) Except for the remarks about skew-symmetric mappings in sec 6.4.

**) Concerning the definition of the characteristic cf. VAN DER WAERDEN, Algebra.

7. Are the vectors $x_1 = (1, 0, 1)$; $x_2 = (i, 1, 0)$; $x_3 = (i, 2, 1 + i)$ linearly independent in C^3? Express $x_4 = (1, 2, 3)$ as a linear combination of x_1, x_2, x_3. Do the same for $x_5 = (i, i, i)$.

§ 2. Linear subspaces

1.9. Definition. A nonempty subset E_1 of a linear space E is called a *linear subspace* if the following conditions hold:

1. If $x \in E_1$ and $y \in E_1$, then $x + y \in E_1$.
2. If $x \in E_1$, then $\lambda x \in E_1$ for every coefficient λ.

The two above conditions are equivalent to the condition that E_1, with any two vectors x and y, contains all linear combinations $\lambda x + \mu y$.

By substituting $\lambda = 0$ in the second condition it follows that every subspace contains the zero-vector. A subspace E_1 is called *improper*, if E_1 consists of the entire space E or if E_1 reduces to the zero-vector; otherwise E_1 is called a *proper* subspace.

Every nonempty set S in E determines a subspace called the *linear closure* of S. It consists of all possible finite linear combinations

$$x = \sum_{\nu} \xi^\nu x_\nu \qquad x_\nu \in S$$

with arbitrary coefficients ξ^ν.

1.10. Intersection and sum. Let E_1 and E_2 be two subspaces of E. Then the set of all vectors contained in E_1 and in E_2 is again a linear subspace. This subspace is called the *intersection* of E_1 and E_2 and is denoted by $E_1 \cap E_2$.

The *sum* of E_1 and E_2, denoted as $E_1 + E_2$, is the set of all vectors

$$x = x_1 + x_2\,, \quad x_1 \in E_1\,, \quad x_2 \in E_2\text{*)}\,.$$

Obviously, $E_1 + E_2$ is a subspace of E, containing E_1 and E_2 as subspaces.

A vector x of the sum $E_1 + E_2$ can generally be decomposed as $x = x_1 + x_2$ ($x_1 \in E_1$, $x_2 \in E_2$) in different ways. Given two decompositions

$$x = x_1 + x_2\,, \quad x_1 \in E_1\,, \quad x_2 \in E_2$$

and

$$x = x_1' + x_2'\,, \quad x_1' \in E_1\,, \quad x_2' \in E_2$$

it follows that

$$x_1 - x_1' = x_2' - x_2\,.$$

Hence, the vector

$$z = x_1 - x_1'$$

is contained in the intersection $E_1 \cap E_2$. Conversely, let $x = x_1 + x_2$ be a

*) The sum $E_1 + E_2$ has to be distinguished from the set-theoretic union $E_1 \cup E_2$ which in general is not a linear space. $E_1 + E_2$ is obviously the linear closure of $E_1 \cup E_2$.

decomposition of x and z be an arbitrary vector of $E_1 \cap E_2$. Then the vectors

$$x_1' = x_1 - z \quad \text{and} \quad x_2' = x_2 + z$$

form again a decomposition of x. It follows from this remark that the vectors x_1 and x_2 are uniquely determined by x if and only if the intersection $E_1 \cap E_2$ reduces to the zero-vector. In this case the space $E_1 + E_2$ is called the *direct sum* of E_1 and E_2 and is denoted by $E_1 \oplus E_2$.

In the same way as it is for two subspaces, the intersection and sum of finitely many subspaces E_i $(i = 1 \ldots p)$ is defined. If any two spaces E_i and E_j $(i \neq j)$ have only the zero-vector in common, the space $E_1 + \ldots + E_p$ is called the *direct sum* of the spaces E_i $(i = 1 \ldots p)$.

1.11. Factor-space. Let E_1 be a subspace of E. Then an equivalence relation among the vectors of E can be defined in the following way: Two vectors x and x' are to be equivalent, $x \sim x'$, if $x' - x \in E_1$. This relation has indeed the three properties of an equivalence:

1. *Reflexivity*: $x \sim x$ for every $x \in E$, since $x - x = 0 \in E_1$.

2. *Commutativity*: $x \sim x'$ implies that $x' \sim x$: If $x' - x \in E_1$, then $x - x' = -(x' - x) \in E_1$.

3. *Transitivity*: $x \sim x'$ and $x' \sim x''$ implies that $x \sim x''$: If $x' - x \in E_1$ and $x'' - x' \in E_1$, then $x'' - x = (x'' - x') + (x' - x) \in E_1$.

An equivalence relation induces a decomposition of the whole space into classes of equivalent vectors. Two vectors x and x' of E are in the same class if and only if they are equivalent. Any two classes C_1 and C_2 are either disjoint or they coincide. In fact, assume that $x \in C_1 \cap C_2$. Then

$$x \sim x_1 \quad \text{and} \quad x \sim x_2$$

for every vector $x_1 \in C_1$ and every vector $x_2 \in C_2$. This implies in view of the transitivity that $x_1 \sim x_2$, whence $C_1 = C_2$.

Thus, every vector $x \in E$ is contained in exactly one class. This class will be denoted by $\bar{x}$. The class $\bar{0}$ containing the zero-vector coincides with the subspace E_1. It should be observed that this is the only class which is itself a linear subspace of E since the other classes do not contain the zero-vector.

To get a geometric picture of the above decomposition let E be a linear space of three dimensions and E_1 be a plane through 0. Then the corresponding classes are the planes parallel to E_1.

1.12. The linear structure of the factor-space. Consider the set of all equivalence classes with respect to E_1. This set can be made into a linear space by defining the linear operations as follows: Let $\bar{x}$ and $\bar{y}$ be two classes. Choose two vectors $x \in \bar{x}$ and $y \in \bar{y}$. Then the vector $x + y$ is contained in a certain class $\overline{x + y}$. This class does not depend on the choice of x and y but only on the classes $\bar{x}$ and $\bar{y}$. In fact, taking two other representatives x' and y', we have the relations

$$x' - x \in E_1 \quad \text{and} \quad y' - y \in E_1$$

and hence,

$$(x' + y') - (x + y) = (x' - x) + (y' - y) \in E_1 .$$

This implies that

$$x' + y' \sim x + y$$

and consequently, that

$$\overline{x' + y'} = \overline{x + y} .$$

The class $\overline{x + y}$ therefore is uniquely determined by the classes $\bar{x}$ and $\bar{y}$ and so it is proper to call it the sum of $\bar{x}$ and $\bar{y}$:

$$\bar{x} + \bar{y} = \overline{x + y} .$$

Similarly, the product $\lambda \bar{x}$ is defined as the equivalence class of the vector λx where x is any representative of $\bar{x}$,

$$\lambda \bar{x} = \overline{\lambda x} .$$

As in the case of the addition, it follows that this class depends only on the class $\bar{x}$.

It is easily verified that the two operations so defined satisfy the axioms listed in sec. 1.1 and 1.2. Thus, the set of all classes becomes a linear space, called the *factor-space* of E with respect to E_1 and denoted by E/E_1. It is also usual to call E/E_1 the *quotient-space* of E with respect to E_1. The zero-vector of the factor-space is the class $\bar{0}$.

If the subspace E_1 coincides with E, all vectors are equivalent and hence there is only the class $\bar{0}$. In this case the factor-space E/E_1 reduces to the zero-vector. If, in the opposite case, E_1 consists only of the zero-vector , two vectors of E are equivalent if and only if they are equal and so each class consists of exactly one vector. In this case the factor-space coincides with E.

Problems: 1. Let (ξ^1, ξ^2, ξ^3) be an arbitrary vector in R^3. Which of the following subsets are subspaces?

a) All vectors with $\xi^1 = \xi^2 = \xi^3$,

b) all vectors with $\xi^3 = 0$,

c) all vectors with $\xi^1 = \xi^2 - \xi^3$,

d) all vectors with $\xi^2 = 1$.

2. Let S be an arbitrary subset of E and $\bar{S}$ its linear closure. Show that $\bar{S}$ is the intersection of all linear subspaces of E containing S.

3. Let $E = E_1 \oplus E_2$ and $F = F_1 \oplus F_2$ be two direct decompositions of the spaces E and F. Show that the Cartesian product $E \times F$ can be directly decomposed as follows:

$$E \times F = E_1 \times F_1 \oplus E_2 \times F_2 .$$

4. Assume a direct decomposition $E = E_1 \oplus E_2$. Show that in each class of E with respect to E_1 there is exactly one vector of E_2.

5. Let E be a plane and E_1 a straight line through the origin. What is the geometrical meaning of the equivalence classes respect to E_1? Give a geometrical interpretation of the fact that $x \sim x'$ and $y \sim y'$ implies $x + y \sim x' + y'$.

§ 3. Linear spaces of finite dimension

1.13. Dimension. In general there are infinitely many linearly independent vectors in a linear space. For instance, in the space of all continuous functions $f(t)$ $(0 \leqq t \leqq 1)$ all the powers $t, t^2, t^3 \ldots$ are linearly independent. We shall be mainly concerned with linear spaces having only finitely many linearly independent vectors. The maximal number of linearly independent vectors of such a space E is called the *dimension* of E and will be denoted by $\dim E$.

A 1-dimensional linear space is called *a straight line* and a 2-dimensional linear space is called a *plane.*

Let E be an n-dimensional linear space and E_1 be a subspace of E. Since every system of linearly independent vectors of E_1 is also linearly independent in E, the dimension of E_1 can be at most equal to the dimension of E,

$$\dim E_1 \leqq \dim E .$$

It will be shown in the next section that the equality holds only if $E_1 = E$.

1.14. Basis. Let E be a linear space of dimension n. A system of n linearly independent vectors is called a *basis* of E. Then every vector x can be uniquely represented as a linear combination

$$x = \sum_\nu \xi^\nu x_\nu .$$

In fact, consider the $n + 1$ vectors $x_1 \ldots x_n, x$. These must be linearly dependent because there can be at most n linearly independent vectors in E. Therefore a relation

$$\sum_\nu \lambda^\nu x_\nu + \lambda x = 0 \tag{1.3}$$

holds with at least one coefficient different from zero. In particular, $\lambda \neq 0$, since otherwise (1.3) reduces to

$$\sum_\nu \lambda^\nu x_\nu = 0$$

which implies that $\lambda^\nu = 0 (\nu = 1 \ldots n)$. Thus the equation (1.3) can be solved with respect to x, yielding

$$x = -\frac{1}{\lambda} \sum_\nu \lambda^\nu x_\nu = \sum_\nu \xi^\nu x_\nu . \tag{1.4}$$

To show that the coefficients ξ^ν are uniquely determined by x, assume that there are two representations

$$x = \sum_\nu \xi^\nu x_\nu \quad \text{and} \quad x = \sum_\nu \eta^\nu x_\nu .$$

Then

$$\sum_\nu (\xi^\nu - \eta^\nu)\, x_\nu = 0$$

and since the vectors x_ν $(\nu = 1 \dots n)$ are linearly independent,

$$\xi^\nu = \eta^\nu \qquad (\nu = 1 \dots n) .$$

The coefficients ξ^ν are called the *components* of x with respect to the basis $(x_1 \dots x_n)$.

Now, let E_1 be a subspace of E having the same dimension as E. Then every basis of E_1 is also a basis of E. Hence, every vector $x \in E$ is contained in E_1 and thus $E_1 = E$.

1.15. The exchange theorem. If the vectors $(x_1 \dots x_n)$ form a basis of the space E, every subsystem is again linearly independent. Conversely, the question arises whether a given system of linearly independent vectors can be extended to a basis. The following *exchange theorem* of Steinitz asserts this. We prove it in the following form:

Let $(a_1 \dots a_p)$ be a system of linearly independent vectors and let $(x_1 \dots x_n)$ be a basis of E. Then it is possible to exchange certain p of the vectors x_ν by the vectors $a_1 \dots a_p$ such that the new system is again a basis of E.

The proof is based on the following *Lemma*: Let

$$a = \sum_\nu \xi^\nu x_\nu \tag{1.5}$$

be a vector of E and assume that its i^{th} component is different from zero. Then the vectors

$$x_1 \dots x_{i-1}, a, x_{i+1} \dots x_n \tag{1.6}$$

again form a basis of E.

To prove this we have to show that the vectors (1.6) are linearly independent.

Assume a relation

$$\sum_{\nu \neq i} \lambda^\nu x_\nu + \lambda a = 0 .$$

Replacing a by the linear combination (1.5) we obtain

$$\sum_{\nu \neq i} (\lambda^\nu + \lambda \xi^\nu)\, x_\nu + \lambda \xi^i x_i = 0$$

and using the linear independence of the vectors x_ν,

$$\lambda^\nu + \lambda \xi^\nu = 0 \quad (\nu \neq i) , \quad \lambda \xi^i = 0 .$$

Since $\xi^i \neq 0$, the last equation implies that $\lambda = 0$ and therefore the first equations reduce to $\lambda^\nu = 0$ $(\nu \neq i)$. Thus, the linear independence of the vectors (1.6) is established.

Now we can proceed with the actual proof of the exchange theorem. In the representation

$$a_1 = \sum_\nu \xi^\nu x_\nu$$

of the vector a_1, at least one $\xi^\nu \neq 0$. It is no restriction to assume that $\xi^1 \neq 0$ since this can be achieved by rearranging the vectors x_ν. It then follows from the above lemma that the vectors $(a_1, x_2 \ldots x_n)$ again form a basis of E. Hence, the vector a_2 can be written as

$$a_2 = \eta^1 a_1 + \sum_{\nu=2}^{n} \eta^\nu x_\nu .$$

Here at least one of the coefficients η^ν $(\nu = 2 \ldots n)$ has to be different from zero because otherwise the vectors a_1 and a_2 would be linearly dependent. We may assume that $\eta^2 \neq 0$. Then, by our lemma, the vectors $(a_1, a_2, x_3 \ldots x_n)$ form a basis of E. Applying this construction p times we finally obtain a basis of E containing all vectors a_ν $(\nu = 1 \ldots p)$.

1.16. Set of generators. Consider a finite set of vectors $(a_1 \ldots a_m)$ of the n-dimensional linear space E. The set $(a_1 \ldots a_m)$ is called a *system of generators* if every vector of E is a linear combination of the vectors a_μ $(\mu = 1 \ldots m)$.

Denote by r the maximal number of the linearly independent vectors among the vectors a_μ. It will be shown that r is equal to the dimension of E. By rearranging the vectors a_μ we can achieve that the vectors $a_1 \ldots a_r$ are linearly independent. Then every vector a_μ $(\mu = 1 \ldots m)$ is a linear combination of the vectors $a_1 \ldots a_r$. Thus, the vectors $a_1 \ldots a_r$ form again a system of generators. The linear independence of the vectors $a_1 \ldots a_r$ implies that $r \leqq n$. To prove the equality we proceed indirectly: Assume that $r < n$. Then the system $(a_1 \ldots a_r)$ can be extended to a basis $(a_1 \ldots a_r, a'_{r+1} \ldots a'_n)$ of E. Now consider the vector a'_n. This vector must be a linear combination of the vectors $(a_1 \ldots a_r)$ because these vectors generate E. Hence, the vectors $a_1 \ldots a_r, a'_n$ are linearly dependent which is a contradiction. Consequently, $r = n$.

It follows from the above result that every system of linearly independent generators forms a basis of E.

In particular, consider the n vectors

$$e_\nu = (\underbrace{0 \ldots 0, 1}_{\nu}, 0 \ldots 0) \quad (\nu = 1 \ldots n)$$

of the number-space R^n. Obviously, these vectors are linearly independent

Furthermore, they generate the space R^n since every vector $x = (\xi^1 \ldots \xi^n)$ can be written as

$$x = \sum_\nu \xi^\nu e_\nu .$$

Therefore the vectors e_ν $(\nu = 1 \ldots n)$ form a basis of R^n. This proves that the number-space R^n has indeed the dimension n.

1.17. Let E be an n-dimensional linear space and assume a direct decomposition

$$E = E_1 \oplus E_2 .$$

Then

$$\dim E = \dim E_1 + \dim E_2 .$$

To prove this, choose a basis $(x_1 \ldots x_p)$ of E_1 and a basis $(x_{p+1} \ldots x_{p+q})$ of E_2. Then the vectors $(x_1 \ldots x_{p+q})$ are again linearly independent. In fact, assume a relation

$$\sum_{\nu=1}^{p+q} \lambda^\nu x_\nu = 0 .$$

Then

$$\sum_{\nu=1}^{p} \lambda^\nu x_\nu = - \sum_{\nu=p+1}^{p+q} \lambda^\nu x_\nu$$

and consequently,

$$\sum_{\nu=1}^{p} \lambda^\nu x_\nu = 0 \quad \text{and} \quad \sum_{\nu=p+1}^{p+q} \lambda^\nu x_\nu = 0 .$$

These relations imply that $\lambda^\nu = 0$ $(\nu = 1 \ldots p + q)$.

At the same time the vectors $x_1 \ldots x_{p+q}$ generate the space E and hence they form a basis of E.

In the same way it is shown that

$$\dim (E_1 \oplus \cdots \oplus E_k) = \sum_{\nu=1}^{k} \dim E_\nu$$

for a direct decomposition into k subspaces.

1.18. The dimension of the factor-space. Consider a subspace E_1 of an n-dimensional linear space E. It will be shown that

$$\dim E/E_1 = \dim E - \dim E_1 .$$

Let $(x_1 \ldots x_p)$ be a basis of E_1 and extend it to a basis $(x_1 \ldots x_p \ldots x_n)$ of E. Then the vectors $(\bar{x}_{p+1} \ldots \bar{x}_n)$ form a basis of E/E_1. First of all, they are linearly independent. In fact, the relation

$$\sum_{\nu=p+1}^{n} \lambda^\nu \bar{x}_\nu = \bar{0}$$

states that the vector

$$\sum_{\nu=p+1}^{n} \lambda^\nu x_\nu$$

is contained in E_1. Therefore it can be represented as a linear combination of the vectors $(x_1 \ldots x_p)$,

$$\sum_{\nu=p+1}^{n} \lambda^\nu x_\nu = \sum_{\nu=1}^{p} \lambda^\nu x_\nu .$$

By the linear independence of the vectors x_ν $(\nu = 1 \ldots n)$ this equation implies that $\lambda^\nu = 0$ $(\nu = 1 \ldots n)$.

It remains to be shown that the vectors $(\bar{x}_{p+1} \ldots \bar{x}_n)$ generate the space E/E_1. Let $\bar{x}$ be an arbitrary vector of E/E_1 and x be a representative of the class $\bar{x}$. Then x can be written as

$$x = \sum_{\nu=1}^{p} \xi^\nu x_\nu + \sum_{\nu=p+1}^{n} \xi^\nu x_\nu .$$

Since the first sum is contained in E_1, it follows that

$$x \sim \sum_{\nu=p+1}^{n} \xi^\nu x_\nu$$

and passing now to the equivalence classes

$$\bar{x} = \overline{\sum_{\nu=p+1}^{n} \xi^\nu x_\nu} = \sum_{\nu=p+1}^{n} \xi^\nu \bar{x}_\nu .$$

Thus, $\bar{x}$ is a linear combination of the vectors $(\bar{x}_{p+1} \ldots \bar{x}_n)$, which completes our proof.

Our result shows that the factor-space has the dimension zero if $E_1 = E$. This is also immediately clear because then there is only one equivalence class, namely the class $\bar{0}$.

If, on the other hand, E_1 reduces to the zero vector, the factor-space has the dimension n. In fact, in this case, the space E/E_1 coincides with E.

Problems: 1. Given two subspaces $E_1 \subset E$ and $E_2 \subset E$ of finite dimension prove the relation

$$\dim E_1 + \dim E_2 = \dim(E_1 + E_2) + \dim(E_1 \cap E_2)$$

2. Prove that the dimension of the Cartesian product $E \times F$ is equal to the sum of the dimensions of E and of F,

$$\dim(E \times F) = \dim E + \dim F .$$

3. Let (x_1, x_2) be a basis of a 2-dimensional linear space. Show that the vectors

$$\bar{x}_1 = x_1 + x_2 , \quad \bar{x}_2 = x_1 - x_2$$

again form a basis. Let (ξ^1, ξ^2) and $(\bar{\xi}^1, \bar{\xi}^2)$ be the components of a vector

x relative to the bases (x_1, x_2) and $(\bar{x}_1, \bar{x}_2)$ respectively. Express the components $(\bar{\xi}^1, \bar{\xi}^2)$ in terms of the components (ξ^1, ξ^2).

4. Consider an n-dimensional complex linear space E. Since the multiplication with real coefficients in particular is defined in E, this space may also be considered as a *real* linear space. Let $(z_1 \ldots z_n)$ be a basis of E. Prove that the vectors $z_1 \ldots z_n, i z_1 \ldots i z_n$ form a basis of E if E is considered as a real linear space.

5. Let E be an n-dimensional real linear space and C the complex linear space as constructed in § 1, Problem 6. If x_ν $(\nu = 1 \ldots n)$ is a basis of E, prove that the vectors $(x_\nu, 0)$ $(\nu = 1 \ldots n)$ form a basis of C.

6. Let S be a set consisting of n elements x_ν $(\nu = 1 \ldots n)$ and $C(S)$ be the linear space generated by S (cf. sec. 1.3, example 3). Prove that the elements x_ν $(\nu = 1 \ldots n)$ form a basis of $C(S)$.

7. Consider the space R^n of n-tuples of numbers. Choose as basis the vectors:

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

Compute the components $\eta^1, \eta^2, \ldots, \eta^n$ of the vector $x = (\xi^1, \xi^2, \ldots, \xi^n)$ relative to the above basis. For which basis in R^n is the connection between the components of x and the scalars $\xi^1, \xi^2, \ldots, \xi^n$ particularly simple?

8. In R^4 consider the subspace T of all vectors $(\xi^1, \xi^2, \xi^3, \xi^4)$ satisfying $\xi^1 + 2\xi^2 = \xi^3 + 2\xi^4$. Show that the vectors: $x_1 = (1, 0, 1, 0)$ and $x_2 = (0, 1, 0, 1)$ are linearly independent and lie in T; then extend this set of two vectors to a basis of T.

9. Let $\alpha_1, \alpha_2, \alpha_3$ be fixed real numbers. Show that all vectors $(\eta^1, \eta^2, \eta^3, \eta^4)$ in R^4 obeying $\eta^4 = \alpha_1\eta^1 + \alpha_2\eta^2 + \alpha_3\eta^3$ form a subspace V. Show that V is generated by

$$x_1 = (1, 0, 0, \alpha_1)\,;\; x_2 = (0, 1, 0, \alpha_2)\,;\; x_3 = (0, 0, 1, \alpha_3)\,.$$

Verify that x_1, x_2, x_3 form a basis of the subspace V.

10. In the space P of all polynomials of degree $\leqq n - 1$ consider the two bases p_ν and q_ν defined by

$$\begin{aligned} p_\nu(t) &= t^\nu \\ q_\nu(t) &= (t - a)^\nu \quad (a, \text{ constant};\; \nu = 0, \ldots, n-1). \end{aligned}$$

Express the vectors q_ν explicitly in terms of the vectors p_ν.

Chapter II

Linear transformations

§ 1. Linear mappings

2.1. Definition. Consider two linear spaces E and F, both real or both complex. Assume that to every vector $x \in E$ a unique vector $y = \varphi x$ of F is assigned. Then φ is called a *mapping* of E into F. For every $x \in E$ the vector φx is called the *image* of x. A mapping φ is called *linear* if

a) $\varphi(x_1 + x_2) = \varphi x_1 + \varphi x_2$.

b) $\varphi(\lambda x) = \lambda \varphi x$.

These two conditions are equivalent to

c) $\varphi(\lambda x_1 + \mu x_2) = \lambda \varphi x_1 + \mu \varphi x_2$.

Insertion of $\lambda = 0$ in (b) yields $\varphi 0 = 0$ showing that the zero-vector of E is mapped into the zero-vector of F. If $(x_1 \ldots x_p)$ is a system of linearly dependent vectors in E, then the image-vectors $(\varphi x_1 \ldots \varphi x_p)$ are again linearly dependent. In fact, assume that

$$\sum_\nu \lambda^\nu x_\nu = 0$$

with at least one $\lambda^\nu \neq 0$. Then, according to (c)

$$\sum_\nu \lambda^\nu \varphi x_\nu = 0$$

and hence the vectors $(\varphi x_1 \ldots \varphi x_p)$ are linearly dependent.

On the other hand, linearly independent vectors of E are not necessarily mapped into linearly independent vectors. For instance, the mapping φ, defined by $\varphi x = 0$ for every $x \in E$, maps every system of vectors into the linearly dependent zero-vector.

2.2. Kernel. The set of all vectors $x \in E$ such that $\varphi x = 0$ forms a linear subspace of E. This subspace is called the *kernel* of φ and will be denoted by $K(\varphi)$ or simply by K. As an example consider a subspace E_1 of E and define a mapping $\pi: E \to E/E_1$ by,

$$\pi x = \bar{x}$$

where $\bar{x}$ is the equivalence-class of x. The linear mapping thus obtained is called the *canonical projection* of E onto E/E_1. The kernel of π is the subspace E_1.

A linear mapping $\varphi: E \to F$ is called *regular* if its kernel reduces to the zero-vector. A regular mapping sends every system of linearly independent vectors of E into a system of linearly independent vectors of F. In fact, let $(x_1 \ldots x_p)$ be such a system and assume a relation

$$\sum_\nu \lambda^\nu \varphi x_\nu = 0 .$$

This can be written as

$$\varphi\left(\sum_\nu \lambda^\nu x_\nu\right) = 0$$

and hence it implies that

$$\sum_\nu \lambda^\nu x_\nu = 0$$

whence

$$\lambda^\nu = 0 \qquad (\nu = 1 \ldots p) .$$

A regular linear mapping is a *one-to-one mapping*, i. e. different vectors of E are mapped into different vectors of F.

2.3. Image-space. The set $\varphi(E)$ of all vectors φx $(x \in E)$ is called the *image-space* of E. If the image-space coincides with the entire space F, φ is called a mapping of E *onto* F.

Now assume that the spaces E and F have finite dimension. Then the dimension of the image-space is called the *rank* of φ and will be denoted by $r(\varphi)$. The inclusion $\varphi(E) \subset F$ shows that

$$r(\varphi) \leqq \dim F .$$

Equality holds if and only if φ is a mapping *onto* F. The fact that linearly dependent vectors of E are mapped into linearly dependent vectors of F implies that

$$r(\varphi) \leqq \dim E .$$

Altogether we see that the rank of a linear mapping $\varphi: E \to F$ never exceeds the dimension of E or the dimension of F.

2.4. Isomorphisms. A regular linear mapping φ of E onto F is also called an *isomorphism* of E onto F. If φ is an isomorphism of E onto F, to every vector $y \in F$ there exists exactly one vector $x \in E$ such that $y = \varphi x$. Hence a linear mapping $\varphi^{-1}: F \to E$ can be defined by

$$\varphi^{-1} y = x .$$

This mapping is obviously again an isomorphism. It is called the *inverse isomorphism* of φ. The inverse isomorphism of φ^{-1} is again the isomorphism φ. Two linear spaces E and F are called *isomorphic* if there exists an isomorphism of E onto F.

Now let E and F be two linear spaces of finite dimension and φ be an isomorphism of E onto F. Then

$$\dim F = \dim \varphi(E) \leqq \dim E$$

and

$$\dim E = \dim \varphi^{-1}(F) \leqq \dim F .$$

These relations imply that

$$\dim E = \dim F$$

showing that isomorphic spaces have the same dimension. Conversely, two linear spaces of the same dimension are isomorphic. It is sufficient to show that every linear space E of dimension n is isomorphic to the number-space R^n. Let $(x_1 \ldots x_n)$ be a basis of E and define the mapping $\varphi: E \to R^n$ by

$$\varphi x = (\xi^1 \ldots \xi^n)$$

where

$$x = \sum_\nu \xi^\nu x_\nu .$$

This mapping is obviously an isomorphism of E onto R^n.

2.5. The induced mapping. Let φ be a linear mapping of E into F and let $K = K(\varphi)$ be the kernel of φ. Then a mapping

$$\bar{\varphi}: E/K \to F$$

can be defined as follows: Consider an equivalence class $\bar{x}$ with respect to the subspace K. Choose a vector $x \in \bar{x}$ and define $\bar{\varphi}$ by

$$\bar{\varphi}\,\bar{x} = \varphi\, x . \tag{2.1}$$

This definition does not depend on the choice of x. In fact, if x_1 and x_2 are two representatives of $\bar{x}$, then $x_2 - x_1 \in K$ whence $\varphi x_2 = \varphi x_1$. The mapping $\bar{\varphi}$, defined by (2.1) is obviously linear. Moreover $\bar{\varphi}$ is regular: $\bar{\varphi}\bar{x} = 0$ implies that $\bar{0} \in K$ and hence $\bar{x} = x$. The image-spaces $\bar{\varphi}(E/K)$ and $\varphi(E)$ coincide, as follows from the definition of $\bar{\varphi}$. Hence, $\bar{\varphi}$ is an isomorphism of E/K onto $\varphi(E)$ and we can state the following theorem: *Every linear mapping $\varphi: E \to F$ induces an isomorphism $\bar{\varphi}$ of the factor-space E/F onto $\varphi(E)$.*

Let us now assume that the spaces E and F have finite dimension. Then it follows from the above isomorphism that

$$\dim \varphi(E) = \dim E/K .$$

But

$$\dim E/K = \dim E - \dim K$$

and we thus obtain the relation

$$\dim \varphi(E) = \dim E - \dim K . \tag{2.2}$$

If E and F have the same dimension the relation (2.2) shows that $\varphi(E) = F$ if and only if K reduces to the zero-vector. In other words:

A linear mapping of an n-dimensional linear space E into an n-dimensional linear space F is a mapping onto F if and only if it is regular.

Problems: 1. Given two subspaces E_1 and E_2 of E define an isomorphism of $(E_1 + E_2)/E_1$ onto $E_2/(E_1 \cap E_2)$.

Now assume that the spaces E_1 and E_2 have finite dimension and derive the relation

$$\dim E_1 + \dim E_2 = \dim (E_1 + E_2) + \dim E_1 \cap E_2 .$$

2. Let x_ν $(\nu=1\ldots n)$ be a basis of E and y_ν $(\nu=1\ldots n)$ be an arbitrary system of n vectors of F. Show that there is exactly one linear mapping $\varphi\colon E\to F$ such that $\varphi x_\nu = y_\nu$ $(\nu = 1\ldots n)$.

3. Assume a decomposition

$$E = E_1 + E_2 .$$

Consider the Cartesian product $E_1 \times E_2$ and define the mapping

$$\varphi\colon E_1 \times E_2 \to E$$

by

$$\varphi(x_1, x_2) = x_1 + x_2, \quad x_1 \in E_1, x_2 \in E_2 .$$

Prove that the kernel of φ is the subspace of $E_1 \times E_2$ consisting of the pairs $(x, -x)$ where $x \in E_1 \cap E_2$. Show that φ is an isomorphism if and only if the decomposition $E = E_1 + E_2$ is direct.

4. Given two linear spaces E and F, consider a subspace E_1 of E and a subspace F_1 of F and the canonical projections

$$\pi\colon E \to E/E_1 \quad \text{and} \quad \pi'\colon F \to F/F_1 .$$

Define the mapping

$$\varphi\colon E \times F \to E/E_1 \times F/F_1$$

by

$$\varphi(x, y) = (\pi x, \pi' y) .$$

Show that $\bar{\varphi}$ induces an isomorphism

$$\bar{\varphi}\colon (E \times F)/(E_1 \times F_1) \to E/E_1 \times F/F_1 .$$

5. Let S and T be two arbitrary sets and φ be an arbitrary mapping of S into T. Prove that φ induces a linear mapping

$$\varphi_*\colon C(S) \to C(T)$$

(cf. sec. 1.3, example 3) defined by

$$\varphi_* \sum_{a\in A} f(a)\cdot a = \sum_{a\in A} f(a)\, \varphi a .$$

6. Let C be a space of all continuous functions and define a linear mapping $\varphi\colon C \to C$ by

$$\varphi\colon f(t) \to \int_0^t f(s)\, ds .$$

Prove that the image-space consists of all continuously differentiable functions.

7. Consider the vector space of all real valued continuous functions defined in the interval $a \leqq t \leqq b$. Show that the mapping

$$\varphi\colon x(t) \to {}^{t\cdot x(t)}$$

is linear.

§ 2. Dual spaces

Throughout § 2 and § 3 all linear spaces are assumed to have finite dimension.

2.6. The scalar-product. Consider two linear spaces E^* and E, both real or both complex. Assume that to every pair $\overset{*}{x}, x$ ($\overset{*}{x} \in E^*$, $x \in E$) a scalar*) is assigned, written as $\langle \overset{*}{x}, x \rangle$, and subject to the conditions

(I) $\langle \lambda \overset{*}{x} + \mu \overset{*}{y}, x \rangle = \lambda \langle \overset{*}{x}, x \rangle + \mu \langle \overset{*}{y}, x \rangle$
$\langle \overset{*}{x}, \lambda x + \mu y \rangle = \lambda \langle \overset{*}{x}, x \rangle + \mu \langle \overset{*}{x}, y \rangle$.

(II) $\langle \overset{*}{x}, x \rangle = 0$ for a fixed $\overset{*}{x} \in E^*$ and all $x \in E$ implies that $\overset{*}{x} = 0$ and conversely, $\langle \overset{*}{x}, x \rangle = 0$ for a fixed $x \in E$ and all $\overset{*}{x} \in E^*$ implies that $x = 0$.

A function in E^* and E having the properties (I) and (II) is called a *scalar-product*. Two vectors $\overset{*}{x} \in E^*$ and $x \in E$ are called *orthogonal* if $\langle \overset{*}{x}, x \rangle = 0$. The condition (II) states that a vector of one space is orthogonal to all vectors of the other space only if it is the zero-vector. Two linear spaces in which a scalar-product is defined are called *dual spaces*.

Dual spaces have the same dimension. To prove this, let $x_\nu (\nu = 1 \ldots n)$ be a basis of E. Define the linear mapping $\varphi: E^* \to R^n$ by

$$\varphi \overset{*}{x} = (\langle \overset{*}{x}, x_1 \rangle \ldots \langle \overset{*}{x}, x_n \rangle) . \tag{2.3}$$

It follows from condition (II) that φ is regular. This implies that

$$\dim E^* \leqq n = \dim E .$$

Interchanging E^* and E we obtain

$$\dim E \leqq \dim E^*$$

whence

$$\dim E = \dim E^*. \tag{2.4}$$

Now it follows that the mapping (2.3) is an isomorphism of E onto R^n. In other words, to every n-tuple of scalars ξ_ν $(\nu = 1 \ldots n)$ there exists exactly one vector $\overset{*}{x} \in E^*$ such that

$$\langle \overset{*}{x}, x_\nu \rangle = \xi_\nu \qquad (\nu = 1 \ldots n) .$$

2.7. Dual bases. Two bases $\overset{*}{x}{}^\nu$ and x_ν of E^* and E respectively are called *dual* if the relations

$$\langle \overset{*}{x}{}^\nu, x_\mu \rangle = \begin{cases} 1 & \text{if} \quad \nu = \mu \\ 0 & \text{if} \quad \nu \neq \mu \end{cases} \quad (\nu, \mu = 1 \ldots n) \tag{2.5}$$

are valid. Introducing the *Kronecker-symbol*

$$\delta^\nu_\mu = \begin{cases} 1 & \text{if} \quad \nu = \mu \\ 0 & \text{if} \quad \nu \neq \mu \end{cases}$$

*) "scalar" means a real or a complex number, depending on whether E^* and E are real or complex spaces.

we can write (2.5) in the form

$$\langle \overset{*}{x}{}^{\nu}, x_{\mu} \rangle = \delta^{\nu}_{\mu} .$$

Given a pair of dual bases $\overset{*}{x}{}^{\nu}, x_{\nu}$ $(\nu = 1 \ldots n)$ the scalar-product of two vectors

$$\overset{*}{x} = \sum_{\nu} \xi_{\nu} \overset{*}{x}{}^{\nu} \quad \text{and} \quad x = \sum_{\nu} \xi^{\nu} x_{\nu}$$

appears as the bilinear form

$$\langle \overset{*}{x}, x \rangle = \sum_{\nu} \xi_{\nu} \overset{*}{\xi}{}^{\nu} .$$

Replacing $\overset{*}{x}$ by the ν-th basis-vector $\overset{*}{x}{}^{\nu}$ we obtain

$$\langle \overset{*}{x}{}^{\nu}, x \rangle = \xi^{\nu} \qquad (\nu = 1 \ldots n) .$$

This formula shows that the components of a vector x relative to the basis x_{ν} are equal to the scalar-products of x with the dual basis-vectors $\overset{*}{x}{}^{\nu}$.

To every given basis x_{ν} $(\nu = 1 \ldots n)$ of E there exists exactly one dual basis. To prove this let μ $(1 \leqq \mu \leqq n)$ be a fixed integer. As it has been shown at the end of sec. 2.6, there exists exactly one vector $\overset{*}{x}{}^{\mu} \in E^*$ such that

$$\langle \overset{*}{x}{}^{\mu}, x_{\nu} \rangle = \delta^{\mu}_{\nu} \qquad (\nu = 1 \ldots n) .$$

The n vectors $\overset{*}{x}{}^{\mu}$ $(\mu = 1 \ldots n)$ thus obtained are linearly independent. In fact, the relation

$$\sum_{\mu} \lambda_{\mu} \overset{*}{x}{}^{\mu} = 0$$

implies that

$$\sum_{\mu} \lambda_{\mu} \langle \overset{*}{x}{}^{\mu}, x_{\nu} \rangle = 0 ,$$

whence $\lambda_{\mu} = 0$ $(\mu = 1 \ldots n)$.

2.8. Orthogonal complement. Let E_1 be a given subspace of E. Then the set of vectors $\overset{*}{x} \in E^*$ orthogonal to all vectors of E_1 is a subspace of E^*. This space is called the *orthogonal complement* of E_1 and will be denoted by $E_1^{\perp}$.

The spaces E_1 and $E_1^{\perp}$ have complementary dimension,

$$\dim E_1 + \dim E_1^{\perp} = \dim E .$$

In fact, let $(x_1 \ldots x_p)$ be a basis of E_1 and extend this basis to a basis $(x_1 \ldots x_p \ldots x_n)$ of E. Then the vectors $(\overset{*}{x}{}^{p+1} \ldots \overset{*}{x}{}^{n})$ of the dual basis form a basis of $E_1^{\perp}$.

The orthogonal complement of $E_1^\perp$ is again the subspace of E which contains E_1. Comparing the dimensions of $(E_1^\perp)^\perp$ and E_1 we find that

$$\dim (E_1^\perp)^\perp = \dim E - \dim (E_1)^\perp = \dim E_1 .$$

Hence $(E_1^\perp)^\perp = E_1$.

Now consider a direct decomposition

$$E = E_1 \oplus E_2 .$$

Then the orthogonal complements $E_1^\perp$ and $E_2^\perp$ obviously form a direct decomposition of E^*,

$$E^* = E_1^\perp \oplus E_2^\perp .$$

Moreover, the spaces $E_1, E_2^\perp$ and $E_2, E_1^\perp$ are dual pairs. In fact, let $y \in E_1$ be a vector such that

$$\langle \overset{*}{z}, y \rangle = 0$$

for all vectors $\overset{*}{z} \in E_2^\perp$. Then $y \in (E_2^\perp)^\perp = E_2$ and, consequently $y \in E_1 \cap E_2$. But $E_1 \cap E_2 = 0$, whence $y = 0$. Conversely, assume that $\overset{*}{z}$ is a vector of $E_2^\perp$ such that

$$\langle \overset{*}{z}, y \rangle = 0$$

for all vectors $y \in E_1$. Then $\overset{*}{z} \in E_1^\perp \cap E_2^\perp$, whence $\overset{*}{z} = 0$. This proves the duality of E_1 and $E_2^\perp$.

2.9. The conjugate space. A *linear function* in the linear space E is a linear mapping of E into the number-space R^1. In other words, a linear function f assigns to every vector $x \in E$ a scalar $f(x)$ such that

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) .$$

The subspace of the vectors $x \in E$ such that $f(x) = 0$ has the dimension $n - 1$ unless f is identically zero. In fact, if $f \neq 0$, the mapping $f\colon E \to R^1$ is a mapping *onto* R^1. Consequently its kernel has the dimension $n - 1$.

Consider the set $L(E)$ of all linear functions in E. Define the sum of two functions f and g by

$$(f + g)(x) = f(x) + g(x)$$

and the function λf by

$$(\lambda f)(x) = \lambda f(x) .$$

Under these operations the set $L(E)$ becomes a linear space called the *conjugate space* of E. The zero-vector of $L(E)$ is the function which is identically zero.

Now consider the bilinear function $\langle f, x\rangle$ defined by

$$\langle f, x\rangle = f(x) \quad f \in L(E), x \in E . \tag{2.6}$$

It will be shown that this bilinear function defines a scalar-product in the spaces $L(E)$ and E. The condition (I) is obviously satisfied. Next, assume that $\langle f, x\rangle = 0$ for a fixed linear function f and all vectors $x \in E$. Then f is the zero-vector of $L(E)$. Finally, let a be a vector of E with the property that $\langle f, a\rangle = 0$ for all functions $f \in L(E)$. To prove that $a = 0$ we proceed indirectly: Assume that $a \neq 0$. Then there exists a basis a_ν $(\nu = 1 \ldots n)$ of E such that $a_1 = a$. Define the linear function f by

$$f(x) = \xi^1$$

where

$$x = \sum_\nu \xi^\nu a_\nu .$$

Then

$$f(a) = f(a_1) = 1$$

which is a contradiction.

Hence, the bilinear function (2.6) is a scalar-product and the spaces $L(E)$ and E are dual. In particular,

$$\dim L(E) = \dim E .$$

2.10. The isomorphism between E^* and $L(E)$. Given an n-dimensional linear space E the space $L(E)$ is essentially the only dual space of E. More precisely, if E^* is an arbitrary dual space of E there exists a natural isomorphism of E^* onto $L(E)$. To construct this isomorphism let a^* be a vector of E^*. Then a linear function f_{a^*} in E is defined by

$$f_{a^*}(x) = \langle a^*, x\rangle .$$

The correspondence $a^* \to f_{a^*}$ defines a linear mapping of E^* into $L(E)$. This mapping is regular: $f_{a^*} = 0$ implies that

$$\langle a^*, x\rangle = 0 \qquad \text{for every} \quad x \in E$$

and hence, according to condition (II) of sec. 2.6., that $a^* = 0$. Since the spaces E^* and $L(E)$ have the same dimension the mapping $a^* \to f_{a^*}$ defines an isomorphism of E^* onto $L(E)$. In other words, every linear function f in E can be written as

$$f(x) = \langle a^*, x\rangle \qquad a^* \in E^*$$

and the vector a^* is uniquely determined by f.

Problems: 1. Given two dual spaces E^* and E consider a subspace E_1 of E. Show that the spaces $E_1^\perp$ and E/E_1 are dual with respect to the scalar product defined by

$$\langle \overset{*}{x}, \bar{x}\rangle = \langle \overset{*}{x}, x\rangle \qquad \overset{*}{x} \in E_1^\perp,\ \bar{x} \in E/E_1,\ x \in \bar{x} .$$

2. Let U be a subspace of E and V^* be a subspace of E^*. Show that the bilinear function $\langle \overset{*}{x}, x\rangle$ defines a scalar product in V^* and U if and only if $U \cap V^{*\perp} = 0$ and $U^\perp \cap V^* = 0$.

3. Given two pairs of dual spaces E^*, E and F^*, F prove that the spaces $E^* \times F^*$ and $E \times F$ are dual with respect to the bilinear function

$$\langle (\overset{*}{x}, \overset{*}{y}), (x, y) \rangle = \langle \overset{*}{x}, x \rangle + \langle \overset{*}{y}, y \rangle .$$

4. Consider two subspaces E_1 and E_2 of E. Establish the relations

$$(E_1 + E_2)^{\perp} = E_1^{\perp} \cap E_2^{\perp}$$

and

$$(E_1 \cap E_2)^{\perp} = E_1^{\perp} + E_2^{\perp} .$$

5. Let U be a subspace of E and V^* be a subspace of E^*. Prove the relation

$$\dim (U^{\perp} \cap V^*) + \dim U = \dim (U \cap V^{*\perp}) + \dim V^* .$$

6. Show that there is a natural isomorphism between a linear space E and its second conjugate space $L(L(E))$. More precisely, associate with every vector $a \in E$ the linear function Φ_a in $L(E)$ which is defined by

$$\Phi_a(f) = f(a) \qquad f \in L(E) .$$

Prove that the correspondence $a \to \Phi_a$ defines an isomorphism of E onto $L(L(E))$.

7. Given a pair of dual bases $\overset{*}{x}{}^{\nu}, x_{\nu}$ $(\nu = 1 \ldots n)$ show that the bases $\left(\overset{*}{x}{}^1 + \sum_{\nu=2}^{n} \lambda_{\nu} \overset{*}{x}{}^{\nu}, \overset{*}{x}{}^2 \ldots \overset{*}{x}{}^n\right)$ and $(x_1, x_2 - \lambda_2 x_1, \ldots x_n - \lambda_n x_1)$ are again dual.

§ 3. Dual mappings

2.11. Definition. Consider two pairs of dual spaces E^*, E and F^*, F. Two linear mappings

$$\varphi: E \to F \quad \text{and} \quad \varphi^*: F^* \to E^*$$

are called *dual,* if

$$\langle \overset{*}{y}, \varphi x \rangle = \langle \varphi^* \overset{*}{y}, x \rangle \quad \text{for all } x \in E \text{ and all } \overset{*}{y} \in F^* .$$

To every linear mapping φ of E into F there exists exactly one dual mapping. We show first the uniqueness of φ^*. Assume two mappings φ_1^* and φ_2^* such that

$$\langle \overset{*}{y}, \varphi x \rangle = \langle \varphi_1^* \overset{*}{y}, x \rangle \quad \text{and} \quad \langle \overset{*}{y}, \varphi x \rangle = \langle \varphi_2^* \overset{*}{y}, x \rangle .$$

Then

$$\langle \varphi_1^* \overset{*}{y} - \varphi_2^* \overset{*}{y}, x \rangle = 0$$

for all $x \in E$. According to condition (II) (sec. 2.6) this implies that

$$\varphi_1^* \overset{*}{y} = \varphi_2^* \overset{*}{y} .$$

To prove the existence of φ^* consider a fixed vector b^* in F^*. Then a linear function f_{b^*} in E is defined by

$$f_{b^*}(x) = \langle b^*, \varphi x\rangle . \tag{2.7}$$

As it has been shown in sec. 2.10 the function f_{b^*} can be written as

$$f_{b^*}(x) = \langle a^*, x\rangle \qquad a^* \in E^* \tag{2.8}$$

where the vector a^* is uniquely determined. Define the mapping φ^* by

$$\varphi^* b^* = a^* . \tag{2.9}$$

Eliminating $f_{b^*}(x)$ and a^* from the equations (2.7), (2.8) and (2.9) we obtain the relation

$$\langle b^*, \varphi x\rangle = \langle \varphi^* b^*, x\rangle .$$

2.12. The relation between kernel and image-space. Consider two dual mappings

$$\varphi: E \to F \quad \text{and} \quad \varphi^*: F^* \to E^* .$$

We shall prove that

$$K(\varphi^*) = \varphi(E)^{\perp} \quad \text{and} \quad K(\varphi) = \varphi^*(F^*)^{\perp} . \tag{2.10}$$

Let b^* be a vector of $K(\varphi^*)$. Then

$$\langle b^*, \varphi x\rangle = \langle \varphi^* b^*, x\rangle = 0 \quad \text{for every} \quad x \in E .$$

Hence, $K(\varphi^*)$ is contained in $\varphi(E)^{\perp}$. Conversely, assume that $b^* \in \varphi(E)^{\perp}$. Then, for every $x \in E$,

$$\langle \varphi^* b^*, x\rangle = \langle b^*, \varphi x\rangle = 0$$

whence $\varphi^* b^* = 0$. So $\varphi(E)^{\perp}$ is contained in $K(\varphi^*)$. This proves the first relation (2.10). The second relation (2.10) is obtained by interchanging φ and φ^*.

Passing over to the orthogonal complements we obtain from the relations (2.10)

$$\varphi(E) = K(\varphi^*)^{\perp} \quad \text{and} \quad \varphi^*(F^*) = K(\varphi)^{\perp}$$

and the following theorem is thereby proved: *A vector $b \in F$ is contained in the image-space $\varphi(E)$ if and only if b is orthogonal to all vectors of kernel $K(\varphi^*)$.*

2.13. The relation between the ranks of φ and φ^*. Comparing the dimensions in the first relation (2.10) we obtain

$$\dim K(\varphi^*) = \dim F - \dim \varphi(E) . \tag{2.11}$$

At the same time, the relation (2.2) applied to the mapping φ^* yields

$$\dim \varphi^*(F^*) = \dim F^* - \dim K(\varphi^*) . \tag{2.12}$$

(2.11) and (2.12) imply the relation

$$\dim \varphi(E) = \dim \varphi^*(F^*)$$

showing that *dual mappings have the same rank.*

Problems: 1. Let a^* and b be two vectors of E^* and F respectively. Define the mapping $\varphi: E \to F$ by

$$\varphi x = \langle a^*, x\rangle\, b \qquad x \in E .$$

Show that

$$\varphi^* \overset{*}{y} = \langle \overset{*}{y}, b\rangle\, a^* \qquad \overset{*}{y} \in F^*.$$

2. Assuming that φ is an isomorphism of E onto F show that φ^* is an isomorphism of F^* onto E^* and that $(\varphi^*)^{-1} = (\varphi^{-1})^*$.

3. Let φ be a linear mapping of E into F. Consider the spaces $L(E)$ and $L(F)$ as the duals of E and F respectively. Prove that the dual mapping φ^* then is defined by

$$(\varphi^* g)(x) = g(\varphi x) \qquad g \in L(F), x \in E .$$

4. Let E and F be two linear spaces of the same dimension. Assume that two linear mappings $\varphi: E \to F$ and $\psi: E^* \to F^*$ are given such that

$$\langle \psi \overset{*}{x}, \varphi x\rangle = \langle \overset{*}{x}, x\rangle \qquad \overset{*}{x} \in E^*, x \in E .$$

Prove that φ is an isomorphism of E onto F and that $\psi = (\varphi^*)^{-1}$.

§ 4. Sum and product of linear mappings

2.14. The linear space $L(E, F)$. Given two linear spaces E and F, consider the set $L(E; F)$ of all linear mappings φ of E into F. By the definitions

$$(\varphi_1 + \varphi_2)\, x = \varphi_1 x + \varphi_2 x$$

and

$$(\lambda \varphi)\, x = \lambda\, \varphi\, x$$

the set $L(E; F)$ is made into a linear space. The zero-vector of this space is the mapping φ defined by $\varphi x = 0$ for every $x \in E$. If E and F have finite dimension, the dimension of the space $L(E; F)$ is given by

$$\dim L(E; F) = \dim E \cdot \dim F .$$

To prove this let $x_\nu\ (\nu = 1 \ldots n)$ and $y_\mu\ (\mu = 1 \ldots m)$ be a basis of E and F respectively. Define the nm mappings $\varphi_\varkappa^\lambda\ (\lambda = 1 \ldots n; \varkappa = 1 \ldots m)$ by

$$\varphi_\varkappa^\lambda\, x_\nu = \delta_\nu^\lambda\, y_\varkappa .$$

We leave it to the reader to show that the mappings $\varphi_\varkappa^\lambda$ form a system of linearly independent generators of the space $L(E; F)$ and hence form a basis.

2.15. Product of linear mappings. Let three linear spaces E, F, G be given. If φ is a linear mapping of E into F and ψ is a linear mapping of F into G, then a linear mapping of E into G, denoted as $\psi \circ \varphi$, can be defined by

$$(\psi \circ \varphi)\, x = \psi(\varphi x) \qquad x \in E$$

The mapping $\psi \circ \varphi$ is called the *product* of the mappings φ and ψ. It follows immediately from the above definition that the product of linear mappings is distributive and associative,

$$(\lambda \psi_1 + \mu \psi_2) \circ \varphi = \lambda \psi_1 \circ \varphi + \mu \psi_2 \circ \varphi$$

$$\psi \circ (\lambda \varphi_1 + \mu \varphi_2) = \lambda \psi \circ \varphi_1 + \mu \psi \circ \varphi_2$$

$$\chi \circ (\psi \circ \varphi) = (\chi \circ \psi) \circ \varphi$$

$$\varphi: E \to F\,, \quad \psi: F \to G\,, \quad \chi: G \to H\,.$$

If φ is an isomorphism of E onto F and ψ is an isomorphism of F onto G then $\psi \circ \varphi$ is an isomorphism of E onto G. It is immediately clear that

$$(\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1}.$$

The dual of the product-mapping $\psi \circ \varphi$ is given by the formula

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^*\,. \tag{2.13}$$

In fact, let E^*, F^* and G^* be dual spaces of E, F and G respectively. Then

$$\langle (\varphi^* \circ \psi^*)\, \overset{*}{z}, x \rangle = \langle \psi^* z, \varphi x \rangle = \langle z, (\psi \circ \varphi) \rangle \qquad \overset{*}{z} \in G^*, x \in E$$

whence (2.13).

2.16. Endomorphisms. A linear mapping φ of E into itself is called an *endomorphism* of E. Since the product of two endomorphisms φ and ψ is again an endomorphism of E, the multiplication $(\varphi, \psi) \to \psi \circ \varphi$ associates with any two endomorphisms of E a third endomorphism. This multiplication has the following properties:

1. Associative law: $\quad \chi \circ (\psi \circ \varphi) = (\chi \circ \psi) \circ \varphi\,.$
2. Distributive laws: $\quad (\lambda \psi_1 + \mu \psi_2) \circ \varphi = \lambda \psi_1 \circ \varphi + \mu \psi_2 \circ \varphi$
   $\quad \psi \circ (\lambda \varphi_1 + \mu \varphi_2) = \lambda \psi \circ \varphi_1 + \mu \psi \circ \varphi_2.$

3. There exists an endomorphism ι (the identity-map) such that $\varphi \circ \iota = \iota \circ \varphi = \varphi$ for every endomorphism φ.

The above product is not commutative, i. e. in general $\psi \circ \varphi \neq \varphi \circ \psi$. As an example, consider the two endomorphisms φ and ψ of a 2-dimensional space E which are defined by

$$\varphi x_1 = \alpha x_1 \quad \psi x_1 = x_2$$

$$\varphi x_2 = \beta x_2 \quad \psi x_2 = x_1$$

where (x_1, x_2) is a basis of E. Then $(\psi \circ \varphi)\, x_1 = \alpha x_2$ and $(\varphi \circ \psi)\, x_1 = \beta\, x_2$ whence $\psi \circ \varphi \neq \varphi \circ \psi$ if $\alpha \neq \beta$.

2.17. Automorphisms. A regular endomorphism of a linear space E is called an *automorphism* of E. The automorphisms form a subset of the space $L(E; E)$. This subset is not a sub-space because the sum of two automorphisms is not necessarily an automorphism. However, the product of two automorphisms is an automorphism and the inverse of every automorphism is again an automorphism.

Denote by $GL(E)$ the set of all automorphisms of E. By assigning to any two automorphisms φ and ψ the automorphism $\psi \circ \varphi$ we obtain a multiplication in the set $GL(E)$ which has the following properties:

1. Associative law: $\chi \circ (\psi \circ \varphi) = (\chi \circ \psi) \circ \varphi$.
2. There exists an element ι (the identity-map) such that $\varphi \circ \iota = \iota \circ \varphi$ for every $\varphi \in GL(E)$.
3. To every φ there exists an element φ^{-1} such that

$$\varphi^{-1} \circ \varphi = \varphi \circ \varphi^{-1} = \iota \,.$$

An arbitrary set in which a multiplication is defined subject to the above conditions is called a *group*. The group $GL(E)$ is non-commutative, i. e. the products $\psi \circ \varphi$ and $\varphi \circ \psi$ are generally different.

2.18. Associative Algebras. Let A be a linear space and assume that to any two vectors $x \in A$ and $y \in A$ a third vector $xy \in A$, called the *product* of x and y is assigned subject to the following conditions:

$$(\lambda x_1 + \mu x_2) y = \lambda x_1 y + \mu x_2 y \,,$$
$$x(\lambda y_1 + \mu y_2) = \lambda x y_1 + \mu x y_2 \,,$$
$$x(yz) = (xy)z \,.$$

Then A is called an *associative algebra*. An element $e \in A$ having the property that

$$ex = xe = x$$

for every vector $x \in A$ is called a *unit-element*. If a unit-element exists in A it is uniquely determined. In fact, assume that e' is another unit-element. Then

$$ex = x \quad \text{and} \quad xe' = x$$

for every vector $x \in A$. Inserting $x = e'$ in the first and $x = e$ in the second equation we find that

$$ee' = e' \quad \text{and} \quad ee' = e$$

whence $e' = e$.

A vector $x \in A$ is called *invertible* if there exists a vector x^{-1} with the property that

$$xx^{-1} = x^{-1}x = e \,.$$

The invertible vectors of A form a group with respect to the multiplication. If all vectors $x \neq 0$ are invertible, A is called a *division-algebra.*

The space $L(E, E)$ of all endomorphisms of a linear space with the multiplication defined as in sec. 2.16 is an associative algebra with the identity-map as unit-element. The invertible elements of this algebra are the automorphisms of E.

Problems: 1. A *projection* is an endomorphism $\pi\colon E \to E$ such that $\pi^2 = \pi$. Show that for every projection the direct decomposition

$$E = \pi(E) \oplus K(\pi)$$

holds.

2. Consider a pair of dual spaces E^* and E and a projection π. Prove that the image-spaces $\pi(E)$ and $\pi^*(E^*)$ are dual.

3. An *involution* is an endomorphism ψ such that $\psi^2 = \iota$. Prove that from every projection π an involution ψ is obtained by $\psi = 2\pi - \iota$ and that every involution can be represented in this way.

4. Assume that φ is an endomorphism of a linear space such that $\psi \circ \varphi = \varphi \circ \psi$ for every endomorphism ψ. Prove that $\varphi = \lambda\,\iota$ where λ is a scalar.

Hint: Show first that, to every vector $x \in E$ there is a scalar $\lambda(x)$ such that $\varphi x = \lambda(x)\, x$. Then prove that $\lambda(x)$ does not depend on x.

5. Let E, F and G be three linear spaces of finite dimension. Given two linear mappings $\varphi\colon E \to F$ and $\psi\colon F \to G$ prove that

$$r(\psi \circ \varphi) \leqq r(\varphi) \quad \text{and} \quad r(\psi \circ \varphi) \leqq r(\psi)\ .$$

If φ is regular show that

$$r(\psi \circ \varphi) = r(\psi)\ .$$

6. Let $\sigma_i\,(i = 1, \ldots, n)$ be a system of n endomorphisms of an n-dimensional linear space E with the property that

$$\sigma_i \circ \sigma_j = \sigma_i \delta_{ij}\ .$$

a) Show that every endomorphism σ_i has rank 1.

b) If $\sigma_i'\,(i = 1, \ldots, n)$ is a second system with the same property, prove that there exists an automorphism τ of E such that

$$\sigma_i' = \tau^{-1} \circ \sigma_i \circ \tau\ .$$

7. Consider two linear mappings

$$\varphi\colon E \to F \quad \text{and} \quad \psi\colon E \to F$$

of rank r and s respectively. Prove that:

$$|r - s| \leqq r(\varphi + \psi) \leqq (r + s)\,.$$

8. A *differential-operator* in a linear space E is an endomorphism $\partial: E \to E$ such that $\partial \circ \partial = 0$. It follows from this property that the image-space ∂E is contained in the kernel K and hence we can form the *factor-space*

$$H(E) = K/\partial E .$$

This space is called the *homology-space* of E with respect to the differential-operator ∂.

Consider two linear spaces E and F with a differential-operator $\partial: E \to E$ and $\partial': F \to F$ defined in E and in F respectively. Let φ be a linear mapping of E into F such that

$$\partial' \circ \varphi = \varphi \circ \partial .$$

Prove that φ induces a linear mapping

$$\varphi_*: H(E) \to H(F) .$$

9. Given a differential-operator ∂ in E, consider a subspace E_1 of E such that $\partial E_1 \subset E_1$. Assume that a linear mapping $\varrho: E \to E_1$, is given subject to the following conditions:

1. $\partial \circ \varrho = \varrho \circ \partial$.
2. $\varrho y = y$ for every $y \in E_1$.
3. $\varrho x - x \in \partial E$ for every x for which $\partial x = 0$.

Show that the linear mapping

$$\varrho_*: H(E) \to H(E_1)$$

induced by ϱ is an isomorphism of $H(E)$ onto $H(E_1)$.

10. Let three linear spaces E_1, E_2, H and linear mappings

$$i_1: E_1 \to H \quad \pi_1: H \to E_1$$
$$i_2: E_2 \to H \quad \pi_2: H \to E_2$$

be given such that

$$\pi_1 \circ i_1 = \pi_2 \circ i_2 = \iota , \quad \pi_1 \circ i_2 = \pi_2 \circ i_1 = 0$$

and

$$i_1 \circ \pi_1 + i_2 \circ \pi_2 = \iota .$$

Prove that H is isomorphic to the Cartesian product $E_1 \times E_2$.

11. A system of linear mappings:

$$\varphi_\nu: E_\nu \to E_{\nu+1} \qquad (\nu = 0, 1, \ldots)$$

is called an *exact sequence* if:

$$\varphi_\nu(E_\nu) = K(\varphi_{\nu+1}) \qquad (\nu = 0, 1, \ldots) .$$

Given a subspace E_1 of E show that the sequence

$$0 \to E_1 \to E \to E/E_1 \to 0$$

is exact. Conversely, assume an exact sequence

$$0 \to E' \to E \to E'' \to 0 .$$

Prove that E' is isomorphic to a subspace E_1 of E and that E'' is isomorphic to the factor-space E/E_1.

12. Consider a system of linear mappings

$$\begin{tikzcd}
E_{00} \arrow[r, "\varphi_{00}"] \arrow[d, "\psi_{00}"'] & E_{01} \arrow[r, "\varphi_{01}"] \arrow[d, "\psi_{01}"'] & E_{02} \arrow[r, "\varphi_{02}"] \arrow[d, "\psi_{02}"'] & {} \\
E_{10} \arrow[r, "\varphi_{10}"] \arrow[d, "\psi_{10}"'] & E_{11} \arrow[r, "\varphi_{11}"] \arrow[d, "\psi_{11}"'] & E_{12} \arrow[r, "\varphi_{12}"] \arrow[d, "\psi_{12}"'] & {} \\
E_{20} \arrow[r, "\varphi_{20}"] \arrow[d, "\psi_{20}"'] & E_{21} \arrow[r, "\varphi_{21}"] \arrow[d, "\psi_{21}"'] & E_{22} \arrow[r, "\varphi_{22}"] \arrow[d, "\psi_{22}"'] & {} \\
{} & {} & {} &
\end{tikzcd}$$

where all vertical and all horizontal sequences are exact. Assume in addition that all the above diagrams are commutative,

$$\psi_{\nu\mu+1} \circ \varphi_{\nu\mu} = \varphi_{\nu+1\mu} \circ \psi_{\nu\mu} \,.$$

Defining the spaces $H_{\nu\mu} (\nu \geqq 1, \mu \geqq 1)$ by

$$H_{\nu,\mu} = (K(\varphi_{\nu\mu}) \cap K(\psi_{\nu\mu})) \,/\, (\psi_{\nu-1\mu} \circ \varphi_{\nu-1\,\mu-1})\,(E_{\nu-1\,\mu-1})$$

prove that $H_{\nu,\mu+1}$ and $H_{\nu+1,\mu}$ are isomorphic.

Chapter III

Matrices

§ 1. Matrices and systems of linear equations

3.1. Definition. A rectangular array

$$A = \begin{pmatrix} \alpha_1^1 \ldots\ldots \alpha_1^m \\ \cdot \qquad\quad \cdot \\ \cdot \qquad\quad \cdot \\ \cdot \qquad\quad \cdot \\ \alpha_n^1 \ldots\ldots \alpha_n^m \end{pmatrix} \tag{3.1}$$

of nm scalars α_ν^μ is called a *matrix* of n rows and m columns or, in brief, an $n \times m$-matrix. The scalars α_ν^μ are called the *entries* or the *elements* of the matrix A. The rows

$$a_\nu = (\alpha_\nu^1 \dots \alpha_\nu^m) \quad (\nu = 1 \dots n)$$

can be considered as vectors of the number-space R^m and therefore are called the *row-vectors* of A. Similarly, the columns

$$b^\mu = (b_1^\mu \dots b_n^\mu) \quad (\mu = 1 \dots m)$$

considered as vectors of the number-space R^n, are called the *column-vectors* of A.

Interchanging rows and columns we obtain from A the *transposed matrix*

$$A^* = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^m & \dots & \alpha_n^m \end{pmatrix}. \tag{3.2}$$

In the following, matrices will rarely be written down explicitly as in (3.1) but rather be abbreviated in the form $A = (\alpha_\nu^\mu)$. This notation has the disadvantage of not identifying which index indicates the rows and which the columns. It has to be mentioned in this connection that it would be very undesirable — as we shall see — to agree once and for all to always let the subscript count the rows, etc. If the above abbreviation is used, it will be stated explicitly which index indicates the rows.

3.2. The matrix of a linear mapping. Consider two linear spaces E and F of dimensions n and m and a linear mapping $\varphi: E \to F$. Employing a basis x_ν $(\nu = 1 \dots n)$ and y_μ $(\mu = 1 \dots m)$ in E and in F respectively, every vector φx_ν can be written as a linear combination of the vectors y_μ $(\mu = 1 \dots m)$,

$$\varphi x_\nu = \sum_\mu \alpha_\nu^\mu y_\mu \quad (\nu = 1 \dots n)\,. \tag{3.3}$$

In this way, the mapping φ determines an $n \times m$-matrix (α_ν^μ), where ν counts the rows and μ counts the columns. This matrix will be denoted by $M(\varphi, x_\nu, y_\mu)$ or simply by $M(\varphi)$ if no ambiguity is possible.

Conversely, every $n \times m$-matrix (α_ν^μ) determines a linear mapping $\varphi: E \to F$ by the equations (3.3). Thus, the operator

$$M: \varphi \to M(\varphi)$$

defines a one-to-one correspondence between all linear mappings $\varphi: E \to F$ and all $n \times m$-matrices.

3.3. The matrix of the dual mapping. Let E^* and F^* be dual spaces of E and F, respectively, and $\varphi: E \to F$, $\varphi^*: F^* \to E^*$ a pair of dual mappings. Consider two pairs of dual bases $\overset{*}{x}{}^\nu, x_\nu$ $(\nu = 1 \dots n)$ and $\overset{*}{y}{}^\mu, y_\mu$ $(\mu = 1 \dots m)$ of E^*, E and F^*, F, respectively. We shall show that the two corresponding matrices $M(\varphi)$ and $M(\varphi^*)$ (relative to these bases) are transposed, i. e., that

$$M(\varphi^*) = M(\varphi)^* . \tag{3.4}$$

The matrices $M(\varphi)$ and $M(\varphi^*)$ are defined by the representations

$$\varphi x_\nu = \sum_\mu \alpha_\nu^\mu y_\mu \quad \text{and} \quad \varphi^* \overset{*}{y}{}^\mu = \sum_\nu \overset{*}{\alpha}{}_\nu^\mu \overset{*}{x}{}^\nu .$$

Note here that the subscript ν indicates in the first formula the rows of the matrix α_ν^μ and in the second the columns of the matrix α_ν^μ. Substituting $x = x_\nu$ and $y = \overset{*}{y}{}^\mu$ in the relation

$$\langle \overset{*}{y}, \varphi x\rangle = \langle \varphi^* \overset{*}{y}, x\rangle \tag{3.5}$$

we obtain

$$\langle \overset{*}{y}{}^\mu, \varphi x_\nu\rangle = \langle \varphi^* \overset{*}{y}{}^\mu, x_\nu\rangle . \tag{3.6}$$

Now

$$\langle \overset{*}{y}{}^\mu, \varphi x_\nu\rangle = \sum_\varkappa \alpha_\nu^\varkappa \langle \overset{*}{y}{}^\mu, y_\varkappa\rangle = \alpha_\nu^\mu \tag{3.7}$$

and

$$\langle \varphi^* \overset{*}{y}{}^\mu, x_\nu\rangle = \sum_\lambda \overset{*}{\alpha}{}_\nu^\mu \langle \overset{*}{x}{}^\lambda, x_\nu\rangle = \overset{*}{\alpha}{}_\nu^\mu . \tag{3.8}$$

The relations (3.6), (3.7) and (3.8) then yield

$$\overset{*}{\alpha}{}_\nu^\mu = \alpha_\nu^\mu .$$

Observing — as stated before — that the subscript ν indicates rows of (α_ν^μ) and columns of $(\overset{*}{\alpha}{}_\nu^\mu)$ we obtain the desired equation (3.4).

3.4. Rank of a matrix. Consider an $n \times m$-matrix A. Denote by r_1 and by r_2 the maximal number of linearly independent row-vectors and column-vectors, respectively. It will be shown that $r_1 = r_2$. To prove this let E and F be two linear spaces of dimensions n and m. Choose a basis x_ν $(\nu = 1 \dots n)$ and y_μ $(\mu = 1 \dots m)$ in E and in F and define the linear mapping $\varphi\colon E \to F$ by

$$\varphi x_\nu = \sum_\mu \alpha_\nu^\mu y_\mu .$$

Besides φ, consider the isomorphism

$$\beta\colon F \to R^m$$

defined by

$$\beta\colon y \to (\eta^1 \dots \eta^m) ,$$

where

$$y = \sum_\mu \eta^\mu y_\mu .$$

Then $\beta \circ \varphi$ is a linear mapping of E into R^m. From the definition of β it follows that $\beta \circ \varphi$ maps x_ν into the ν-th row-vector,

$$\beta\, \varphi x_\nu = a_\nu .$$

Consequently, the rank of $\beta \circ \varphi$ is equal to the maximal number r_1 of linearly independent row-vectors. Since β is an isomorphism, $\beta \circ \varphi$ has the same rank as φ and hence r_1 is equal to the rank r of φ.

Replacing φ by φ^* we see that the maximal number r_2 of linearly independent column-vectors is equal to the rank of φ^*. But φ^* has the same rank as φ and thus $r_1 = r_2 = r$. The number r is called the *rank* of the matrix A.

3.5. Systems of linear equations. Matrices play an important role in the discussion of systems of linear equations. Such a system

$$\sum_\nu \alpha_\nu^\mu \xi^\nu = \eta^\mu \quad (\mu = 1 \dots m) \tag{3.9}$$

of m equations with n unknowns is called *inhomogeneous* if at least one η^μ is different from zero. Otherwise it is called *homogeneous*.

From the results of Chapter II it is easy to obtain theorems about the existence and uniqueness of solutions of the system (3.9). Let E and F be two linear spaces of dimensions n and m. Choose a basis x_ν $(\nu = 1 \dots n)$ of E as well as a basis y_μ $(\mu = 1 \dots m)$ of F and define the linear mapping $\varphi: E \to F$ by

$$\varphi x_\nu = \sum_\mu \alpha_\nu^\mu y_\mu .$$

Consider two vectors

$$x = \sum_\nu \xi^\nu x_\nu \tag{3.10}$$

and

$$y = \sum_\mu \eta^\mu y_\mu . \tag{3.11}$$

Then

$$\varphi x = \sum_\nu \xi^\nu \varphi x_\nu = \sum_{\nu, \mu} \alpha_\nu^\mu \xi^\nu y_\mu . \tag{3.12}$$

Comparing the representations (3.9) and (3.12) we see that the system (3.9) is equivalent to the vector-equation

$$\varphi x = y .$$

Consequently, the system (3.9) has a solution if and only if the vector y is contained in the image-space $\varphi(E)$. Moreover, this solution is uniquely determined if and only if the kernel of φ consists only of the zero-vector.

3.6. The homogeneous system. Consider the homogeneous system

$$\sum_\nu \alpha_\nu^\mu \xi^\nu = 0 \quad (\mu = 1 \dots m) . \tag{3.13}$$

From the foregoing discussion it is immediately clear that $(\xi^1 \dots \xi^n)$ is a solution of this system if and only if the vector x defined by (3.10) is contained in the kernel $K(\varphi)$ of the linear mapping φ. In sec. 2.5 we have shown that the dimension of $K(\varphi)$ equals $n - r$ where r denotes the rank of φ.

Since the rank of φ is equal to the rank of the matrix (α_ν^μ), we therefore obtain the following theorem:

A homogeneous system of m equations with n unknowns whose coefficient-matrix is of rank r has exactly $n - r$ linearly independent solutions. In the special case that the number m of equations is less than the number n of unknowns we have $n - r \geqq n - m \geqq 1$. Hence the theorem asserts that the system (3.13) always has non-trivial solutions if m is less than n.

3.7. The alternative-theorem. Let us assume that the number of equations is equal to the number of unknowns,

$$\sum_\nu \alpha_\nu^\mu \xi^\nu = \eta^\mu \quad (\mu = 1 \ldots n). \tag{3.14}$$

Besides (3.14) consider the so-called "corresponding" homogeneous system

$$\sum_\nu \alpha_\nu^\mu \xi^\nu = 0 \quad (\mu = 1 \ldots n). \tag{3.15}$$

The mapping φ introduced in sec. 3.5 is now a linear mapping of the n-dimensional space E into a space of the same dimension. Hence we may apply the result of sec. 2.5 and obtain the following *alternative-theorem:*

If the homogeneous system possesses only the trivial solution $(0 \ldots 0)$, *the inhomogeneous system has a solution* $(\xi^1 \ldots \xi^n)$ *for every choice of the right-hand side. If the homogeneous system has non-trivial solutions, then the inhomogeneous is not solvable for every choice of the* η^ν $(\nu = 1 \ldots n)$.

From the last statement of section 3.5 it follows immediately that in the first case the solution of (3.14) is uniquely determined while in the second case the system (3.14) has – if it is solvable at all – infinitely many solutions.

3.8. The main-theorem. We now proceed to the general case of an arbitrary system

$$\sum_\mu \alpha_\nu^\mu \xi^\nu = \eta^\mu \quad (\mu = 1 \ldots m) \tag{3.16}$$

of m linear equations in n unknowns. As stated before, this system has a solution if and only if the vector

$$y = \sum_\mu \eta^\mu y_\mu$$

is contained in the image-space $\varphi(E)$. In sec. 2.12 it has been shown that the space $\varphi(E)$ is the orthogonal complement of the kernel $K(\varphi^*)$ of the dual mapping $\varphi^*: F^* \to E^*$. In other words, the system (3.16) is solvable if and only if the right-hand side η^μ $(\mu = 1 \ldots m)$ satisfies the conditions

$$\sum_\mu \eta^\mu \overset{*}{\eta}_\mu = 0 \tag{3.17}$$

for all solutions $\overset{*}{\eta}_\mu$ $(\mu = 1 \dots m)$ of the system

$$\sum_\mu \alpha_\nu^\mu \overset{*}{\eta}_\mu = 0 \quad (\nu = 1 \dots n)\,. \tag{3.18}$$

We formulate this result in the following

Main-theorem: An inhomogeneous system of n equations in m unknowns has a solution if and only if every solution $\overset{*}{\eta}_\mu (\mu = 1 \dots m)$ *of the transposed homogeneous system* (3.18) *satisfies the orthogonality-relation* (3.17).

Problems: 1. Find the matrices corresponding to the following mappings:

a) $\varphi x = 0$.
b) $\varphi x = x$.
c) $\varphi x = \lambda x$.
d) $\varphi x = \sum_{\nu=1}^{m} \xi^\nu e_\nu$ where $e_\nu (\nu = 1, \dots, n)$ is a given basis and $m \leqq n$ is a given number.

2. Consider a system of two equations and n unknowns

$$\sum_{\nu=1}^{n} \alpha_\nu \xi^\nu = \alpha \qquad \sum_{\nu=1}^{n} \beta_\nu \xi^\nu = \beta\,.$$

Find the solutions of the corresponding transposed homogeneous system.

3. Prove the following statement:

The general solution of the inhomogeneous system is equal to the sum of any particular solution of this system and the general solution of the corresponding homogeneous system.

4. Let x_ν and $\bar{x}_\nu$ be two bases of E and A be the matrix of the basis-transformation $x_\nu \to \bar{x}_\nu$. Define the automorphism α of E by $\alpha x_\nu = \bar{x}_\nu$. Prove that A is the matrix of α as well with respect to the basis x_ν as with respect to the basis $\bar{x}_\nu$.

5. Show that a necessary and sufficient condition for the $n \times n$-matrix $A = (\alpha_\nu^\mu)$ to have rank $\leqq 1$ is that there exist numbers $\alpha_1, \alpha_2, \dots, \alpha_n$ and $\beta^1, \beta^2, \dots, \beta^n$ such that

$$\alpha_\nu^\mu = \alpha_\nu \beta^\mu \quad (\nu = 1, 2, \dots, n;\ \mu = 1, 2, \dots, n)\,.$$

If $A \neq 0$, show that the numbers α_ν and β_ν are uniquely determined up to a constant factor λ and μ respectively, where $\lambda\mu = 1$.

6. Given a basis a_ν of a linear space E, define the mapping $\varphi: E \to E$ as

$$\varphi a_\nu = \sum_\mu a_\mu\,.$$

Find the matrix of the dual mapping relative to the dual basis.

7. Verify that the system of three equations:

$$\xi + \eta + \zeta = 3\,,$$
$$\xi - \eta - \zeta = 4\,,$$
$$\xi + 3\eta + 3\zeta = 1\,,$$

has no solution. Find a solution of the transposed homogeneous system which is not orthogonal to the vector (3, 4, 1). Replace the number 1 on the right-hand side of the third equation in such a way that the resulting system is solvable.

8. Let an inhomogeneous system of linear equations be given,

$$\sum_{\nu} \alpha_{\nu}^{\mu} \xi^{\nu} = \eta^{\mu} \qquad (\mu = 1, \ldots, m)\,.$$

The *augmented matrix* of the system is defined as the $m \times (n+1)$-matrix obtained from the matrix α_{ν}^{μ} by adding the column $(\eta^1, \ldots, \eta^m)$. Prove that the above system has a solution if and only if the augmented matrix has the same rank as the matrix (α_{ν}^{μ}).

§ 2. Multiplication of matrices

3.9. The linear space of the $\mathbf{n \times m}$ matrices. Consider the space $L(E;F)$ of all linear mappings $\varphi: E \to F$ and the set $M^{n\times m}$ of all $n \times m$-matrices. Once a basis has been chosen in E and in F there is a 1-1 correspondence between the mappings $\varphi: E \to F$ and the $n \times m$-matrices defined by

$$\varphi \to M\,(\varphi, x_{\nu}, y_{\mu})\,. \tag{3.19}$$

This correspondence suggests defining a linear structure in the set $M^{n\times m}$ such that the mapping (3.19) becomes an isomorphism.

We define the *sum* of two $n \times m$-matrices

$$A = (\alpha_{\nu}^{\mu}) \quad \text{and} \quad B = (\beta_{\nu}^{\mu})$$

as the $n \times m$-matrix

$$A + B = (\alpha_{\nu}^{\mu} + \beta_{\nu}^{\mu})$$

and the product of a scalar λ and a matrix A as the matrix

$$\lambda A = (\lambda a_{\nu}^{\mu})\,.$$

It is immediately apparent that with these operations the set $M^{n\times m}$ is a linear space. The zero-vector in this linear space is the matrix which has only zero-entries.

Furthermore, it follows from the above definitions that

$$M\,(\lambda\varphi + \mu\psi) = \lambda M(\varphi) + \mu M(\psi) \quad \varphi, \psi \in L(E;F)$$

i. e., that the mapping (3.19) defines an isomorphism between $L(E;F)$ and the space $M^{n\times m}$.

3.10. Product of matrices. Assume that

$$\varphi: E \to F \quad \text{and} \quad \psi: F \to G$$

are two linear mappings between three linear spaces E, F, G of dimensions n, m and l, respectively. Then $\psi \circ \varphi$ is a linear mapping of E into G. Select a basis x_ν $(\nu = 1 \dots n)$, y_μ $(\mu = 1 \dots m)$ and z_λ $(\lambda = 1 \dots l)$ in each of the three spaces. Then the mappings φ and ψ determine two matrices (α_ν^μ) and (β_μ^λ) by the relations

$$\varphi x_\nu = \sum_\mu \alpha_\nu^\mu y_\mu$$

and

$$\psi y_\mu = \sum_\lambda \beta_\mu^\lambda z_\lambda .$$

These two equations yield

$$(\psi \circ \varphi)\, x_\nu = \sum_{\mu, \lambda} \alpha_\nu^\mu \, \beta_\mu^\lambda \, z_\lambda .$$

Consequently, the matrix of the mapping $\psi \circ \varphi$ relative to the bases x_ν and z_λ is given by

$$\gamma_\nu^\lambda = \sum_\mu \alpha_\nu^\mu \, \beta_\mu^\lambda . \tag{3.20}$$

The $n \times l$-matrix (3.20) is called the *product* of the $n \times m$-matrix $A = (\alpha_\nu^\mu)$ and the $m \times l$-matrix $B = (\beta_\mu^\lambda)$ and is denoted by AB. It follows immediately from this definition that

$$M(\psi \circ \varphi) = M(\varphi)\, M(\psi) . \tag{3.21}$$

Note that the matrix $M(\psi \circ \varphi)$ of the product-mapping $\psi \circ \varphi$ is the product of the matrices $M(\varphi)$ and $M(\psi)$ in reversed order of the factors.

It follows immediately from (3.21) and the formulas of sec. 2.16 that the matrix-multiplication has the following properties:

$$A\,(\lambda B_1 + \mu B_2) = \lambda\, A\, B_1 + \mu\, A\, B_2$$
$$(\lambda A_1 + \mu A_2)\, B = \lambda\, A_1 B + \mu\, A_2 B$$
$$(A\,B)\, C = A\,(B C)$$
$$(A\,B)^* = B^* A^* .$$

3.11. Endomorphisms and square-matrices. Let A and B be two $n \times n$-matrices. Then the product AB is again an $n \times n$-matrix. Hence the matrix-product defines a multiplication in the linear space $M^{n \times n}$ of all $n \times n$-matrixes. Since this multiplication is distributive and associative the space $M^{n \times n}$ becomes an algebra. The unit-element of this algebra is the unit-matrix J defined by the Kronecker-symbol

$$J = (\delta_\nu^\mu) .$$

Now let E be an n-dimensional linear space. If a basis x_ν $(\nu = 1 \ldots n)$ is chosen in E the mapping

$$M: \varphi \to M(\varphi; x_\nu, x_\mu)$$

defines an isomorphism between the spaces $L(E; E)$ and $M^{n \times n}$. The relation

$$M(\psi \circ \varphi) = M(\varphi)\, M(\psi)$$

shows that the isomorphism M sends the product of two endomorphisms into the product of the corresponding matrices taken in the reverse order.

3.12. Automorphisms and regular matrices. An $n \times n$-matrix A is called *regular* if it has the maximal rank n. Let φ be an automorphism of the n-dimensional linear space E and $A = M(\varphi)$ the corresponding $n \times n$-matrix relative to a basis x_ν $(\nu = 1 \ldots n)$. By the result of section 3.4 the rank of φ is equal to the rank of the matrix A. Consequently, the matrix A is regular. Conversely, every endomorphism $\varphi: E \to E$ having a regular matrix is an automorphism.

To every regular matrix A there exists an *inverse matrix*, i. e., a matrix A^{-1} such that

$$A A^{-1} = A^{-1} A = J\,.$$

In fact, let φ be the automorphism of E such that $M(\varphi) = A$ and let φ^{-1} be the inverse automorphism. Then

$$\varphi^{-1} \circ \varphi = \varphi \circ \varphi^{-1} = \iota\,,$$

whence

$$M(\varphi)\, M(\varphi)^{-1} = M(\varphi^{-1} \circ \varphi) = M(\iota) = J$$

and

$$M(\varphi^{-1})\, M(\varphi) = M(\varphi \circ \varphi^{-1}) = M(\iota) = J\,.$$

These equations show that the matrix

$$A^{-1} = M(\varphi^{-1})$$

is the inverse of the matrix A.

Problems: 1. Verify the following properties:

a) $(A + B)^* = A^* + B^*$.

b) $(\lambda A)^* = \lambda A^*$.

c) $(A^{-1})^* = (A^*)^{-1}$.

2. A square-matrix is called *upper (lower) triangular* if all the elements below (above) the main diagonal are zero. Prove that all the triangular matrices form an associative algebra.

3. Let φ be an endomorphism such that $\varphi^2 = \varphi$. Show that there exists a basis in which φ is represented by a matrix of the form:

$$\underbrace{\begin{pmatrix} 1 & & & 0 & \dots & 0 \\ & 1 & & \cdot & & \cdot \\ & & \ddots & \vdots & & \vdots \\ & & 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}}_{n} \begin{array}{l} \left.\vphantom{\begin{matrix} 1 \\ 1 \\ \ddots \\ 1 \end{matrix}}\right\} m \\ \left.\vphantom{\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}}\right\} n - m \end{array}$$

4. Denote by A_{ij} the matrix having the entry 1 at the place (i, j) and zero elsewhere. Verify the formula

$$A_{ij} \cdot A_{jk} = A_{ik}\,.$$

§ 3. Basis-transformation

3.13. Definition. Consider two bases x_ν and $\bar{x}_\nu$ $(\nu = 1 \dots n)$ of the space E. Then every vector $\bar{x}_\nu$ $(\nu = 1 \dots n)$ can be written as

$$\bar{x}_\nu = \sum_\mu \alpha_\nu^\mu x_\mu\,. \tag{3.22}$$

Similarly,

$$x_\nu = \sum_\mu \check{\alpha}_\nu^\mu \bar{x}_\mu\,. \tag{3.23}$$

The two $n \times n$-matrices defined by (3.22) and (3.23) are inverse to each other. In fact, combining (3.22) and (3.23) we obtain

$$\bar{x}_\nu = \sum_{\mu, \lambda} \alpha_\nu^\mu \check{\alpha}_\mu^\lambda \bar{x}_\lambda\,.$$

This is equivalent with

$$\sum_\lambda \left(\sum_\mu \alpha_\nu^\mu \check{\alpha}_\mu^\lambda - \delta_\nu^\lambda\right) \bar{x}_\lambda = 0$$

and hence it implies that

$$\sum_\mu \alpha_\nu^\mu \check{\alpha}_\mu^\lambda = \delta_\nu^\lambda\,.$$

In a similar way the relations

$$\sum_\mu \check{\alpha}_\nu^\mu \alpha_\mu^\lambda = \delta_\nu^\lambda$$

are proved. Thus, any two bases of E are connected by a pair of inverse matrices.

Conversely, given a basis x_ν $(\nu = 1 \dots n)$ and a regular $n \times n$-matrix (α_ν^μ), another basis can be obtained by

$$\bar{x}_\nu = \sum_\mu \alpha_\nu^\mu x_\mu\,.$$

To show that the vectors $\bar{x}_\nu$ are linearly independent, assume that

$$\sum_\nu \lambda^\nu \bar{x}_\nu = 0 .$$

Then

$$\sum_{\nu,\mu} \lambda^\nu \alpha_\nu^\mu x_\mu = 0$$

and hence, in view of the linear independence of the vectors x_μ,

$$\sum_\nu \lambda^\nu \alpha_\nu^\mu = 0 \quad (\mu = 1 \ldots n) .$$

Multiplication with the inverse matrix $\check{\alpha}_\mu^\varkappa$ yields

$$\sum_{\nu,\mu} \lambda^\nu \alpha_\nu^\mu \check{\alpha}_\mu^\varkappa = \sum_\nu \lambda^\nu \delta_\nu^\varkappa = \lambda^\varkappa = 0 \quad (\varkappa = 1 \ldots n) .$$

3.14. Transformation of the dual basis. Let E^* be a dual space of E, $\overset{*}{x}{}^\nu$ the dual basis of x_ν and $\overset{*}{\bar{x}}_\nu$ the dual of the basis $\bar{x}_\nu$ $(\nu = 1 \ldots n)$. Then

$$\overset{*}{\bar{x}}{}^\varrho = \sum_\sigma \beta_\sigma^\varrho \overset{*}{x}{}^\sigma , \tag{3.24}$$

where β_σ^ϱ is a regular $n \times n$-matrix. The relations (3.23) and (3.24) yield

$$\sum_\sigma \beta_\sigma^\varrho \langle \overset{*}{x}{}^\sigma, x_\nu \rangle = \sum_\mu \check{\alpha}_\nu^\mu \langle \overset{*}{\bar{x}}{}^\varrho, \bar{x}_\mu \rangle . \tag{3.25}$$

Now

$$\langle \overset{*}{x}{}^\sigma, x_\nu \rangle = \delta_\nu^\sigma \quad \text{and} \quad \langle \overset{*}{\bar{x}}{}^\varrho, \bar{x}_\mu \rangle = \delta_\mu^\varrho .$$

Substituting this in (3.25) we obtain

$$\beta_\nu^\varrho = \check{\alpha}_\nu^\varrho .$$

This shows that the matrix of the basis-transformation $\overset{*}{x}{}^\nu \to \overset{*}{\bar{x}}{}^\nu$ is the inverse of the matrix of the transformation $x_\nu \to \bar{x}_\nu$. The two basis-transformations

$$\bar{x}_\nu = \sum_\mu \alpha_\nu^\mu x_\mu \quad \text{and} \quad \overset{*}{\bar{x}}{}^\nu = \sum_\mu \check{\alpha}_\mu^\nu \overset{*}{x}{}^\mu \tag{3.26}$$

are called *contragredient* to each other.

The relations (3.26) permit the derivation of the transformation-law for the components of a vector $x \in E$ under the basis-transformation $x_\nu \to \bar{x}_\nu$. Decomposing x relative to the bases x_ν and $\bar{x}_\nu$ we obtain

$$x = \sum_\nu \xi^\nu x_\nu \quad \text{and} \quad x = \sum_\nu \bar{\xi}^\nu \bar{x}_\nu .$$

From the two above equations we obtain in view of (3.26).

$$\bar{\xi}^\nu = \sum_\mu \check{\alpha}_\mu^\nu \langle \overset{*}{x}{}^\mu, x \rangle = \sum_\mu \check{\alpha}_\mu^\nu \xi^\mu . \tag{3.27}$$

Comparing (3.27) with the second equation (3.26) we see that the components of a vector are transformed exactly in the same way as the vectors of the dual basis.

3.15. The transformation of the matrix of a linear mapping. In this section it will be investigated how the matrix of a linear mapping $\varphi: E \to F$ is changed under a basis-transformation in E as well as in F. Let $M(\varphi; x_\nu, y_\mu) = (\gamma_\nu^\mu)$ and $M(\varphi; \bar{x}_\nu, \bar{y}_\mu) = (\bar{\gamma}_\nu^\mu)$ be the $n \times m$-matrices of φ relative to the bases x_ν, y_μ and $\bar{x}_\nu, \bar{y}_\mu$ ($\nu = 1 \ldots n$, $\mu = 1 \ldots m$), respectively. Then

$$\varphi x_\nu = \sum_\mu \gamma_\nu^\mu y_\mu \quad \text{and} \quad \varphi \bar{x}_\nu = \sum_\mu \bar{\gamma}_\nu^\mu \bar{y}_\mu \quad (\nu = 1 \ldots n). \tag{3.28}$$

Introducing the matrices

$$A = (\alpha_\nu^\lambda) \quad \text{and} \quad B = (\beta_\mu^\varkappa)$$

of the basis-transformations $x_\nu \to \bar{x}_\nu$ and $y_\mu \to \bar{y}_\mu$ and their inverse matrices, we then have the relations

$$\begin{aligned} \bar{x}_\nu &= \sum_\lambda \alpha_\nu^\lambda x_\lambda \qquad & x_\nu &= \sum_\lambda \check{\alpha}_\nu^\lambda \bar{x}_\lambda \\ \bar{y}_\mu &= \sum_\varkappa \beta_\mu^\varkappa y_\varkappa \qquad & y_\mu &= \sum_\varkappa \check{\beta}_\mu^\varkappa \bar{y}_\varkappa . \end{aligned} \tag{3.29}$$

The equations (3.28) and (3.29) yield

$$\varphi \bar{x}_\nu = \sum_\lambda \alpha_\nu^\lambda \varphi x_\lambda = \sum_{\lambda,\mu} \alpha_\nu^\lambda \gamma_\lambda^\mu y_\mu = \sum_{\lambda,\mu,\varkappa} \alpha_\nu^\lambda \gamma_\lambda^\mu \check{\beta}_\mu^\varkappa \bar{y}_\varkappa$$

and we obtain the following relation between the matrices (γ_ν^μ) and $(\bar{\gamma}_\nu^\mu)$:

$$\bar{\gamma}_\nu^\varkappa = \sum_{\lambda,\mu} \alpha_\nu^\lambda \gamma_\lambda^\mu \check{\beta}_\mu^\varkappa . \tag{3.30}$$

Using capital letters for the matrices, the transformation formula (3.30) can be written in the form

$$M(\varphi; \bar{x}_\nu, \bar{y}_\mu) = A\, M(\varphi; x_\nu, y_\mu)\, B^{-1}.$$

It shows that all possible matrices of the endomorphism φ are obtained from a particular matrix by left-multiplication with a regular $n \times n$-matrix and right-multiplication with a regular $m \times m$-matrix.

Problems: 1. Let f be a function defined in the set of all $n \times n$-matrices such that

$$f(TAT^{-1}) = f(A)$$

for every regular matrix T. Define the function F in the space $L(E; E)$ by

$$F(\varphi) = f(M(\varphi; x_\nu, x_\mu))$$

where E is an n-dimensional linear space and x_ν ($\nu = 1 \ldots n$) is a basis of E. Prove that the function F does not depend on the choice of the basis x_ν.

2. Assume that φ is an endomorphism of E having the same matrix relative to every basis x_ν $(\nu = 1 \dots n)$. Prove that $\varphi = \lambda\iota$ where λ is a scalar.

3. Given the basis transformation

$$\begin{aligned} \bar{x}_1 &= 2x_1 - x_2 - x_3 \\ \bar{x}_2 &= -x_2 \\ \bar{x}_3 &= 2x_2 + x_3 \end{aligned}$$

find all the vectors which have the same components with respect to the bases x_μ and $\bar{x}_\mu$. $(\mu = 1, 2, 3)$.

§ 4. Elementary transformations

3.16. Definition. Consider a linear mapping $\varphi: E \to F$. Then there exists a basis $a_\nu (\nu = 1, \dots, n)$ of E and a basis $b_\mu (\mu = 1, \dots, n)$ of F such that the corresponding matrix of φ has the following *normal-form:*

$$\begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{r} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} \tag{3.31}$$

where r is the rank of φ. In fact, let $a_\nu (\nu = 1, \dots, n)$ be a basis of E such that the vectors $a_{r+1} \dots a_n$ form a basis of the kernel. Then the vectors $b_\varrho = \varphi a_\varrho (\varrho = 1, \dots, r)$ are linearly independent and hence this system can be extended to a basis $(b_1, \dots, b_m)$ of F. It follows from the construction of the bases a_ν and b_μ that the matrix of φ has the form (3.31).

Now let $x_\nu (\nu = 1, \dots, n)$ and $y_\mu (\mu = 1, \dots, m)$ be two arbitrary bases of E and F. It will be shown that the corresponding matrix $M(\varphi; x_\nu, y_\mu)$ can be converted into the normal-form (3.31) by a number of elementary basis-transformations. These transformations are:

(I.1.) Interchange of two vectors x_i and x_j $(i \neq j)$.

(I.2.) Interchange of two vectors y_k and y_l $(k \neq l)$.

(II.1.) Adding to a vector x_i an arbitrary multiple of a vector x_j $(j \neq i)$.

(II.2.) Adding to a vector y_k an arbitrary multiple of a vector y_l $(l \neq k)$.

It is easy to see that the four above transformations have the following effect on the matrix $M(\varphi)$:

(I.1.) Interchange of the rows i and j.

(I.2.) Interchange of the columns k and l.

(II.1.) Replacement of the row-vector a_i by $a_i + \lambda a_j$ $(j \neq i)$.

(II.2.) Replacement of the column-vector b_k by $b_k + \lambda b_l$ $(l \neq k)$.

It remains to be shown that every $n \times m$-matrix can be converted into the normal form (3.31) by a sequence of these elementary matrix-transformations.

3.17. Transformation into the normal-form. Let (γ_ν^μ) be the given $n \times m$-matrix. It is no restriction to assume that at least one $\gamma_\nu^\mu \neq 0$, otherwise the matrix is already in the normal-form. By the operations (I.1.) and (I.2.) this element can be moved to the place (1.1.). Then $\gamma_1^1 \neq 0$ and it is no restriction to assume that $\gamma_1^1 = 1$. Now, by adding proper multiples of the first row to the other rows we can obtain a matrix whose first column consists of zeros except for γ_1^1. Next, by adding certain multiples of the first column to the other columns this matrix can be converted into the form

$$\begin{pmatrix} 1 & 0 \dots 0 \\ 0 & * \quad\; * \\ \cdot & \\ \cdot & \\ 0 & * \quad\; * \end{pmatrix} \tag{3.32}$$

If all the elements γ_ν^μ $(\nu = 2 \dots n, \mu = 2 \dots m)$ are zero, (3.32) is the normal-form. Otherwise there is an element $\gamma_\nu^\mu \neq 0$ $(2 \leqq \nu \leqq m, 2 \leqq \gamma \leqq m)$. This can be moved to the place (2.2) by the operations (I.1.) and (I.2.). Hereby the first row and the first column are not changed. Dividing the second row by γ_2^2 and applying the operations (II.1.) and (II.2.) we can obtain a matrix of the form

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 \dots & 0 \\ \cdot & 0 & * & * \\ \cdot & & & \\ 0 & 0 & * & * \end{pmatrix}$$

In this way the original matrix is ultimately converted into the form (3.31.).

3.18. The Gaussian elimination. The technique described in sec. 3.17 can be used to solve a system of linear equations by successive elimination. Let

$$\begin{aligned} \alpha_1^1 \xi^1 + \cdots \alpha_n^1 \xi^n &= \eta^1 \\ &\cdot \\ &\cdot \\ &\cdot \\ \alpha_1^m \xi^1 + \cdots \alpha_n^m \xi^n &= \eta^m \end{aligned} \tag{3.33}$$

be a system of m linear equations in n unknowns. Before starting the elimination we perform the following reductions:

If all coefficients in a certain row, say in the i-th row, are zero, consider the corresponding number η^i on the right hand-side. If $\eta^i \neq 0$, the i-th equation contains a contradiction and the system (3.33) has no solution. If $\eta^i = 0$, the i-th equation is an identity and can be omitted.

Hence, we can assume that at least one coefficient in every equation is different from zero. Rearranging the unknowns we can achieve that $\alpha_1^1 \neq 0$. Multiplying the first equation by $-\frac{\alpha_1^\mu}{\alpha_1^1}$ and adding it to the other equations we obtain a system of the form

$$\begin{aligned} \alpha_1^1 \xi^1 + \alpha_2^1 \xi^2 + \cdots \alpha_n^1 \xi^n &= \zeta^1 \\ \beta_2^2 \xi^2 + \cdots \beta_n^2 \xi^n &= \zeta^2 \\ &\vdots \\ \beta_2^m \xi^2 + \cdots \beta_n^m \xi^n &= \zeta^m \end{aligned} \tag{3.34}$$

which is equivalent to the system (3.33).

Now apply the above reduction to the $(m-1)$ last equations of the system (3.34). If one of these equations contains a contradiction, the system (3.34) has no solutions. Then the equivalent system (3.33) does not have a solution either. Otherwise eliminate the next unknown, say ξ^2, from the reduced system.

Continue this process until either a contradiction arises at a certain step or until no equations are left after the reduction. In the first case, (3.33) does not have a solution. In the second case we finally obtain a triangular system

$$\begin{aligned} \alpha_1^1 \xi^1 + \alpha_2^1 \xi^2 + \cdots \alpha_n^1 \xi^n &= \omega^1 \qquad \alpha_1^1 \neq 0 \\ \beta_2^2 \xi^2 + \cdots \beta_n^2 \xi^n &= \omega^2 \qquad \beta_2^2 \neq 0 \\ &\vdots \\ \varkappa_r^r \xi^r + \cdots \varkappa_n^r \xi^n &= \omega^r \qquad \varkappa_r^r \neq 0 \end{aligned} \tag{3.35}$$

which is equivalent to the original system*).

The system (3.35) can be solved in a step by step manner beginning with ξ^r,

$$\xi^r = -\frac{1}{\varkappa_r^r}\left(\omega^r - \sum_{\nu=r+1}^{n} \varkappa_\nu^r \xi^\nu\right). \tag{3.36}$$

Inserting (3.36) into the first $(r-1)$ equations we can reduce the system to a triangular one of $r-1$ equations. Continuing this way we

---

*) If no equations are left after the reduction, then every n-tuple $(\xi^1 \ldots \xi^n)$ is a solution of (3.33).

finally obtain the solution of (3.33) in the form

$$\xi^\nu = \sum_{\mu=r+1}^{n} \lambda^\nu_\mu \xi^\mu + \varrho^\nu \quad (\nu = 1 \cdots r)$$

where the ξ^ν $(\nu = r+1 \dots n)$ are arbitrary parameters.

Problems: 1. Two $n \times m$-matrices C and C' are called *equivalent* if there exists a regular $n \times n$-matrix A and a regular $m \times m$-matrix B such that $C' = ACB$. Prove that two matrices are equivalent if and only if they have the same rank.

2. Apply the Gauss elimination to the following systems:

a)
$$\begin{aligned} \xi^1 - \xi^2 + 2\xi^3 &= 1\,, \\ 2\xi^1 \qquad + 2\xi^3 &= 1\,, \\ \xi^1 - 3\xi^2 + 4\xi^3 &= 2\,. \end{aligned}$$

b)
$$\begin{aligned} \eta^1 + 2\eta^2 + 3\eta^3 + 4\eta^4 &= 5\,, \\ 2\eta^1 + \eta^2 + 4\eta^3 + \eta^4 &= 2\,, \\ 3\eta^1 + 4\eta^2 + \eta^3 + 5\eta^4 &= 6\,, \\ 2\eta^1 + 3\eta^2 + 5\eta^3 + 2\eta^4 &= 3\,. \end{aligned}$$

c)
$$\begin{aligned} \varepsilon^1 + \varepsilon^2 + \varepsilon^3 &= 1\,, \\ 3\varepsilon^1 + \varepsilon^2 - \varepsilon^3 &= 0\,, \\ 2\varepsilon^1 + \varepsilon^2 \qquad &= 1\,. \end{aligned}$$

Chapter IV

Determinants

§ 1. Determinant-functions

4.1. Definition. Consider a linear space E of dimension n $(n \geqq 1)$. A *determinant-function* Δ is a function of n vectors subject to the following conditions:

1. Δ is linear with respect to every argument,

$$\Delta(x_1 \dots \lambda x_i + \mu y_i \dots x_n) = \lambda \Delta(x_1 \dots x_i \dots x_n) + \mu \Delta(x_1 \dots y_i \dots x_n) \qquad (i = 1 \dots n)\,.$$

2. Δ is skew-symmetric with respect to all arguments. More precisely, if σ is any permutation of the numbers $(1 \dots n)$, then

$$\Delta(x_{\sigma(1)} \dots x_{\sigma(n)}) = \varepsilon_\sigma \Delta(x_1 \dots x_n)\,,$$

where

$$\varepsilon_\sigma = \begin{cases} +1 \text{ for an even permutation } \sigma \\ -1 \text{ for an odd permutation } \sigma. \end{cases}$$

It will be shown in sec. 4.4 that there exist non-trivial determinant-functions in every linear space. First of all, a few consequences of the above conditions will be derived.

Since the interchange of any two numbers (i, j) is an odd permutation, we obtain from the second condition

$$\Delta\,(x_1 \dots x_i \dots x_j \dots x_n) = -\Delta\,(x_1 \dots x_j \dots x_i \dots x_n)\,.$$

In particular, if $x_i = x_j = x$,

$$\Delta\,(x_1 \dots x \dots x \dots x_n) = 0\,. \tag{4.1}$$

Thus, a determinant-function assumes zero whenever two arguments coincide. More generally, it will be shown that

$$\Delta\,(x_1 \dots x_n) = 0$$

if the arguments are linearly dependent. In fact, assume that

$$x_n = \sum_{\nu=1}^{n-1} \lambda^\nu x_\nu\,.$$

Then in view of (4.1),

$$\Delta\,(x_1 \dots x_n) = \sum_{\nu=1}^{n-1} \lambda^\nu \Delta\,(x_1 \dots x_\nu \dots x_{n-1}, x_\nu) = 0\,.$$

As another consequence of (4.1) we note that the value of a determinant-function is not changed if a multiple of an argument x_j is added to another argument x_i $(i \neq j)$,

$$\Delta\,(x_1 \dots x_i + \lambda x_j \dots x_n) = \Delta\,(x_1 \dots x_n) \quad (i \neq j)\,.$$

4.2. Representation in a basis. Let e_ν $(\nu = 1 \dots n)$ be a basis of E. Then every vector x_ν can be written as

$$x_\nu = \sum_\lambda \xi_\nu^\lambda e_\lambda \quad (\nu = 1 \dots n)\,.$$

Inserting these linear combinations into Δ we obtain

$$\Delta\,(x_1 \dots x_n) = \sum_{(\lambda)} \xi_1^{\lambda_1} \dots \xi_n^{\lambda_n} \Delta\,(e_{\lambda_1} \dots e_{\lambda_n}) \tag{4.2}$$

the summation being taken over all systems $(\lambda_1 \dots \lambda_n)$ $(1 \leqq \lambda_\nu \leqq n)$. It follows from (4.1) that all terms for which at least two indices λ_i and λ_j coincide, are zero. Therefore we can restrict ourselves to those systems $(\lambda_1 \dots \lambda_n)$ for which any two λ_i are different. In other words, we have to sum over all permutations σ of the set $(1 \dots n)$. Hence (4.2) can be written as

$$\Delta\,(x_1 \dots x_n) = \sum_\sigma \xi_1^{\sigma(1)} \dots \xi_n^{\sigma(n)} \Delta\,(e_{\sigma(1)} \dots e_{\sigma(n)})\,. \tag{4.3}$$

Next we observe that

$$\Delta\,(e_{\sigma(1)} \dots e_{\sigma(n)}) = \varepsilon_\sigma \Delta\,(e_1 \dots e_n)$$

for every permutation σ. We thus obtain from (4.3)

$$\Delta(x_1 \dots x_n) = \Delta(e_1 \dots e_n) \sum_\sigma \varepsilon_\sigma \xi_1^{\sigma(1)} \dots \xi_n^{\sigma(n)} .$$

This equation shows that a determinant-function is identically zero if it assumes the value zero at a basis of E. In other words, a non-trivial determinant-function is different from zero on every basis of E.

Altogether we have shown that a non-trivial determinant-function assumes the value zero for a system of vectors x_ν $(\nu = 1 \dots n)$ if and only if the vectors x_ν are linearly dependent.

4.3. Uniqueness. Let Δ and Δ_1 be two determinant-functions in E and assume that Δ_1 is non-trivial. Employing a basis e_ν $(\nu = 1 \dots n)$ we have the relations

$$\Delta(x_1 \dots x_n) = \Delta(e_1 \dots e_n) \sum_\sigma \varepsilon_\sigma \xi_1^{\sigma(1)} \dots \xi_n^{\sigma(n)} \tag{4.4}$$

and

$$\Delta_1(x_1 \dots x_n) = \Delta_1(e_1 \dots e_n) \sum_\sigma \varepsilon_\sigma \xi_1^{\sigma(1)} \dots \xi_n^{\sigma(n)} . \tag{4.5}$$

Since Δ_1 is non-trivial,

$$\Delta_1(e_1 \dots e_n) \neq 0 .$$

Defining the scalar λ as the quotient

$$\lambda = \frac{\Delta(e_1 \dots e_n)}{\Delta_1(e_1 \dots e_n)}$$

we obtain from (4.4) and (4.5)

$$\Delta(x_1 \dots x_n) = \lambda \Delta_1(x_1 \dots x_n) .$$

Since this is an identity with respect to all vectors $x_1 \dots x_n$, it can be written as

$$\Delta = \lambda \Delta_1 .$$

This formula shows that every determinant-function Δ is a constant multiple of a fixed non-trivial determinant-function Δ_1.

4.4. Existence. To prove that there exist non-trivial determinant-functions in E define the function Δ by

$$\Delta(x_1 \dots x_n) = \sum_\sigma \varepsilon_\sigma \xi_1^{\sigma(1)} \dots \xi_n^{\sigma(n)} . \tag{4.6}$$

It is immediately clear that Δ is linear with respect to every argument. Furthermore, Δ is not identically zero since

$$\Delta(e_1 \dots e_n) = 1 .$$

It remains to be shown that Δ is skew-symmetric with respect to all arguments.

Consider a fixed permutation τ of $(1 \ldots n)$. Then

$$\Delta\,(x_{\tau(1)} \ldots x_{\tau(n)}) = \sum_\sigma \varepsilon_\sigma\, \xi^{\sigma(\tau)}_{\tau(1)} \ldots \xi^{\sigma(n)}_{\tau(n)}\,.$$

Rearranging the factors in every term so that the subscripts appear in the natural order we can write

$$\Delta\,(x_{\tau(1)} \ldots x_{\tau(n)}) = \sum_\sigma \varepsilon_\sigma\, \xi_1^{\sigma\tau^{-1}(1)} \ldots \xi_n^{\sigma\tau^{-1}(n)}\,.$$

Now, if σ runs over all the permutations of $(1 \ldots n)$, the same holds for the permutations $\varrho = \sigma\tau^{-1}$. Therefore we can introduce ϱ as a new "index of summation" and obtain

$$\Delta\,(x_{\tau(1)} \ldots x_{\tau(n)}) = \sum_\varrho \varepsilon_{\varrho\tau}\, \xi_1^{\varrho(1)} \ldots \xi_n^{\varrho(n)}\,. \tag{4.7}$$

Since

$$\varepsilon_{\varrho\tau} = \varepsilon_\varrho\, \varepsilon_\tau$$

we finally obtain from (4.7)

$$\Delta\,(x_{\tau(1)} \ldots x_{\tau(n)}) = \varepsilon_\tau \sum_\varrho \varepsilon_\varrho \xi_1^{\varrho(1)} \ldots \xi_n^{\varrho(n)} = \varepsilon_\tau \Delta\,(x_1 \ldots x_n)\,.$$

Thus the equation (4.6) defines a non-trivial determinant-function.

4.5. Dual determinant-functions. Let E^* and E be two dual spaces, Δ^* a determinant-function in E^* and Δ a determinant-function in E. It will be shown that the product $\Delta^* \cdot \Delta$ attains the same value on every two systems of dual bases $\overset{*}{e}{}^\nu$, e_ν and $\overset{*}{\bar{e}}{}^\nu$, $\bar{e}_\nu$ $(\nu = 1 \ldots n)$.

Denote by α_ν^μ the matrix of the basis-transformation $\bar{e}_\nu \to e_\nu$. Then

$$e_\nu = \sum_\mu \alpha_\nu^\mu\, \bar{e}_\mu \quad \text{and} \quad \overset{*}{\bar{e}}{}^\nu = \sum_\mu \alpha_\mu^\nu\, \overset{*}{e}{}^\mu\,.$$

Inserting these expressions into $\Delta^*(\overset{*}{\bar{e}}{}^1 \ldots \overset{*}{\bar{e}}{}^n)$ and $\Delta\,(e_1 \ldots e_n)$ we obtain

$$\Delta^*\,(\overset{*}{\bar{e}}{}^1 \ldots \overset{*}{\bar{e}}{}^n) = \sum_\sigma \varepsilon_\sigma\, \alpha^1_{\sigma(1)} \ldots \alpha^n_{\sigma(n)}\, \Delta^*(\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n) \tag{4.8}$$

and

$$\Delta\,(e_1 \ldots e_n) = \sum_\sigma \varepsilon_\sigma \alpha_1^{\sigma(1)} \ldots \alpha_n^{\sigma(n)}\, \Delta\,(\bar{e}_1 \ldots \bar{e}_n)\,. \tag{4.9}$$

Rearranging the factors on the right hand side of (4.9) so that the superscripts appear in the natural order, we obtain

$$\begin{aligned}\Delta\,(e_1 \ldots e_n) &= \sum_\sigma \varepsilon_\sigma \alpha^1_{\sigma^{-1}(1)} \ldots \alpha^n_{\sigma^{-1}(n)}\, \Delta\,(\bar{e}_1 \ldots \bar{e}_n) \\ &= \sum_\varrho \varepsilon_{\varrho^{-1}}\, \alpha^1_{\varrho(1)} \ldots \alpha^n_{\varrho(n)}\, \Delta\,(\bar{e}_1 \ldots \bar{e}_n)\,.\end{aligned}$$

Since $\varepsilon_{\varrho^{-1}} = \varepsilon_\varrho$, this can be written as

$$\Delta\,(e_1 \ldots e_n) = \sum_\varrho \varepsilon_\varrho \alpha^1_{\varrho(1)} \ldots \alpha^n_{\varrho(n)}\, \Delta\,(\bar{e}_1 \ldots \bar{e}_n)\,, \tag{4.10}$$

Comparison of the equations (4.8) and (4.10) yields

$$\Delta^* (\bar{\overset{*}{e}}{}^1 \ldots \bar{\overset{*}{e}}{}^n)\, \Delta\, (\bar{e}_1 \ldots \bar{e}_n) = \Delta^* (\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n)\, \Delta\, (e_1 \ldots e_n)\,.$$

Hence, the product $\Delta^* \cdot \Delta$ assumes a constant value λ for all dual bases $\overset{*}{e}{}^\nu$, e_ν. Now assume that $\Delta \neq 0$ and $\Delta^* \neq 0$. Then $\lambda \neq 0$ and hence Δ^* can be replaced by $\frac{1}{\lambda}\Delta^*$. We then have

$$\Delta^* (\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n)\, \Delta\, (e_1 \ldots e_n) = 1 \quad \text{whenever} \quad \langle \overset{*}{e}{}^\nu, e_\mu \rangle = \delta^\nu_\mu\,. \tag{4.11}$$

Two determinant-functions in E^* and in E satisfying the relation (4.11) are called *dual* to each other. It follows from the above result that to every non-trivial determinant-function in E there exists exactly one dual determinant-function in E^*.

Problem: 1. Let E^*, E be a pair of dual spaces and $\Delta \neq 0$ be a determinant-function in E. Define the function Δ^* of n vectors in E^* as follows:

If the vectors $\overset{*}{x}{}^\nu\,(\nu = 1 \ldots n)$ are linearly dependent, then $\Delta^* (\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^1) = 0$.

If the vectors $\overset{*}{x}{}^\nu (\nu = 1 \ldots n)$ are linearly independent, then $\Delta^* (\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n) = \dfrac{1}{\Delta\, (\overset{*}{x}{}^1 \ldots x_n)}$ where $x_\nu\,(\nu = 1 \ldots n)$ is the dual basis. Prove that Δ^* is a determinant-function in E^* and that the determinant-functions Δ and Δ^* are dual.

§ 2. The determinant of an endomorphism

4.6. Definition. Let φ be an endomorphism of the n-dimensional linear space E. To define the determinant of φ choose a non-trivial determinant-function Δ. Then the function Δ_φ, defined by

$$\Delta_\varphi\, (x_1 \ldots x_n) = \Delta\, (\varphi x_1 \ldots \varphi x_n)$$

obviously is again a determinant-function. Hence, by the uniqueness-theorem of section 4.3,

$$\Delta_\varphi = \alpha \Delta\,,$$

where α is a scalar. This scalar does not depend on the choice of Δ. In fact, if Δ' is another non-trivial determinant-function, then $\Delta' = \lambda \Delta$ and consequently

$$\Delta'_\varphi = \lambda \Delta_\varphi = \lambda \alpha \Delta = \alpha \Delta'\,.$$

Thus, the scalar α is uniquely determined by the endomorphism φ. It is called the *determinant of* φ and it will be denoted by $\det \varphi$. So we have the following equation of definition:

$$\Delta_\varphi = \det \varphi \cdot \Delta\,,$$

where Δ is an arbitrary non-trivial determinant-function. In a less condensed form this equation reads

$$\Delta\,(\varphi x_1 \ldots \varphi x_n) = \det\varphi\,\Delta\,(x_1 \ldots x_n)\,. \tag{4.12}$$

In particular, if $\varphi = \lambda \cdot \iota$, then

$$\Delta_\varphi = \lambda^n \Delta$$

and hence

$$\det\,(\lambda \cdot \iota) = \lambda^n\,.$$

It follows from the above equation that the determinant of the identity-map is 1 and the determinant of the zero-map is zero.

4.7. Properties of the determinant. An endomorphism φ is regular if and only if its determinant is different from zero. To prove this, select a basis e_ν $(\nu = 1 \ldots n)$ of E. Then

$$\Delta\,(\varphi e_1 \ldots \varphi e_n) = \det\varphi\,\Delta\,(e_1 \ldots e_n)\,. \tag{4.13}$$

If φ is regular, the vectors φe_ν $(\nu = 1 \ldots n)$ are linearly independent; hence

$$\Delta\,(\varphi e_1 \ldots \varphi e_n) \neq 0\,. \tag{4.14}$$

The relations (4.13) and (4.14) imply that

$$\det\varphi \neq 0\,.$$

Conversely, assume that $\det\varphi \neq 0$. Then it follows from (4.13) that

$$\Delta\,(\varphi e_1 \ldots \varphi e_n) \neq 0\,.$$

Hence the vectors $\varphi e_\nu (\nu = 1 \ldots n)$ are linearly independent and φ is regular.

Consider two endomorphisms φ and ψ of E. Then

$$\det(\psi \circ \varphi) = \det\psi\,\det\varphi\,. \tag{4.15}$$

In fact,

$$\begin{aligned}\Delta\,(\psi\,\varphi x_1 \ldots \psi\,\varphi\,x_n) &= \det\psi\,\Delta\,(\varphi x_1 \ldots \varphi x_n)\\ &= \det\psi\,\det\varphi\,\Delta\,(x_1 \ldots x_n)\,,\end{aligned}$$

whence (4.15). In particular, if φ is an automorphism and φ^{-1} is the inverse automorphism, we obtain

$$\det\varphi^{-1}\det\varphi = \det\iota = 1\,.$$

4.8. The determinant of the dual endomorphism. Let E^* be a dual space of E. It will be shown that

$$\det\varphi^* = \det\varphi \tag{4.16}$$

for two dual mappings φ and φ^*. First, assume that φ is not regular. Then φ^* is also not regular and hence

$$\det\,\varphi^* = \det\varphi = 0\,.$$

Thus, we may assume that φ is regular. Choosing non-trivial determinant-functions in E and in E^* we have the relations

$$\Delta_\varphi = \det\varphi \cdot \Delta$$

and

$$\Delta^*{}_{\varphi^*} = \det\varphi^* \cdot \Delta^*$$

whence

$$\Delta^* \cdot \Delta_\varphi - \Delta \cdot \Delta^*_{\varphi^*} = (\det\varphi - \det\varphi^*)\, \Delta \cdot \Delta^* .$$

We thus obtain the following identity:

$$\begin{aligned} &\Delta^* (\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n)\, \Delta\, (\varphi x_1 \ldots \varphi x_n) - \Delta\, (x_1 \ldots x_n)\, \Delta^* (\varphi^* \overset{*}{x}{}^1 \ldots \varphi^* \overset{*}{x}{}^n) \\ &= (\det\varphi - \det\varphi^*)\, \Delta\, (x_1 \ldots x_n)\, \Delta^* (\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n) \qquad (4.17) \\ &\qquad \overset{*}{x}{}^\nu \in E^*,\ x_\nu \in E \qquad (\nu = 1 \ldots n) . \end{aligned}$$

Now let $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$ be a pair of dual bases. Substituting $\overset{*}{x}{}^\nu = \overset{*}{e}{}^\nu$ and $x_\nu = \varphi^{-1} e_\nu$ in (4. 17) we obtain

$$\begin{aligned} &\Delta^* (\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n)\, \Delta\, (e_1 \ldots e_n) - \Delta^* (\varphi^* \overset{*}{e}{}^1 \ldots \varphi^* \overset{*}{e}{}^n)\, \Delta\, (\varphi^{-1} e_1 \ldots \varphi^{-1} e_n) \\ &\qquad = (\det\varphi - \det\varphi^*)\, \Delta\, (e_1 \ldots e_n)\, \Delta^* (\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n) . \qquad (4.18) \end{aligned}$$

Since the mappings φ and φ^* are dual, it follows that

$$\langle \varphi^* \overset{*}{e}{}^\nu, \varphi^{-1} e_\mu \rangle = \langle \overset{*}{e}{}^\nu, e_\mu \rangle = \delta^\nu_\mu .$$

This relation shows that the bases $\varphi^* \overset{*}{e}{}^\nu$ and $\varphi^{-1} e_\nu$ are dual again. As it has been proved in sec. 4.5, the product $\Delta^* \cdot \Delta$ assumes the same value for every two dual bases. Consequently the left-hand side of (4.18) is zero. Thus

$$(\det\varphi - \det\varphi^*)\, \Delta\, (e_1 \ldots e_n)\, \Delta^* (\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n) = 0$$

whence

$$\det\varphi^* = \det\varphi .$$

Problems: 1. Consider the endomorphism φ defined by

$$\varphi e_\nu = \lambda_\nu e_\nu$$

where e_ν $(\nu = 1 \ldots n)$ is a basis of E. Show that

$$\det\varphi = \lambda_1 \ldots \lambda_n .$$

2. Assume a direct decomposition $E = E_1 \oplus E_2$. Let φ_1 be an endomorphism of E_1 and φ_2 be an endomorphism of E_2. Define the endomorphism φ of E by

$$\varphi x = \varphi_1 x_1 + \varphi_2 x_2 \quad \text{where} \quad x = x_1 + x_2 \quad (x_1 \in E_1, x_2 \in E_2) .$$

Prove that

$$\det\varphi = \det\varphi_1 \cdot \det\varphi_2 .$$

3. Consider a function F in the space $L(E; E)$ of all endomorphisms such that

$$F(\psi \circ \varphi) = F(\psi) \cdot F(\varphi)$$

and

$$F(\lambda \iota) = \lambda^n .$$

Prove that

$$F(\varphi) = \det \varphi .$$

Hint: Select a basis e_ν $(\nu = 1 \dots n)$ and consider the endomorphisms φ_i $(i = 1 \dots n)$ and ψ_{ij} $(i \neq j)$ defined by

$$\varphi_i e_\nu = \lambda \, \delta_{i\nu} e_\nu \quad (\nu = 1 \dots n)$$

and

$$\psi_{ij} e_\nu = \delta_{i\nu} (e_i + \lambda e_j) \quad (\nu = 1 \dots n) .$$

Show first that

$$F(\varphi_i) = \lambda \quad \text{and} \quad F(\psi_{ij}) = 1 .$$

§ 3. The determinant of a matrix

4.9. Definition. Let φ be an endomorphism of E and (α_ν^μ) the corresponding matrix relative to a basis e_ν $(\nu = 1 \dots n)$. Then

$$\varphi e_\nu = \sum_\mu \alpha_\nu^\mu e_\mu .$$

Substituting $x_\nu = e_\nu$ in (4.12) we obtain

$$\Delta(\varphi e_1, \dots \varphi e_n) = \det \varphi \, \Delta(e_1 \dots e_n) .$$

The left-hand side of this equation can be written as

$$\begin{aligned} \Delta(\varphi e_1, \dots \varphi e_n) &= \Delta\left(\sum_\mu \alpha_1^\mu e_\mu, \dots \sum_\mu \alpha_n^\mu e_\mu\right) \\ &= \sum_\sigma \varepsilon_\sigma \, \alpha_1^{\sigma(1)} \dots \alpha_n^{\sigma(n)} \cdot \Delta(e_1 \dots e_n) . \end{aligned}$$

We thus obtain

$$\det \varphi = \sum_\sigma \varepsilon_\sigma \, \alpha_1^{\sigma(1)} \dots \alpha_n^{\sigma(n)} . \tag{4.19}$$

This formula shows how the determinant of φ is expressed in terms of the corresponding matrix.

We now define the *determinant of an $n \times n$-matrix* $A = (\alpha_\nu^\mu)$ by

$$\det A = \sum_\sigma \varepsilon_\sigma \, \alpha_1^{\sigma(1)} \dots \alpha_n^{\sigma(n)} . \tag{4.20}$$

Then the equation (4.19) can be written as

$$\det \varphi = \det M(\varphi) . \tag{4.21}$$

It is also usual to designate the determinant of an $n \times n$-matrix by two bars,

$$\det A = \begin{vmatrix} \alpha_1^1 & \dots & \alpha_1^n \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha_n^1 & \dots & \alpha_n^n \end{vmatrix}$$

In the case $n = 2$ the formula (4.20) yields

$$\begin{vmatrix} \alpha_1^1 & \alpha_1^2 \\ \alpha_2^1 & \alpha_2^2 \end{vmatrix} = \alpha_1^1 \alpha_2^2 - \alpha_1^2 \alpha_2^1 .$$

4.10. Properties of the determinant of a matrix. Using the relation (4.21) we obtain for each of the formulas of sec. 4.7. and 4.8. a corresponding one. The equation (4.15) yields

$$\det (A B) = \det A \det B \tag{4.22}$$

for any two $n \times n$-matrices A and B. In fact, let E be an n-dimensional linear space and φ and ψ be two endomorphisms of E such that (with respect to a given basis)

$$M(\varphi) = A \quad \text{and} \quad M(\psi) = B .$$

Then

$$\begin{aligned} \det (A B) &= \det M(\varphi) M(\psi) = \det M (\psi \circ \varphi) = \det (\psi \circ \varphi) \\ &= \det \varphi \cdot \det \psi = \det M(\varphi) \det M(\psi) = \det A \cdot \det B . \end{aligned}$$

The formula (4.22) yields for two inverse matrices

$$\det A \cdot \det (A^{-1}) = \det J = 1 \quad (J \text{ unit-matrix})$$

showing that

$$\det (A^{-1}) = (\det A)^{-1}.$$

The formula (4.16) implies that transposed matrices have the same determinant,

$$\det A^* = \det A .$$

Now the expansion-formula (4.20) can be also written in the form

$$\det A = \sum_\sigma \varepsilon_\sigma \alpha^1_{\sigma(1)} \dots \alpha^n_{\sigma(n)} .$$

In other words, the sum can be extended over the subscripts instead of over the superscripts.

We finally note that an $n \times n$-matrix A is regular if and only if $\det A \neq 0$. This follows from (4.21) and from the corresponding property of the determinant of φ.

4.11. The determinant considered as a function of its rows. If the rows $a_\nu = (\alpha_\nu^1 \dots \alpha_\nu^n)$ of the matrix A are considered as vectors of the number-space R^n the determinant $\det A$ appears as a function of the n vectors

a_ν $(\nu = 1 \dots n)$. To investigate this function define an endomorphism $\varphi: R^n \to R^n$ by

$$\varphi e_\nu = a_\nu \quad (\nu = 1 \dots n)$$

where the vectors e_ν are the n-tuples

$$e_\nu = (\underbrace{0 \dots 1}_{\nu} \dots 0) \quad (\nu = 1 \dots n) .$$

Then A is the matrix of φ relative to the basis e_ν. Now let Δ be the determinant-function in R^n which assumes the value one at the basis e_ν $(\nu = 1 \dots n)$,

$$\Delta\,(e_1 \dots e_n) = 1 .$$

Then

$$\Delta\,(a_1 \dots a_n) = \Delta\,(\varphi e_1 \dots \varphi e_n) = \det \varphi\, \Delta\,(e_1 \dots e_n) = \det \varphi$$

and hence

$$\det A = \Delta\,(a_1 \dots a_n) . \tag{4.23}$$

This formula shows that the determinant of A considered as a function of the row-vectors has the following properties:

1. The determinant is linear with respect to every row-vector.
2. If two row-vectors are interchanged the determinant changes the sign.
3. The determinant does not change if to a row-vector a multiple of another row-vector is added.
4. The determinant is different from zero if and only if the row-vectors are linearly independent.

It is obvious that the same properties hold if the determinant is considered as a function of the column-vectors.

4.12. A relation between determinant-functions in E* and E. We know from sec. 4.5, that the product of two determinant-functions in two dual spaces E^* and E assumes the same value on every pair of dual bases. Generalizing this result we shall now prove that

$$\Delta^*\,(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^n)\,\Delta\,(x_1 \dots x_n) = \alpha \begin{vmatrix} \langle \overset{*}{x}{}^1, x_1\rangle \cdots \langle \overset{*}{x}{}^n, x_1\rangle \\ \cdot \\ \cdot \\ \cdot \\ \langle \overset{*}{x}{}^1, x_n\rangle \cdots \langle \overset{*}{x}{}^n, x_n\rangle \end{vmatrix} \tag{4.24}$$

where α is a constant. Denote by Ω the function of the $2n$ vectors $\overset{*}{x}{}^\nu$ and x_ν on the right hand-side of (4.24). It follows from the properties of a determinant listed in sec. 4.11 that the function Ω is linear and skew-symmetric with respect to the vectors x_ν and with respect to the vectors $\overset{*}{x}{}^\nu$. Hence, for every fixed system of vectors $\overset{*}{x}{}^\nu$ $(\nu = 1 \dots n)$,

Ω is a determinant-function in E. Now the uniqueness-theorem of sec. 4.3 implies that

$$\Omega(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n; x_1 \ldots x_n) = \Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n)\, \Delta(x_1 \ldots x_n) \tag{4.25}$$

where Φ depends only on the vectors $\overset{*}{x}{}^\nu$ $(\nu = 1 \ldots n)$. Inserting a fixed basis for the vectors x_ν in (4.25) we see, that the function Φ is linear and skew-symmetric with respect to all n arguments. Applying the uniqueness-theorem again we find that

$$\Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n) = \alpha \Delta(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n) \tag{4.26}$$

where α is a constant. The equations (4.25) and (4.26) yield (4.24).

Now assume that the determinant-functions Δ^* and Δ are dual. Inserting a pair of dual bases into (4.24) we obtain $\alpha = 1$. Then the identity (4.24) reads

$$\Delta^*(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n)\, \Delta(x_1 \ldots x_n) = \begin{vmatrix} \langle \overset{*}{x}{}^1, x_1 \rangle \cdots \langle \overset{*}{x}{}^n, x_1 \rangle \\ \cdot \\ \cdot \\ \cdot \\ \langle \overset{*}{x}{}^1, x_n \rangle \cdots \langle \overset{*}{x}{}^n, x_n \rangle \end{vmatrix} \tag{4.27}$$

Problems: 1. Let $A = (\alpha_\nu^\mu)$ be a matrix such that $\alpha_\nu^\mu = 0$ if $\nu < \mu$. Prove that

$$\det A = \alpha_1^1 \ldots \alpha_n^n .$$

2. Prove that the determinant of the $n \times n$-matrix

$$\alpha_\nu^\mu = 1 - \delta_\nu^\mu$$

is equal to $(n-1)(-1)^{n-1}$.

Hint: Consider the mapping $\varphi: E \to E$ defined by

$$\varphi e_\nu = \sum_\mu e_\mu - e_\nu \quad (\nu = 1 \ldots n)$$

3. Using the expansion formula (4.20), prove that

$$\det A^* = \det A .$$

4. Given an $n \times n$-matrix $A = (\alpha_\nu^\mu)$ define the matrix $B = (\beta_\nu^\mu)$ by

$$\beta_\nu^\mu = (-1)^{\nu+\mu} \alpha_\nu^\mu .$$

Prove that

$$\det B = \det A .$$

5. How many operations are necessary to evaluate a determinant of order n using the definition? How many steps are necessary to evaluate the same determinant using elementary row or column operations?

6. Given n complex numbers α_ν prove that

$$\begin{vmatrix} \alpha_1 & \alpha_2 \dots \alpha_{n-1} & \alpha_n \\ \alpha_2 & \alpha_3 \dots \alpha_n & \alpha_1 \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \alpha_n & \alpha_1 \dots \alpha_{n-2} & \alpha_{n-1} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} \beta_1 \dots \beta_n$$

where the numbers β_k are defined by

$$\beta_k = \sum_\nu \varepsilon_k^\nu \alpha_\nu \qquad \varepsilon_k = \cos\frac{2\pi k}{n} + i \sin\frac{2\pi k}{n} \qquad (k = 1 \dots n)\,.$$

Hint: Multiply the above matrix by the matrix

$$\begin{pmatrix} \varepsilon_1 \dots \varepsilon_n \\ \varepsilon_1^2 \dots \varepsilon_n^2 \\ \dots\dots \\ \dots\dots \\ \varepsilon_1^n \dots \varepsilon_n^n \end{pmatrix}.$$

§ 4. Cofactors

4.13. Definition. Consider an $n \times n$-matrix $A = (\alpha_\nu^\mu)$. Replacing the element α_i^j by 1 and all other elements of row i and column j by zero, we obtain the matrix

$$C_i^j = \begin{pmatrix} \alpha_1^1 & \dots \alpha_1^{j-1} & 0 & \alpha_1^{j+1} \dots \alpha_1^n \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \alpha_{i-1}^1 & \dots \alpha_{i-1}^{j-1} & 0 & \alpha_{i-1}^{j+1} \dots \alpha_{i-1}^n \\ 0 & \dots\ 0 & 1 & 0\ \dots\ 0 \\ \alpha_{i+1}^1 & \dots \alpha_{i+1}^{j-1} & 0 & \alpha_{i+1}^{j+1} \dots \alpha_{i+1}^n \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \alpha_n^1 & \dots \alpha_n^{j-1} & 0 & \alpha_n^{j+1} \dots \alpha_n^n \end{pmatrix}.$$

The determinant of C_i^j is called the *cofactor* of the element α_i^j and the $n \times n$-matrix (β_j^i) defined by the determinants

$$\beta_j^i = \det C_i^j$$

is called the *adjoint matrix of* A *). In other words, the adjoint of a matrix is the transpose of the matrix formed by the cofactors. Applying the formula (4.23) to C_i^j we obtain

$$\det C_i^j = \Delta\,(a_1 \dots a_{i-1}, e_j, a_{i+1} \dots a_n)\,.$$

*) In the above equation j counts the row and i counts the column.

Multiplication by α^j_k $(1 \leqq k \leqq n)$ and summation over j yields

$$\sum_j \alpha^j_k \beta^i_j = \Delta\,(a_1 \ldots a_{i-1}, \sum_j \alpha^j_k e_j, a_{i+1} \ldots a_n)$$
$$= \Delta\,(a_1 \ldots a_{i-1}, a_k, a_{i+1} \ldots a_n)\,.$$

If $k \neq i$, the vector a_k appears twice on the right hand-side whence

$$\sum_j \alpha^j_k \beta^i_j = 0 \quad \text{if} \quad i \neq k\,. \tag{4.28}$$

Now assume that $k = i$. Then

$$\Delta\,(a_1 \ldots a_{i-1}, a_i, a_{i+1} \ldots a_n) = \det A$$

and we thus obtain

$$\sum_j \alpha^j_i \beta^i_j = \det A \quad (i = 1 \ldots n)\,. \tag{4.29}$$

The relations (4.28) and (4.29) can be combined in the formula

$$\sum_j \alpha^j_k \beta^i_j = \delta^i_k \cdot \det A \quad (i, k = 1 \ldots n)\,. \tag{4.30}$$

Denoting the adjoint matrix by $\operatorname{ad} A$ we can write the equation (4.30) as

$$A \cdot \operatorname{ad} A = J \cdot \det A\,.$$

4.14. The inverse matrix. Assume that $\det A \neq 0$. Then the equations (4.30) can be divided by $\det A$ yielding

$$\frac{1}{\det A} \sum_j \alpha^j_k \beta^i_j = \delta^i_k \quad (i, k = 1 \ldots n)\,.$$

This equation shows that the matrix

$$\check{\alpha}^i_j = \frac{1}{\det A}\,\beta^i_j \tag{4.31}$$

is the inverse of (α^i_j).

Applying the relation (4.30) to a system of n linear equations with n unknowns we obtain the *Cramer's solution* formula. Let

$$\sum_k \alpha^j_k \xi^k = \eta^j \quad (j = 1 \ldots n) \tag{4.32}$$

be the given system and assume that the determinant of the matrix (α^j_k) is different from zero. Multiplying the j^{th} equation by β^i_j and summing with respect to j we obtain in view of (4.30)

$$\xi^i \det A = \sum_j \beta^i_j \eta^j$$

whence

$$\xi^i = \frac{1}{\det A} \sum_j \beta^i_j \eta^j = \frac{1}{\det A} \sum_j \det C^j_i\, \eta^j\,. \tag{4.33}$$

In this formula the solution of the system (4.32) is expressed in term of the numbers $(\eta^1 \dots \eta^n)$ and the cofactors of the matrix A.

4.15. The submatrices S^j_i. Denote by S^j_i the $(n-1) \times (n-1)$-matrix obtained from A by deleting the row i and the column j. It will be shown that

$$\det C^j_i = (-1)^{i+j} \det S^j_i . \tag{4.34}$$

Assume first that $i = 1$ and $j = 1$. Then, by (4.20)

$$\det S^1_1 = \sum_{\varrho} \varepsilon_{\varrho} \alpha^{\varrho(2)}_2 \dots \alpha^{\varrho(n)}_n \tag{4.35}$$

where the summation is taken over all the permutations of the numbers $(2 \dots n)$. The elements of C^1_1 are given by $\beta^{\mu}_1 = \delta^{\mu}_1$ and $\beta^{\mu}_{\nu} = \alpha^{\mu}_{\nu} - \delta^{\mu}_1 \alpha^1_{\nu}$ $(\nu = 2 \dots n)$. Hence, the expansion-formula (4.20) yields

$$\det C^1_1 = \sum_{\sigma} \varepsilon_{\sigma} \left(\alpha^{\sigma(2)}_2 - \delta^{\sigma(2)}_1 \alpha^1_2\right) \dots \left(\alpha^{\sigma(n)}_n - \delta^{\sigma(n)}_1 \alpha^1_n\right) . \tag{4.36}$$

In this sum all terms are zero for which $\sigma(1) \neq 1$. Consequently, (4.36) can be written as

$$\det C^1_1 = \sum_{\sigma} \varepsilon_{\sigma} \alpha^{\sigma(2)}_2 \dots \alpha^{\sigma(n)}_n$$

where the summation is taken over all permutations leaving the number 1 fixed. Every such permutation σ induces a permutation ϱ of the numbers $(2 \dots n)$. Since σ and τ have the same parity, it follows that

$$\det C^1_1 = \sum_{\varrho} \varepsilon_{\varrho} \alpha^{\varrho(2)}_2 \dots \alpha^{\varrho(n)}_n \tag{4.37}$$

where ϱ runs over all the permutations of $(2 \dots n)$. The equations (4.35) and (4.37) yield (4.34) for the case $i = 1, j = 1$.

Now we proceed to the general case. Interchanging the row i with all the preceding rows, and the column j with all the preceding columns, the matrix C^j_i is converted into the matrix

$$B = \begin{pmatrix} 1 & 0 \; . \; . \; . \; . \; 0 \\ 0 & \\ \vdots & \boxed{\quad S^j_i \quad} \\ 0 & \end{pmatrix} .$$

The determinants of B and C^j_i are obviously related by

$$\det B = (-1)^{i+j} \det C^j_i . \tag{4.38}$$

Now S^j_i is obtained from B by deleting the first row and the first column and hence, as it has been shown above,

$$\det B = \det S^j_i . \tag{4.39}$$

The equations (4.38) and (4.39) yield (4.34).

4.16. Expansion by cofactors. From the relations (4.34) and (4.29) we obtain the *expansion-formula* of the determinant with respect to the i^{th} row,

$$\det A = \sum_j (-1)^{i+j} \alpha_i^j \det S_i^j \qquad (i = 1 \ldots n) . \tag{4.40}$$

By this formula the evaluation of the determinant of n rows is reduced to the evaluation of n determinants of $n-1$ rows.

In the same way the expansion-formula with respect to the j^{th} column is proved:

$$\det A = \sum_i (-1)^{i+j} \alpha_i^j \det S_i^j \quad (j = 1 \ldots n) . \tag{4.41}$$

In the case $n = 3$, $i = 1, j = 1$ the formula (4.40) reads

$$\begin{vmatrix} \alpha_1^1 & \alpha_1^2 & \alpha_1^3 \\ \alpha_2^1 & \alpha_2^2 & \alpha_2^3 \\ \alpha_3^1 & \alpha_3^2 & \alpha_3^3 \end{vmatrix} = \alpha_1^1 \begin{vmatrix} \alpha_2^2 & \alpha_2^3 \\ \alpha_3^2 & \alpha_3^3 \end{vmatrix} - \alpha_1^2 \begin{vmatrix} \alpha_2^1 & \alpha_2^3 \\ \alpha_3^1 & \alpha_3^3 \end{vmatrix} + \alpha_1^3 \begin{vmatrix} \alpha_2^1 & \alpha_2^2 \\ \alpha_3^1 & \alpha_3^2 \end{vmatrix} .$$

4.17. Minors. Let $A = (\alpha_\nu^\mu)$ be a given $n \times m$-matrix. For every system of indices

$$1 \leqq i_1 < i_2 < \cdots < i_k \leqq n \quad \text{and} \quad 1 \leqq j_1 < j_2 < \cdots < j_k \leqq m$$

denote by $A_{i_1 \ldots i_k}^{j_1 \ldots j_k}$ the submatrix of A, consisting of the rows $i_1 \ldots i_k$ and the columns $j_1 \ldots j_k$. The determinant of $A_{i_1 \ldots i_k}^{j_1 \ldots j_k}$ is called a *minor of order k* of the matrix A. It will be shown that in a matrix of rank r there is always a minor of order r which is different from zero, whereas all minors of order $k > r$ are zero. Let $A_{i_1 \ldots i_k}^{j_1 \ldots j_k}$ be a minor of order $k > r$. Then the row-vectors $a_{i_1} \ldots a_{i_k}$ of A are linearly dependent. This implies that the rows of the matrix $A_{i_1 \ldots i_k}^{j_1 \ldots j_k}$ are also linearly dependent and thus the determinant must be zero.

It remains to be shown that there is a minor of order r which is different from zero. Since A has the rank r, there are r linearly independent row-vectors $a_{i_1} \ldots a_{i_r}$. The submatrix consisting of these row-vectors has again the rank r. Therefore it must contain r linearly independent column-vectors $b^{j_1} \ldots b^{j_r}$. Consider the matrix $A_{i_1 \ldots i_r}^{j_1 \ldots j_r}$. Its column-vectors are linearly independent, whence

$$\det A_{i_1 \ldots i_r}^{j_1 \ldots j_r} \neq 0 .$$

If A is a square-matrix, the minors

$$\det A_{i_1 \ldots i_k}^{i_1 \ldots i_k}$$

are called the *principal minors* of order k.

Problems: 1. Compute the inverse of the following matrices.

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

$$B = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \qquad (\lambda \neq 0)$$

2. Show that

a) $$\begin{vmatrix} x & a & a \dots a \\ a & x & a \dots a \\ \vdots & & \vdots \\ a & \dots & a\ x \end{vmatrix} = [x + (n-1)a](x-a)^{n-1}$$

and that

b) $$\begin{vmatrix} 1 & 1 & \dots 1 \\ \lambda_1 & \lambda_2 & \dots \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \dots \lambda_n^2 \\ \dots & \dots & \dots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots \lambda_n^{n-1} \end{vmatrix} = \prod_{i>j} (\lambda_i - \lambda_j)\,.$$

(*Vandermonde* determinant.)

3. Define

$$\Delta_n = \begin{vmatrix} x_1 & 1 & & & & \\ -1 & x_2 & 1 & & & \\ & -1 & x_3 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & & x_{n-1} & 1 \\ & & & & -1 & x_n \end{vmatrix}$$

Show that $\Delta_n = x_n \Delta_{n-1} + \Delta_{n-2} \qquad (n > 2)$

$$\Delta_1 = x_1; \quad \Delta_2 = x_1 x_2 + 1\,.$$

4. Verify the following formula for *quasi-triangular* determinant:

$$\begin{vmatrix} x_{11} \dots x_{1p} & 0 \dots 0 \\ \vdots \quad \vdots & \vdots \quad \vdots \\ x_{p1} \dots x_{pp} & 0 \dots 0 \\ x_{p+1\,1} \dots & \dots x_{p+1\,n} \\ \vdots & \vdots \\ x_{n1} \dots & \dots x_{nn} \end{vmatrix} = \begin{vmatrix} x_{11} \dots x_{1p} \\ \vdots \quad \vdots \\ x_{p1} \dots x_{pp} \end{vmatrix} \cdot \begin{vmatrix} x_{p+1\,p+1} \dots x_{p+1\,n} \\ \vdots \quad \vdots \\ x_{n1} \dots x_{nn} \end{vmatrix}$$

5. Prove that the operation $A \to \operatorname{ad} A$ has the following properties provided that the matrix A is regular*)

a) $\operatorname{ad}(AB) = \operatorname{ad} A \cdot \operatorname{ad} B$

b) $\det \operatorname{ad} A = (\det A)^{n-1}$.

c) $\operatorname{ad} \operatorname{ad} A = (\det A)^{n-2} \cdot A$.

d) $\det \operatorname{ad} \operatorname{ad} A = (\det A)^{(n-1)^2}$.

§ 5. The characteristic polynomial

4.18. Eigenvectors. Consider an endomorphism φ of an n-dimensional linear space E. A vector $a \neq 0$ of E is called an *eigenvector* of φ if

$$\varphi a = \lambda a . \tag{4.42}$$

The scalar λ determined by (4.42) is called the corresponding *eigenvalue.* An endomorphism φ need not to have eigenvectors. As an example let E be a real linear space of two dimensions and define φ by

$$\varphi x_1 = x_2 \quad \varphi x_2 = - x_1$$

where the vectors x_1 and x_2 form a basis of E. This endomorphism does not have to have eigenvectors. In fact, assume that

$$a = \xi^1 x_1 + \xi^2 x_2$$

is an eigenvector. Then $\varphi a = \lambda a$ and hence

$$\xi^1 = \lambda \xi^2, \qquad \xi^2 = - \lambda \xi^1 .$$

These equations yield

$$(\xi^1)^2 + (\xi^2)^2 = 0$$

whence $\xi^1 = 0$ and $\xi^2 = 0$.

4.19. The characteristic equation. Assume that a is an eigenvector of φ and λ is the corresponding eigenvalue. Then

$$\varphi a = \lambda a , \quad a \neq 0 .$$

This equation can be written as

$$(\varphi - \lambda \iota)\, a = 0 \tag{4.43}$$

showing that the endomorphism $\varphi - \lambda \iota$ is not regular. This implies that

$$\det(\varphi - \lambda \iota) = 0 . \tag{4.44}$$

Hence, every eigenvalue of φ satisfies the equation (4.44). Conversely, assume that λ is a solution of the equation (4.44). Then the endomorphism $\varphi - \lambda \iota$ is not regular. Consequently there is a vector $a \neq 0$ such that

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

whence $\varphi a = \lambda a$.

*) It will be shown in Chap. IX, § 4, that these relations are also valid for a non-regular matrix.

Thus, the eigenvalues of φ are the solutions of the equation (4.44). This equation is called the *characteristic equation* of the endomorphism φ.

4.20. The characteristic polynomial. To obtain a more explicit expression for the characteristic equation choose a determinant-function $\Delta \neq 0$ in E. Then

$$\begin{gathered}\Delta\,(\varphi x_1 - \lambda x_1 \dots \varphi x_n - \lambda x_n) = \det\,(\varphi - \lambda\,\iota)\,\Delta\,(x_1 \dots x_n)\\ x_\nu \in E\ (\nu = 1 \dots n)\,.\end{gathered} \tag{4.45}$$

Expanding the left hand-side we obtain a sum of 2^n terms of the form

$$\Delta\,(z_1 \dots z_n)\,,$$

where every argument is either φx_ν or $-\lambda x_\nu$. Denote by S_p $(0 \leqq p \leqq n)$ the sum of all terms in which p arguments are equal to φx_ν and $n - p$ arguments are equal to $-\lambda x_\nu$. Collect in each term of S_p the indices $\nu_1 \dots \nu_p$ $(\nu_1 < \dots < \nu_p)$ such that

$$z_{\nu_1} = \varphi x_{\nu_1} \dots z_{\nu_p} = \varphi x_{\nu_p}$$

and the indices $\nu_{p+1} \dots \nu_n$ $(\nu_{p+1} < \dots < \nu_n)$ such that

$$z_{\nu_{p+1}} = -\lambda x_{\nu_{p+1}} \dots z_{\nu_n} = -\lambda x_{\nu_n}\,.$$

Introducing the permutation σ by

$$\sigma(i) = \nu_i \quad (i = 1 \dots n)$$

we can write

$$\begin{aligned}\Delta\,(z_1 \dots z_n) &= \varepsilon_\sigma \Delta\,(z_{\sigma(1)} \dots z_{\sigma(n)})\\ &= \varepsilon_\sigma \Delta\,(\varphi x_{\sigma(1)} \dots \varphi x_{\sigma(p)}\,,\ -\lambda x_{\sigma(p+1)} \dots -\lambda x_{\sigma(n)})\\ &= (-\lambda)^{n-p}\,\varepsilon_\sigma \Delta\,(\varphi x_{\sigma(1)} \dots \varphi x_{\sigma(p)}\,,\ x_{\sigma(p+1)} \dots x_{\sigma(n)})\,.\end{aligned}$$

Thus,

$$S_p = (-\lambda)^{n-p} \sum_\sigma \varepsilon_\sigma \Delta\,(\varphi x_{\sigma(1)} \dots \varphi x_{\sigma(p)}\,,\ x_{\sigma(p+1)} \dots x_{\sigma(n)}) \tag{4.46}$$

where the sum is extended over all permutations σ subject to the conditions

$$\sigma(1) < \dots < \sigma(p) \quad \text{and} \quad \sigma(p+1) < \dots < \sigma(n)\,.$$

Observing the skew symmetry of Δ we obtain from (4.46)

$$S_p = \frac{(-\lambda)^{n-p}}{p!\,(n-p)!} \sum_\sigma \varepsilon_\sigma \Delta\,(\varphi x_{\sigma(1)} \dots \varphi x_{\sigma(p)},\ x_{\sigma(p+1)} \dots x_{\sigma(n)}) \tag{4.47}$$

where the sum on the right hand-side is taken over all permutations. Let Φ_p be the function defined by

$$\Phi_p\,(x_1 \dots x_n) = \sum_\sigma \varepsilon_\sigma \Delta\,(\varphi x_{\sigma(1)} \dots \varphi x_{\sigma(p)},\ x_{\sigma(p+1)} \dots x_{\sigma(n)}) \quad (0 \leqq p \leqq n)$$

and τ be an arbitrary permutation of $(1 \ldots n)$. Then

$$\begin{aligned}\Phi_p(x_{\tau(1)} \ldots x_{\tau(n)}) &= \sum_\sigma \varepsilon_\sigma \Delta(\varphi x_{\tau\sigma(1)} \ldots \varphi x_{\tau\sigma(p)}, x_{\tau\sigma(p+1)} \ldots x_{\tau\sigma(n)}) \\ &= \varepsilon_\tau \sum_\sigma \varepsilon_{\tau\sigma} \Delta(\varphi x_{\tau\sigma(1)} \ldots \varphi x_{\tau\sigma(p)}, x_{\tau\sigma(p+1)} \ldots x_{\tau\sigma(n)}) \\ &= \varepsilon_\tau \sum_\sigma \varepsilon_\varrho \Delta(\varphi x_{\varrho(1)} \ldots \varphi x_{\varrho(p)}, x_{\varrho(p+1)} \ldots x_{\varrho(n)}) \\ &= \varepsilon_\tau \Phi_p(x_1 \ldots x_n).\end{aligned}$$

This equation shows that Φ_p is skew-symmetric with respect to all arguments. This implies that

$$\Phi_p = (-1)^{n-p} p!\,(n-p)!\,\alpha_p \cdot \Delta \tag{4.48}$$

where α_p is a scalar. Inserting (4.48) into (4.47) we obtain

$$S_p = \alpha_p \lambda^{n-p} \cdot \Delta .$$

Hence, the left hand-side of (4.45) can be written as

$$\Delta(\varphi x_1 - \lambda x_1, \ldots \varphi x_n - \lambda x_n) = \Delta(x_1 \ldots x_n) \sum_{p=0}^{n} \alpha_p \lambda^{n-p} . \tag{4.49}$$

Now the equations (4.45) and (4.49) yield

$$\det(\varphi - \lambda\,\iota) = \sum_{p=0}^{n} \alpha_p \lambda^{n-p}$$

showing that the determinant of $\varphi - \lambda\,\iota$ is a polynomial of degree n in λ. This polynomial is called the *characteristic polynomial* of the endomorphism φ. The coefficients of the characteristic polynomial are determined by the equation (4.48).

These relations yield for $p = 0$ and $p = n$

$$\alpha_0 = (-1)^n \quad \text{and} \quad \alpha_n = \det\varphi$$

respectively.

4.21. Existence of eigenvalues. Combining the results of sec. 4.19 and 4.20, we see that the eigenvalues of the endomorphism φ are the roots of the characteristic polynomial

$$f(\lambda) = \sum_{\nu=0}^{n} \alpha_\nu \lambda^{n-\nu} .$$

This shows that *an endomorphism of an n-dimensional linear space has at most n different eigenvalues.*

Assume that E is a complex linear space. Then, according to the fundamental theorem of algebra, the polynomial f has at least one zero. Consequently, *every endomorphism of a complex linear space has at least one eigenvalue.*

If E is a real linear space, this does not generally hold, as it has been shown in the beginning of this paragraph.

Now assume that the dimension of E is odd. Then

$$\lim_{\lambda\to\infty} f(\lambda) = -\infty \quad \text{and} \quad \lim_{\lambda\to-\infty} f(\lambda) = +\infty$$

and thus the polynomial $f(\lambda)$ must have at least one zero. This proves that an *endomorphism of an odd-dimensional linear space has at least one eigenvalue*. Observing that

$$f(0) = \alpha_n = \det \varphi$$

we see that an endomorphism of positive determinant has at least one positive eigenvalue and an endomorphism of negative determinant has at least one negative eigenvalue, provided that E has odd dimension.

If the dimension of E is even we have the relations

$$\lim_{\lambda\to\infty} f(\lambda) = \infty \quad \text{and} \quad \lim_{\lambda\to-\infty} f(\lambda) = \infty$$

and hence nothing can be said if $\det\varphi > 0$. However, if $\det\varphi < 0$, there exists at least one positive and one negative eigenvalue.

4.22. The characteristic polynomial of the inverse mapping. Assume that φ is an automorphism and that φ^{-1} is the inverse automorphism. The characteristic polynomial of φ^{-1} is defined by

$$F(\lambda) = \det(\varphi^{-1} - \lambda\,\iota)\,.$$

Now,

$$\varphi^{-1} - \lambda\,\iota = \varphi^{-1}\circ(\iota - \lambda\,\varphi) = -\lambda\,\varphi^{-1}\circ\left(\varphi - \frac{1}{\lambda}\,\iota\right),$$

whence

$$\det(\varphi^{-1} - \lambda\,\iota) = (-\lambda)^n \det\varphi^{-1}\cdot\det\left(\varphi - \frac{1}{\lambda}\,\iota\right).$$

This equation shows that the characteristic polynomials of φ and of φ^{-1} are related by

$$F(\lambda) = (-\lambda)^n \det\varphi^{-1} f\left(\frac{1}{\lambda}\right).$$

Expanding $F(\lambda)$ as

$$F(\lambda) = \sum_{\nu=0}^{n} \beta_\nu \lambda^{n-\nu}$$

we obtain the following relations between the coefficients of f and of F:

$$\beta_\nu = (-1)^n \det\varphi^{-1}\,\alpha_{n-\nu} \quad (\nu = 0\ldots n)\,.$$

4.23. The characteristic polynomial of a matrix. Let e_ν $(\nu = 1\ldots n)$ be a basis of E and $A = M(\varphi)$ the matrix of the endomorphism φ relative to this basis. Then

$$M(\varphi - \lambda\iota) = M(\varphi) - \lambda M(\iota) = A - \lambda J$$

whence

$$\det(\varphi - \lambda\iota) = \det M(\varphi - \lambda\iota) = \det(A - \lambda J)\,.$$

Thus, the characteristic polynomial of φ can be written as

$$f(\lambda) = \det(A - \lambda J)\,. \tag{4.50}$$

The polynomial (4.50) is called the *characteristic polynomial of the matrix* A. The roots of the polynomial f are called the *eigenvalues of the matrix* A.

Problems: 1. Compute the eigenvalues of the matrix

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

2. Show that the eigenvalues of the matrix

$$\begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix} \quad \text{are real.}$$

3. Prove that the characteristic polynomial of a projection $\pi: E \to E_1$ (see Chapter II, § 4, Problem 1) is given by

$$f(\lambda) = (-1)^n \lambda^{n-p} (1 - \lambda)^p$$

where $n = \dim E$ and $p = \dim E_1$.

4. Show that the coefficients of the characteristic polynomial of an involution satisfy the relations

$$\alpha_p = \varepsilon\, \alpha_{n-p} \qquad \varepsilon = \pm 1 \qquad (p = 0 \ldots n)\,.$$

5. Consider a direct decomposition $E = E_1 \oplus E_2$. Given an endomorphism φ_1 of E_1 and an endomorphism φ_2 of E_2 define the endomorphism φ of E by

$$\varphi x = \varphi_1 x_1 + \varphi_2 x_2$$

where $x = x_1 + x_2$ with $x_1 \in E_1$ and $x_2 \in E_2$. Prove that the characteristic polynomial of φ is the product of the characteristic polynomials of φ_1 and of φ_2.

6. Given two endomorphisms φ and ψ show that the products $\psi \circ \varphi$ and $\varphi \circ \psi$ have the same characteristic polynomial provided that at least one of the endomorphisms is regular*).

7. Show that dual endomorphisms have the same characteristic polynomial.

*) It will be shown in see. 9.5 that this hypothesis is not necessary.

§ 6. The Trace

4.24. The trace of an endomorphism. In a similar way as the determinant, another scalar can be associated with a given endomorphism φ. Let $\Delta \neq 0$ be a determinant-function in E. Consider the sum

$$\sum_{i=1}^{n} \Delta\,(x_1 \ldots \varphi x_i \ldots x_n)\,.$$

This sum obviously is again a determinant-function and thus it can be written as

$$\sum_{i=0}^{n} \Delta\,(x_1 \ldots \varphi x_i \ldots x_n) = \alpha \cdot \Delta\,(x_1 \ldots x_n) \tag{4.51}$$

where α is a scalar. This scalar which is uniquely determined by φ is called the *trace* of the endomorphism φ and will be denoted by $\operatorname{tr} \varphi$. It follows immediately that the trace depends linearly on the endomorphism φ,

$$\operatorname{tr}\,(\lambda\varphi + \mu\psi) = \lambda \operatorname{tr}\varphi + \mu \operatorname{tr}\psi\,.$$

Next we show that

$$\operatorname{tr}\,(\psi \circ \varphi) = \operatorname{tr}\,(\varphi \circ \psi) \tag{4.52}$$

for any two endomorphisms φ and ψ. The trace of $\psi \circ \varphi$ is defined by the equation

$$\sum_i (x_1 \ldots (\psi \circ \varphi)\, x_i \ldots x_n) = \operatorname{tr}\,(\psi \circ \varphi)\, \Delta\,(x_1 \ldots x_n) \quad x_\nu \in E\,.$$

Replacing the vectors x_ν by ψx_ν $(\nu = 1 \ldots n)$ we obtain

$$\begin{aligned} &\sum_i \Delta\,(\psi x_1 \ldots (\psi \circ \varphi \circ \psi)\, x_i \ldots \psi x_n) \\ &= \operatorname{tr}\,(\psi \circ \varphi)\, \Delta\,(\psi x_1 \ldots \psi x_n) = \operatorname{tr}\,(\psi \circ \varphi) \det\psi\, \Delta\,(x_1 \ldots x_n). \end{aligned} \tag{4.53}$$

The left hand-side of this equation can be written as

$$\begin{aligned} \sum_i \Delta\,(\psi x_1 \ldots (\psi \circ \varphi \circ \psi)\, x_i \ldots \psi x_n) &= \det\psi \sum_i \Delta\,(x_1 \ldots (\varphi \circ \psi)\, x_i \ldots x_n) \\ &= \det\psi \cdot \operatorname{tr}\,(\varphi \circ \psi)\, \Delta\,(x_1 \ldots x_n) \end{aligned}$$

and thus (4.53) implies that

$$\det\psi \operatorname{tr}\,(\varphi \circ \psi) = \operatorname{tr}\,(\psi \circ \varphi) \det\psi\,. \tag{4.54}$$

If ψ is regular, this equation may be divided by $\det\psi$ yielding (4.52). If ψ is non-regular, consider the mapping $\psi - \lambda\iota$ where λ is different from all eigenvalues of ψ. Then $\psi - \lambda\iota$ is regular, whence

$$\operatorname{tr}\,[(\psi - \lambda\iota) \circ \varphi] = \operatorname{tr}\,[\varphi \circ (\psi - \lambda\iota)]\,.$$

In view of the linearity of the trace-operator this equation yields

$$\operatorname{tr}\,(\psi \circ \varphi) - \lambda \operatorname{tr}\varphi = \operatorname{tr}\,(\varphi \circ \psi) - \lambda \operatorname{tr}\varphi$$

whence (4.52).

Finally it will be shown that the coefficient of λ^{n-1} in the characteristic polynomial of φ can be written as

$$\alpha_1 = (-1)^{n-1} \operatorname{tr} \varphi . \tag{4.55}$$

The formula (4.48) yields for $p = 1$

$$\sum_\sigma \varepsilon_\sigma \Delta (\varphi x_{\sigma(1)}, x_{\sigma(2)} \dots x_{\sigma(n)}) = (-1)^{n-1} \alpha_1 \Delta (x_1 \dots x_n) \tag{4.56}$$

the sum being taken over all permutations σ subject to the restrictions

$$\sigma(2) < \dots < \sigma(n) .$$

This sum can be written as

$$\sum_{i=1}^{n} (-1)^{i-1} \Delta (\varphi x_i, x_1 \dots \hat{x}_i \dots x_n) = \sum_{i=1}^{n} \Delta (x_1 \dots x_{i-1}, \varphi x_i, x_{i+1} \dots x_n) .$$

We thus obtain from (4.56)

$$\sum_i \Delta (x_1 \dots \varphi x_i \dots x_n) = (-1)^{n-1} \alpha_1 \Delta (x_1 \dots x_n) . \tag{4.57}$$

Comparing the relations (4.57) and (4.51) we find (4.55).

4.25. The trace of a matrix. Let e_ν $(\nu = 1 \dots n)$ be a basis of E. Then the endomorphism φ determines an $n \times n$-matrix α_ν^μ by the equations

$$\varphi e_\nu = \sum_\mu \alpha_\nu^\mu e_\mu . \tag{4.58}$$

Inserting $x_\nu = e_\nu$ $(\nu = 1 \dots n)$ in (4.51) we find

$$\sum_i \Delta (e_1 \dots \varphi e_i \dots e_n) = \operatorname{tr} \varphi \Delta (e_1 \dots e_n) . \tag{4.59}$$

The equations (4.58) and (4.59) imply that

$$\Delta (e_1 \dots e_n) \sum_{i=1}^{n} \alpha_i^i = \Delta (e_1 \dots e_n) \operatorname{tr} \varphi$$

whence

$$\operatorname{tr} \varphi = \sum_i \alpha_i^i . \tag{4.60}$$

Observing that

$$\alpha_\nu^\mu = \langle \overset{*}{e}{}^\mu, \varphi e_\nu \rangle ,$$

where $\overset{*}{e}{}^\nu$ $(\nu = 1 \dots n)$ is the dual basis of e_ν, we can write the equation (4.60) as

$$\operatorname{tr} \varphi = \sum_i \langle \overset{*}{e}{}^i, \varphi e_i \rangle . \tag{4.61}$$

The formula (4.60) shows that the trace of an endomorphism is equal to the sum of all entries in the main-diagonal of the corresponding matrix. For any $n \times n$-matrix $A = (\alpha_\nu^\mu)$ this sum is called the *trace* of A and will be denoted by $\operatorname{tr} A$,

$$\operatorname{tr} A = \sum_i \alpha_i^i . \tag{4.62}$$

Now the equation (4.60) can be written in the form

$$\operatorname{tr}\varphi = \operatorname{tr} M(\varphi) .$$

4.26. The duality of $L(E; F)$ and $L(F; E)$. Now consider two linear spaces E and F and the spaces $L(E; F)$ and $L(F; E)$ of all linear mappings $\varphi: E \to F$ and $\psi: F \to E$. With the help of the trace a scalar-product can be introduced in these spaces in the following way:

$$\langle \varphi, \psi \rangle = \operatorname{tr}(\psi \circ \varphi) \quad \varphi \in L(E; F), \quad \psi \in L(F; E) . \tag{4.63}$$

The function defined by (4.63) is obviously bilinear. Now assume that

$$\langle \varphi, \psi \rangle = 0 \tag{4.64}$$

for a fixed mapping $\varphi \in L(E; F)$ and all linear mappings $\psi \in L(F; E)$. It has to be shown that this implies that $\varphi = 0$. Assume that $\varphi \neq 0$. Then there exists a vector $a \in E$ such that $\varphi a \neq 0$. Extend the vector $b_1 = \varphi a$ to a basis $(b_1 \ldots b_m)$ of F and define the linear mapping $\psi: F \to E$ by

$$\psi b_1 = a, \quad \psi b_\mu = 0 \quad (\mu = 2 \ldots m) .$$

Then

$$(\varphi \circ \psi)\, b_1 = b_1, \quad (\varphi \circ \psi)\, b_\mu = 0 \quad (\mu = 2 \ldots m) ,$$

whence

$$\langle \varphi, \psi \rangle = \operatorname{tr}(\psi \circ \varphi) = \operatorname{tr}(\varphi \circ \psi) = 1 .$$

This is in contradiction with (4.64). Interchanging E and F we see that the relation

$$\langle \varphi, \psi \rangle = 0$$

for a fixed mapping $\psi \in L(F; E)$ and all mappings $\varphi \in L(E; F)$ implies that $\psi = 0$. Hence, a scalar-product is defined in $L(E; F)$ and $L(F; E)$ by (4.63).

Problems: 1. Show that the characteristic polynomial of an endomorphism φ of a 2-dimensional linear space can be written as

$$f(\lambda) = \lambda^2 - \lambda \operatorname{tr}\varphi + \det\varphi .$$

Verify that every endomorphism φ satisfies its characteristic equation,

$$\varphi^2 - \varphi \cdot \operatorname{tr}\varphi + \iota \cdot \det\varphi = 0 .$$

2. Given three endomorphisms φ, ψ, χ of E show that

$$\operatorname{tr}(\chi \circ \psi \circ \varphi) \neq \operatorname{tr}(\chi \circ \varphi \circ \psi)$$

in general.

3. Assume that φ is an endomorphism of E such that $\varphi^2 = 0$. Prove that $\operatorname{tr}\varphi = 0$.

4. Show that the trace of a projection $\pi: E \to E_1$ (see Chapter II § 4 prob. 1) is equal to the dimension of E_1.

5. Consider two pairs of dual spaces E^*, E and F^*, F. Prove that the spaces $L(E;F)$ and $L(E^*;F^*)$ are dual with respect to the scalar-product defined by

$$\langle\varphi,\psi\rangle = \operatorname{tr}(\varphi^*\circ\psi) \quad \varphi\in L(E;F) \quad \psi\in L(E^*;F^*)\,.$$

6. Let f be a linear function in the space $L(E;E)$. Show that f can be written as

$$f(\varphi) = \operatorname{tr}(\varphi\circ\alpha)$$

where α is a fixed endomorphism of E. Prove that α is uniquely determined by f.

7. Assume that f is a linear function in the space $L(E;E)$ such that

$$f(\psi\circ\varphi) = f(\varphi\circ\psi)\,.$$

Prove that

$$f(\varphi) = \lambda\cdot\operatorname{tr}\varphi$$

where λ is a scalar.

8. Let φ and ψ be two endomorphisms of E. Consider the sum

$$\sum_{i\neq j}\Delta(x_1\ldots\varphi x_i\ldots\psi x_j\ldots x_n)$$

where $\Delta\neq 0$ is a determinant-function in E. This sum is again a determinant-function and hence it can be written as

$$\sum_{i\neq j}\Delta(x_1\ldots\varphi x_i\ldots\psi x_j\ldots x_n) = B(\varphi,\psi)\,\Delta(x_1\ldots x_n)\,.$$

By the above relation a bilinear function B is defined in the space $L(E;E)$. Prove:

a) $B(\varphi,\psi) = \operatorname{tr}\varphi\operatorname{tr}\psi - \operatorname{tr}(\psi\circ\varphi)$.

b) $\frac{1}{2}B(\varphi,\varphi) = (-1)^n\alpha_2$ where α_2 is the coefficient of λ^{n-2} in the characteristic polynomial of φ.

c) $\alpha_2 = \frac{(-1)^n}{2}[(\operatorname{tr}\varphi)^2 - \operatorname{tr}(\varphi^2)]$.

9. Consider two $n\times n$-matrices A and B. Prove the relation

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)\,.$$

a) by direct computation.

b) using the relation $\operatorname{tr}\varphi = \operatorname{tr}M(\varphi)$.

10. If φ and ψ are two endomorphisms of a 2-dimensional linear space prove the relation

$$\psi\circ\varphi + \varphi\circ\psi = \varphi\operatorname{tr}\psi + \psi\operatorname{tr}\varphi + \iota(\operatorname{tr}(\psi\circ\varphi) - \operatorname{tr}\varphi\operatorname{tr}\psi)\,.$$

11. Assume that A is an endomorphism of the space $L(E; E)$ subject to the following conditions:

$$A(\sigma \circ \tau) = A(\sigma)\, A(\tau)$$

and

$$A(\iota) = \iota .$$

Prove that $\operatorname{tr} A(\sigma) = \operatorname{tr} \sigma$.

Chapter V

Oriented linear spaces

In the present chapter, except for the last section, we shall concern ourselves with *real* linear spaces.

§ 1. Orientation by a determinant-function

5.1. Definition. Consider a real linear space E of dimension $n \geqq 1$. If $\Delta_1 \neq 0$ and $\Delta_2 \neq 0$ are two determinant-functions in E, then $\Delta_2 = \lambda \Delta_1$ where $\lambda \neq 0$ is a real number. So we can introduce an equivalence-relation in the set of all determinant-functions $\Delta \neq 0$ as follows:

$$\Delta_1 \sim \Delta_2 \quad \text{if} \quad \lambda > 0 .$$

This relation is obviously an equivalence. It induces a decomposition of all determinant-functions $\Delta \neq 0$ into two equivalence-classes. Each of these classes is called an *orientation* of E. Thus, a real linear space can be oriented in two different ways. If (Δ) is the class of determinant-functions which defines the orientation, then every determinant-function $\Delta \in (\Delta)$ is called a *representing determinant-function.*

A basis e_ν $(\nu = 1 \ldots n)$ of an oriented linear space is called *positive,* if

$$\Delta(e_1 \ldots e_n) > 0$$

for a representing determinant-function Δ. If $(e_1 \ldots e_n)$ is a positive basis and σ is a permutation of the numbers $(1 \ldots n)$, then the basis $(e_{\sigma(1)} \ldots e_{\sigma(n)})$ is again positive if and only if the permutation σ is even.

5.2. Isomorphisms of oriented linear spaces. Let E and F be two oriented linear spaces of the same dimension and φ be an isomorphism of E onto F. Choose two representing determinant-functions Δ and Δ' in E and in F. Then a second determinant-function Δ_φ is defined in E by

$$\Delta_\varphi(x_1 \ldots x_n) = \Delta'(\varphi x_1 \ldots \varphi x_n) \quad x_\nu \in E \ (\nu = 1 \ldots n) .$$

If Δ' is multiplied by a factor $\lambda \neq 0$ then Δ_φ is multiplied by the same factor. Hence, the equivalence-class (Δ_φ) depends only on the class of Δ'.

The isomorphism φ is said to *preserve the orientation* if

$$\Delta_\varphi \sim \Delta .$$

Otherwise we say that φ *reverses the orientation.*

In particular, assume that $F = E$ and that φ is an automorphism of E. Then $\Delta' = \Delta$ and hence

$$\Delta_\varphi (x_1 \dots x_n) = \Delta (\varphi x_1 \dots \varphi x_n) = \det \varphi\, \Delta (x_1 \dots x_n) .$$

It follows that $\Delta_\varphi \sim \Delta$ if and only if $\det \varphi > 0$. Thus, an automorphism of an oriented linear space preserves the orientation if and only if its determinant is positive.

As an example, consider the automorphism $\varphi = -\iota$. Since

$$\det (-\iota) = (-1)^n,$$

it follows that the automorphism φ preserves the orientation if the dimension of E is even and it reverses the orientation if the dimension is odd.

5.3. Induced orientation. Assume a direct decomposition of E

$$E = E_1 \oplus E_2 \quad \dim E_i \geqq 1 \quad (i = 1, 2) .$$

If an orientation is defined in E and in E_1, then an orientation in E_2 is induced in the following way: Let Δ be a representing determinant-function in E and $(e_1 \dots e_p)$ be a positive basis of E_1. Then a determinant-function in E_2 is defined by

$$\Delta_2 (x_{p+1} \dots x_n) = \Delta (e_1 \dots e_p, x_{p+1} \dots x_n) . \tag{5.1}$$

It will be shown that the equivalence class of Δ_2 depends only on the orientations of E and of E_1. First of all, if Δ is replaced by $\lambda\Delta$, then Δ_2 is replaced by $\lambda\Delta_2$. Next, consider another positive basis $(\bar{e}_1 \dots \bar{e}_p)$ of E_1. The corresponding determinant-function $\bar{\Delta}_2$ in E_2 then is defined by

$$\bar{\Delta}_2 (x_{p+1} \dots x_n) = \Delta (\bar{e}_1 \dots \bar{e}_p, x_{p+1} \dots x_n) . \tag{5.2}$$

Let φ_1 be the automorphism of E_1 defined by setting

$$\varphi_1 : e_\nu \to \bar{e}_\nu \quad (\nu = 1 \dots p) .$$

Since both bases e_ν and $\bar{e}_\nu$ $(\nu = 1 \dots p)$ are positive, it follows that

$$\det \varphi_1 > 0 .$$

Extend φ_1 to an automorphism φ of E by

$$\varphi x = \begin{cases} \varphi_1 x & \text{if } x \in E_1 \\ x & \text{if } x \in E_2 . \end{cases}$$

Then

$$\det \varphi = \det \varphi_1 > 0 .$$

The equations (5.1) and (5.2) yield

$$\begin{aligned}\bar{\Delta}_2(x_{p+1}\dots x_n) &= \Delta(\bar{e}_1\dots\bar{e}_p, x_{p+1}\dots x_n)\\ &= \Delta(\varphi e_1\dots\varphi e_p, \varphi x_{p+1}\dots\varphi x_n)\\ &= \det\varphi\,\Delta(e_1\dots e_p, x_{p+1}\dots x_n)\\ &= \det\varphi\,\Delta_2(x_{p+1}\dots x_n)\,, \qquad x_\nu\in E_2 \quad (\nu = p+1\dots n)\end{aligned}$$

whence $\bar{\Delta}_2 \sim \Delta_2$.

The space E_2 in turn induces an orientation in E_1. It will be shown that this orientation coincides with the original orientation of E_1 if and only if $p\,(n-p)$ is even. The induced orientation of E_1 is represented by the determinant-function

$$\Delta_1(x_1\dots x_p) = \Delta(e_{p+1}\dots e_n, x_1\dots x_p) \tag{5.3}$$

where e_λ $(\lambda = p+1\dots n)$ is a positive basis of E_2. Substituting $x_\nu = e_\nu$ $(\nu = 1\dots p)$ in the equation (5.3) we find that

$$\Delta_1(e_1\dots e_p) = \Delta(e_{p+1}\dots e_n, e_1\dots e_p) = (-1)^{p(n-p)}\Delta_2(e_{p+1}\dots e_n)\,. \tag{5.4}$$

But e_λ $(\lambda = 1\dots p)$ is a positive basis of E_2, whence

$$\Delta_2(e_{p+1}\dots e_n) > 0\,. \tag{5.5}$$

It follows from (5.4) and (5.5) that

$$\Delta_1(e_1\dots e_p)\begin{cases} >0 \text{ if } p\,(n-p) \text{ is even}\\ <0 \text{ if } p\,(n-p) \text{ is odd.}\end{cases} \tag{5.6}$$

Since the basis $(e_1\dots e_p)$ of E_1 is positive with respect to the original orientation, the relation (5.6) shows that the induced orientation coincides with the original orientation if and only if $p\,(n-p)$ is even.

5.4. Example. Consider a 2-dimensional linear space E. Given a basis (e_1, e_2) we choose the orientation of E in which the basis (e_1, e_2) is positive. Then the determinant-function Δ, defined by

$$\Delta(e_1, e_2) = 1$$

represents this orientation. Now consider the subspaces E_j $(j = 1, 2)$ generated by e_j $(j = 1, 2)$ with the orientation defined by e_j. Then E_1 induces in E_2 the given orientation, but E_2 induces in E_1 the inverse orientation. In fact, defining the determinant-functions Δ_1 and Δ_2 in E_1 and in E_2 by

$$\Delta_1(x) = \Delta(e_2, x) \quad x\in E_1 \quad\text{and}\quad \Delta_2(x) = \Delta(e_1, x) \quad x\in E_2$$

we find that

$$\Delta_2(e_2) = \Delta(e_1, e_2) = 1 \quad\text{and}\quad \Delta_1(e_1) = \Delta(e_2, e_1) = -1\,.$$

Problems: 1. Let E be an oriented n-dimensional linear space and $x_\nu (\nu = 1, \ldots, n)$ be a positive basis; denote by E_ν the subspace generated by the vectors $x_1 \ldots \hat{x}_\nu \ldots x_n$*). Prove that the basis $x_1 \ldots \ldots \hat{x}_\nu \ldots x_n$ is positive with respect to the orientation induced by the vector $(-1)^{i-1} x_i$.

2. Let E be a linear space and E_1 be a subspace. Introduce an orientation in E by the determinant-function Δ and an orientation in E_1 by the basis $(e_1 \ldots e_p)$. Define the determinant-function $\bar{\Delta}$ in the factor-space in the following way:

$$\bar{\Delta}(\bar{x}_{p+1} \ldots \bar{x}_n) = \Delta(e_1 \ldots e_p, x_{p+1} \ldots x_n) \text{ where } x_\nu \in \bar{x}_\nu .$$

Show that the orientation of E/E_1 represented by $\bar{\Delta}$ depends only on the orientations of E and E_1.

§ 2. The Topology of a linear space

5.5. Neighborhoods. Let E be a linear space and e_ν $(\nu = 1 \ldots n)$ be a basis of E. A *neighborhood* U_a of a vector $a \in E$ is the set of all vectors x defined by

$$U_a \colon x - a = \sum_\nu \xi^\nu e_\nu, \quad |\xi^\nu| < \delta \quad (\nu = 1 \ldots n) \tag{5.7}$$

where δ is a positive number. If $n = 1$, U_a is an open interval of length 2δ with a as midpoint; in the case $n = 2$ we obtain an open parallelogramm with the center a.

Now consider the system of subsets U_a for all vectors $a \in E$ and all positive numbers δ. This system has the following properties which are immediate consequences of the definition:

(H_1): Every vector a is contained in all its neighborhoods.

(H_2): To every vector $b \in U_a$ there exists a neighborhood U_b which is contained in U_a.

(H_3): Given two neighborhoods U_a and V_a there exists a neighborhood W_a which is contained in the intersection $U_a \cap V_a$.

(H_4): Every two different vectors a and b have disjoint neighborhoods.

5.6. The topology of a linear space. An arbitrary set E in which to every point a certain subsets U_a are assigned subject to the conditions (H_1), (H_2) and (H_3) is called a *topological space*. If, in addition (H_4) is satisfied, E is called a *Hausdorff-space*.

Two systems of neighborhoods (U_a) and (V_a) in a set E are called *equivalent*, if every neighborhood U_a contains a neighborhood V_a and conversely, every neighborhood V_a contains a neighborhood U_a. Two topologies of E are considered as equal if the defining systems of neighborhoods are equivalent.

*) $\hat{x}_\nu$ indicates that the vector x_ν is deleted.

The system of neighborhoods in a linear space E defined in sec. 5.5 depends on a basis of E.

Now it will be shown that every two bases of E define equivalent systems of neighborhoods. Let $\bar{e}_\nu$ $(\nu = 1 \ldots n)$ be another basis of E and α_ν^μ $(\nu, \mu = 1 \ldots n)$ the matrix of the basis-transformation $e_\nu \to \bar{e}_\nu$. Given a neighborhood

$$U_a \colon x - a = \sum_\nu \xi^\nu e_\nu \quad |\xi^\nu| < \delta$$

of a with respect to the first system consider the neighborhood

$$V_a \colon x - a = \sum_\nu \bar{\xi}^\nu \bar{e}_\nu \quad |\bar{\xi}^\nu| < \frac{\delta}{n\alpha}$$

where

$$\alpha = \max |\alpha_\nu^\mu| .$$

Let x be a vector of V_a. Then the transformation-formulas

$$\xi^\nu = \sum_\mu \alpha_\mu^\nu \bar{\xi}^\mu$$

imply that

$$|\xi^\nu| \leqq \alpha \sum_\mu |\bar{\xi}^\mu| < \delta$$

showing that x is contained in U_a. In the same way it follows that every neighborhood V_a contains a neighborhood U_a. The above result shows that a linear space of finite dimension carries a natural topology not depending on a basis.

5.7. Continuity. Consider an arbitrary real-valued function f in E. The function f is called *continuous* at $x = a$ if to every number $\varepsilon > 0$ there is a neighborhood U_a, such that

$$|f(x) - f(a)| < \varepsilon \quad \text{for all} \quad x \in U_a .$$

A linear function f is everywhere continuous. In fact, given a number $\varepsilon > 0$, select a basis e_ν $(\nu = 1 \ldots n)$ of E and choose a number $\delta > 0$ such that

$$\delta \sum_\nu |f(e_\nu)| < \varepsilon .$$

Define the neighborhood U_a by

$$x - a = \sum_\nu \xi^\nu e_\nu, \quad |\xi^\nu| < \delta \quad (\nu = 1 \ldots n) .$$

Then, for every $x \in U_a$,

$$f(x) - f(a) = f(x - a) = \sum_\nu \xi^\nu f(e_\nu)$$

whence

$$|f(x) - f(a)| < \delta \sum_\nu |f(e_\nu)| < \varepsilon .$$

Similarly, the continuity of a function of serveral arguments is defined. A function Φ of p arguments is called *continuous* at $x_\nu = a_\nu (\nu = 1 \ldots p)$ if to every number $\varepsilon > 0$ there is a neighborhood U_{a_ν} of a_ν such that

$$|\Phi(x_1 \ldots x_p) - \Phi(a_1 \ldots a_p)| < \varepsilon \quad \text{if} \quad x_\nu \in U_{a_\nu} \quad (\nu = 1 \ldots p) .$$

A function of several arguments which is linear with respect to every argument is everywhere continuous.

A mapping $t \to x(t)$ of a closed interval $t_0 \leqq t \leqq t_1$ into the linear space E is called continuous at $t = \tau$ if to every neighborhood $U_{x(\tau)}$ there is a positive number δ such that

$$x(t) \subset U_{x(\tau)} \quad \text{if} \quad |t - \tau| < \delta \quad \text{and} \quad t_0 \leqq t \leqq t_1 .$$

5.8. Deformation of a basis. Consider two bases a_ν and b_ν $(\nu = 1 \ldots n)$ of the linear space E. A *deformation* of the basis a_ν into the basis b_ν is a system of n continuous mappings $x_\nu(t)$ of a closed interval $t_0 \leqq t \leqq t_1$ into E, subject to the following conditions:

1. $x_\nu(t_0) = a_\nu$ and $x_\nu(t_1) = b_\nu$ $(\nu = 1 \ldots n)$.

2. The n vectors $x_\nu(t)$ $(\nu = 1 \ldots n)$ are linearly independent for every t $(t_0 \leqq t \leqq t_1)$.

The deformability of two bases is obviously an equivalence-relation. Thus, the set of all bases of E is decomposed into classes of deformable bases. We shall prove that there are exactly two such classes. This is a consequence of the following theorem:

The basis a_ν $(\nu = 1 \ldots n)$ is deformable into the basis b_ν $(\nu = 1 \ldots n)$ if and only if the automorphism $\varphi: a_\nu \to b_\nu$ $(\nu = 1 \ldots n)$ has a positive determinant.

Proof: Assume that the mappings $x_\nu(t)$ $(t_0 \leqq t \leqq t_1)$ define a deformation of the basis a_ν into the basis b_ν $(\nu = 1 \ldots n)$. Choose a determinant-function $\Delta \neq 0$ in E and define the real-valued function Φ by

$$\Phi(t) = \Delta(x_1(t) \ldots x_n(t)) \quad (t_0 \leqq t \leqq t_1) .$$

The continuity of the function Δ and the mappings $t \to x_\nu(t)$ implies that the function Φ is continuous. Furthermore,

$$\Phi(t) \neq 0 \quad (t_0 \leqq t \leqq t_1)$$

because the vectors $x_\nu(t)$ $(\nu = 1 \ldots n)$ are linearly independent. Thus the function Φ assumes the same sign at $t = t_0$ and at $t = t_1$. But

$$\Phi(t_1) = \Delta(b_1 \ldots b_n) = \Delta(\varphi a_1 \ldots \varphi a_n) = \det\varphi\, \Delta(a_1 \ldots a_n) = \det\varphi \cdot \Phi(t_0)$$

whence

$$\det\varphi > 0$$

and so the first part of the theorem is proved.

5.9. Conversely, assume that the automorphism $\varphi\colon a_\nu \to b_\nu$ has a positive determinant. To construct a deformation $(a_1 \ldots a_n) \to (b_1 \ldots b_n)$ assume first that the vector n-tuple

$$(a_1 \ldots a_i, b_{i+1} \ldots b_n) \tag{5.8}$$

is linearly independent for every $i (1 \leqq i \leqq n-1)$. Then consider the decomposition

$$b_n = \sum_\nu \beta^\nu a_\nu .$$

By the above assumption the vectors $(a_1 \ldots a_{n-1}, b_n)$ are linearly independent, whence $\beta^n \neq 0$. Define the number ε_n by

$$\varepsilon_n = \begin{cases} +1 & \text{if} \quad \beta_n > 0 \\ -1 & \text{if} \quad \beta_n < 0 . \end{cases}$$

It will be shown that the n mappings

$$\begin{cases} x_\nu(t) = a_\nu \ (\nu = 1 \ldots n-1) \\ x_n(t) = (1-t)\, a_n + t\varepsilon_n\, b_n \end{cases} \qquad (0 \leqq t \leqq 1)$$

define a deformation

$$(a_1 \ldots a_n) \to (a_1 \ldots a_{n-1}, \varepsilon_n\, b_n) .$$

Let $\Delta \neq 0$ be a determinant-function in E. Then

$$\Delta(x_1(t) \ldots x_n(t)) = ((1-t) + \varepsilon_n\, \beta_n t)\, \Delta(a_1 \ldots a_n) .$$

Since $\varepsilon_n\, \beta_n > 0$, it follows that

$$1 - t + \varepsilon_n\, \beta_n\, t > 0 \quad (0 \leqq t \leqq 1)$$

whence

$$\Delta\, (x_1(t) \ldots x_n(t)) \neq 0 \quad (0 \leqq t \leqq 1) .$$

This implies the linear independence of the vectors $x_\nu(t)$ $(\nu = 1 \ldots n)$ for every t.

In the same way a deformation

$$(a_1 \ldots a_{n-1}, \varepsilon_n\, b_n) \to (a_1 \ldots a_{n-2}, \varepsilon_{n-1}\, b_{n-1}, \varepsilon_n\, b_n)$$

can be constructed where $\varepsilon_{n-1} = \pm 1$. Continuing this way we finally obtain a deformation

$$(a_1 \ldots a_n) \to (\varepsilon_1 b_1 \ldots \varepsilon_n b_n) \quad \varepsilon_\nu = \pm 1 \quad (\nu = 1 \ldots n) .$$

To construct a deformation

$$(\varepsilon_1\, b_1 \ldots \varepsilon_n\, b_n) \to (b_1 \ldots b_n)$$

consider the automorphisms

$$\varphi\colon a_\nu \to \varepsilon_\nu\, b_\nu \quad (\nu = 1 \ldots n)$$

and

$$\psi\colon \varepsilon_\nu\, b_\nu \to b_\nu \quad (\nu = 1 \ldots n) .$$

The product of these automorphisms is the automorphism

$$\psi \circ \varphi : a_\nu \to b_\nu \quad (\nu = 1 \ldots n) .$$

By hypothesis,

$$\det (\psi \circ \varphi) > 0 \tag{5.9}$$

and by the result of sec. 5.8

$$\det \varphi > 0 . \tag{5.10}$$

The relations (5.9), and (5.10) imply that

$$\det \psi > 0 .$$

But

$$\det \psi = \varepsilon_1 \ldots \varepsilon_n$$

whence

$$\varepsilon_1 \ldots \varepsilon_n = +1 .$$

Thus, the number of ε_ν equal to -1 is even. Rearranging the vectors b_ν $(\nu = 1 \ldots n)$ we can achieve that

$$\varepsilon_\nu = \begin{cases} -1 & (\nu = 1 \ldots 2p) \\ +1 & (\nu = 2p + 1 \ldots n). \end{cases}$$

Then a deformation

$$(\varepsilon_1 b_1 \ldots \varepsilon_n b_n) \to (b_1 \ldots b_n)$$

is defined by the mappings

$$\left.\begin{aligned} x_{2\nu-1}(t) &= -\, b_{2\nu-1} \cos t + b_{2\nu} \sin t \\ x_{2\nu}(t) &= -\, b_{2\nu-1} \sin t - b_{2\nu} \cos t \end{aligned}\right\} (\nu = 1 \ldots p) \qquad 0 \leqq t \leqq \pi.$$
$$x_\nu(t) = b_\nu \qquad (\nu = 2p + 1 \ldots n)$$

5.10. The case remains to be considered that not all the vector n-tuples (5.8) are linearly independent. Let $\Delta \neq 0$ be a determinant-function. The linear independence of the vectors a_ν $(\nu = 1 \ldots n)$ implies that

$$\Delta (a_1 \ldots a_n) \neq 0 .$$

Since Δ is a continuous function, there exists a neighborhood U_{a_ν} of a_ν $(\nu = 1 \ldots n)$ such that

$$\Delta (x_1 \ldots x_n) \neq 0 \quad \text{if} \quad x_\nu \in U_{a_\nu} \quad (\nu = 1 \ldots n) .$$

Choose a vector $a_1' \in U_{a_1}$ which is not contained in the $(n-1)$-dimensional subspace generated by the vectors $(b_2 \ldots b_n)$. Then the vectors $(a_1', b_2 \ldots b_n)$ are linearly independent. Next, choose a vector $a_2' \in U_{a_2}$ which is not contained in the $(n-1)$-dimensional subspace generated by the vectors $(a_1', b_2 \ldots b_n)$. Then the vectors $(a_1', a_2', b_3 \ldots b_n)$ are linearly independent. Going on this way we finally obtain a system of n

vectors a'_ν $(\nu = 1 \dots n)$ such that every n-tuple

$$(a'_1 \dots a'_i, b_{i+1} \dots b_n) \quad (i = 1 \dots n-1)$$

is linearly independent. Since $a'_\nu \in U_{a_\nu}$, it follows that

$$\Delta(a'_1 \dots a'_n) \neq 0 .$$

Hence the vectors a'_ν $(\nu = 1 \dots n)$ form a basis of E. The n mappings

$$x_\nu(t) = (1-t)\, a_\nu + t a'_\nu \quad (0 \leqq t \leqq 1)$$

define a deformation

$$(a_1 \dots a_n) \to (a'_1 \dots a'_n) . \tag{5.11}$$

In fact, $x_\nu(t)$ $(0 \leqq t \leqq 1)$ is contained in U_{a_ν} whence

$$\Delta(x_1(t) \dots x_n(t)) \neq 0 \quad (0 \leqq t \leqq 1) .$$

This implies the linear independence of the vectors $x_\nu(t)$ $(\nu = 1 \dots n)$.

By the result of sec. 5.9. there exists a deformation

$$(a'_1 \dots a'_n) \to (b_1 \dots b_n). \tag{5.12}$$

The two deformations (5.11) and (5.12) yield a deformation

$$(a_1 \dots a_n) \to (b_1 \dots b_n) .$$

This completes the proof of the theorem in sec. 5.8.

5.11. Basis-deformation in an oriented linear space. If an orientation is given in the linear space E, the theorem of 5.8. can be formulated as follows: Two bases a_ν and b_ν $(\nu = 1 \dots n)$ can be deformed into each other if and only if they are both positive or both negative with respect to the given orientation. In fact, the automorphism

$$\varphi: a_\nu \to b_\nu \quad (\nu = 1 \dots n)$$

has a positive determinant if and only if the bases a_ν and b_ν $(\nu = 1 \dots n)$ are both positive or both negative.

Thus the two classes of deformable bases consist of all positive bases and all negative bases.

5.12. Complex linear spaces. The existence of the two orientations in a real linear space is based upon the fact that every real number $\lambda \neq 0$ is either positive or negative. Therefore it is not possible to distinguish two orientations of a complex linear space. In this context the question arises whether any two bases of a complex linear space can be deformed into each other. It will be shown that this is indeed always possible.

Consider two bases a_ν and b_ν $(\nu = 1 \dots n)$ of the complex linear space E. As in sec. 5.9. we can assume that the vector n-tuples

$$(a_1 \dots a_i, b_{i+1} \dots b_n)$$

are linearly independent for every i $(1 \leqq i \leqq n-1)$. It follows from the above assumption that the coefficient β^n in the decomposition

$$b_n = \sum_\nu \beta^\nu a_\nu$$

is different from zero. The complex number β^n can be written as

$$\beta^n = r e^{i\vartheta} \quad (r > 0, 0 \leqq \vartheta < 2\pi) .$$

Now choose a continuous function $r(t)$ $(0 \leqq t \leqq 1)$ such that

$$r(0) = 1, \quad r(1) = r \tag{5.13}$$

and a continuous function $\vartheta(t)$ $(0 \leqq t \leqq 1)$ such that

$$\vartheta(0) = 0, \quad \vartheta(1) = \vartheta . \tag{5.14}$$

Define the mappings $x_\nu(t)$ $(0 \leqq t \leqq 1)$ by

$$\left.\begin{aligned} x_\nu(t) &= a_\nu \quad (\nu = 1 \ldots n-1) \\ \text{and} \qquad x_n(t) &= t \sum_{\nu=1}^{n-1} \beta^\nu a_\nu + r(t)\, e^{i\vartheta(t)} a_n . \end{aligned}\right\} \; 0 \leqq t \leqq 1 \tag{5.15}$$

Then the vectors $x_\nu(t)$ $(\nu = 1 \ldots n)$ are linearly independent for every t. In fact, assume a relation

$$\sum_{\nu=1}^{n} \lambda^\nu x_\nu(t) = 0 .$$

Then

$$\sum_{\nu=1}^{n-1} \lambda^\nu a_\nu + \lambda^n t \sum_{\nu=1}^{n-1} \beta^\nu a_\nu + \lambda^n r(t)\, e^{i\vartheta(t)} a_n = 0$$

whence

$$\lambda^\nu + \lambda^n t \beta^\nu = 0 \qquad (\nu = 1 \ldots n-1)$$

and

$$\lambda^n r(t)\, e^{i\vartheta(t)} = 0 .$$

Since $r(t) \neq 0$ for $0 \leqq t \leqq 1$, the last equation implies that $\lambda^n = 0$. Hence the first $(n-1)$ equations reduce to $\lambda^\nu = 0$ $(\nu = 1 \ldots n-1)$.

It follows from (5.13) and (5.14) that

$$x_n(0) = a_n \quad \text{and} \quad x_n(1) = b_n .$$

Thus the mappings (5.15) define a deformation

$$(a_1 \ldots a_{n-1}, a_n) \to (a_1 \ldots a_{n-1}, b_n) .$$

Continuing this way we obtain after n steps a deformation of the basis a_ν into the basis $b_\nu (\nu = 1 \ldots n)$.

Problems: 1. Let f be a real valued continuous function in the real n-dimensional linear space E such that

$$f(x+y) = f(x) + f(y) \quad x, y \in E .$$

Prove that f is linear.

2. Let E be a real n-dimensional linear space and E_1 be an $(n-1)$-dimensional subspace. Denote by E' the set of all vectors $x \in E$ which are not contained in E_1. Define an equivalence relation in E' as follows: Two vectors $a \in E'$ and $b \in E'$ are equivalent, if the straight segment

$$x(t) = (1-t)\,a + t\,b \quad (0 \leqq t \leqq 1)$$

is disjoint to E_1. Prove that there are exactly two equivalence classes.

Chapter VI

Multilinear mappings

§ 1. Basic properties

6.1. Definition. Let three linear spaces E, F, G be given. Assume that to every pair of vectors $x \in E$ and $y \in F$ a vector $\varphi(x, y) \in G$ is assigned subject to the conditions

$$\varphi(\lambda x_1 + \mu x_2, y) = \lambda\,\varphi(x_1, y) + \mu\,\varphi(x_2, y)$$

and

$$\varphi(x, \lambda y_1 + \mu y_2) = \lambda\,\varphi(x, y_1) + \mu\,\varphi(x, y_2) .$$

Then φ is called a *bilinear mapping* of E and F into G. A bilinear mapping of E and F into the number-space R^1 is called a *bilinear function.*

The set of all vectors $\varphi(x, y)$ $(x \in E, y \in F)$ is generally not a linear subspace of G. As an example, let $E = F$ be a 2-dimensional linear space and G be a 4-dimensional linear space. Select a basis a_1, a_2 in E and a basis c_ν $(\nu = 1 \ldots 4)$ in G and define the bilinear mapping φ by

$$\varphi(x, y) = \xi^1\eta^1 c_1 + \xi^1\eta^2 c_2 + \xi^2\eta^1 c_3 + \xi^2\eta^2 c_4$$

where $x = \xi^1 a_1 + \xi^2 a_2$ and $y = \eta^1 a_1 + \eta^2 a_2$. Then it is easy to see that a vector

$$z = \sum_\nu \zeta^\nu c_\nu$$

is of the form $z = \varphi(x, y)$ if and only if the components satisfy the relation

$$\zeta^1\zeta^4 - \zeta^2\zeta^3 = 0 .$$

In the following sections we shall denote by $\varphi(E, F)$ the subspace of G *generated* by the set $\varphi(x, y)$ $(x \in E, y \in F)$.

Consider the set $B(E, F; G)$ of all bilinear functions $\varphi: (E, F) \to G$. By defining the sum of two bilinear mappings φ_1 and φ_2 by

$$(\varphi_1 + \varphi_2)(x, y) = \varphi_1(x, y) + \varphi_2(x, y)$$

and the mapping $\lambda\varphi$ by

$$(\lambda\varphi)(x, y) = \lambda\varphi(x, y)$$

we can introduce a linear structure in the set $B(E, F; G)$. The space of all bilinear functions in E and F will be simply denoted by $B(E, F)$.

6.2. Bilinear mappings of linear spaces of finite dimension. Let E and F be two linear spaces of dimension n and m and φ be a bilinear mapping of E and F into G. Employing a basis a_ν $(\nu = 1 \dots n)$ and b_μ $(\mu = 1 \dots m)$ of E and F respectively we can write every vector $\varphi(x, y)$ as

$$\varphi(x, y) = \sum_{\nu, \mu} \xi^\nu \eta^\mu \varphi(a_\nu, b_\mu) \tag{6.1}$$

where

$$x = \sum_\nu \xi^\nu a_\nu \quad \text{and} \quad y = \sum_\mu \eta^\mu b_\mu .$$

This equation shows that the space $\varphi(E, F)$ is generated by the vectors $\varphi(a_\nu, b_\mu)$ and hence it has at most the dimension $n\,m$,

$$\dim \varphi(E, F) \leqq \dim E \cdot \dim F .$$

Now assume that G has also a finite dimension and let z_λ $(\lambda = 1 \dots l)$ be a basis of G. Then the vector $\varphi(a_\nu, b_\mu)$ can be written as

$$\varphi(a_\nu, b_\mu) = \sum_\lambda \gamma^\lambda_{\nu\mu} z_\lambda \tag{6.2}$$

where the $\gamma^\lambda_{\nu\mu}$ are certain scalars. The equations (6.1) and (6.2) yield

$$\varphi(x, y) = \sum_{\nu, \mu, \lambda} \gamma^\lambda_{\nu\mu} \xi^\nu \eta^\mu z_\lambda$$

showing that the bilinear mapping φ is uniquely determined by the scalars $\gamma^\lambda_{\nu\mu}$.

The dimension of the space $B(E, F; G)$ is given by

$$\dim B(E, F; G) = \dim E \cdot \dim F \cdot \dim G .$$

In fact, it is easy to show that the bilinear mappings $\varphi^{\nu\mu}_\lambda$ defined by

$$\varphi^{\nu\mu}_\lambda(a_\alpha, b_\beta) = \delta^\nu_\alpha \delta^\mu_\beta c_\lambda \quad (\alpha = 1 \dots n,\ \beta = 1 \dots m,\ \lambda = 1 \dots l)$$

form a basis of $B(E, F; G)$.

6.3. Multilinear mappings. Let $(p + 1)$ linear spaces E_ν $(\nu = 1 \dots p)$ and G be given. A *p-linear mapping* of the spaces E_ν $(\nu = 1 \dots p)$ into G

assigns to every p-tuple of vectors $x_\nu \in E_\nu$ a vector $\varphi(x_1 \ldots x_p) \in G$ subject to the condition

$$\begin{aligned}&\varphi(x_1 \ldots x_{i-1}, \lambda x_i + \mu y_i, x_{i+1} \ldots x_p)\\ &= \lambda\varphi(x_1 \ldots x_i \ldots x_p) + \mu\varphi(x_1 \ldots y_i \ldots x_p) \qquad (i = 1 \ldots p).\end{aligned}$$

A p-linear mapping into the number-space R^1 is called a *p-linear function.*

The sum of two p-linear mappings

$$\varphi\colon (E_1 \ldots E_p) \to G \quad \text{and} \quad \psi\colon (E_1 \ldots E_p) \to G$$

is defined as the p-linear mapping

$$(\varphi + \psi)(x_1 \ldots x_p) = \varphi(x_1 \ldots x_p) + \psi(x_1 \ldots x_p)$$

and the mapping $\lambda\varphi$ is defined by

$$(\lambda\varphi)(x_1 \ldots x_p) = \lambda\varphi(x_1 \ldots x_p).$$

Under these operations the set of all p-linear mappings becomes a linear space.

6.4. Skew-symmetric mappings. If all the linear spaces E_ν $(\nu = 1 \ldots p)$ coincide, $E_\nu = E$ $(\nu = 1 \ldots p)$, a p-linear mapping $\varphi\colon (E \ldots E) \to G$ will be simply called a p-linear mapping of E into G. A p-linear mapping is called *skew-symmetric*, if

$$\varphi(x_{\sigma(1)}, \ldots x_{\sigma(p)}) = \varepsilon_\sigma \varphi(x_1, \ldots x_p) \qquad x_\nu \in E$$

for every permutation σ, where the symbol ε_σ is defined by

$$\varepsilon_\sigma = \begin{cases} +1, & \text{if } \sigma \text{ is even} \\ -1, & \text{if } \sigma \text{ is odd.}\end{cases}$$

It follows from the above condition that

$$\varphi(x_1 \ldots x_i \ldots x_j \ldots x_p) = -\varphi(x_1 \ldots x_j \ldots x_i \ldots x_p) \tag{6.3}$$

for every pair $i \neq j$.

Inserting $x_i = x_j = x$ into (6.3) we find that

$$\varphi(x_1 \ldots x \ldots x \ldots x_p) = 0.$$

More generally, every system of linearly dependent vectors is mapped into the zero-vector. In fact, assume that

$$x_p = \sum_{\nu=1}^{p-1} \lambda^\nu x_\nu.$$

Then

$$\varphi(x_1 \ldots x_p) = \sum_{\nu=1}^{p-1} \lambda^\nu \varphi(x_1 \ldots x_{p-1}, x_\nu) = 0.$$

Problems: 1. If $\dim E = n$ and $\dim G = l$ show that the space of all skew-symmetric p-linear mappings of E into G has the dimension $\binom{n}{p}\, l$.

2. Given a bilinear mapping $\varphi: (E, F) \to G$ define a mapping $\psi: E \times F \to G$ by

$$\psi z = \varphi(\pi_1 z, \pi_2 z) \quad z \in E \times F$$

where π_1 and π_2 are the projections

$$\pi_1(x, y) = x \quad \text{and} \quad \pi_2(x, y) = y .$$

Prove the relations

$$\psi(z_1 + z_2) + \psi(z_1 - z_2) = 2\psi(z_1) + 2\psi(z_2)$$

and

$$\psi(\lambda z) = \lambda^2 \psi(z) .$$

3. Let $\Delta \neq 0$ be a determinant-function in the n-dimensional linear space E and φ be a skew-symmetric n-linear mapping of E into a linear space G. Show that φ can be written as

$$\varphi(x_1, \ldots x_n) = \Delta(x_1 \ldots x_n) \cdot b$$

where b is a fixed vector of G.

4. A *Lie-product* in a linear space E is a skew-symmetric bilinear mapping $(x, y) \to [x, y]$ of E into itself satisfying the *Jacobi-identity:*

$$[[x, y], z] + [[y, z], x] + [[z, x], y] = 0 .$$

a) Assume that an associative multiplication $(x, y) \to xy$ is defined in E. Show that then a Lie-product is given by $[x, y] = xy - yx$.

b) If a Lie-product is defined in E, denote for every vector $a \in E$ by $\mathrm{Ad}(a)$ the endomorphism given by

$$\mathrm{Ad}(a)\, x = [a, x] .$$

Prove that

$$\mathrm{Ad}([a, b]) = \mathrm{Ad}(a) \circ \mathrm{Ad}(b) - \mathrm{Ad}(b) \circ \mathrm{Ad}(a) .$$

5. Given a pair of dual spaces E^*, E define a bilinear mapping $(E, E) \to L(E^*, E)$ as follows:

$$\varphi_{a,b} \overset{*}{x} = \langle \overset{*}{x}, a \rangle\, b - \langle \overset{*}{x}, b \rangle\, a .$$

Similarly, define a bilinear mapping $(E^*, E^*) \to L(E, E^*)$ by

$$\varphi_{\overset{*}{a}, \overset{*}{b}}\, x = \langle \overset{*}{a}, x \rangle\, \overset{*}{b} - \langle \overset{*}{b}, x \rangle\, \overset{*}{a} .$$

Prove that $\langle \varphi_{a,b}, \varphi_{\overset{*}{a}, \overset{*}{b}} \rangle = -2 \begin{vmatrix} \langle \overset{*}{a}, a \rangle & \langle \overset{*}{a}, b \rangle \\ \langle \overset{*}{b}, a \rangle & \langle \overset{*}{b}, b \rangle \end{vmatrix}$

where the scalar-product on the left hand side is defined as (4.63).

§ 2. Tensor-product

6.5. Definition. Let E, F and G be three linear spaces. A bilinear mapping $(x, y) \to x \otimes y$ of E and F into G is called a *tensor-product*, if it has the following properties:

T_1: The linear closure of the set $x \otimes y$ $(x \in E, y \in F)$ is the whole space G.

T_2: Every bilinear mapping φ of E and F into any linear space H can be written in the form

$$\varphi(x, y) = \chi(x \otimes y) \tag{6.4}$$

where χ is a linear mapping of G into H (factorization-property).

The existence-proof of the tensor-product will be given in the next section. First of all it will be shown that the space G is determined by E and F up to an isomorphism. Assume two linear spaces G_1 and G_2 and two linear mappings

$$(x, y) \to x \otimes_1 y \quad \text{and} \quad (x, y) \to x \otimes_2 y$$

satisfying the conditions T_1 and T_2. Then there exist linear mappings

$$\chi_1 \colon G_1 \to G_2 \quad \text{and} \quad \chi_2 \colon G_2 \to G_1$$

such that

$$x \otimes_2 y = \chi_1(x \otimes_1 y) \quad \text{and} \quad x \otimes_1 y = \chi_2(x \otimes_2 y) .$$

These equations yield

$$(\chi_2 \circ \chi_1)(x \otimes_1 y) = x \otimes_1 y \quad \text{and} \quad (\chi_1 \circ \chi_2)(x \otimes_2 y) = x \otimes_2 y . \tag{6.5}$$

Observing that the spaces G_1 and G_2 are generated by the products $x \otimes_1 y$ and $x \otimes_2 y$ respectively we conclude from (6.5) that

$$\chi_2 \circ \chi_1 = \iota \quad \text{and} \quad \chi_1 \circ \chi_2 = \iota .$$

Hence, χ_1 is an isomorphism of G_1 onto G_2 and χ_2 is the inverse isomorphism.

This proves that the space G is determined by E and F up to an isomorphism. G is called the *tensor-product* of the spaces E and F and is denoted by $E \otimes F$.

Let φ be a bilinear mapping of E and F into a linear space H. By virtue of T_2 there exists a linear mapping $\chi \colon E \otimes F \to H$ such that

$$\varphi(x, y) = \chi(x \otimes y) .$$

The mapping χ is uniquely determined by the bilinear mapping φ. In fact, assume that χ_1 and χ_2 are two linear mappings of $E \otimes F$ into H such that

$$\varphi(x, y) = \chi_1(x \otimes y) \quad \text{and} \quad \varphi(x, y) = \chi_2(x \otimes y) .$$

Then

$$\chi_1(x \otimes y) = \chi_2(x \otimes y) .$$

Since $E \otimes F$ is generated by the products $x \otimes y$ this equation implies that $\chi_1 = \chi_2$. It follows from this remark that there is a one-to-one correspondence between all bilinear mappings of E and F into H and all linear mappings of $E \otimes F$ into H. This correspondence can be expressed by the diagramm

$$\begin{tikzcd}
(E,F) \arrow[r, "\otimes"] \arrow[dr, "\varphi"'] & E \otimes F \arrow[d, "\chi"] \\
& H
\end{tikzcd}$$

6.6. Existence. Given the linear spaces E and F, consider the set $E \times F$ of all pairs (x, y) with $x \in E$ and $y \in F$. This set generates a linear space $C\,(E \times F)$ (cf. sec. 1.3, example 3). Let N be the subspace of $C\,(E \times F)$ generated by the elements

$$(\lambda x_1 + \mu x_2, y) - \lambda\,(x_1, y) - \mu\,(x_2, y) \quad x_1, x_2 \in E\,, \quad y \in F$$

and

$$(x, \lambda y_1 + \mu y_2) - \lambda\,(x, y_1) - \mu\,(x, y_2) \quad x \in E\,, \quad y_1, y_2 \in F\,. \tag{6.6}$$

It will be shown that the factor-space $C(E \times F)/N$ is the tensor-product of E and F.

Define the mapping $(E, F) \to C(E \times F)/N$ by

$$x \otimes y = \pi\,(x, y) \tag{6.7}$$

where π denotes the canonical projection

$$\pi : C(E \times F) \to C(E \times F)/N\,.$$

To prove that the mapping (6.7) is bilinear observe that

$$(\lambda x_1 + \mu x_2, y) - \lambda\,(x_1, y) - \mu\,(x_2, y) \in N$$

whence

$$\pi\,(\lambda x_1 + \mu x_2, y) = \lambda\pi\,(x_1, y) + \mu\pi\,(x_2, y)\,.$$

This implies that

$$(\lambda x_1 + \mu x_2) \otimes y = \lambda x_1 \otimes y + \mu x_2 \otimes y\,.$$

In the same way it follows that

$$x \otimes (\lambda y_1 + \mu y_2) = \lambda x \otimes y_1 + \mu x \otimes y_2\,.$$

It remains to be shown that the bilinear mapping (6.7) satisfies the conditions T_1 and T_2. Let

$$z = \pi \sum_{\nu,\mu} \lambda^{\nu\mu}\,(x_\nu, y_\mu)$$

be an arbitrary element of the space $C(E \times F)/N$. Then

$$z = \sum_{\nu, \mu} \lambda^{\nu\mu} \pi (x_\nu, y_\mu) = \sum_{\nu, \mu} \lambda^{\nu\mu} x_\nu \otimes y_\mu$$

and hence z is a linear combination of products $x \otimes y$.

To verify T_2 consider a bilinear mapping φ of E and F into a third linear space H. Define the linear mapping

$$\psi: C(E \times F) \rightarrow H$$

by

$$\psi \sum_{\nu, \mu} \lambda^{\nu\mu} (x_\nu, y_\mu) = \sum_{\nu, \mu} \lambda^{\nu\mu} \varphi (x_\nu, y_\mu) . \tag{6.8}$$

Then N is contained in the kernel of ψ. It is sufficient to show that the generators (6.6) of N are mapped into zero. The equation (6.8) yields

$$\psi \{(\lambda x_1 + \mu x_2, y) - \lambda (x_1, y) - \mu (x_2, y)\}$$
$$= \varphi (\lambda x_1 + \mu x_2, y) - \lambda \varphi (x_1, y) - \mu \varphi (x_2, y) = 0 .$$

In the same way it follows that

$$\psi \{(x, \lambda y_1 + \mu y_2) - \lambda (x, y_1) - \mu (x, y_2)\} = 0 .$$

Since N is contained in the kernel of ψ, the mapping ψ induces a linear mapping

$$\chi: C(E \times F)/N \rightarrow H$$

which is connected with ψ by the equation

$$\psi = \chi \circ \pi . \tag{6.9}$$

Now the relations (6.7), (6.9) and (6.8) yield

$$\chi (x \otimes y) = (\chi \circ \pi) (x, y) = \psi (x, y) = \varphi (x, y)$$

showing that the mapping (6.7) has the property T_2.

6.7. The property T_2'. In this section it will be shown that the factorization-property is a consequence of the following condition:

T_2': Given two arbitrary systems $(a_1 \dots a_p)$ and $(b_1 \dots b_q)$ of linearly independent vectors in E and F respectively, the vectors $a_\nu \otimes b_\mu$ are again linearly independent.

Let $(x, y) \rightarrow x \otimes y$ be a bilinear mapping of E and F into G satisfying the conditions T_1 and T_2'. We first prove the following

Lemma: The relation

$$\sum_\alpha x_\alpha \otimes y_\alpha = 0 \quad x_\alpha \in E, \quad y_\alpha \in F \quad (\alpha = 1 \dots r) \tag{6.10}$$

implies that

$$\sum_\alpha \varphi (x_\alpha, y_\alpha) = 0 . \tag{6.11}$$

Denote by E_1 and F_1 the subspaces of E and F generated by the vectors x_α and y_α $(\alpha = 1 \dots r)$. Employing a basis a_ν $(\nu = 1 \dots p)$ of E_1 and a basis b_μ $(\mu = 1 \dots q)$ of F_1 we can write the vectors x_α and y_α as

$$x_\alpha = \sum_\nu \xi_\alpha^\nu a_\nu \quad \text{and} \quad y_\alpha = \sum_\mu \eta_\alpha^\mu b_\mu .$$

Then

$$\sum_\alpha \varphi(x_\alpha, y_\alpha) = \sum_{\nu,\mu} \zeta^{\nu\mu} \varphi(a_\nu, b_\mu) \tag{6.12}$$

where the coefficients $\zeta^{\nu\mu}$ are given by

$$\zeta^{\nu\mu} = \sum_\alpha \xi_\alpha^\nu \eta_\alpha^\mu .$$

In the same way we find that

$$\sum_\alpha x_\alpha \otimes y_\alpha = \sum_{\nu,\mu} \zeta^{\nu\mu} a_\nu \otimes b_\mu . \tag{6.13}$$

The equations (6.10) and (6.13) yield

$$\sum_{\nu,\mu} \zeta^{\nu\mu} a_\nu \otimes b_\mu = 0$$

whence, by T_2'

$$\zeta^{\nu\mu} = 0 \quad (\nu = 1 \dots p, \mu = 1 \dots q) . \tag{6.14}$$

Inserting (6.14) into (6.12) we obtain the equation (6.11).

Now let $\varphi: (E, F) \to H$ be a given bilinear mapping. It has to be shown that φ can be factored by the bilinear mapping $\otimes: (E, F) \to G$. Let z be an arbitrary vector of G. In view of T_1 this vector can be written (in different ways) as

$$z = \sum_\alpha x_\alpha \otimes y_\alpha \qquad x_\alpha \in E, y_\alpha \in F .$$

Let

$$z = \sum_\beta x'_\beta \otimes y'_\beta \qquad x'_\beta \in E, y'_\beta \in F$$

be a different representation of z. Then

$$\sum_\alpha x_\alpha \otimes y_\alpha - \sum_\beta x'_\beta \otimes y'_\beta = 0$$

and hence, by the above lemma,

$$\sum_\alpha \varphi(x_\alpha, y_\alpha) = \sum_\beta \varphi(x'_\beta, y'_\beta) .$$

Consequently, a mapping $\chi: G \to H$ can be defined by

$$\chi z = \sum_\alpha \varphi(x_\alpha, y_\alpha) . \tag{6.15}$$

To show that χ is linear, consider two vectors

$$z = \sum_\alpha x_\alpha \otimes y_\alpha \quad \text{and} \quad z' = \sum_\beta x'_\beta \otimes y'_\beta .$$

Then

$$\lambda z + \mu z' = \sum_\alpha \lambda x_\alpha \otimes y_\alpha + \sum_\beta \mu x'_\beta \otimes y'_\beta$$

whence

$$\chi(\lambda z + \mu z') = \sum_\alpha \varphi(\lambda x_\alpha, y_\alpha) + \sum_\beta \varphi(\mu x'_\beta, y'_\beta)$$
$$= \lambda \sum_\alpha \varphi(x_\alpha, y_\alpha) + \mu \sum_\beta \varphi(x'_\beta, y'_\beta) = \lambda\, \chi z + \mu\, \chi z' .$$

Inserting $z = x \otimes y$ in (6.15) we find the relation

$$\chi(x \otimes y) = \varphi(x, y)$$

showing that the bilinear mapping φ has the property T_2.

Here the question arises whether, conversely, T'_2 is a consequence of T_2. In the next section it will be shown that this is always true provided that E and F have finite dimension.*)

6.8. The tensor-product of linear spaces of finite dimension. Let E and F be two linear spaces of dimension n and m respectively and suppose that a bilinear mapping $(x, y) \to x \otimes y$ is given which has the properties T_1 and T_2. To prove T'_2, denote by E_1 and F_1 the subspaces of E and F generated by the vectors a_ν $(\nu = 1 \ldots p)$ and b_μ $(\mu = 1 \ldots q)$ respectively. Select an arbitrary system of pq scalars $\gamma_{\nu\mu}$ and define the bilinear function Φ_1 in E_1 and F_1 by

$$\Phi_1(a_\nu, b_\mu) = \gamma_{\nu\mu} .$$

With the help of the exchange theorem (cf. sec. 1.15) the function Φ_1 can be extended to a bilinear function Φ in E and F. In view of T_2 there exists a linear function f in $E \otimes F$ such that

$$\Phi(x, y) = f(x \otimes y) .$$

Now assume a relation

$$\sum_{\nu,\mu} \lambda^{\nu\mu} a_\nu \otimes b_\mu = 0 .$$

Then

$$\sum_{\nu,\mu} \lambda^{\nu\mu} \gamma_{\nu\mu} = \sum_{\nu,\mu} \lambda^{\nu\mu} \Phi_1(a_\nu, b_\mu) = \sum_{\nu,\mu} \lambda^{\nu\mu} \Phi(a_\nu, b_\mu)$$
$$= \sum_{\nu,\mu} \lambda^{\nu\mu} f(a_\nu \otimes b_\mu) = f\Big(\sum_{\nu,\mu} \lambda^{\nu\mu} a_\nu \otimes b_\mu\Big) = 0 .$$

Since the scalars $\gamma_{\nu\mu}$ are arbitrary this equation implies that $\lambda^{\nu\mu} = 0$. It follows from the above result that

$$\dim E \otimes F = \dim E \cdot \dim F . \tag{6.16}$$

*) This assertion is also true for spaces of infinite dimension. But the proof is based upon the fact that every linear space possesses a basis (i. e. a system of linearly independent generators). Since we did not prove this theorem we restrict ourselves to spaces of finite dimension.

In fact, let a_ν $(\nu = 1 \ldots n)$ be a basis of E and b_μ $(\mu = 1 \ldots m)$ be a basis of F. Then the products $a_\nu \otimes b_\mu$ form a system of linearly independent generators of $E \otimes F$ and hence they form a basis.

Conversely, let φ be a bilinear mapping of E and F into a linear space G of dimension $n\,m$ such that the vectors $\varphi(x, y)$ $(x \in E, y \in F)$ generate the entire space G. Then φ is a tensor-product. In view of the result of sec. 6.7 it is sufficient to establish the property T_2'. Consider two systems of linearly independent vectors a_ν $(\nu = 1 \ldots p)$ and b_μ $(\mu = 1 \ldots q)$ in E and F respectively. Extend these systems to a basis a_ν $(\nu = 1 \ldots n)$ of E and a basis $b_\mu (\mu = 1 \ldots m)$ of F. By hypothesis, the nm vectors $\varphi(a_\nu, b_\mu)$ $(\nu = 1 \ldots n, \mu = 1 \ldots m)$ generate the space G. Since G has the dimension $n\,m$ these products must be linearly independent. Consequently, the subsystem $\varphi(a_\nu, b_\mu)$ $(\nu = 1 \ldots p, \mu = 1 \ldots q)$ is also linearly independent.

6.9. The tensor-product of factor-spaces. Consider two arbitrary linear spaces E and F and two subspaces $E_1 \subset E$ and $F_1 \subset F$. Then the products $E \otimes F_1$ and $E_1 \otimes F$ are subspaces of $E \otimes F$. Let π be the canonical projection

$$\pi : E \otimes F \to E \otimes F/(E \otimes F_1 + E_1 \otimes F)$$

and define the bilinear mapping

$$\beta : (E, F) \to E \otimes F/(E \otimes F_1 + E_1 \otimes F)$$

by

$$\beta(x, y) = \pi(x \otimes y) .$$

Then

$$\beta(x_1, y) = 0 \quad \text{for all pairs } x_1 \in E_1, y \in F$$

and

$$\beta(x, y_1) = 0 \quad \text{for all pairs } x \in E, y \in F_1 .$$

Hence, a bilinear mapping

$$\bar{\beta} : (E/E_1, F/F_1) \to E \otimes F/(E \otimes F_1 + E_1 \otimes F)$$

can be defined by

$$\bar{\beta}(\bar{x}, \bar{y}) = \beta(x, y) \quad \bar{x} \in E/E_1, \quad \bar{y} \in F/F_1. \tag{6.17}$$

It will be shown that this bilinear mapping is a tensor-product. The property T_1 is an immediate consequence of the definition (6.17). To prove T_2, let $\bar{\varphi}$ a bilinear mapping of E/E_1 and F/F_1 into some linear space H. Define the bilinear mapping

$$\varphi : (E, F) \to H$$

by

$$\varphi(x, y) = \bar{\varphi}(\bar{x}, \bar{y}) . \tag{6.18}$$

In view of T_2 there exists a linear mapping $\chi: E \otimes F \rightarrow H$ such that

$$\varphi(x, y) = \chi(x \otimes y) . \tag{6.19}$$

Next we show that the subspaces $E \otimes F_1$ and $E_1 \otimes F$ are contained in the kernel of χ. It is sufficient to do this for the generators $x \otimes y_1$ ($x \in E$, $y_1 \in F_1$) and $x_1 \otimes y$ ($x_1 \in E_1, y \in F$). The relations (6.19) and (6.18) yield

$$\chi(x \otimes y_1) = \varphi(x, y_1) = \bar{\varphi}(\bar{x}, \bar{0}) = 0 .$$

In the same way it follows that

$$\chi(x_1 \otimes y) = 0 .$$

Let

$$\bar{\chi}: E \otimes F/(E \otimes F_1 + E_1 \otimes F) \rightarrow H$$

be the linear mapping induced by χ. Then

$$\bar{\chi} \circ \pi = \chi$$

and consequently

$$\bar{\chi}\,\bar{\beta}(\bar{x}, \bar{y}) = \bar{\chi}\,\beta(x, y) = \bar{\chi}\,\pi(x \otimes y) = \chi(x \otimes y) = \varphi(x, y) = \bar{\varphi}(\bar{x}, \bar{y}).$$

This equation shows that every bilinear mapping $\bar{\varphi}$ can be factored over the mapping $\bar{\beta}$. Hence, we can write

$$\beta(\bar{x}, \bar{y}) = \bar{x} \otimes \bar{y} .$$

In other words, the space $E/F_1 \otimes F/F_1$ is isomorphic to the factor-space $E \otimes F/(E \otimes F_1 + E_1 \otimes F)$.

6.10. Bilinear functions. Let four linear spaces E', E, F', F be given and consider a bilinear function Φ in E', E and a bilinear function Ψ in F', F. Then there exists exactly one bilinear function X in $E' \otimes F'$, $E \otimes F$ such that

$$X(x' \otimes y', x \otimes y) = \Phi(x', x)\,\Psi(y', y) . \tag{6.20}$$

To prove this, let $x' \in E'$, $y' \in F'$ be a fixed pair of vectors. Then a bilinear function $\Omega_{(x', y')}$ is defined in E and F by the equation

$$\Omega_{(x', y')}(x, y) = \Phi(x', x)\,\Psi(y', y) . \tag{6.21}$$

The bilinear function $\Omega_{(x', y')}$ can be written as

$$\Omega_{(x', y')}(x, y) = f_{(x', y')}(x \otimes y) \tag{6.22}$$

where $f_{(x', y')}$ is a linear function in $E \otimes F$. The function $f_{(x', y')}$ depends bilinearly on x' and y'. Hence, the correspondence

$$(x', y') \rightarrow f_{(x', y')}$$

defines a bilinear mapping of E' and F' into the space $L(E \otimes F)$ of all linear functions in $E \otimes F$. In view of T_2 there exists a linear mapping

$$\chi: E' \otimes F' \to L(E \otimes F)$$

such that

$$f_{(x', y')} = \chi_{x' \otimes y'} \,. \tag{6.23}$$

We now define the bilinear function X in the spaces $E' \otimes F'$ and $E \otimes F$ by the equation

$$X(z', z) = \chi_{z'}(z) \quad z' \in E' \otimes F', \quad z \in E \otimes F\,, \tag{6.24}$$

Then it follows from (6.24), (6.23), (6.22) and (6.21) that

$$\begin{aligned} X(x' \otimes y', x \otimes y) &= \chi_{x' \otimes y'}(x \otimes y) = f_{(x', y')}(x \otimes y) \\ &= \Omega_{(x', y')}(x, y) = \Phi(x', x)\, \Psi(y', y)\,. \end{aligned}$$

To prove the uniqueness of X, assume that X_1 and X_2 are two bilinear functions in the spaces $E' \otimes F'$, $E \otimes F$ such that the relation (6.20) holds. Then

$$X_1(x' \otimes y', x \otimes y) = X_2(x' \otimes y', x \otimes y)\,.$$

This implies that $X_1 = X_2$ because the spaces $E' \otimes F'$ and $E \otimes F$ are generated by the products $x' \otimes y'$ and $x \otimes y$ respectively.

6.11. The tensor-product of linear mappings. Given four linear spaces E, F, E', F' consider two linear mappings

$$\varphi: E \to E' \quad \text{and} \quad \psi: F \to F'\,.$$

Then a bilinear mapping $(E, F) \to E' \otimes F'$ is defined by

$$(x, y) \to \varphi x \otimes \psi y\,.$$

In view of the factorization-property there exists a linear mapping

$$\chi: E \otimes F \to E' \otimes F'$$

such that

$$\chi(x \otimes y) = \varphi x \otimes \psi y \tag{6.25}$$

and this mapping is uniquely determined by φ and ψ.

The correspondence $(\varphi, \psi) \to \chi$ defines a bilinear mapping

$$(L(E; E'), L(F; F')) \to L(E \otimes F; E' \otimes F')\,. \tag{6.26}$$

Now it will be shown that the bilinear mapping (6.26) is a tensor-product provided that the linear spaces E, F, E', F' have finite dimension. In this case we can write $\varphi \otimes \psi$ instead of χ and the formula (6.25) reads

$$(\varphi \otimes \psi)(x \otimes y) = \varphi x \otimes \psi y\,.$$

To prove this assertion select a basis a_ν $(\nu = 1 \dots n)$, b_μ $(\mu = 1 \dots m)$, a'_ϱ $(\varrho = 1 \dots r)$ and b'_σ $(\sigma = 1 \dots s)$ in E, F, E' and F' respectively define the linear mappings

$$\varphi^\lambda_\varrho : E \to E' \quad \text{and} \quad \psi^\varkappa_\sigma : F \to F'$$

by

$$\varphi^\lambda_\varrho a_\nu = \delta^\lambda_\nu a'_\varrho \quad \text{and} \quad \psi^\varkappa_\sigma b_\mu = \delta^\varkappa_\mu b'_\sigma .$$

Then

$$(\varphi^\lambda_\varrho \otimes \psi^\varkappa_\sigma)(a_\nu \otimes b_\mu) = \delta^\lambda_\nu \delta^\varkappa_\mu a'_\varrho \otimes b'_\sigma .$$

Observing that the products $a_\nu \otimes b_\mu$ and $a'_\varrho \otimes b'_\sigma$ form a basis of $E \otimes F$ and $E' \otimes F'$ respectively we see that the space $L(E \otimes F; E' \otimes F')$ is generated by the products $\varphi^\lambda_\varrho \otimes \psi^\varkappa_\sigma$. Hence, the mapping (6.26) satisfies the condition T_1. The property T_2 follows from the fact that

$$\dim L(E \otimes F; E' \otimes F') = n\, m\, r\, s = \dim L(E; E') \cdot \dim L(F; F').$$

Problems: 1. Given an arbitrary linear space E prove that the bilinear mapping

$$\lambda \otimes x = \lambda x$$

is a tensor-product.

2. Let E^*, E and F^*, F be two pairs of dual spaces. Define a bilinear mapping of E and F into the space $B(E^*, F^*)$ of all bilinear functions in E^* and F^* by

$$(a \otimes b)(\overset{*}{x}, \overset{*}{y}) = \langle \overset{*}{x}, a\rangle \langle \overset{*}{y}, b\rangle \quad a \in E,\ b \in F .$$

Prove that this bilinear mapping is a tensor-product.

3. Consider a pair of dual spaces E^*, E. Show that the bilinear mapping $(E^*, E) \to L(E; E)$ defined by

$$(\overset{*}{a} \otimes a)\, x = \langle \overset{*}{a}, x\rangle\, a$$

is a tensor-product.

Prove that the bilinear function $\langle \overset{*}{a}, a\rangle$ can be written as

$$\langle \overset{*}{a}, a\rangle = \operatorname{tr}(\overset{*}{a} \otimes a) .$$

4. Prove that the condition T'_2 (see sec. 6.7) is equivalent to the following one: Let x_ν $(\nu = 1 \dots p)$ be an arbitrary system of linearly independent vectors in E. Then the relation

$$\sum_\nu x_\nu \otimes y_\nu = 0$$

implies that $y_\nu = 0$ $(\nu = 1 \dots p)$.

5. Consider two pairs of dual spaces E^*, E and F^*, F. Prove that the bilinear function in $E^* \otimes F^*$, $E \otimes F$ defined by

$$\langle \overset{*}{x} \otimes \overset{*}{y}, x \otimes y\rangle = \langle \overset{*}{x}, x\rangle \langle \overset{*}{y}, y\rangle$$

(see sec. 6.10) defines a scalar-product in $E^* \otimes F^*$ and $E \otimes F$.

6. Consider two linear spaces E and F of finite dimension and two endomorphisms $\varphi: E \to E$ and $\psi: F \to F$. Prove the relation

$$\operatorname{tr}(\varphi \otimes \psi) = \operatorname{tr}\varphi \cdot \operatorname{tr}\psi .$$

7. Let E, F, E', F' be four linear spaces of finite dimension and $\varphi: E \to E'$ and $\psi: F \to F'$ be two linear mappings. Show that the rank of $\varphi \otimes \psi$ is the product of the ranks of φ and ψ.

8. Let E and F be two linear spaces of finite dimension.

a) Given two subspaces $E_1 \subset E$ and $E_2 \subset E$ and two subspaces $F_1 \subset F$ and $F_2 \subset F$ prove the relation

$$(E_1 \otimes F_1) \cap (E_2 \otimes F_2) = (E_1 \cap E_2) \otimes (F_1 \cap F_2) .$$

b) Given two direct decompositions

$$E = \sum_i E_i \quad \text{and} \quad F = \sum_j F_j$$

prove that

$$E \otimes F = \sum_{i,j} E_i \otimes F_j .$$

9. Define a bilinear mapping of the number-spaces R^n and R^m into the linear space of all $n \times m$-matrices by

$$(\xi^1 \ldots \xi^n) \otimes (\eta^1 \ldots \eta^m) = \begin{pmatrix} \xi^1 \eta^1 \ldots \xi^1 \eta^m \\ \cdot \qquad\quad \cdot \\ \vdots \qquad\quad \vdots \\ \xi^n \eta^1 \ldots \xi^n \eta^m \end{pmatrix}.$$

Prove that this bilinear mapping is a tensor-product.

10. Let E be an arbitrary linear space. In $E \otimes E$ consider the subspace S generated by all products $x \otimes x$. Define the space $E \wedge E$ by

$$E \wedge E = (E \otimes E)/S$$

and the bilinear mapping $\wedge$ of E into $E \wedge E$ by

$$x \wedge y = \pi(x \otimes y)$$

where π is the canonical projection

$$\pi: E \otimes E \to (E \otimes E)/S .$$

a) Show that the mapping $x \wedge y$ is skew-symmetric.

b) Prove that every skew-symmetric bilinear mapping φ of E into any linear space H can be written as

$$\varphi(x, y) = \chi(x \wedge y)$$

where χ is a linear mapping of $E \wedge E$ into H. Show that χ is uniquely determined by φ.

c) If E has the dimension n prove that

$$\dim E \wedge E = \frac{n(n-1)}{2} .$$

11. Let E and F be two linear spaces of dimension n and m respectively and Δ be a determinant-function in $E \otimes F$. Select a basis a_ν $(\nu = 1 \dots n)$ in E and a basis b_μ $(\mu = 1 \dots m)$ in F and show that the functions

$$\Delta_1(x_1 \dots x_n) = \Delta \begin{pmatrix} x_1 \otimes b_1, a_1 \otimes b_2 \dots a_1 \otimes b_m \\ \cdot \qquad\qquad\qquad \cdot \\ \cdot \qquad\qquad\qquad \cdot \\ \cdot \qquad\qquad\qquad \cdot \\ x_n \otimes b_1, a_n \otimes b_2 \dots a_n \otimes b_m \end{pmatrix}$$

and

$$\Delta_2(y_1 \dots y_m) = \Delta \begin{pmatrix} a_1 \otimes y_1 \dots a_1 \otimes y_m \\ a_2 \otimes b_1 \dots a_2 \otimes b_m \\ \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \\ a_n \otimes b_1 \dots a_n \otimes b_m \end{pmatrix}$$

are determinant-functions in E and F respectively.

12. Using the determinant-functions of prob. 11 show that

$$\det(\varphi \otimes \psi) = (\det \varphi)^m (\det \psi)^n \qquad \begin{array}{l} n = \dim E \\ m = \dim F \,, \end{array}$$

where φ is an endomorphism of E and ψ is an endomorphism of F.

Hint: Observe that the mapping $\varphi \otimes \psi$ can be written as

$$\varphi \otimes \psi = (\varphi \otimes \iota) \circ (\iota \otimes \psi) \,.$$

13. Given an automorphism α of E define the linear mapping

$$A : L(E, E) \to L(E, E)$$

by

$$A\varphi = \alpha^{-1} \circ \varphi \circ \alpha \,.$$

Using the formula in problem 12, prove that $\det A = 1$.

14. Let S and T be two arbitrary sets and $C(S)$ and $C(T)$ be the linear spaces generated by S and T respectively. Prove that the tensor-product $C(S) \otimes C(T)$ is isomorphic to the space $C(S \times T)$.

15. Consider three linear spaces E, F, G and the tensor-products $E \otimes F$ and $F \otimes G$. Then a bilinear mapping

$$\varphi : (E \otimes F, G) \to E \otimes (F \otimes G)$$

is defined by

$$\varphi(x \otimes y, z) = x \otimes (y \otimes z) \,.$$

In view of the factorization-property the bilinear mapping φ induces a linear mapping

$$\chi : (E \otimes F) \otimes G \to E \otimes (F \otimes G) \,.$$

Prove that χ is an isomorphism of $(E \otimes F) \otimes G$ onto $E \otimes (F \otimes G)$.

Chapter VII

Tensor-Algebra

§ 1. Tensors

7.1. Definition. Consider a pair of dual spaces E^* and E. A *tensor of order* (p, q) in E is a $(p+q)$-linear function in the spaces E^* and E with p arguments in E^* and q arguments in E.

A tensor of order $(p, 0)$ is called a *contravariant tensor of order* p, a tensor of order $(0, q)$ is called a *covariant tensor of order* q. If $p \geqq 1$ and $q \geqq 1$, Φ is called a *mixed tensor*. It will be convenient to regard the scalars as tensors of order $(0, 0)$.

For every fixed pair (p, q) the tensors of order (p, q) form a linear space which will be denoted by $T_q^p(E)$. The zero-vector of this space is called the *zero-tensor* of order (p, q). The spaces $T_0^p(E)$ and $T_q^0(E)$ will be denoted simply by $T^p(E)$ and by $T_q(E)$.

7.2. Examples. Let Φ be a contravariant tensor of order 1. Then Φ is a linear function in E^*. As it has been shown in sec. 2.10. this linear function can be written as

$$\Phi(\overset{*}{x}) = \langle \overset{*}{x}, a \rangle ,$$

where a is a vector of E and is uniquely determined by Φ. The one-to-one correspondence $\Phi \leftrightarrow a$ defines an isomorphism

$$T^1(E) \rightleftarrows E .$$

By this isomorphism the contravariant tensors of order one can be identified with the vectors of E.

In the same way we have a one-to-one correspondence between the covariant tensors of order one and the vectors of E^* yielding an isomorphism

$$T_1(E) \rightleftarrows E^* .$$

Next, consider a tensor Φ of order $(1, 1)$, i. e., a bilinear function in E^* and E. Let a be a fixed vector of E. Then a linear function f_a is defined in E^* by

$$f_a(\overset{*}{x}) = \Phi(\overset{*}{x}, a) \qquad \overset{*}{x} \in E^* .$$

This function can be written in the form

$$f_a(\overset{*}{x}) = \langle \overset{*}{x}, b \rangle ,$$

where b is a vector of E. The vector b is uniquely determined by f_a and hence by a. Thus, an endomorphism $\varphi: E \to E$ is defined by $\varphi a = b$. In this way, every tensor Φ of order $(1,1)$ determines an endomorphism φ

of E. The tensor Φ and the endomorphism φ are connected by the equation

$$\Phi(\overset{*}{x}, x) = \langle \overset{*}{x}, \varphi x\rangle \qquad \overset{*}{x} \in E^*,\ x \in E\,. \tag{7.1}$$

Conversely, every endomorphism φ of E determines a tensor $\Phi \in T_1^1(E)$ by the equation (7.1). The correspondence $\varphi \to \Phi$ defines an isomorphism

$$\mathscr{A}: L(E; E) \to T_1^1(E)\,.$$

Now the relation (7.1) can be written as

$$(\mathscr{A}\varphi)(\overset{*}{x}, x) = \langle \overset{*}{x}, \varphi x\rangle\,.$$

The identity-map ι corresponds to the *unit-tensor*

$$J(\overset{*}{x}, x) = \langle \overset{*}{x}, x\rangle\,.$$

In the same way as above, an isomorphism between $T_1^1(E)$ and the space of all endomorphisms of E^* can be defined. The tensor Φ and the corresponding endomorphism ψ of E^* are related by

$$\Phi(\overset{*}{x}, x) = \langle \psi\overset{*}{x}, x\rangle \qquad \overset{*}{x} \in E^*,\ x \in E\,. \tag{7.2}$$

Comparing the equations (7.1) and (7.2) we see that the two endomorphisms φ and ψ are dual to each other, $\psi = \varphi^*$.

7.3. Multiplication of tensors. The product of two tensors $\Phi \in T_q^p(E)$ and $\Psi \in T_s^r(E)$, denoted as $\Phi \otimes \Psi$, is defined by the equation

$$(\Phi \otimes \Psi)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p+r}, x_1 \ldots x_{q+s}) = \Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q)\,.$$
$$\Psi(\overset{*}{x}{}^{p+1} \ldots \overset{*}{x}{}^{p+r}; x_{q+1} \ldots x_{q+s})\,. \tag{7.3}$$

It is immediately clear from the above definition that the multiplication of tensors has the following properties:

1. $\left.\begin{array}{l} \Phi \otimes (\Psi_1 + \Psi_2) = \Phi \otimes \Psi_1 + \Phi \otimes \Psi_2 \\ (\Phi_1 + \Phi_2) \otimes \Psi = \Phi_1 \otimes \Psi + \Phi_2 \otimes \Psi \end{array}\right\}$ (distributive laws).

2. $(\Phi \otimes \Psi) \otimes X = \Phi \otimes (\Psi \otimes X)$ (associative law).

3. $\lambda \otimes \Phi = \Phi \otimes \lambda = \lambda\Phi$ where the scalar λ is considered as a tensor of order (0, 0).

It follows from the associative law that the product of several tensors Φ_ν $(\nu = 1 \ldots p)$ can be simply written as $\Phi_1 \otimes \cdots \otimes \Phi_p$.

The commutative law does not hold generally. Consider for instance two covariant tensors Φ and Ψ of order 1. Then

$$(\Phi \otimes \Psi)(x_1, x_2) = \Phi(x_1)\Psi(x_2)$$

and

$$(\Psi \otimes \Phi)(x_1, x_2) = \Psi(x_1)\Phi(x_2)\,,$$

whence $\Psi\Phi \neq \Phi\Psi$ unless Ψ is a multiple of Φ.

7.4. Decomposable tensors. Consider p vectors $a_\nu \in E$. Every vector a_ν determines a contravariant tensor of order 1. This tensor will be also denoted by a_ν. Forming the product

$$\Phi = a_1 \otimes \cdots \otimes a_p \tag{7.4}$$

we obtain a contravariant tensor of order p. Similarly, the product of q vectors $\overset{*}{a}{}^\mu \in E^*$ defines a covariant tensor

$$\Psi = \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^q$$

of order q. The product

$$\Phi \otimes \Psi = a_1 \otimes \cdots \otimes a_p \otimes \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^q$$

defines a tensor of order (p, q). A tensor of this form is called *decomposable.* The product of two decomposable tensors is obviously again decomposable.

7.5. Basis of $T^p_q(E)$. Let $\overset{*}{e}{}^\nu$, e_ν $(\nu = 1 \dots n)$ be a pair of dual bases of E^* and E. It will be shown that the n^{p+q} tensors

$$e_{\nu_1} \otimes \cdots \otimes e_{\nu_p} \otimes \overset{*}{e}{}^{\mu_1} \otimes \cdots \otimes \overset{*}{e}{}^{\mu_p} \quad (1 \leqq \nu_i \leqq n, \quad 1 \leqq \mu_j \leqq n) \tag{7.5}$$

form a basis of the space $T^p_q(E)$.

To prove that the tensors (7.5) are linearly independent, assume a relation

$$\sum_{(\nu)(\mu)} \lambda^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} e_{\nu_1} \otimes \cdots \otimes e_{\nu_p} \otimes \overset{*}{e}{}^{\mu_1} \otimes \cdots \otimes \overset{*}{e}{}^{\mu_q} = 0 .$$

Then

$$\sum_{(\nu)(\mu)} \lambda^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} \langle \overset{*}{x}{}^1, e_{\nu_1} \rangle \cdots \langle \overset{*}{x}{}^p, e_{\nu_p} \rangle \langle \overset{*}{e}{}^{\mu_1}, x_1 \rangle \cdots \langle \overset{*}{e}{}^{\mu_q}, x_q \rangle = 0 \tag{7.6}$$

for all vectors $\overset{*}{x}{}^i \in E^*$ $(i = 1 \dots p)$ and all vectors $x_j \in E$ $(j = 1 \dots q)$. Select two systems of indices $(\alpha_1 \dots a_p)$ and $(\beta_1 \dots \beta_q)$ and substitute in (7.6)

$$\overset{*}{x}{}^i = e^{\alpha_i} \; (i = 1 \dots p) \quad \text{and} \quad x_j = e_{\beta_j} \; (j = 1 \dots q) .$$

Then

$$\sum_{(\nu)(\mu)} \lambda^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} \delta^{\alpha_1}_{\nu_1} \dots \delta^{\alpha_p}_{\nu_p} \delta^{\mu_1}_{\beta_1} \dots \delta^{\mu_q}_{\beta} = 0 ,$$

whence

$$\lambda^{\alpha_1 \dots \alpha_p}_{\beta_1 \dots \beta_q} = 0 .$$

It remains to be shown that the tensors (7.5) generate the whole space $T^p_q(E)$. Let $\Phi \in T^p_q(E)$ be an arbitrary tensor and define n^{p+q} numbers by

$$\Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} = \Phi \, (\overset{*}{e}{}^{\nu_1} \dots \overset{*}{e}{}^{\nu_p}; e_{\mu_1} \dots e_{\mu_q}) . \tag{7.7}$$

Then

$$\Phi = \sum_{(\nu)(\mu)} \Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} e_{\nu_1} \otimes \cdots \otimes e_{\nu_p} \otimes \overset{*}{e}{}^{\mu_1} \otimes \cdots \otimes \overset{*}{e}{}^{\mu_q} . \tag{7.8}$$

In fact, the tensors on both sides of (7.8) assume the same value for every system of arguments $(\overset{*}{e}^{\alpha_1} \dots \overset{*}{e}^{\alpha_p}, e_{\beta_1} \dots e_{\beta_q})$. Hence these two tensors are equal.

Consequently, the tensors (7.5) form a basis of the space $T^p_q(E)$. This implies that

$$\dim T^p_q(E) = n^{p+q}.$$

The components of a tensor $\Phi \in T^p_q(E)$ relative to the basis (7.5) of $T^p_q(E)$ are given by the equations (7.7). These numbers will be briefly called the *components of* Φ *relative to the basis* $e_\nu (\nu = 1 \dots n)$ of E.

7.6. The transformation-formula for the components. Consider another pair of dual bases $\overset{*}{\bar{e}}^\nu$, $\bar{e}_\nu$ $(\nu = 1 \dots n)$ of E^* and E. Then

$$\bar{e}_\nu = \sum_\mu \alpha^\mu_\nu \, e_\mu \tag{7.9}$$

and

$$\overset{*}{\bar{e}}^\nu = \sum_\mu \check{\alpha}^\nu_\mu \, \overset{*}{e}^\mu , \tag{7.10}$$

where $(\check{\alpha}^\nu_\mu)$ is the inverse of the matrix (α^μ_ν). The components of Φ relative to the basis $\bar{e}_\nu$ $(\nu = 1 \dots n)$ are given by

$$\bar{\Phi}^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} = \Phi\,(\overset{*}{\bar{e}}^{\nu_1} \dots \overset{*}{\bar{e}}^{\nu_p}; \bar{e}_{\mu_1} \dots \bar{e}_{\mu_q}) . \tag{7.11}$$

Substituting the expressions (7.9) and (7.10) in (7.11) we obtain the following transformation-formula for the components of Φ:

$$\bar{\Phi}^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} = \sum_{(\lambda)(\varkappa)} \check{\alpha}^{\nu_1}_{\lambda_1} \dots \check{\alpha}^{\nu_p}_{\lambda_p} \alpha^{\varkappa_1}_{\mu_1} \dots \alpha^{\varkappa_q}_{\mu_q} \Phi^{\lambda_1 \dots \lambda_p}_{\varkappa_1 \dots \varkappa_q} . \tag{7.12}$$

7.7. The tensor-product $T^p_q(E) \otimes T^r_s(E)$. Now we are ready to show that the bilinear mapping

$$(T^p_q(E), T^r_s(E)) \to T^{p+r}_{q+s}(E)$$

defined by the multiplication is indeed a tensor-product in the sense of sec. 6.5. To prove the property T_1 let $X \in T^{p+r}_{q+s}(E)$ be an arbitrary tensor. It follows from (7.8) that X is a linear combination of products

$$\begin{aligned} X &= a_1 \otimes \cdots \otimes a_{p+r} \otimes \overset{*}{a}^1 \otimes \cdots \otimes \overset{*}{a}^{q+s} \\ &= (a_1 \otimes \cdots \otimes a_p \otimes \overset{*}{a}^1 \otimes \cdots \otimes \overset{*}{a}^q) \otimes (a_{p+1} \otimes \cdots \otimes a_{p+r} \otimes \\ &\quad \otimes \overset{*}{a}^{q+1} \otimes \cdots \otimes \overset{*}{a}^{q+s}) . \end{aligned}$$

This implies that the products $\Phi \otimes \Psi$ $(\Phi \in T^p_q(E), \Psi \in T^r_s(E))$ generate the whole space $T^{p+r}_{q+s}(E)$. The property T_2 follows from the fact that

$$\dim T^{p+r}_{q+s}(E) = n^{p+r+q+s} = \dim T^p_q(E) \cdot \dim T^r_s(E) ,$$

see sec. 6.8.

Hence, the notation $\Phi \otimes \Psi$ for the product of two tensors is justified and we can write

$$T_{q+s}^{p+r}(E) = T_q^p(E) \otimes T_s^r(E) .$$

7.8. Tensors and linear mappings. Consider two linear spaces E and F and a linear mapping $\vartheta: E \to F$. Then every contravariant tensor $\Phi \in T^p(E)$ determines a contravariant tensor $\Phi' \in T^p(F)$ by the equation

$$\Phi'(\overset{*}{y}{}^1 \dots \overset{*}{y}{}^p) = \Phi(\vartheta^* \overset{*}{y}{}^1 \dots \vartheta^* \overset{*}{y}{}^p) \quad \overset{*}{y}{}^\nu \in F^* \quad (\nu = 1 \dots p) .$$

Hence, a linear mapping

$$(\vartheta)^p: T^p(E) \to T^p(F)\,^{*)}$$

is induced by ϑ. Now the above equation can be written as

$$(\vartheta)^p \Phi(\overset{*}{y}{}^1 \dots \overset{*}{y}{}^p) = \Phi(\vartheta^* \overset{*}{y}{}^1 \dots \vartheta^* \overset{*}{y}{}^p) .$$

The mapping $(\vartheta)^1: T^1(E) \to T^1(F)$ coincides with ϑ if the spaces $T^1(E)$ and $T^1(F)$ are identified with E and F respectively. It follows immediately from the definition of $(\vartheta)^p$ that

$$(\vartheta)^{p+q}(\Phi_1 \otimes \Phi_2) = (\vartheta)^p \Phi_1 \otimes (\vartheta)^q \Phi_2 \tag{7.13}$$

and

$$(\vartheta_2 \circ \vartheta_1)^p = (\vartheta_2)^p \circ (\vartheta_1)^p \qquad \vartheta_1: E \to F, \quad \vartheta_2: F \to G . \tag{7.14}$$

The formula (7.13) implies that

$$(\vartheta)^p (a_1 \otimes \cdots \otimes a_p) = \vartheta a_1 \otimes \cdots \otimes \vartheta a_p, \qquad a_\nu \in E .$$

Let us now consider the covariant tensor-spaces $T_p(E)$ and $T_p(F)$. The mapping ϑ induces a linear mapping

$$(\vartheta)_p: T_p(F) \to T_p(E)$$

in the inverse direction defined by

$$(\vartheta)_p \Psi(x_1 \dots x_p) = \Psi(\vartheta x_1 \dots \vartheta x_p) \quad x_\nu \in E \qquad (\nu = 1 \dots p) .$$

The mapping

$$(\vartheta)_1: T_1(F) \to T_1(E)$$

coincides with the dual mapping ϑ^* if the spaces $T_1(E)$ and $T_1(F)$ are identified with E^* and F^* respectively. The formulas

$$(\vartheta)_{p+q}(\Psi_1 \otimes \Psi_2) = (\vartheta)_p \Psi_1 \otimes (\vartheta)_q \Psi_2 \tag{7.15}$$

and

$$(\vartheta_2 \circ \vartheta_1)_p = (\vartheta_1)_p \circ (\vartheta_2)_p \tag{7.16}$$

are immediate consequences of the definition of $(\vartheta)_p$. It follows from (7.15) that

$$(\vartheta)_p (\overset{*}{b}{}^1 \otimes \cdots \otimes \overset{*}{b}{}^p) = \vartheta^* \overset{*}{b}{}^1 \otimes \cdots \otimes \vartheta^* \overset{*}{b}{}^p, \qquad \overset{*}{b}{}^\nu \in F .$$

---

*) The notation $(\vartheta)^p$ is used to avoid confusion with the product ϑ^p in the case $F = E$.

7.9. Tensorial mappings. Let α be an automorphism of E. Then every tensor $\Phi \in T_q^p(E)$ determines another tensor $\alpha\Phi$ of the same order, defined by

$$(\alpha\Phi)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) = \Phi(\alpha^* \overset{*}{x}{}^1 \ldots \alpha^* \overset{*}{x}{}^p; \alpha^{-1}x_1 \ldots \alpha^{-1}x_q).$$

In this way, a multiplication between the automorphisms α of E and the tensors $\Phi \in T_q^p(E)$ is defined. This multiplication has the following properties:

1. $\alpha(\Phi_1 + \Phi_2) = \alpha\Phi_1 + \alpha\Phi_2$.
2. $\alpha(\Phi \otimes \Psi) = \alpha\Phi \otimes \alpha\Psi$.
3. $(\alpha\beta)\Phi = \alpha(\beta\Phi)$.
4. $\iota\Phi = \Phi$ (ι identity-map).

Now consider a linear mapping

$$F: T_q^p(E) \to T_s^r(E).$$

The mapping F is called *tensorial*, if

$$F(\alpha\Phi) = \alpha F\Phi$$

for every automorphism α of E and every tensor $\Phi \in T_q^p(E)$. For example, the mapping

$$F\Phi = \Phi \otimes J,$$

where J denotes the unit-tensor, is a tensorial mapping of $T_q^p(E)$ into $T_{q+1}^{p+1}(E)$. Another example will be given in sec. 7.13.

7.10. Vector-valued tensors. The notion of a tensor can be generalized in an obvious way to the case of *vector-valued tensors*. Given a linear space H, an H-valued tensor of order (p, q) in E is a $(p+q)$-linear mapping of E^* and E into H with p arguments in E^* and q arguments in E. If H is the number-space, the H-valued tensors reduce to the scalar-valued tensors defined in sec. 7.1. An H-valued tensor of order (1,0) is a linear mapping of E into H and an H-valued tensor of order (0,1) is a linear mapping of E^* into H.

The set of all H-valued tensors of order (p, q) forms a linear space $T_q^p(E; H)$. It is convenient to extend this definition to the case $p = 0$, $q = 0$ by setting $T_0^0(E; H) = H$. The dimension of the space $T_q^p(E; H)$ is given by

$$\dim T_q^p(E; H) = (\dim E)^{p+q} \dim H.$$

Suppose that two linear spaces H_1 and H_2 and a linear mapping $\varphi: H_1 \to H_2$ are given. Then every H_1-valued tensor Φ determines an H_2-valued tensor $\varphi_{\#}\Phi$ of the same order. The tensor $\varphi_{\#}\Phi$ is defined by

$$(\varphi_{\#}\Phi)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) = \varphi\Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q). \quad (7.17)$$

In this way, a linear mapping

$$\varphi_{\#}: T^p_q(E; H_1) \to T^p_q(E; H_2)$$

is induced by φ.

As in the case of scalar-valued tensors, a multiplication can be defined between the automorphism α of E and the H-valued tensors in the following way:

$$(\alpha\Phi)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p, x_1 \ldots x_q) = \Phi(\alpha^* \overset{*}{x}{}^1 \ldots \alpha^* \overset{*}{x}{}^p; \alpha^{-1} x_1 \ldots \alpha^{-1} x_q). \quad (7.18)$$

The equations (7.17) and (7.18) imply that

$$\varphi_{\#}\, \alpha\Phi = \alpha\, \varphi_{\#}\Phi \qquad \Phi \in T^p_q(E; H_1) \quad (7.19)$$

for every automorphism α of E and every linear mapping $\varphi: H_1 \to H_2$.

Problems: 1. Consider two decomposable tensors

$$\Phi_1 = a_1 \otimes b_1 \quad \text{and} \quad \Phi_2 = a_2 \otimes b_2 \quad a_i, b_i \in E \quad (i = 1, 2)$$

and assume that $\Phi_1 \neq 0$. Show that the tensor $\Phi_1 + \Phi_2$ is decomposable if and only if $a_2 = \lambda a_1$ or $b_2 = \lambda b_1$.

2. Find all the endomorphisms φ of E such that the corresponding tensors of order (1,1) are decomposable.

3. Given a covariant tensor Φ of order p, define the subspaces U_i $(i = 1 \ldots p)$ of E as follows: A vector $u \in E$ is contained in U_i if and only if

$$\Phi(x_1 \ldots x_{i-1}, u, x_{i+1} \ldots x_p) = 0 \quad \text{for all vectors } x_\nu \in E \; (\nu \neq i).$$

Prove: The tensor Φ is decomposable if and only if

$$\dim U_i = n - 1 \quad (i = 1 \ldots p).$$

4. Assume that a covariant tensor $\Phi \neq 0$ is represented in two different ways as a product of vectors,

$$\Phi = \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^p, \quad \Phi = \overset{*}{b}{}^1 \otimes \cdots \otimes \overset{*}{b}{}^p.$$

Prove that

$$\overset{*}{b}{}^i = \lambda^i \overset{*}{a}{}^i \quad (i = 1 \ldots p),$$

where λ^i $(i = 1 \ldots p)$ are scalars and

$$\lambda^1 \ldots \lambda^p = 1.$$

5. A bilinear mapping

$$B: (T^p_q(E), T^r_s(E)) \to T^k_l(E)$$

is called *tensorial*, if

$$B(\alpha\Phi, \alpha\Psi) = \alpha B(\Phi, \Psi) \quad \Phi \in T^p_q(E), \Psi \in T^r_s(E)$$

for every automorphism α of E. Prove that the multiplication of tensors is a bilinear tensorial mapping.

6. Denote by $T(E)$ the set of all infinite sequences

$$\Phi: (\Phi_0, \Phi_1, \ldots \Phi_n \ldots) \quad \text{with} \quad \Phi_\nu \in T_\nu(E) \quad (\nu = 0, 1 \ldots).$$

Define the sum of two sequences

$$\Phi: (\Phi_0, \Phi_1 \ldots) \quad \text{and} \quad \Psi: (\Psi_0, \Psi_1 \ldots)$$

as the sequence

$$\Phi + \Psi: (\Phi_0 + \Psi_0, \Phi_1 + \Psi_1 \ldots)$$

and the product of two sequences Φ and Ψ as the sequence

$$X: (\chi_0, \chi_1 \ldots),$$

where $X_i = \sum\limits_{\nu+\mu=i} \Phi_\nu \otimes \Psi_\mu \quad (i = 0, 1 \ldots)$.

Prove that these two operations define an algebra (of dimension infinity). Show that the unit-element of this algebra is the sequence $(1, 0 \ldots)$.

7. Given two linear spaces H_1 and H_2 define the bilinear mapping

$$(T_q^p(E; H_1), T_s^r(E; H_2)) \to T_{q+s}^{p+r}(E; H_1 \otimes H_2)$$

by

$$\begin{aligned}&(\Phi \otimes \Psi)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p+r}; x_1 \ldots x_{q+s})\\ &= \Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) \otimes \Psi(\overset{*}{x}{}^{p+1} \ldots \overset{*}{x}{}^{p+r}; x_{q+1} \ldots x_{q+s})\end{aligned}$$

where $\Phi \in T_q^p(E; H_1)$ and $\Psi \in T_s^r(E; H_2)$. Prove that this bilinear mapping is a tensor-product.

8. Let E and F be linear spaces of finite dimension.

a) Given two tensors $\Phi \in T_q^p(E)$ and $\Psi \in T_q^p(F)$, prove that there exists exactly one tensor $X \in T_q^p(E \otimes F)$ such that

$$\begin{aligned}&X(\overset{*}{x}{}^1 \otimes \overset{*}{y}{}^1 \ldots \overset{*}{x}{}^p \otimes \overset{*}{y}{}^p; x_1 \otimes y_1 \ldots x_q \otimes y_q)\\ &= \Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) \cdot \Psi(\overset{*}{y}{}^1 \ldots \overset{*}{y}{}^p; y_1 \ldots y_q).\end{aligned}$$

b) Show that the bilinear mapping

$$(T_q^p(E), T_q^p(F)) \to T_q^p(E \otimes F)$$

defined by $(\Phi, \Psi) \to X$ is a tensor-product.

c) If φ is an endomorphism of E and ψ is an endomorphism of F prove that

$$\mathscr{A}(\varphi \otimes \psi) = \mathscr{A}(\varphi) \otimes \mathscr{A}(\psi)$$

where $\mathscr{A}$ is the operator defined in sec. 7.2.

9. A tensor $\Phi \in T_q^p(E)$ is called *invariant*, if $\alpha\Phi = \Phi$ for every automorphism α of E (see sec. 7.9).

a) Show that the tensor

$$J(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_p) = \langle \overset{*}{x}{}^1, x_1 \rangle \cdots \langle \overset{*}{x}{}^p, x_p \rangle$$

is invariant.

b) Prove that every invariant tensor $\Phi \in T^p_q(E)$ is the zero-tensor if $p \neq q$.

10. Let φ be an endomorphism of E. Denote by $\theta^p(\varphi)$ and by $\theta_p(\varphi)$ the endomorphisms of $T^p(E)$ and $T_p(E)$ which are defined by

$$\theta^p(\varphi)\,\Phi(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^p) = \sum_{i=1}^{p} \Phi(\overset{*}{x}{}^1 \dots \varphi^* \overset{*}{x}{}^i \dots \overset{*}{x}{}^p)$$

and

$$\theta_p(\varphi)\,\Psi(x_1 \dots x_p) = -\sum_{i=1}^{p} \Psi(x_1 \dots \varphi x_i \dots x_p)$$

respectively. Prove the following formulas:

$$\theta^{p+q}(\varphi)\,(\Phi_1 \otimes \Phi_2) = \theta^p(\varphi)\,\Phi_1 \otimes \Phi_2 + \Phi_1 \otimes \theta^q(\varphi)\,\Phi_2\,,$$

$$\theta_{p+q}(\varphi)\,(\Psi_1 \otimes \Psi_2) = \theta_p(\varphi)\,\Psi_1 \otimes \Psi_2 + \Psi_1 \otimes \theta_q(\varphi)\,\Psi_2\,,$$

$$\theta^p(\varphi_1 \circ \varphi_2 - \varphi_2 \circ \varphi_1) = \theta^p(\varphi_1) \circ \theta^p(\varphi_2) - \theta^p(\varphi_2) \circ \theta^p(\varphi_1)\,,$$

$$\theta_p(\varphi_1 \circ \varphi_2 - \varphi_2 \circ \varphi_1) = \theta_p(\varphi_1) \circ \theta_p(\varphi_2) - \theta_p(\varphi_2) \circ \theta_p(\varphi_1)\,.$$

11. With the help of problem 10 prove again that

$$\operatorname{tr}(\varphi_1 \circ \varphi_2) = \operatorname{tr}(\varphi_2 \circ \varphi_1)\,.$$

§ 2. Contraction

7.11. Definition. The contraction is an operation which associates with every tensor of order (p, q) $(p \geqq 1, q \geqq 1)$ a tensor of order $(p-1, q-1)$. Consider first a tensor θ of order $(1,1)$. Then $\mathscr{A}^{-1}\theta$ is an endomorphism of E (cf. sec. 7.2). We now define the contracted tensor $C\theta$ as the scalar $\operatorname{tr}(\mathscr{A}^{-1}\theta)$,

$$C\theta = \operatorname{tr}(\mathscr{A}^{-1}\theta)\,. \tag{7.20}$$

In this way we obtain a linear mapping

$$C\colon T^1_1(E) \to T^0_0(E)\,.$$

Passing over to the general case assume that a tensor $\Phi \in T^p_q(E)$ $(p \geqq 1, q \geqq 1)$ and a pair of integers (i, j) $(1 \leqq i \leqq p, 1 \leqq j \leqq q)$ is given. Define the $T^1_1(E)$-valued tensor Φ^i_j of order $(p-1, q-1)$ by

$$\begin{aligned}&\Phi^i_j\,(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^{p-1}; x_1 \dots x_{q-1})\,(\overset{*}{x}, x)\\ &= \Phi\,(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^{i-1}, \overset{*}{x}, \overset{*}{x}{}^i \dots \overset{*}{x}{}^p; x_1 \dots x_{j-1}, x, x_j \dots x_{q-1})\,.\end{aligned} \tag{7.21}$$

The operator C induces a linear mapping

$$C_{\#}\colon T^{p-1}_{q-1}(E; T^1_1(E)) \to T^{p-1}_{q-1}(E)$$

(cf. sec. 7.10). Applying this operation to Φ^i_j we obtain a scalar-valued tensor $C_{\#}\Phi^i_j$ of order $(p-1, q-1)$. This tensor is called the *contraction* of Φ with respect to the pair (i, j) and will be denoted by $C_j^i\Phi$,

$$C_j^i\Phi = C_{\#}\Phi_j^i\,. \tag{7.22}$$

We thus obtain for every pair of integers (i, j) a linear mapping

$$C_j^i \colon T_q^p(E) \to T_{q-1}^{p-1}(E)$$

called the *contraction* with respect to the pair (i, j).

The operator

$$C_1^1 \colon T_1^1(E) \to T_0^0(E)$$

coincides with the operator C defined by (7.20). In fact, let θ be a tensor of order (1,1). Then θ_1^1 is a $T_1^1(E)$-valued tensor of order (0,0) i.e. a vector of $T_1^1(E)$. It follows from (7.21) that $\theta_1^1 = \theta$. This implies that

$$C_1^1\,\theta = C_{\#}\theta_1^1 = C\,\theta$$

whence $C_1^1 = C$.

7.12. Explicit formulas. In this section we shall derive a more explicit expression for the contraction. Let us again begin with a tensor θ of order (1,1). Employing a pair of dual bases $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$ we obtain from (7.20) and (4.61)

$$C\,\theta = \operatorname{tr}\,(\mathscr{A}^{-1}\theta) = \sum_\nu \langle \overset{*}{e}{}^\nu, (\mathscr{A}^{-1}\theta)\,e_\nu\rangle = \sum_\nu \theta\,(\overset{*}{e}{}^\nu, e_\nu)$$

whence

$$C\,\theta = \sum_\nu \theta\,(\overset{*}{e}{}^\nu, e_\nu)\,. \tag{7.23}$$

In the general case it follows from (7.22) and (7.23) that

$$\begin{aligned}
&(C_j^i\Phi)\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{q-1})\\
&= C\Phi_j^i\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{q-1})\\
&= \sum_\nu \Phi_j^i\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{q-1})\,(\overset{*}{e}{}^\nu, e_\nu)\,.
\end{aligned} \tag{7.24}$$

But in view of (7.21)

$$\begin{aligned}
&\sum_\nu \Phi_j^i\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{q-1})\,(\overset{*}{e}{}^\nu, e_\nu)\\
&= \sum_\nu \Phi\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{i-1}, \overset{*}{e}{}^\nu, \overset{*}{x}{}^i \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{j-1}, e_\nu, x_j \ldots x_q).
\end{aligned} \tag{7.25}$$

Combining (7.24) and (7.25) we obtain the formula

$$\begin{aligned}
&(C_j^i\Phi)\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{q-1})\\
&= \sum_\nu \Phi\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{i-1}, \overset{*}{e}{}^\nu, \overset{*}{x}{}^i \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{j-1}, e_\nu, x_j \ldots x_{q-1}).
\end{aligned} \tag{7.26}$$

In particular, this relation shows that the sum on the right hand-side does not depend on the choice of the dual bases $\overset{*}{e}{}^\nu, e_\nu$.

It follows from (7.26) that the contraction of a decomposable tensor

$$\Phi = a_1 \otimes \cdots \otimes a_p \otimes \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^q \qquad a_\nu \in E,\ \overset{*}{a}{}^\mu \in E^*$$

is given by

$$C_j^i \Phi = \langle \overset{*}{a}{}^j, a_i \rangle\, a_1 \otimes \cdots \otimes \hat{a}_i \otimes \cdots a_p \otimes \overset{*}{a}{}^1 \otimes \cdots \otimes \hat{\overset{*}{a}}{}^j \otimes \cdots \otimes \overset{*}{a}{}^q \tag{7.27}$$

where the symbol $\hat{a}_i$ indicates that the factor a_i is deleted. In fact,

$$\begin{aligned}(C_j^i \Phi)\,&(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{q-1}) \\ = \sum_\nu &\langle \overset{*}{x}{}^1, a_1 \rangle \cdots \langle \overset{*}{x}{}^{i-1}, a_{i-1} \rangle \langle \overset{*}{e}{}^\nu, a_i \rangle \langle \overset{*}{x}{}^i, a_{i+1} \rangle \ldots \langle \overset{*}{x}{}^{p-1} a_p \rangle \\ &\langle \overset{*}{b}{}^1, x_1 \rangle \cdots \langle \overset{*}{b}{}^{j-1}, x_{j-1} \rangle \langle \overset{*}{b}{}^j, e_\nu \rangle \langle \overset{*}{b}{}^{j+1}, x_j \rangle \cdots \langle \overset{*}{b}{}^q, x_{q-1} \rangle .\end{aligned} \tag{7.28}$$

Observing that

$$\sum_\nu \langle \overset{*}{e}{}^\nu, a_i \rangle \langle \overset{*}{b}{}^j, e_\nu \rangle = \langle \overset{*}{b}{}^j, a_i \rangle$$

we obtain (7.27) from (7.28).

7.13. The tensorial character of the contraction. In concluding this paragraph it will be shown that the operator

$$C_j^i\colon T_q^p(E) \to T_{q-1}^{p-1}(E)$$

defines a tensorial mapping, i. e. that

$$\alpha\, C_j^i \Phi = C_j^i\, \alpha \Phi \qquad \Phi \in T_q^p(E)$$

for every automorphism α of E (cf. sec. 7.9).

Denote by T_α the endomorphism of the space $T_1^1(E)$ defined by

$$(T_\alpha\, \theta)\, (\overset{*}{x}, x) = \theta(\overset{*}{\alpha} \overset{*}{x}, \alpha^{-1} x)\, . \tag{7.29}$$

We show first that

$$C \circ T_\alpha = T_\alpha \tag{7.30}$$

where C is the operator defined by (7.20). Let $\theta \in T_1^1(E)$ be an arbitrary tensor. Then

$$C\theta = \operatorname{tr}(\mathscr{A}^{-1}\theta)$$

and

$$C(T_\alpha \theta) = \operatorname{tr} \mathscr{A}^{-1}(T_\alpha \theta)\, .$$

Now it follows from (7.29) that

$$\mathscr{A}^{-1}(T_\alpha \theta) = \alpha \circ (\mathscr{A}^{-1}\theta) \circ \alpha^{-1}$$

and we thus obtain

$$C(T_\alpha \theta) = \operatorname{tr}(\alpha \circ \mathscr{A}^{-1}\theta \circ \alpha^{-1}) = \operatorname{tr}(\mathscr{A}^{-1}\theta) = C\theta$$

whence (7.30).

Now consider an arbitrary tensor $\Phi \in T^p_q(E)$. The tensors $(\alpha\Phi^i_j)$ and $\alpha(\Phi^i_j)$ are connected by the relation

$$(\alpha\Phi)^i_j = (T_\alpha)_{\#}\,\alpha(\Phi^i_j) \tag{7.31}$$

where $(T_\alpha)_{\#}$ is the endomorphism of $T^{p-1}_{q-1}(E, T^1_1(E))$ induced by T_α (cf. sec. 7.10). The equation (7.31) follows immediately from (7.29) and (7.21). Next, we observe that in view of (7.30)

$$C_{\#} \circ (T_\alpha)_{\#} = C_{\#}\,. \tag{7.32}$$

Combining the formulas (7.22), (7.31), (7.32) and (7.19) we obtain

$$C^i_j(\alpha\Phi) = C_{\#}(\alpha\Phi)^i_j = C_{\#}(T_\alpha)_{\#}\,\alpha(\Phi^i_j) = C_{\#}\alpha\Phi^i_j = \alpha C^i_j\Phi\,. \tag{7.33}$$

Problems: 1. Given two endomorphisms $\varphi: E \to E$ and $\psi: E \to E$, prove the formulas

$$\mathscr{A}(\psi \circ \varphi) = C^1_2(\mathscr{A}\varphi \otimes \mathscr{A}\psi) \quad \text{and} \quad \mathscr{A}(\varphi \circ \psi) = C^2_1(\mathscr{A}\varphi \otimes \mathscr{A}\psi)\,.$$

2. Given two vectors $a \in E$ and $\overset{*}{a} \in E^*$, define the *substitution-operators*

$$i(a): T^p_q(E) \to T^p_{q-1}(E) \quad \text{and} \quad i(\overset{*}{a}): T^p_q(E) \to T^{p-1}_q(E)$$

by

$$i(a)\,\Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_{q-1}) = \Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; a, x_1 \ldots x_{q-1})$$

and

$$i(\overset{*}{a})\,\Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_p) = \Phi(\overset{*}{a}, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_q)\,.$$

Prove the formulas

$$C^1_1(a \otimes \Phi) = i(a)\,\Phi \quad \text{and} \quad C^1_1(\overset{*}{a} \otimes \Phi) = i(\overset{*}{a})\,\Phi\,.$$

3. Using the transformation-formula (7.12), show that the sum

$$\sum_\nu \Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{i-1}, \overset{*}{e}{}^\nu, \overset{*}{x}{}^i \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{j-1}, e_\nu, x_j \ldots x_{q-1})$$

is independent of the choice of the dual bases $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$.

4. Employing the explicit formula (7.26), prove that the mapping C^i_j is a tensorial mapping.

5. Prove that every tensorial mapping

$$F: T^1_1(E) \to T^0_0(E)$$

has the form

$$F\Phi = \lambda \cdot C\Phi$$

where λ is a scalar.

6. Verify the relations

$$C^i_j \circ C^k_l = \begin{cases} C^{k-1}_{l-1} \circ C^i_j & \text{if } \ i < k, j < l \\ C^{k-1}_{l} \circ C^i_{j-1} & \text{if } \ i < k, j \geqq l \\ C^k_{l-1} \circ C^{i+1}_j & \text{if } \ i \geqq k, j < l\,. \end{cases}$$

7. Given two pairs of dual spaces E^*, E and F^*, F define the bilinear mapping

$$(T_q^p(E),\, T_q^p(F)) \to T_q^p\,(E \otimes F)$$

as in § 1, prob. 8. Show that

$$C_j^i\, X = C_j^i \Phi \otimes C_j^i \Psi\,.$$

8. Assume that a bilinear mapping

$$\Phi\colon (E^*, E) \to H$$

is given.

a) Prove that the sum $\sum_\nu \Phi\,(\overset{*}{e}{}^\nu, e_\nu)$ does not depend on the choice of the dual bases $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$.

b) Show that the vector

$$h = \sum_\nu \Phi\,(\overset{*}{e}{}^\nu, e_\nu)$$

has the following property: Let H^* be a dual space of H and define the linear mapping $H^* \to L(E; E)$ by

$$\langle x^*, \varphi_{h^*}\, x\rangle = \langle h^*, \Phi\,(\overset{*}{x}, x)\rangle\,.$$

Then

$$\operatorname{tr} \varphi_{h^*} = \langle h^*, h\rangle\,.$$

§ 3. The duality of $T_q^p(E)$ and $T_p^q(E)$

7.14. The bilinear function $\langle \Phi, \Psi\rangle$. With the help of the contraction a scalar-product can be defined in any two spaces $T_p^q(E)$ and $T_q^p(E)$. Consider two tensors $\Phi \in T_q^p(E)$ and $\Psi \in T_p^q(E)$. Then $\Phi \otimes \Psi$ is a tensor of order $(p+q, p+q)$. Define the *"complete contraction-operator"* C by

$$C = \underbrace{C_1^1 \ldots C_1^1}_{q}\, \underbrace{C_{q+1}^1 \ldots C_{q+1}^1}_{p}\,.$$

Then, a bilinear function is given in the spaces $T_q^p(E)$ and $T_p^q(E)$ by

$$\langle \Phi, \Psi\rangle = C\,(\Phi \otimes \Psi) \qquad \Phi \in T_q^p(E),\ \Psi \in T_p^q(E)\,. \tag{7.34}$$

Employing the formula (7.26) $(p+q)$ times we see that the bilinear function (7.34) can be written as

$$\langle \Phi, \Psi\rangle = \sum_{(\nu)(\mu)} \Phi^{\nu_1 \ldots \nu_p}_{\mu_1 \ldots \mu_q}\, \Psi^{\mu_1 \ldots \mu_q}_{\nu_1 \ldots \nu_p}\,. \tag{7.35}$$

It will be shown that the bilinear function (7.34) defines a scalar-product in $T_q^p(E)$ and $T_p^q(E)$.

First of all we exhibit the following properties:

$$\langle \Phi \otimes \Phi', \Psi \otimes \Psi' \rangle = \langle \Phi, \Psi \rangle \langle \Phi', \Psi' \rangle \tag{7.36}$$
$$\Phi \in T_q^p(E), \Phi' \in T_s^r(E), \Psi \in T_p^q(E), \Psi' \in T_r^s(E)$$

$$\langle \Phi_1 \otimes \Phi_2, \Psi_1 \otimes \Psi_2 \rangle = \langle \Phi_1, \Psi_2 \rangle \langle \Phi_2, \Psi_1 \rangle \tag{7.37}$$
$$\Phi_1 \in T^p(E), \Phi_2 \in T_q(E), \Psi_1 \in T^q(E), \Psi_2 \in T_p(E)$$

$$\begin{aligned} &\langle \Phi, x_1 \otimes \cdots \otimes x_q \otimes \overset{*}{x}{}^1 \otimes \cdots \otimes \overset{*}{x}{}^p \rangle \\ &= \Phi(\overset{*}{x}{}^1, \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) \quad \Phi \in T_q^p(E) \end{aligned} \tag{7.38}$$

$$\langle a_1 \otimes \cdots \otimes a_p, \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^p \rangle = \langle \overset{*}{a}{}^1, a_1 \rangle \cdots \langle \overset{*}{a}{}^p, a_p \rangle . \tag{7.39}$$

All these formulas are immediate consequences of (7.35).

7.15. The duality of $T_q^p(E)$ and $T_p^q(E)$. Now we are ready to prove that the bilinear function (7.34) is a scalar-product (cf. sec. 2.6). In view of the symmetry it is sufficient to show: If

$$\langle \Phi, \Psi \rangle = 0$$

for a fixed tensor $\Phi \in T_q^p(E)$ and all tensors $\Psi \in T_q^p(E)$, then $\Phi = 0$. Consider p vectors $\overset{*}{x}{}^\nu \in E$ and q vectors $x_\mu \in E$. Using (7.38) we find that

$$\Phi(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) = \langle \Phi, x_1 \otimes \cdots \otimes x_q \otimes \overset{*}{x}{}^1 \otimes \cdots \otimes \overset{*}{x}{}^p \rangle = 0 ,$$

whence $\Phi = 0$.

Thus, any two spaces $T_q^p(E)$ and $T_p^q(E)$ are dual to each other with respect to the bilinear function (7.34).

Given a pair of dual bases $e_\nu, \overset{*}{e}{}^\nu$ $(\nu = 1 \ldots n)$ the formulas (7.37) and (7.39) yield

$$\begin{aligned} &\langle e_{\nu_1} \otimes \cdots \otimes e_{\nu_p} \otimes \overset{*}{e}{}^{\mu_1} \otimes \cdots \otimes \overset{*}{e}{}^{\mu_q}, e_{\varkappa_1} \otimes \cdots \otimes e_{\varkappa_q} \otimes \overset{*}{e}{}^{\lambda_1} \otimes \cdots \otimes \overset{*}{e}{}^{\lambda_p} \rangle \\ &\qquad = \delta_{\nu_1}^{\lambda_1} \ldots \delta_{\nu_p}^{\lambda_p} \delta_{\varkappa_1}^{\mu_1} \ldots \delta_{\varkappa_q}^{\mu_q} \end{aligned}$$

showing that the products

$$e_{\nu_1} \otimes \cdots \otimes e_{\nu_p} \otimes \overset{*}{e}{}^{\mu_1} \otimes \cdots \otimes \overset{*}{e}{}^{\mu_q}$$

and $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (7.40)

$$e_{\varkappa_1} \otimes \cdots \otimes e_{\varkappa_q} \otimes \overset{*}{e}{}^{\lambda_1} \otimes \cdots \otimes \overset{*}{e}{}^{\lambda_p}$$

form dual bases of $T_q^p(E)$ and $T_p^q(E)$.

7.16. The duality of $(\vartheta)^p$ and $(\vartheta)_p$. Let two linear spaces E and F and a linear mapping $\vartheta: E \to F$ be given. As it has been shown in sec. 7.8, the mapping ϑ induces linear mappings

$$(\vartheta)^p\colon T^p(E) \to T^p(F)$$

and

$$(\vartheta)_p\colon T_p(F) \to T_p(E)\,.$$

These mappings are dual to each other,

$$\langle \Psi, (\vartheta)^p \Phi\rangle = \langle (\vartheta)_p \Psi, \Phi\rangle \quad \Phi \in T^p(E), \Psi \in T_p(F). \tag{7.41}$$

It is sufficient to prove the relation (7.41) for two decomposable tensors

$$\Phi = a_1 \otimes \cdots \otimes a_p \qquad a_\nu \in E$$

and

$$\Psi = \overset{*}{b}{}^1 \otimes \cdots \otimes \overset{*}{b}{}^p \qquad \overset{*}{b}{}^\nu \in F^* \quad (\nu = 1 \ldots p)$$

because every tensor can be written as a linear combination of decomposable tensors. Employing the formulas (7.13), (7.15) and (7.39) we obtain

$$\langle \Psi, (\vartheta)^p \Phi\rangle = \langle \overset{*}{b}{}^1 \otimes \cdots \otimes \overset{*}{b}{}^p, \vartheta a_1 \otimes \cdots \otimes \vartheta a_p\rangle$$

$$= \langle \overset{*}{b}{}^1, \vartheta a_1\rangle \cdots \langle \overset{*}{b}{}^p, \vartheta a_p\rangle = \langle \vartheta^* \overset{*}{b}{}^1, a_1\rangle \cdots \langle \vartheta^* \overset{*}{b}{}^p, a_p\rangle$$

$$= \langle \vartheta^* \overset{*}{b}{}^1 \otimes \cdots \otimes \vartheta^* \overset{*}{b}{}^p, a_1 \otimes \cdots \otimes a_p\rangle = \langle (\vartheta)_p \Psi, \Phi\rangle\,.$$

Problems: 1. Given two endomorphisms φ and ψ of E prove the formula

$$\langle \mathscr{A}\varphi, \mathscr{A}\psi\rangle = \operatorname{tr}(\varphi \circ \psi)\,.$$

(cf. sec. 7.2).

2. Establish the relation

$$\langle \alpha\Phi, \alpha\Psi\rangle = \langle \Phi, \Psi\rangle \qquad \Phi \in T^p_q(E), \Psi \in T^q_p(E)$$

where α is an automorphism of E.

3. Define the linear mapping

$$F^i_j\colon T^{p-1}_{q-1}(E) \to T^p_q(E)$$

as follows*):

$$(F^i_j \Phi)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) = \langle \overset{*}{x}{}^i, x_j\rangle F(\overset{*}{x}{}^1 \ldots \widehat{\overset{*}{x}{}^i} \ldots \overset{*}{x}{}^p; x_1 \ldots \hat{x}_j \ldots x_q)$$

Show that the mapping F^i_j is dual to the contraction

$$C^i_j\colon T^q_p(E) \to T^{q-1}_{p-1}(E)\,.$$

*) The symbol $\widehat{\ }$ indicates that the corresponding arguments are deleted.

4. Define a multiplication in the space $T^p_q(E)$ in the following way:

$$\Phi \circ \Psi = \underbrace{C^1_{p+1} \ldots C^1_{p+1}}_{p} (\Phi \otimes \Psi) .$$

Prove: a) The space $T^p_q(E)$ becomes an algebra under this multiplication.

b) The tensor I^p_p, defined by

$$I^p_p(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_p) = \langle \overset{*}{x}{}^1, x_1 \rangle \cdots \langle \overset{*}{x}{}^p, x_p \rangle$$

is the unit-element of this algebra.

c) Given two endomorphisms $\varphi: E \to E$ and $\psi: E \to E$ prove that

$$\mathscr{A}(\psi \circ \varphi) = \mathscr{A}\varphi \circ \mathscr{A}\psi$$

where $\mathscr{A}$ is the operator defined in sec. 7.2.

5. *The dual product.* Consider a fixed tensor $\Phi \in T^p_q(E)$. This tensor defines a linear mapping

$$A_\Phi: T^{k-p}_{l-q}(E) \to T^k_l(E) \quad (k \geqq p, l \geqq q)$$

by

$$A_\Phi \Psi = \Phi \otimes \Psi .$$

Define the dual product $X \llcorner \Phi$ by

$$X \llcorner \Phi = A^*_\Phi X \qquad\qquad X \in T^l_k(E)$$

where A^*_Φ is the dual operator of A_Φ. Prove the following properties of the dual product:

1. $\langle X, \Phi \otimes \Psi \rangle = \langle X \llcorner \Phi, \Psi \rangle \qquad X \in T^l_k(E), \Phi \in T^p_q(E), \Psi \in T^{k-p}_{l-q}(E)$.
2. $X \llcorner (\Phi_1 \otimes \Phi_2) = (X \llcorner \Phi_1) \llcorner \Phi_2 \qquad X \in T^l_k(E), \Phi_1 \in T^p_q(E), \Phi_2 \in T^r_s(E)$. $(l \geqq q + s, k \geqq p + r)$.
3. $X \llcorner \Phi = \langle X, \Phi \rangle \quad X \in T^q_p(E), \quad \Phi \in T^p_q(E)$.
4. $X \llcorner 1 = X \quad X \in T^l_k(E)$.

6. Let φ be an endomorphism of E and

$$\theta^p(\varphi): T^p(E) \to T^p(E) \text{ and } \theta_p(\varphi): T_p(E) \to T_p(E)$$

the endomorphisms defined in § 1, prob. 10. Prove that

$$\langle \theta^p(\varphi)\Phi, \Psi \rangle = -\langle \Phi, \theta_p(\varphi)\Psi \rangle \qquad \Phi \in T^p(E), \Psi \in T_p(E) .$$

Chapter VIII

Exterior Algebra

§ 1. Covariant skew-symmetric tensors

8.1. Definition. Consider a covariant tensor Φ of order $p \geqq 1$. For every permutation σ of the numbers $(1 \ldots p)$ define the tensor $\sigma\Phi$ by

$$\sigma\Phi\,(x_1 \ldots x_p) = \Phi\,(x_{\sigma^{-1}(1)} \ldots x_{\sigma^{-1}(p)})\,.$$

The following formulas are immediate consequences of the above definition:

$$\sigma\,(\lambda\Phi_1 + \mu\Phi_2) = \lambda\,\sigma\Phi_1 + \mu\,\sigma\Phi_2$$

$$(\tau\sigma)\,\Phi = \tau\,(\sigma\Phi)$$

$$\iota\Phi = \Phi \quad (\iota \text{ identity permutation}).$$

A tensor Φ is called *skew-symmetric* if $\sigma\Phi = \varepsilon_\sigma\Phi$ for every permutation σ. A determinant-function, for instance, is a skew-symmetric tensor of order n. If Φ_1 and Φ_2 are skew-symmetric tensors of the same order, then every linear combination $\lambda\Phi_1 + \mu\Phi_2$ is again skew-symmetric. Thus, the set of all skew-symmetric tensors of order p forms a subspace of the space $T_p(E)$. This subspace will be denoted by $S_p(E)$. If $p > n$, the subspace $S_p(E)$ reduces to the zero-tensor.

8.2. The antisymmetry-operator. If Φ is an arbitrary tensor of order p we define the skew-symmetric tensor $A\Phi$ by

$$A\Phi = \frac{1}{p!}\sum_\sigma \varepsilon_\sigma\,\sigma\Phi \tag{8.1}$$

where σ runs over all permutations of the numbers $(1 \ldots p)$. The tensor $A\Phi$ is indeed skew-symmetric. To prove this, consider a fixed permutation τ. Then

$$\tau\,(A\Phi) = \frac{1}{p!}\sum_\sigma \varepsilon_\sigma\,\tau\,(\sigma\Phi) = \frac{1}{p!}\sum_\sigma \varepsilon_\sigma\,(\tau\sigma)\,\Phi = \frac{1}{p!}\,\varepsilon_\tau\sum_\sigma \varepsilon_{\tau\sigma}\,(\tau\sigma)\,\Phi\,. \tag{8.2}$$

Introducing the permutation $\varrho = \tau\sigma$ as a new "index of summation" we obtain from (8.2) the equation

$$\tau\,(A\Phi) = \frac{1}{p!}\,\varepsilon_\tau\sum_\varrho \varepsilon_\varrho\,\varrho\Phi = \varepsilon_\tau A\Phi$$

showing that $A\Phi$ is skew-symmetric. The tensor $A\Phi$ is called the *skew-symmetric* part of Φ.

Replacing Φ by $\tau\Phi$ in the equation (8.1), where τ is a fixed permutation, we obtain

$$A\,(\tau\Phi) = \frac{1}{p!}\sum_\sigma \varepsilon_\sigma\,\sigma\tau\Phi = \frac{1}{p!}\,\varepsilon_\tau\sum_\sigma \varepsilon_{\sigma\tau}\,\sigma\tau\Phi = \frac{1}{p!}\,\varepsilon_\tau\sum_\varrho \varepsilon_\varrho\,\varrho\Phi = \varepsilon_\tau\,A\Phi\,.$$

This yields the relation

$$A\,(\tau\Phi) = \varepsilon_\tau\, A\Phi \tag{8.3}$$

for every permutation τ. Multiplying (8.3) by ε_τ and summing over all permutations τ we find

$$A^2\Phi = A\Phi\,.$$

Written as an operator-identity this formula reads

$$A^2 = A$$

showing that A reduces to the identity in the subspace $S_p(E)$. In other words, A in a *projection-operator*.

8.3. The kernel of the operator A. Denote by $K_p(E)$ the kernel of the linear mapping

$$A: T_p(E) \to S_p(E)\,.$$

This kernel contains every tensor which is symmetric with respect to at least two arguments. In fact, assume that Φ is symmetric relative to the i^{th} and the j^{th} argument. Then $\tau\Phi = \Phi$, where τ is the transposition $i \leftrightarrow j$. Since $\varepsilon_\tau = -1$, the equation (8.3) yields $A\Phi = -A\Phi$, whence $A\Phi = 0$.

Denote by $Q_p(E)$ the subspace generated by all tensors which are symmetric with respect to at least two arguments. Then $Q_p(E)$ is contained in $K_p(E)$. It will be shown that $Q_p(E) = K_p(E)$. The equation (8.1) can be written as

$$\Phi - A\Phi = \frac{1}{p!}\sum_\sigma (\Phi - \varepsilon_\sigma\,\sigma\Phi)\,.$$

Now assume that Φ is contained in $K_p(E)$. Then $A\Phi = 0$, whence

$$\Phi = \frac{1}{p!}\sum_\sigma (\Phi - \varepsilon_\sigma\,\sigma\Phi)\,.$$

Thus, it is sufficient to show that

$$\Phi - \varepsilon_\sigma\,\sigma\Phi \in Q_p(E) \tag{8.4}$$

for every tensor $\Phi \in T_p(E)$ and every permutation σ.

Every permutation σ can be written as a product of N transpositions. If $N = 1$, σ is a transposition itself, $\sigma: i \leftrightarrow j$. Then

$$\sigma\Phi\,(x_1 \dots x_i \dots x_j \dots x_p) = \Phi\,(x_1 \dots x_j \dots x_i \dots x_p)$$

whence

$$(\Phi - \varepsilon_\sigma\,\sigma\Phi)\,(x_1 \dots x_p) = \Phi\,(x_1 \dots x_i \dots x_j \dots x_p) + \\ + \Phi\,(x_1 \dots x_j \dots x_i \dots x_p) = 0\,.$$

This equation shows that the tensor $\Phi - \varepsilon_\sigma\,\sigma\Phi$ is symmetric with respect to the i^{th} and the j^{th} argument and hence it is contained in $Q_p(E)$.

Now assume, by induction, that (8.4) is correct for a permutation σ which is the product of N transpositions. Consider the permutation $\varrho = \tau\sigma$, where τ is an arbitrary transposition. Then, by induction

$$\Phi - \varepsilon_\sigma \sigma\Phi \in Q_p(E) . \tag{8.5}$$

Furthermore,

$$\Phi + \tau\Phi \in Q_p(E) . \tag{8.6}$$

The equation (8.5) yields

$$\tau\Phi - \varepsilon_\sigma \tau \sigma\Phi \in Q_p(E)$$

which can be written as

$$\tau\Phi + \varepsilon_\varrho \varrho\Phi \in Q_p(E) . \tag{8.7}$$

From (8.6) and (8.7) we obtain

$$\Phi - \varepsilon_\varrho \varrho \Phi \in Q_p(E) .$$

Thus, the induction is closed.

8.4. The skew-symmetric part of a product. The skew-symmetric part of the product $\Phi \otimes \Psi$ ($\Phi \in T_p(E), \Psi \in T_q(E)$) can be written as

$$A\,(\Phi \otimes \Psi) = A\,(\Phi \otimes A\Psi) = A\,(A\Phi \otimes \Psi) . \tag{8.8}$$

To prove the formula (8.8), consider a fixed permuatation τ of the numbers $(p + 1 \ldots p + q)$. Defining the permutation τ' of the numbers $(1 \ldots p + q)$ by

$$\tau'(\nu) = \nu \ (\nu = 1 \ldots p) \quad \text{and} \quad \tau'(\nu) = \tau(\nu) \ (\nu = p + 1 \ldots p + q) ,$$

we can write the product $\Phi \otimes \tau\Psi$ as

$$\Phi \otimes \tau\Psi = \tau'\,(\Phi \otimes \Psi) .$$

Applying the operator A and using (8.3) we obtain from this equation

$$A\,(\Phi \otimes \tau\Psi) = \varepsilon_{\tau'}\, A\,(\Phi \otimes \Psi) .$$

Since τ' has the same parity as τ, this can be written as

$$A\,(\Phi \otimes \varepsilon_\tau \tau\Psi) = A\,(\Phi \otimes \Psi) .$$

Summation over all permutations τ of $(1 \ldots p + q)$ and division by $q!$ yields

$$A\,(\Phi \otimes A\Psi) = A\,(\Phi \otimes \Psi) .$$

The second relation (8.8) is proved in the same way. Combining the two formulas (8.8) we obtain

$$A\,(\Phi \otimes \Psi) = A\,(A\Phi \otimes A\Psi) \quad \Phi \in T_p(E), \Psi \in T_q(E). \tag{8.9}$$

8.5. The skew-symmetric product. Given two skew-symmetric tensors Φ and Ψ, the product $\Phi \otimes \Psi$ is not skew-symmetric generally. Consider for instance two tensors Φ and Ψ of order one. Then

$$(\Phi \otimes \Psi)\,(x_1, x_2) = \Phi(x_1)\,\Psi(x_2)$$

and

$$(\Phi \otimes \Psi)\,(x_2, x_1) = \Phi(x_2)\,\Psi(x_1)\,.$$

To obtain a multiplication yielding a skew-symmetric tensor, we define the *skew-symmetric product* $\Phi \wedge \Psi$ of two tensors $\Phi \in S_p(E)$ and $\Psi \in S_q(E)$ by

$$\Phi \wedge \Psi = \binom{p+q}{p} A\,(\Phi \otimes \Psi)\,. \tag{8.10}$$

Written in a less condensed form, this formula reads

$$\begin{aligned}&(\Phi \wedge \Psi)\,(x_1 \ldots x_{p+q})\\ &= \frac{1}{p!\,q!} \sum_{\varrho} \varepsilon_{\varrho}\, \Phi\,(x_{\varrho(1)} \ldots x_{\varrho(p)})\, \Psi\,(x_{\varrho(p+1)} \ldots x_{\varrho(p+q)})\end{aligned} \tag{8.11}$$

where the sum is extended over all permutations of the numbers $(1 \ldots p+q)$. For instance, if $p = 1$ and $q = 1$,

$$(\Phi \wedge \Psi)\,(x_1, x_2) = \Phi(x_1)\,\Psi(x_2) - \Phi(x_2)\,\Psi(x_1)$$

and if $p = 1$ and $q = 2$,

$$(\Phi \wedge \Psi)\,(x_1, x_2, x_3) = \Phi(x_1)\,\Psi\,(x_2, x_3) + \Phi(x_2)\,\Psi(x_3, x_1) + \Phi(x_3)\,\Psi(x_1, x_2).$$

Since every skew-symmetric tensor of order $r > n$ is the zero-tensor, it follows that

$$\Phi \wedge \Psi = 0 \quad \text{if} \quad p + q > n \quad \Phi \in S_p(E), \Psi \in S_q(E)\,.$$

The distributivity of the product $\Phi \otimes \Psi$ and the linearity of the operator A imply that the skew-symmetric product is distributive:

$$(\lambda \Phi_1 + \mu \Phi_2) \wedge \Psi = \lambda \Phi_1 \wedge \Psi + \mu \Phi_2 \wedge \Psi$$

$$\Phi \wedge (\lambda \Psi_1 + \mu \Psi_2) = \lambda \Phi \wedge \Psi_1 + \mu \Phi \wedge \Psi_2\,.$$

The skew-symmetric product is also associative,

$$(\Phi \wedge \Psi) \wedge X = \Phi \wedge (\Psi \wedge X) \quad \Phi \in S_p(E), \Psi \in S_q(E), X \in S_r(E)\,. \tag{8.12}$$

To prove the formula (8.12), we note first that the relation (8.9) can be written in the form

$$A\Phi \wedge A\Psi = \binom{p+q}{p} A(\Phi \otimes \Psi) \quad \Phi \in T_p(E), \Psi \in T_q(E)\,.$$

We thus obtain

$$(\Phi \wedge \Psi) \wedge X = (A\Phi \wedge A\Psi) \wedge A X = \binom{p+q}{p} A\,(\Phi \otimes \Psi) \wedge A X$$
$$= \frac{(p+q+r)!}{p!\,q!\,r!} A\,(\Phi \otimes \Psi \otimes X)$$

and

$$\Phi \wedge (\Psi \wedge X) = A\Phi \wedge (A\Psi \wedge A X) = \binom{q+r}{q} A\Phi \wedge A\,(\Psi \otimes X)$$
$$= \frac{(p+q+r)!}{p!\,q!\,r!} A\,(\Phi \otimes \Psi \otimes X)$$

whence (8.12).

Finally, it will be shown that the commutative law holds in the following form:

$$\Phi \wedge \Psi = (-1)^{pq}\, \Psi \wedge \Phi \quad \Phi \in S_p(E), \Psi \in S_q(E)\,. \tag{8.13}$$

If $\Phi \in T_p(E)$ and $\Psi \in T_q(E)$ are two arbitrary tensors, the products $\Phi \otimes \Psi$ and $\Psi \otimes \Phi$ are related by

$$\Psi \otimes \Phi = \varrho\,(\Phi \otimes \Psi) \tag{8.14}$$

where ϱ is the permutation

$$\varrho\colon (1 \ldots q, q+1 \ldots q+p) \to (p+1 \ldots p+q, 1 \ldots p)\,.$$

Applying the anti-symmetry-operator to (8.14) and using (8.3) we obtain

$$A\,(\Psi \otimes \Phi) = \varepsilon_\varrho A\,(\Phi \otimes \Psi) = (-1)^{pq} A\,(\Phi \otimes \Psi)$$

whence (8.13).

Substituting $\Psi = \Phi$ in the commutative law we find the relation

$$\Phi \wedge \Phi = (-1)^{p^2}\, \Phi \wedge \Phi = (-1)^p\, \Phi \wedge \Phi\,.$$

showing that

$$\Phi \wedge \Phi = 0 \quad \text{if } p \text{ is odd.}$$

8.6. The skew-symmetric product of vectors. Let $\overset{*}{a}{}^\nu$ $(\nu = 1 \ldots p)$ be p given vectors of E^*. If these vectors are considered as covariant tensors of order one, we can form their skew-symmetric product. This yields a skew-symmetric tensor

$$\Phi = \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p$$

of order p. Using (8.11) we obtain the explicit formula

$$\Phi\,(x_1, \ldots x_p) = \sum_\sigma \varepsilon_\sigma \langle \overset{*}{a}{}^1, x_{\sigma^{-1}(1)} \rangle \cdots \langle \overset{*}{a}{}^p, x_{\sigma^{-1}(p)} \rangle$$

whence

$$\Phi(x_1, \ldots x_p) = \begin{vmatrix} \langle \overset{*}{a}{}^1, x_1 \rangle & \cdots & \langle \overset{*}{a}{}^p, x_1 \rangle \\ \vdots & & \vdots \\ \langle \overset{*}{a}{}^1, x_p \rangle & \cdots & \langle \overset{*}{a}{}^p, x_p \rangle \end{vmatrix}\,. \tag{8.15}$$

The skew-symmetric product of the vectors $\overset{*}{a}{}^\nu$ ($\nu = 1 \ldots p$) can also be written in the form

$$\overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p = \sum_\sigma \varepsilon_\sigma \overset{*}{a}{}^{\sigma(1)} \otimes \cdots \otimes \overset{*}{a}{}^{\sigma(p)} . \tag{8.16}$$

In fact, expanding the determinant (8.15) by its rows we obtain

$$\Phi(x_1, \ldots x_p) = \sum_\sigma \varepsilon_\sigma \langle \overset{*}{a}{}^{\sigma(1)}, x_1 \rangle \cdots \langle \overset{*}{a}{}^{\sigma(p)}, x_p \rangle$$

whence

$$\Phi = \sum_\sigma \varepsilon_\sigma \overset{*}{a}{}^{\sigma(1)} \otimes \cdots \otimes \overset{*}{a}{}^{\sigma(p)} .$$

8.7. Basis of $S_p(E)$. Consider a pair of dual bases $\overset{*}{e}{}^\nu$, e_ν of E^* and E. It will be shown that the $\binom{n}{p}$ tensors

$$\overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p} \qquad (\nu_1 < \cdots < \nu_p) \tag{8.17}$$

form a basis of $S_p(E)$. First of all, the tensors (8.17) are linearly independent. In fact, assume a relation

$$\sum_< \lambda_{\nu_1 \ldots \nu_p} \overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p} = 0 \tag{8.18}$$

where the symbol $<$ indicates that the indices $(\nu_1 \ldots \nu_p)$ are subject to the condition $\nu_1 < \cdots < \nu_p$. The equation (8.18) implies that

$$\sum_< \lambda_{\nu_1 \ldots \nu_p} \begin{vmatrix} \langle \overset{*}{e}{}^{\nu_1}, x_1 \rangle & \cdots & \langle \overset{*}{e}{}^{\nu_p}, x_1 \rangle \\ \vdots & & \vdots \\ \langle \overset{*}{e}{}^{\nu_1}, x_p \rangle & \cdots & \langle \overset{*}{e}{}^{\nu_p}, x_p \rangle \end{vmatrix} = 0 \tag{8.19}$$

for all vectors $x_\nu \in E$. Now choose a fixed system of indices $\mu_1 < \cdots < \mu_p$ and substitute $x_i = e_{\mu_i}$ ($i = 1 \ldots p$) in (8.19). This yields

$$\sum_< \lambda_{\nu_1 \ldots \nu_p} \begin{vmatrix} \delta_{\mu_1}^{\nu_1} & \cdots & \delta_{\mu_1}^{\nu_p} \\ \vdots & & \vdots \\ \delta_{\mu_p}^{\nu_1} & \cdots & \delta_{\mu_p}^{\nu_p} \end{vmatrix} = 0 . \tag{8.20}$$

The above determinant is equal to one if $\nu_i = \mu_i$ ($i = 1 \ldots p$) and equal to zero othencise. Hence (8.19) implies that

$$\lambda_{\mu_1 \ldots \mu_p} = 0 \quad (\mu_1 < \cdots < \mu_p)$$

showing that the tensors (8.17) are linearly independent.

To prove that the tensors (8.17) generate the whole space $S_p(E)$, consider an arbitrary tensors $\Phi \in S_p(E)$. This tensor can be written as

$$\Phi = \sum_{(\nu)} \Phi_{\nu_1 \ldots \nu_p} \overset{*}{e}{}^{\nu_1} \otimes \cdots \otimes \overset{*}{e}{}^{\nu_p} \tag{8.21}$$

where the numbers $\Phi_{\nu_1 \ldots \nu_p}$ are defined by

$$\Phi_{\nu_1 \ldots \nu_p} = \Phi(e_{\nu_1} \ldots e_{\nu_p}) . \tag{8.22}$$

Applying the antisymmetry-operator to (8.21) and observing that $A\Phi = \Phi$, we obtain

$$\Phi = \frac{1}{p!}\sum_{(\nu)} \Phi_{\nu_1 \ldots \nu_p} \overset{*}{e}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}^{\nu_p} . \tag{8.23}$$

It follows from the skew symmetry of Φ that the components (8.22) are skew-symmetric with respect to all indices. The same is true for the products $e_{\nu_1} \wedge \cdots \wedge e_{\nu_p}$. Hence the equation (8.23) can be written as

$$\Phi = \sum_{<} \Phi_{\nu_1 \ldots \nu_p} \overset{*}{e}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}^{\nu_p}$$

showing that the tensors (8.17) generate the space $S_p(E)$. Hence they form a basis of $S_p(E)$. This implies that

$$\dim S_p(E) = \binom{n}{p} . \tag{8.24}$$

8.8. Skew-symmetric tensors and linear mappings. It has been shown in sec. 7.8 that a linear mapping $\vartheta: E \to F$ induces a linear mapping

$$(\vartheta)_p : T_p(F) \to T_p(E) \tag{8.25}$$

defined by

$$(\vartheta)_p \Psi(x_1 \ldots x_p) = \Psi(\vartheta x_1 \ldots \vartheta x_p) .$$

This equation shows that

$$(\vartheta)_p (\sigma\Psi) = \sigma(\vartheta)_p \Psi \tag{8.26}$$

for every permutation σ of $(1 \ldots p)$. Multiplying the equation (8.26) by ε_σ and summing over all permutations σ we obtain the formula

$$(\vartheta)_p A\Psi = A(\vartheta)_p \Psi \qquad \Psi \in T_p(F) \tag{8.27}$$

showing that the operator $(\vartheta)_p$ commutes with the antisymmetry-operator. It follows from (8.27) that the mapping $(\vartheta)_p$ transforms skew-symmetric tensors into skew-symmetric tensors and hence it induces a linear mapping

$$(\vartheta)_p : S_p(F) \to S_p(E) .$$

This mapping preserves the skew-symmetric product,

$$(\vartheta)_{p+q}(\Psi_1 \wedge \Psi_2) = (\vartheta)_p \Psi_1 \wedge (\vartheta)_q \Psi_2 . \tag{8.28}$$

In fact,

$$\begin{aligned}(\vartheta)_{p+q}(\Psi_1 \wedge \Psi_2) &= \binom{p+q}{p} (\vartheta)_{p+q} A(\Psi_1 \otimes \Psi_2) \\ &= \binom{p+q}{p} A(\vartheta)_{p+q}(\Psi_1 \otimes \Psi_2) \\ &= \binom{p+q}{p} A((\vartheta)_p \Psi_1 \otimes (\vartheta)_q \Psi_2) \\ &= (\vartheta)_p \Psi_1 \wedge (\vartheta)_q \Psi_2 .\end{aligned}$$

Applying the formula (8.28) to p covariant tensors of order 1 we obtain the relation

$$(\vartheta)_p\,(\overset{*}{y}^1 \wedge \dots \overset{*}{y}^p) = \overset{*}{\vartheta}\overset{*}{y}^1 \wedge \cdots \wedge \overset{*}{\vartheta}\overset{*}{y}^p \quad \overset{*}{y}^\nu \in F^* \; (\nu = 1 \dots p)\,. \tag{8.29}$$

Note: All definitions and formulas of this paragraph can be carried over in an obvious way to contravariant tensors.

Problems: 1. Assume a linear mapping

$$F\colon T_p(E) \to S_p(E)$$

having the following properties:

1. $F\Phi = \Phi$ if $\Phi \in S_p(E)$.

2. $F(\sigma\Phi) = \varepsilon_\sigma F\Phi$ for every permutation σ.

Prove that F is the antisymmetry-operator.

2. Consider the subspace $S_p(E) \otimes S_q(E)$ of $T_{p+q}(E)$ generated by all products $\Phi \otimes \Psi$ ($\Phi \in S_p(E), \Psi \in S_q(E)$). Show that $S_p(E) \otimes S_q(E)$ consists of all covariant tensors which are skew-symmetric with respect to the first p arguments and the last q arguments.

3. Denote by $S(E)$ the Cartesian product of all spaces $S_\nu(E)$ ($\nu = 1 \dots n$). Define a multiplication in $S(E)$ as follows: The product of two n-tuples

$$(\Phi) = (\Phi_0, \Phi_1 \dots \Phi_n) \quad \Phi_\nu \in S_\nu(E)$$

and $\qquad\qquad (\nu = 1 \dots n)$

$$(\Psi) = (\Psi_0, \Psi_1 \dots \Psi_n) \quad \Psi_\nu \in S_\nu(E)$$

is the n-tuple

$$(X) = (X_0, X_1 \dots X_n)$$

where

$$X_\lambda = \sum_{\nu+\mu=\lambda} \Phi_\nu \wedge \Psi_\mu\,.$$

a) Prove that the space $S(E)$ becomes an algebra under this multiplication and that the scalar 1 is the unit-element of this algebra.

b) Show that

$$\dim S(E) = 2^n.$$

4. Let E and F be two linear spaces of finite dimension and consider the projections

$$\pi_1\colon (x, y) \to x \quad \text{and} \quad \pi_2\colon (x, y) \to y$$

of $E \times F$ onto E and F respectively. Given two tensors $\Phi \in S_p(E)$ and $\Psi \in S_q(F)$ define the tensor $\Phi \times \Psi \in S_{p+q}(E \times F)$ by

$$\Phi \times \Psi = \pi_1^* \Phi \wedge \pi_2^* \Psi\,.$$

a) Prove the formulas

$$\Phi \times \Psi = (-1)^{pq}\, \Psi \times \Phi$$

and

$$\Phi \times 1 = \pi_1^* \Phi \quad 1 \times \Psi = \pi_2^* \Psi .$$

b) Denote by $S_p(E) \times S_q(E)$ the subspace of $S_{p+q}(E \times F)$ which is generated by the products $\Phi \times \Psi (\Phi \in S_p(E),\ \Psi \in S_q(E))$. Prove the relation

$$S_r(E \times F) = \sum_{p+q=r} S_p(E) \times S_q(F) .$$

c) Consider the bilinear mapping

$$(S(E), S(F)) \to S\,(E \times F)$$

defined by

$$(\Phi_0 \ldots \Phi_n) \times (\Psi_0 \ldots \Psi_m) = (X_0 \ldots X_{n+m})$$

where

$$X_\lambda = \sum_{\nu+\mu=\lambda} \Phi_\nu \times \Psi_\mu \quad (\lambda = 0 \ldots n+m) .$$

Prove that this bilinear mapping is a tensor-product.

d) Put $F = E$ and define the linear mapping $\Delta: E \to E \times E$ by

$$\Delta x = (x, x) .$$

Show that

$$\Delta_{p+q}(\Phi \times \Psi) = \Phi \wedge \Psi \qquad \Phi \in S_p(E), \Psi \in S_q(E) .$$

5. Prove that the rank of the linear mapping

$$(\vartheta)_p : S_p(F) \to S_p(E)$$

(see sec. 8.8) is equal to $\binom{r}{p}$ where r denotes the rank of ϑ.

6. Assume that a system of linear mappings

$$d_p : T_p(E) \to T_{p+1}(E) \quad (p = 0, 1 \ldots)$$

is given such that

$$d_{p+q}(\Phi \otimes \Psi) = d_p \Phi \otimes \Psi + \tau\,(\Phi \otimes d_q \Psi) \quad \Phi \in T_p(E), \quad \Psi \in T_q(E)$$

where τ is the permutation defined by

$$\tau: (1 \ldots p, p+1, p+2 \ldots p+q) \to (p+1, 1 \ldots p, p+2 \ldots p+q) .$$

Define a new system of linear mappings

$$\delta_p : S_p(E) \to S_{p+1}(E) \quad (p = 0, 1 \ldots n)$$

by

$$\delta_p = (p+1)\, A \circ d_p .$$

Prove that

$$\delta_{p+q}(\Phi \wedge \Psi) = \delta_p \Phi \wedge \Psi + (-1)^p \Phi \wedge \delta_q \Psi \qquad \begin{matrix} \Phi \in S_p(E) \\ \Psi \in S_q(E) . \end{matrix}$$

7. Assume that the operators d_p in prob. 6 have the following additional properties:

1. $d_p K_p(E) \subset K_{p+1}(E)$ where $K_p(E)$ denotes the kernel of the antisymmetry-operator.

2. $d_{r+1} d_p T_p(E) \subset K_{p+2}(E)$.

Prove that then $\delta_{p+1} \delta_p = 0$.

8. Given a fixed vector $a \in E$ define the *substitution-operator*

$$i(a): S_p(E) \to S_{p-1}(E)$$

as follows:

$$\begin{aligned} i(a)\Phi(x_1 \ldots x_{p-1}) &= \Phi(a, x_1 \ldots x_{p-1}) \quad \text{if} \quad p \geqq 1 \\ i(a)\Phi &= 0 \qquad\qquad\qquad\qquad\quad \text{if} \quad p = 0 \end{aligned}$$

Prove the formula

$$i(a)(\Phi \wedge \Psi) = i(a)\Phi \wedge \Psi + (-1)^p \Phi \wedge i(a)\Psi.$$

9. A tensor $\Phi \in T_p(E)$ is called *symmetric* if $\sigma\Phi = \Phi$ for every permutation σ. Denote by $Y_p(E)$ the space of all symmetric tensors of order p.

1. Prove that $\dim Y_p(E) = \binom{n+p-1}{p}$.

2. Define the product of two symmetric tensors

$$\Phi \in Y_p(E) \text{ and } \Psi \in Y_q(E) \text{ by } (\Phi \circ \Psi)(x_1 \ldots x_{p+q})$$
$$= \frac{1}{p!\,q!} \sum_\sigma \Phi(x_{\sigma(1)} \ldots x_{\sigma(p)}) \Psi(x_{\sigma(p+1)} \ldots x_{\sigma(p+q)})$$

Prove that $\Phi \circ \Psi = \Psi \circ \Phi$.

10. Let E be a linear space of dimension $n = 2m$ and $\Phi_\mu \in S_2(E)$ be m given tensors. Define the substitution-operator $i(a)$ as in prob. 8. Prove that for every two vectors $\overset{*}{a} \in E^*$ and $a \in E$

$$\sum_{\mu=1}^{m} \overset{*}{a} \wedge i(a)\Phi_\mu \wedge \Phi_1 \wedge \ldots \wedge \hat{\Phi}_\mu \ldots \wedge \Phi_m = \langle \overset{*}{a}, a \rangle \Phi_1 \wedge \ldots \wedge \Phi_m.$$

§ 2. Decomposable skew-symmetric tensors

8.9. **Definition.** A skew-symmetric tensor $\Phi \in S_p(E)$ is called *decomposable* if it can be written as the skew-symmetric product of p vectors $\overset{*}{a}{}^\nu \in E^*$. To obtain a criterion for the decomposability, associate with every tensor $\Phi \in S_p(E)$ a subspace $U(\Phi) \subset E$ in the following way: A vector $u \in E$ is contained in $U(\Phi)$ if and only if

$$\Phi(u, x_2 \ldots x_p) = 0 \quad \text{for all vectors } x_\nu \in E \ (\nu = 2 \ldots p).$$

If $\Phi = 0$, the space $U(\Phi)$ coincides with E. It will be shown that

$$\dim U(\Phi) \leqq n - p \quad \text{if} \quad \Phi \neq 0. \tag{8.30}$$

Choose p vectors $a_\nu \in E$ such that

$$\Phi(a_1 \dots a_p) \neq 0 .$$

These vectors are linearly independent and thus they generate a p-dimensional subspace E_1 of E. Consider the intersection $E_1 \cap U(\Phi)$. Then

$$\Phi(a, x_2 \dots x_p) = 0 \quad \text{for} \quad a \in E_1 \cap U(\Phi) \quad \text{and} \quad x_\nu \in E \; (\nu = 2 \dots p)$$

whence

$$\Phi(a, a_1 \dots \hat{a}_i \dots a_p) = 0 \quad (i = 1 \dots p)^{*)}. \tag{8.31}$$

At the same time the vector a can be written as

$$a = \sum_\nu \lambda^\nu a_\nu . \tag{8.32}$$

The equations (8.31) and (8.32) and the skew symmetry of Φ imply that

$$\lambda^i \Phi(a_1 \dots a_p) = 0 \qquad (i = 1 \dots p)$$

whence $\lambda^i = 0$ $(i = 1 \dots p)$. Consequently,

$$E_1 \cap U(\Phi) = 0 .$$

This implies the relation (8.30).

Now the following criterion can be proved: *A skew-symmetric tensor $\Phi \neq 0$ of order p is decomposable if and only if*

$$\dim U(\Phi) = n - p . \tag{8.33}$$

Assume that Φ is decomposable,

$$\Phi = \overset{*}{a}{}^1 \wedge \dots \wedge \overset{*}{a}{}^p .$$

Denote by U^* the subspace of E^* which is generated by the vectors $\overset{*}{a}{}^\nu$ $(\nu = 1 \dots p)$. The equation (8.15) shows that

$$\Phi(u, x_2 \dots x_p) = 0$$

for every vector $u \in (U^*)^\perp$ and all vectors $x_\nu \in E$ $(\nu = 1 \dots p)$. This implies that

$$(U^*)^\perp \subset U(\Phi) . \tag{8.34}$$

At the same time we know from (8.30) that

$$\dim U(\Phi) \leqq n - p = \dim (U^*)^\perp . \tag{8.35}$$

The relations (8.34) and (8.35) imply that

$$(U^*)^\perp = U(\Phi) \tag{8.36}$$

whence (8.33).

Conversely, assume that the equation (8.33) is valid. Then $U(\Phi)^\perp$ is a p-dimensional subspace of E^*. Let $\overset{*}{a}{}^\nu$ $(\nu = 1 \dots p)$ be a basis of $U(\Phi)^\perp$.

---

*) The symbol $\hat{a}_i$ indicates that the i-th argument is deleted.

Define the tensor $\Psi \in S_p(E)$ by

$$\Psi = \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p .$$

Then

$$U(\Psi) = U(\Phi)^{\perp\perp} = U(\Phi) . \tag{8.37}$$

It will be shown that

$$\Phi = \lambda \Psi$$

where λ is a scalar. Let E_1 be a subspace of E such that

$$E = E_1 \otimes U(\Phi)$$

and denote by Φ_1 and by Ψ_1 the restrictions of Φ and Ψ to the subspace E_1. Then Φ_1 and Ψ_1 are skew-symmetric tensors of order p in the p-dimensional space E_1. This implies that

$$\Phi_1 = \lambda \Psi_1 \tag{8.38}$$

where λ is a scalar.

Now consider p arbitrary vectors $x_\nu \in E$. Every vector x_ν can be written as

$$x_\nu = y_\nu + z_\nu \quad \text{with} \quad y_\nu \in E_1 \quad \text{and} \quad z_\nu \in U(\Phi) .$$

Since $z_\nu \in U(\Phi)$, it follows that

$$\Phi(x_1 \ldots x_p) = \Phi(y_1 + z_1 \ldots y_p + z_p) = \Phi(y_1 \ldots y_p) = \Phi_1(y_1 \ldots y_p). \tag{8.39}$$

Using the relation (8.37) we obtain in the same way

$$\Psi(x_1 \ldots x_p) = \Psi_1(y_1 \ldots y_p) . \tag{8.40}$$

The equations (8.38), (8.39) and (8.40) imply that

$$\Phi = \lambda \Psi = \lambda \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p$$

showing that Φ is decomposable.

8.10. Uniqueness of the product-representation. Assume that a tensor $\Phi \neq 0$ is represented in two different ways as a skew-symmetric product

$$\Phi = \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p = \overset{*}{b}{}^1 \wedge \cdots \wedge \overset{*}{b}{}^p . \tag{8.41}$$

Denote by U^* and V^* the subspaces of E^* generated by the vectors $\overset{*}{a}{}^\nu$ and $\overset{*}{b}{}^\nu$ $(\nu = 1 \ldots p)$ respectively. Then, by (8.36),

$$U^* = U(\Phi)^\perp \quad \text{and} \quad V^* = U(\Phi)^\perp ,$$

whence $V^* = U^*$. Thus, every vector $\overset{*}{b}{}^\nu$ can be written as

$$\overset{*}{b}{}^\nu = \sum_\mu \beta^\nu_\mu \overset{*}{a}{}^\mu .$$

This implies that

$$\overset{*}{b}{}^1 \wedge \cdots \wedge \overset{*}{b}{}^p = \begin{vmatrix} \beta_1^1 \cdots \beta_1^p \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \beta_p^1 \cdots \beta_p^p \end{vmatrix} \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p . \tag{8.42}$$

The equations (8.41) and (8.42) yield

$$\begin{vmatrix} \beta_1^1 \cdots \beta_1^p \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \beta_p^1 \cdots \beta_p^p \end{vmatrix} = 1 .$$

Thus, the vectors $\overset{*}{a}{}^\nu$ and $\overset{*}{b}{}^\nu$ are related by a $p \times p$-matrix whose determinant is equal to one.

Problems: 1. Prove that a skew-symmetric tensor $\Phi \neq 0$ of order two is decomposable if and only if the matrix of its components has the rank one.

2. Show that a skew-symmetric tensor Φ of order two is decomposable if and only if

$$\Phi \wedge \Phi = 0 .$$

Hint: To prove that the above condition is sufficient, we can assume that $\Phi \neq 0$. Choose two vectors $a \in E$ and $b \in E$ such that $\Phi(a, b) \neq 0$, and define the vectors $\overset{*}{a} \in E^*$ and $\overset{*}{b} \in E^*$ by the equations

$$\Phi(a, y) = \langle \overset{*}{a}, y \rangle \quad \text{and} \quad \Phi(b, y) = \langle \overset{*}{b}, y \rangle \quad \text{for all } y \in E .$$

Then show that

$$\Phi = \frac{\overset{*}{a} \wedge \overset{*}{b}}{\Phi(a, b)} .$$

3. Let $\Delta \neq 0$ be a determinant-function in the n-dimensional linear space E. Show that every tensor $\Phi \in S_{n-1}(E)$ can be written as

$$\Phi(x_1 \ldots x_{n-1}) = \Delta(a, x_1 \ldots x_{n-1}) \quad a \in E .$$

Employing this result prove that every tensor $\Phi \in S_{n-1}(E)$ is decomposable.

Hint: Consider the n-linear mapping φ of E into itself, which is defined by

$$\varphi(x_1 \ldots x_n) = \sum_\nu (-1)^{n-\nu} \Phi(x_1 \ldots \hat{x}_\nu \ldots x_n) x_\nu$$

and represent φ in the form

$$\varphi(x_1 \ldots x_n) = a\Delta(x_1 \ldots x_n)$$

(cf. Chap. VI § 1 problem 3).

§ 3. Mixed skew-symmetric tensors

8.11. Definition. Consider a tensor Φ of order (p, q) where $p \geqq 1$ and $q \geqq 1$. If σ is a permutation of the numbers $(1 \ldots p)$, and τ is a permutation of the numbers $(1 \ldots q)$, we define the tensor $(\sigma, \tau)\Phi$ by:

$$(\sigma, \tau)\, \Phi\, (\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_q) = \Phi\, (\overset{*}{x}{}^{\sigma^{-1}(1)} \ldots \overset{*}{x}{}^{\sigma^{-1}(p)}; x_{\tau^{-1}(1)} \ldots x_{\tau^{-1}(q)}).$$

The tensor Φ is called *skew-symmetric*, if

$$(\sigma, \tau)\, \Phi = \varepsilon_\sigma\, \varepsilon_\tau \Phi \quad \text{for every pair } (\sigma, \tau)\,.$$

The set of all skew-symmetric tensors of order (p, q) forms a linear subspace $S^p_q(E)$ of $T^p_q(E)$. If $p > n$ or $q > n$, this subspace reduces to the zero-tensor.

The *skew-symmetric part* $A\Phi$ of a tensor $\Phi \in T^p_q(E)$ is defined by

$$A\Phi = \frac{1}{p!\, q!} \sum_{\sigma, \tau} \varepsilon_\sigma\, \varepsilon_\tau\, (\sigma, \tau)\, \Phi \tag{8.43}$$

where the sum is extended over all permutations σ of $(1 \ldots p)$ and all permutations τ of $(1 \ldots q)$. The operator A defines a projection of the space $T^p_q(E)$ onto the subspace $S^p_q(E)$,

$$A\colon T^p_q(E) \to S^p_q(E)\,.$$

It follows from (8.43) that

$$A\, (\Phi \otimes \Psi) = A\Phi \otimes A\Psi \quad \Phi \in S^p(E), \quad \Psi \in S_q(E)\,. \tag{8.44}$$

In fact, consider two fixed permutations σ and τ. Then

$$(\sigma, \tau)\, (\Phi \otimes \Psi) = \sigma\Phi \otimes \tau\Psi$$

whence

$$A\, (\Phi \otimes \Psi) = \frac{1}{p!\, q!} \sum_{\sigma, \tau} \varepsilon_\sigma\, \varepsilon_\tau\, (\sigma, \tau)\, (\Phi \otimes \Psi)$$
$$= \frac{1}{p!} \sum_\sigma \varepsilon_\sigma\, \sigma\Phi\, .\ \frac{1}{q!} \sum_\tau \varepsilon_\tau\, \tau\Psi = A\Phi \otimes A\Psi\,.$$

As a consequence of (8.44) we shall prove that the bilinear mapping

$$(S^p(E), S_q(E)) \to S^p_q(E)$$

defined by

$$(\Phi, \Psi) \to \Phi \otimes \Psi \quad \Phi \in S^p(E),\ \Psi \in S_q(E) \tag{8.45}$$

satisfies the conditions T_1 and T_2' (cf. sec. 6.5 and sec. 6.7) and hence is a tensor product.

As it has been shown in sec. 7.7 every tensor $X \in T^p_q(E)$ can be written as a linear combination

$$X = \sum_{\nu, \mu} \lambda^{\nu\mu}\, \Phi_\nu \otimes \Psi_\mu \qquad \Phi \in T^p(E),\ \Psi \in T_q(E)\,.$$

Applying the antisymmetry-operator and using the formula (8.44) we obtain

$$A X = \sum_{\nu,\mu} \lambda^{\nu\mu} A\Phi_\nu \otimes A\Psi_\mu .$$

In particular, assume that $X \in S_q^p(E)$. Then $AX = X$ and the above relation can be written as

$$X = \sum_{\nu,\mu} \lambda^{\nu\mu} A\Phi_\nu \otimes A\Psi_\mu .$$

This proves that the bilinear mapping (8.45) satisfies the condition T_1.

The property T_2' follows immediately from the fact that the bilinear mapping

$$(\Phi, \Psi) \to \Phi \otimes \Psi \qquad \Phi \in T^p(E), \quad \Psi \in T_q(E)$$

has this property (cf. sec. 6.7).

The above result shows that $S_q^p(E)$ is the tensor-product of the spaces $S^p(E)$ and $S_q(E)$,

$$S_q^p(E) = S^p(E) \otimes S_q(E) . \tag{8.46}$$

Comparing the dimensions in the relation (8.46) we obtain the formula

$$\dim S_q^p(E) = \binom{n}{p}\binom{n}{q} . \tag{8.47}$$

8.12. The skew-symmetric product. We define the skew-symmetric product of two tensors $\Phi \in S_q^p(E)$ and $\Psi \in S_s^r(E)$ by

$$\Phi \wedge \Psi = \binom{p+r}{q}\binom{q+s}{q} A\,(\Phi \otimes \Psi) .$$

In the same way as for covariant tensors it is shown that the skew-symmetric product is distributive and associative. The commutative law has the form

$$\Phi \wedge \Psi = (-1)^{pr+qs}\Psi \wedge \Phi \quad \Phi \in S_q^p(E), \quad \Psi \in S_s^r(E) . \tag{8.48}$$

This yields for $\Psi = \Phi$

$$\Phi \wedge \Phi = (-1)^{p^2+q^2}\Phi \wedge \Phi = (-1)^{p+q}\Phi \wedge \Phi$$

showing that

$$\Phi \wedge \Phi = 0 \quad \text{if } p+q \text{ is odd.}$$

We finally note that

$$(\Phi_1 \otimes \Phi_2) \wedge (\Psi_1 \otimes \Psi_2) = (\Phi_1 \wedge \Psi_1) \otimes (\Phi_2 \otimes \Psi_2) \quad \begin{array}{l} \Phi_1 \in S^p(E), \Psi_1 \in S^r(E) \\ \Phi_2 \in S_q(E), \Psi_2 \in S_s(E). \end{array} \tag{8.49}$$

In fact,

$$(\Phi_1 \otimes \Phi_2) \wedge (\Psi_1 \otimes \Psi_2) = \binom{p+r}{p}\binom{q+s}{q} A\,(\Phi_1 \otimes \Phi_2 \otimes \Psi_1 \otimes \Psi_2)$$

and, in view of (8.10) and (8.44)

$$(\Phi_1 \wedge \Psi_1) \otimes (\Phi_2 \wedge \Psi_2) = \binom{p+r}{p}\binom{q+s}{q} A\,(\Phi_1 \otimes \Psi_1) \otimes A\,(\Phi_2 \otimes \Psi_2)$$
$$= \binom{p+r}{p}\binom{q+s}{q} A\,(\Phi_1 \otimes \Psi_1 \otimes \Phi_2 \otimes \Psi_2)\,.$$

Comparing these two equations and observing that $\Psi_1 \otimes \Phi_2 = \Phi_2 \otimes \Psi_1$ we obtain the formula (8.49).

8.13. Skew-symmetric powers. Given a tensor $\Phi \in S_q^p(E)$ and a positive integer k we define the tensor Φ^k by

$$\Phi^k = \frac{1}{k!}\underbrace{\Phi \wedge \cdots \wedge \Phi}_{k}. \tag{8.50}$$

It is convenient to extend this definition to the case $k = 0$ by setting $\Phi^0 = 1$.

It follows from the associativity of the skew-symmetric product that

$$\Phi^k \wedge \Phi^l = \binom{k+l}{k}\Phi^{k+l}\,. \tag{8.51}$$

Now consider two tensors $\Phi \in S_q^p(E)$ and $\Psi \in S_q^p(E)$ and assume that $p + q$ is even. Then, by (8.48),

$$\Phi \wedge \Psi = \Psi \wedge \Phi\,.$$

Hence, the power $(\Phi + \Psi)^k$ can be expanded by the binomial theorem yielding

$$(\Phi + \Psi)^k = \frac{1}{k!}(\Phi + \Psi) \wedge \cdots \wedge (\Phi + \Psi)$$
$$= \frac{1}{k!}\sum_{\nu=0}^{k}\binom{k}{\nu}(\underbrace{\Phi \wedge \cdots \wedge \Phi}_{\nu}) \wedge (\underbrace{\Psi \wedge \cdots \wedge \Psi}_{k-\nu})$$
$$= \frac{1}{k!}\sum_{\nu=0}^{k}\binom{k}{\nu}\nu!\,(k-\nu)!\,\Phi^\nu \wedge \Psi^{k-\nu}$$
$$= \sum_{\nu=0}^{k}\Phi^\nu \wedge \Psi^{k-\nu}\,.$$

We thus obtain the following binomial formula:

$$(\Phi + \Psi)^k = \sum_{\nu=0}^{k}\Phi^\nu \wedge \Psi^{k-\nu} \quad \Phi, \Psi \in S_q^p(E), \quad p + q \text{ even}. \tag{8.52}$$

The k-th skew-symmetric power of a (1—1)-tensor θ can be written as the determinant

$$\theta^k(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^k; x_1 \ldots x_k) = \begin{vmatrix} \theta(\overset{*}{x}{}^1, x_1) \ldots \theta(\overset{*}{x}{}^k, x_1) \\ \vdots \qquad\qquad \vdots \\ \theta(\overset{*}{x}{}^1, x_k) \ldots \theta(\overset{*}{x}{}^k, x_k) \end{vmatrix}. \tag{8.53}$$

In fact,

$$\theta^k(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^k; x_1 \dots x_k) = \frac{1}{k!} \sum_{\sigma,\tau} \varepsilon_\sigma \varepsilon_\tau \,\theta(\overset{*}{x}{}^{\sigma(1)}, x_{\tau(1)}) \dots \theta(\overset{*}{x}{}^{\sigma(k)}, x_{\tau(k)})$$

$$= \sum_\sigma \varepsilon_\sigma \,\theta(\overset{*}{x}{}^{\sigma(1)}, x_1) \dots \theta(\overset{*}{x}{}^{\sigma(k)}, x_k)\,.$$

Problems: 1. Prove that the contraction-operator C^i_j induces a linear mapping of $S^p_q(E)$ into $S^{p-1}_{q-1}(E)$ $(1 \leqq i \leqq p,\ 1 \leqq j \leqq q)$.

2. Given two linear spaces E and H, denote by $S_q(E; H)$ the space of all q-linear skew-symmetric mappings of E into H. Establish a natural isomorphism of $S_q(E; S^p(E))$ onto $S(E^p_q)$.

§ 4. The duality of $S^p_q(E)$ and $S^q_p(E)$

8.14. The scalar-product in the spaces $S^p_q(E)$ and $S^q_p(E)$. As it has been shown in sec. 7.15 the two spaces $T^p_q(E)$ and $T^q_p(E)$ are dual with respect to the scalar-product $\langle \Phi, \Psi \rangle$ defined by (7.34). Now we shall prove that the two subspaces $S^p_q(E)$ and $S^q_p(E)$ are again dual with respect to the bilinear function $\langle \Phi, \Psi \rangle$.

First of all it will be shown that the two endomorphisms

$$A\colon T^p_q(E) \to T^p_q(E)$$

and

$$A\colon T^q_p(E) \to T^q_p(E)$$

are dual,

$$\langle A\Phi, \Psi \rangle = \langle \Phi, A\Psi \rangle \quad \Phi \in T^p_q(E), \Psi \in T^q_p(E)\,. \tag{8.54}$$

According to (7.35), the scalar-product $\langle \Phi, \Psi \rangle$ is given by

$$\langle \Phi, \Psi \rangle = \sum_{(\nu)(\mu)} \Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} \Psi^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_p}\,. \tag{8.55}$$

Replacing Φ by $(\sigma, \tau)\Phi$ we obtain from (8.55)

$$\langle (\sigma,\tau)\Phi, \Psi \rangle = \sum_{(\nu)(\mu)} \Phi^{\nu_{\sigma^{-1}(1)} \dots \nu_{\sigma^{-1}(p)}}_{\mu_{\tau^{-1}(1)} \dots \mu_{\tau^{-1}(q)}} \Psi^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_p}\,.$$

Introducing new indices of summation by

$$\nu_i = \alpha_{\sigma(i)}\ (i = 1 \dots p) \quad \text{and} \quad \mu_j = \beta_{\tau(j)}\ (j = 1 \dots q)$$

we see that

$$\langle (\sigma,\tau)\Phi, \Psi \rangle = \sum_{(\alpha)(\beta)} \Phi^{\alpha_1 \dots \alpha_p}_{\beta_1 \dots \beta_q} \Psi^{\beta_{\tau(1)} \dots \beta_{\tau(q)}}_{\alpha_{\sigma(1)} \dots \alpha_{\sigma(p)}} = \langle \Phi, (\tau^{-1}, \sigma^{-1})\Psi \rangle\,.$$

Multiplication by $\varepsilon_\sigma \varepsilon_\tau$ and summation over all permutations σ and τ gives

$$\left\langle \sum_{\sigma,\tau} \varepsilon_\sigma \varepsilon_\tau (\sigma,\tau)\Phi, \Psi \right\rangle = \left\langle \Phi, \sum_\sigma \varepsilon_\sigma \varepsilon_\tau (\tau^{-1}, \sigma^{-1})\Psi \right\rangle$$

$$= \left\langle \Phi, \sum_{\sigma,\tau} \varepsilon_{\sigma^{-1}} \varepsilon_{\tau^{-1}} (\tau^{-1}, \sigma^{-1})\Psi \right\rangle$$

$$= \left\langle \Phi, \sum_{\sigma,\tau} \varepsilon_\sigma \varepsilon_\tau (\tau, \sigma)\Psi \right\rangle$$

whence (8.54).

To prove the duality of $S_q^p(E)$ and $S_p^q(E)$ assume that

$$\langle \Phi, \Psi \rangle = 0$$

for a fixed tensor $\Phi \in S_q^p(E)$ and all tensors $\Psi \in S_p^q(E)$. Let $X \in T_p^q(E)$ be an arbitrary tensor. Then, in view of (8.54)

$$\langle \Phi, X \rangle = \langle A\Phi, X \rangle = \langle \Phi, AX \rangle = 0 .$$

In view of the duality of $T_q^p(E)$ and $T_p^q(E)$ this implies that $\Phi = 0$.

To avoid dimension-factors in our formulas as far as possible it will be convenient to introduce a new scalar-product $\langle\!\langle \Phi, \Psi \rangle\!\rangle$ defined by

$$\langle\!\langle \Phi, \Psi \rangle\!\rangle = \frac{1}{p!\,q!} \langle \Phi, \Psi \rangle \qquad \Phi \in S_q^p(E), \Psi \in S_p^q(E) . \tag{8.56}$$

From (8.55) and (8.56) we obtain the formula

$$\langle\!\langle \Phi, \Psi \rangle\!\rangle = \sum_{\nu_1 < \cdots < \nu_p} \sum_{\mu_1 < \cdots < \mu_q} \Phi^{\nu_1 \ldots \nu_p}_{\mu_1 \ldots \mu_q} \Psi^{\mu_1 \ldots \mu_q}_{\nu_1 \ldots \nu_p}$$

or, in brief

$$\langle\!\langle \Phi, \Psi \rangle\!\rangle = \sum_{<} \Phi^{\nu_1 \ldots \nu_p}_{\mu_1 \ldots \mu_q} \Psi^{\mu_1 \ldots \mu_q}_{\nu_1 \ldots \nu_p} . \tag{8.57}$$

Here the symbol $<$ indicates that the indices $(\nu_1 \ldots \nu_p)$ and $(\mu_1 \ldots \mu_q)$ are subject to the conditions $\nu_1 < \cdots < \nu_p$ and $\mu_1 < \cdots < \mu_q$.

8.15. The scalar-product of decomposable tensors. Assume that one of the tensors Φ and Ψ, say Ψ, is decomposable,

$$\Psi = (a_1 \wedge \cdots \wedge a_q) \otimes (\overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p) \quad a_\mu \in E, \overset{*}{a}{}^\nu \in E^* .$$

Then

$$\langle\!\langle \Phi, \Psi \rangle\!\rangle = \Phi(\overset{*}{a}{}^1 \ldots \overset{*}{a}{}^p; a_1 \ldots a_q) . \tag{8.58}$$

In fact, the formulas (8.54) and (7.38) yield

$$\begin{aligned} \langle\!\langle \Phi, \Psi \rangle\!\rangle = \frac{1}{p!\,q!} \langle \Phi, \Psi \rangle &= \langle \Phi, A(a_1 \otimes \cdots \otimes a_q \otimes \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^p) \\ &= \langle A\Phi, a_1 \otimes \cdots \otimes a_q \otimes \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^p \rangle \\ &= \langle \Phi, a_1 \otimes \cdots \otimes a_q \otimes \overset{*}{a}{}^1 \otimes \cdots \otimes \overset{*}{a}{}^p \rangle \\ &= \Phi(\overset{*}{a}{}^1 \ldots \overset{*}{a}{}^p; a_1 \ldots a_q) . \end{aligned}$$

The scalar-product of a decomposable contravariant tensor and a decomposable covariant tensor can be written as the determinant

$$\langle\!\langle a_1 \wedge \cdots \wedge a_p, \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p \rangle\!\rangle = \begin{vmatrix} \langle \overset{*}{a}{}^1, a_1 \rangle \ldots \langle \overset{*}{a}{}^p, a_1 \rangle \\ \cdot \\ \cdot \\ \cdot \\ \langle \overset{*}{a}{}^1, a_p \rangle \cdots \langle \overset{*}{a}{}^p, a_p \rangle \end{vmatrix} . \tag{8.59}$$

The equation (8.59) is a consequence of the formulas (8.16) and (7.39):

$$\langle\!\langle a_1 \wedge \cdots \wedge a_p, \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^p \rangle\!\rangle = \frac{1}{p!} \sum_{\sigma,\tau} \varepsilon_\sigma \varepsilon_\tau \langle a_{\sigma(1)} \otimes \cdots \otimes a_{\sigma(p)}, \overset{*}{a}{}^{\tau(1)} \otimes \cdots \otimes \overset{*}{a}{}^{\tau(p)} \rangle = \begin{vmatrix} \langle \overset{*}{a}{}^1, a_1 \rangle \cdots \langle \overset{*}{a}{}^p, a_1 \rangle \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \langle \overset{*}{a}{}^1, a_p \rangle \cdots \langle \overset{*}{a}{}^p, a_p \rangle \end{vmatrix}.$$

Now consider a pair of dual bases $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$. Inserting $a_i = e_{\nu_i}$ and $\overset{*}{a}{}^i = \overset{*}{e}{}^{\mu_i}$ $(i = 1 \ldots p)$ in (8.59) we obtain

$$\langle\!\langle e_{\nu_1} \wedge \cdots \wedge e_{\nu_p}, \overset{*}{e}{}^{\mu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\mu_p} \rangle\!\rangle = \begin{vmatrix} \delta_{\nu_1}^{\mu_1} \ldots \delta_{\nu_1}^{\mu_p} \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \delta_{\nu_p}^{\mu_1} \cdots \delta_{\nu_p}^{\mu_p} \end{vmatrix} \quad (\nu_1 < \cdots < \nu_p, \mu_1 < \cdots < \mu_p).$$

The inequalities $\nu_1 < \cdots < \nu_p$ and $\mu_1 < \cdots < \mu_p$ imply that this determinant is equal to 1 if $\nu_i = \mu_i$ $(i = 1 \ldots p)$ and equal to zero otherwise. So we have the relations

$$\langle\!\langle e_{\nu_1} \wedge \cdots \wedge e_{\nu_p}, \overset{*}{e}{}^{\mu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\mu_p} \rangle\!\rangle = \begin{cases} 1 \text{ if } \nu_i = \mu_i \ (i = 1 \ldots p) \\ 0 \text{ otherwise} \end{cases}$$

showing that the bases

$$e_{\nu_1} \wedge \cdots \wedge e_{\nu_p} \quad \text{and} \quad \overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p}$$

of $S^p(E)$ and $S_p(E)$ are dual with respect to the scalar-product (8.57). In the case $p = n$ the above equations yield

$$\langle\!\langle e_1 \wedge \cdots \wedge e_n, \overset{*}{e}{}^1 \wedge \cdots \wedge \overset{*}{e}{}^n \rangle\!\rangle = 1 .$$

This implies that the two determinant-functions

$$\Delta^* = e_1 \wedge \cdots \wedge e_n \quad \text{and} \quad \Delta = \overset{*}{e}{}^1 \wedge \cdots \wedge \overset{*}{e}{}^n$$

are dual (see sec. 4.5). In fact, the product $\Delta^* \otimes \Delta$ assumes the value 1 on the dual bases $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$:

$$(\Delta^* \otimes \Delta)\,(\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n; e_1 \ldots e_n) = \Delta^*(\overset{*}{e}{}^1 \ldots \overset{*}{e}{}^n)\,\Delta\,(e_1 \ldots e_n) = \langle\!\langle \Delta^*, \Delta \rangle\!\rangle = 1 .$$

8.16. The isomorphism $\mathscr{A}$. The duality of the spaces $S^p(E)$ and $S_p(E)$ leads in a natural way to an isomorphism of the space $L(S^p(E); S^p(E))$ of all endomorphisms of $S^p(E)$ onto the space $S^p_p(E)$. Let α be a given endomorphism of $S^p(E)$. Then a bilinear function F_α is defined in the spaces $S^p(E)$ and $S_p(E)$ by

$$F_\alpha(\Phi, \Phi^*) = \langle\!\langle \alpha\Phi, \Phi^* \rangle\!\rangle \quad \Phi \in S^p(E), \Phi^* \in S_p(E) . \tag{8.60}$$

The bilinear function F_α in turn determines a linear function f_α in $S^p_p(E)$ such that

$$F_\alpha(\Phi, \Phi^*) = f_\alpha(\Phi \otimes \Phi^*) . \tag{8.61}$$

Finally, the linear function f_α can be written as

$$f_\alpha(\Phi \otimes \Phi^*) = \langle\!\langle X_\alpha, \Phi \otimes \Phi^* \rangle\!\rangle \quad X_\alpha \in S^p_p(E)\,, \tag{8.62}$$

where X_α is a fixed tensor. The equations (8.60), (8.61) and (8.62) yield

$$\langle\!\langle X_\alpha, \Phi \otimes \Phi^* \rangle\!\rangle = \langle\!\langle \alpha\Phi, \Phi^* \rangle\!\rangle\,. \tag{8.63}$$

The correspondence $\alpha \to X_\alpha$ defines a linear mapping

$$\mathscr{A}: L\,(S^p(E); S^p(E)) \to S^p_p(E)\,.$$

The endomorphism α and the corresponding tensor $\mathscr{A}\alpha$ are connected by the relation

$$\langle\!\langle \mathscr{A}\alpha, \Phi \otimes \Phi^* \rangle\!\rangle = \langle\!\langle \alpha\Phi, \Phi^* \rangle\!\rangle \quad \Phi \in S^p(E), \Phi^* \in S_p(E)\,. \tag{8.64}$$

The mapping $\mathscr{A}$ is regular and consequently it is an isomorphism of the space $L\,(S^p(E); S^p(E))$ onto the space $S^p_p(E)$. For $p = 1$ the operator $\mathscr{A}$ coincides with the isomorphism $L(E; E) \to T^1_1(E)$ defined in sec. 7.2 if $S^1(E)$ is identified with E.

In the same way an isomorphism

$$\mathscr{A}: L\,(S_p(E); S_p(E)) \to S^p_p(E)$$

is defined by the relation

$$\langle\!\langle \mathscr{A}\beta, \Phi^* \otimes \Phi \rangle\!\rangle = \langle\!\langle \beta\Phi^*, \Phi \rangle\!\rangle \quad \Phi^* \in S_p(E), \Phi \in S^p(E)\,. \tag{8.65}$$

Inserting for β the dual isomorphism α^* we obtain

$$\langle\!\langle \mathscr{A}\alpha^*, \Phi^* \otimes \Phi \rangle\!\rangle = \langle\!\langle \alpha^*\Phi^*, \Phi \rangle\!\rangle = \langle\!\langle \Phi^*, \alpha\Phi \rangle\!\rangle = \langle\!\langle \mathscr{A}\alpha, \Phi^* \otimes \Phi \rangle\!\rangle$$

whence

$$\mathscr{A}\alpha^* = \mathscr{A}\alpha\,.$$

This equation shows that dual endomorphisms have the same image-tensor in $S^p_p(E)$.

8.17. The relation between the scalar-products. Now consider the scalar-product in $L\,(S^p(E); S^p(E))$ defined by

$$\langle \alpha, \beta \rangle = \operatorname{tr}\,(\alpha \circ \beta)$$

(cf. sec. 4.26). It will be shown that

$$\langle\!\langle \mathscr{A}\alpha, \mathscr{A}\beta \rangle\!\rangle = \langle \alpha, \beta \rangle\,. \tag{8.66}$$

It is sufficient to prove the formula (8.66) for two endomorphisms of the special form

$$\alpha\Phi = \langle\!\langle \Phi, \Phi_1^* \rangle\!\rangle \Phi_1 \quad \text{and} \quad \beta\Phi = \langle\!\langle \Phi, \Phi_2^* \rangle\!\rangle \Phi_2 \tag{8.67}$$

$$\Phi_i \in S^p(E) \quad \text{and} \quad \Phi_i^* \in S_p(E) \qquad (i = 1, 2)$$

because every endomorphism is a linear combination of such endomorphisms. In view of (7.37) we obtain form the equations (8.64) and the first equation (8.67)

$$\langle\!\langle \mathscr{A}\alpha, \Phi \otimes \Phi^* \rangle\!\rangle = \langle\!\langle \Phi, \Phi_1^* \rangle\!\rangle \langle\!\langle \Phi_1, \Phi^* \rangle\!\rangle = \langle\!\langle \Phi_1 \otimes \Phi_1^*, \Phi \otimes \Phi^* \rangle\!\rangle$$

whence

$$\mathscr{A}\alpha = \Phi_1 \otimes \Phi_1^* . \tag{8.68}$$

Similarly,

$$\mathscr{A}\beta = \Phi_2 \otimes \Phi_2^* . \tag{8.69}$$

The equations (8.69), (8.70) and (7.37) yield

$$\langle\!\langle \mathscr{A}\alpha, \mathscr{A}\beta \rangle\!\rangle = \langle\!\langle \Phi_1, \Phi_2^* \rangle\!\rangle \langle\!\langle \Phi_2, \Phi_1^* \rangle\!\rangle . \tag{8.70}$$

At the same time it follows from (8.67) that

$$(\alpha \circ \beta)\Phi = \langle\!\langle \Phi, \Phi_2^* \rangle\!\rangle \langle\!\langle \Phi_2, \Phi_1^* \rangle\!\rangle \Phi_1$$

whence

$$\operatorname{tr}(\alpha \circ \beta) = \langle\!\langle \Phi_1, \Phi_2^* \rangle\!\rangle \langle\!\langle \Phi_2, \Phi_1^* \rangle\!\rangle . \tag{8.71}$$

Comparing (8.70) and (8.71) we find (8.66).

8.18. Induced endomorphisms. Let us now consider the case that the endomorphism α of $S^p(E)$ is induced by an endomorphism ϑ of E, $\alpha = (\vartheta)^p$. It will be shown, that then $\mathscr{A}\alpha$ is the p-th skew-symmetric power of the tensor $\mathscr{A}\vartheta$,

$$\mathscr{A}(\vartheta)^p = (\mathscr{A}\vartheta)^p . \tag{8.72}$$

Let

$$\Phi = x_1 \wedge \cdots \wedge x_p \quad \text{and} \quad \Phi^* = \overset{*}{x}{}^1 \wedge \cdots \wedge \overset{*}{x}{}^p$$

be two decomposable skew-symmetric tensors. Then

$$\langle\!\langle \mathscr{A}(\vartheta)^p, \Phi \otimes \Phi^* \rangle\!\rangle = \langle\!\langle (\vartheta)^p \Phi, \Phi^* \rangle\!\rangle = \langle\!\langle \vartheta x_1 \wedge \cdots \wedge \vartheta x_p, \overset{*}{x}{}^1 \wedge \cdots \wedge \overset{*}{x}{}^p \rangle\!\rangle$$

$$= \begin{vmatrix} \langle \overset{*}{x}{}^1, \vartheta x_1 \rangle & \cdots & \langle \overset{*}{x}{}^p, \vartheta x_1 \rangle \\ \vdots & & \\ \langle \overset{*}{x}{}^1, \vartheta x_p \rangle & \cdots & \langle \overset{*}{x}{}^p, \vartheta x_p \rangle \end{vmatrix} . \tag{8.73}$$

Now consider the right hand-side of (8.72). Employing the formulas (8.58) and (8.53) we obtain

$$\langle\!\langle (\mathscr{A}\vartheta)^p, \Phi \otimes \Phi^* \rangle\!\rangle = \langle\!\langle (\mathscr{A}\vartheta)^p, (x_1 \wedge \cdots \wedge x_p) \otimes (\overset{*}{x}{}^1 \wedge \cdots \wedge \overset{*}{x}{}^p) \rangle\!\rangle$$

$$= (\mathscr{A}\vartheta)^p (\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_p) \tag{8.74}$$

$$= \begin{vmatrix} (\mathscr{A}\vartheta)(\overset{*}{x}{}^1, x_1) \ldots (\mathscr{A}\vartheta)(\overset{*}{x}{}^p, x_1) \\ \vdots \\ (\mathscr{A}\vartheta)(\overset{*}{x}{}^1, x_p) \ldots (\mathscr{A}\vartheta)(\overset{*}{x}{}^p, x_p) \end{vmatrix} = \begin{vmatrix} \langle \overset{*}{x}{}^1, \vartheta x_1 \rangle \ldots \langle \overset{*}{x}{}^p, \vartheta x_1 \rangle \\ \vdots \\ \langle \overset{*}{x}{}^1, \vartheta x_p \rangle \ldots \langle \overset{*}{x}{}^p, \vartheta x_p \rangle \end{vmatrix} .$$

Comparing (8.73) and (8.74) we obtain (8.72).

In the same way it is shown that

$$\mathscr{A}(\vartheta)_p = (\mathscr{A}\vartheta)^p$$

for every endomorphism ϑ of E.

Problem: 1. Define a multiplication in the space $S_p^p(E)$ by

$$\Phi \circ \Psi = \frac{1}{p!} C_{p+1}^1 \ldots C_{p+1}^1 (\Phi \otimes \Psi)$$

where C_{p+1}^1 designates the contraction-operator.

a) Prove that

$$(\Phi \circ \Psi)_{\beta_1 \ldots \beta_p}^{\alpha_1 \ldots \alpha_p} = \sum_{<} \Phi_{\beta_1 \ldots \beta_p}^{\nu_1 \ldots \nu_p} \Psi_{\nu_1 \ldots \nu_p}^{\alpha_1 \ldots \alpha_p}.$$

b) Show that the space $S_p^p(E)$ becomes an algebra under this multiplication. What is the unit element of this algebra?

c) Prove that

$$\mathscr{A}(\alpha \circ \beta) = \mathscr{A}(\alpha) \circ \mathscr{A}(\beta)$$

for any two endomorphisms α and β of $S^p(E)$.

Chapter IX

Duality in exterior algebra

Comparing the dimensions of the two spaces $S^p(E)$ and $S_{n-p}(E)$ we see that these spaces are isomorphic. For the same reason every space $S_q^p(E)$ is isomorphic to $S_{n-p}^{n-q}(E)$. The purpose of the present chapter is to show that there are canonical isomorphisms

$$S^p(E) \to S_{n-p}(E) \quad \text{and} \quad S_q^p(E) \to S_{n-p}^{n-q}(E).$$

To define these isomorphisms we introduce first a dual operation to the skew-symmetric product.

§ 1. The dual product

9.1. Definition. In the present paragraph it will be shown that the duality of the spaces $S_q^p(E)$ and $S_p^q(E)$ leads to a dual operation of the skew-symmetric product.

Consider a fixed tensor $\Phi \in S_q^p(E)$. Then, for every pair of integers k, l $(k \geqq p, l \geqq q)$ a linear mapping

$$F_\Phi : S_{l-q}^{k-p}(E) \to S_l^k(E)$$

is defined by

$$F_\Phi \Psi = \Phi \wedge \Psi \qquad \Psi \in S_{l-q}^{k-p}(E). \tag{9.1}$$

F_Φ^* determines a dual mapping

$$F_\Phi^*\colon S_k^l(E) \to S_{k-p}^{l-q}(E)$$

which is connected with F_Φ by the relation

$$\langle\!\langle X, F_\Phi \Psi\rangle\!\rangle = \langle\!\langle F_\Phi^* X, \Psi\rangle\!\rangle \quad X \in S_k^l(E), \Psi \in S_{l-q}^{k-p}(E)\,. \tag{9.2}$$

We now define the *dual product* of the tensors X and Φ, denoted as $X \vee \Phi$, by

$$X \vee \Phi = F_\Phi^* X\,. \tag{9.3}$$

This product determines a bilinear mapping

$$(S_k^l(E), S_q^p(E)) \to S_{k-p}^{l-q}(E) \quad (l \geqq q, k \geqq p)\,.$$

The equations (9.1), (9.2) and (9.3) imply that the skew-symmetric product and the dual product are connected by the relation

$$\langle\!\langle X, \Phi \wedge \Psi\rangle\!\rangle = \langle\!\langle X \vee \Phi, \Psi\rangle\!\rangle \tag{9.4}$$

$$X \in S_k^l(E), \Phi \in S_q^p(E), \Psi \in S_{l-q}^{k-p}(E)\,.$$

9.2. Properties. The dual product has the following properties:

$$X \vee \Phi = \langle\!\langle X, \Phi\rangle\!\rangle \quad X \in S_p^q(E) \quad \Phi \in S_q^p(E) \tag{9.5}$$

$$X \vee 1 = X \qquad X \in S_k^l(E) \tag{9.6}$$

$$\left.\begin{aligned} (X \vee \Phi_1) \vee \Phi_2 &= X \vee (\Phi_1 \wedge \Phi_2) \\ (X \vee \Phi_1) \vee \Phi_2 &= (-1)^{pr+qs} (X \vee \Phi_2) \vee \Phi_1 \end{aligned}\right\} \begin{aligned} &X \in S_k^l(E), \Phi_1 \in S_q^p(E)\,, &(9.7)\\ &\Phi_2 \in S_s^r(E), &(9.8)\end{aligned}$$

$$l \geqq q + s,\ k \geqq p + r$$

The formulas (9.5) and (9.6) follow from (9.4) by inserting $\Psi = 1$ and $\Phi = 1$ respectively. To prove (9.7), let $\Psi \in S_{k-p-r}^{l-q-s}(E)$ be an arbitrary tensor. Then,

$$\begin{aligned} \langle\!\langle (X \vee \Phi_1) \vee \Phi_2, \Psi\rangle\!\rangle &= \langle\!\langle X \vee \Phi_1, \Phi_2 \wedge \Psi\rangle\!\rangle \\ = \langle\!\langle X, \Phi_1 \wedge \Phi_2 \wedge \Psi\rangle\!\rangle &= \langle\!\langle X \vee (\Phi_1 \wedge \Phi_2), \Psi\rangle\!\rangle \end{aligned}$$

whence (9.7). The relation (9.8) follows from (9.7) and the commutative law of the skew-symmetric product.

Employing a pair of dual bases, we can write the product-tensor $X \vee \Phi$ as

$$\begin{aligned} &(X \vee \Phi)(\overset{*}{x}{}^{q+1} \dots \overset{*}{x}{}^l; x_{p+1} \dots x_k) \\ &= \sum_< \Phi_{\mu_1 \dots \mu_q}^{\nu_1 \dots \nu_p} X(\overset{*}{e}{}^{\mu_1} \dots \overset{*}{e}{}^{\mu_q}, \overset{*}{x}{}^{q+1} \dots \overset{*}{x}{}^l; e_{\nu_1} \dots e_{\nu_p}, x_{p+1} \dots x_k)\,. \end{aligned} \tag{9.9}$$

To prove this formula, denote by Ω the tensor on the right-hand side of (9.9). Choose an arbitrary tensor $\Psi \in S^{k-p}_{l-q}(E)$. Then

$$\begin{aligned}\langle \Omega, \Psi \rangle &= \sum_{(\nu)(\mu)} \Omega^{\mu_{q+1}\dots\mu_l}_{\nu_{p+1}\dots\nu_k} \Psi^{\nu_{p+1}\dots\nu_k}_{\mu_{q+1}\dots\mu_l} \\ &= \frac{1}{p!\,q!} \sum_{(\nu)(\mu)} \Phi^{\nu_1\dots\nu_p}_{\mu_1\dots\mu_q} \Psi^{\nu_{p+1}\dots\nu_k}_{\mu_{q+1}\dots\mu_l} X^{\mu_1\dots\mu_q\dots\mu_l}_{\nu_1\dots\nu_p\dots\nu_k} \\ &= \frac{1}{p!\,q!} \langle \Phi \otimes \Psi, X \rangle = \frac{1}{p!\,q!} \langle \Phi \otimes \Psi, A\,X \rangle = \frac{1}{p!\,q!} \langle A\,(\Phi \otimes \Psi), X \rangle \\ &= (k-p)!\,(l-q)!\,\langle\!\langle \Phi \wedge \Psi, X \rangle\!\rangle = (k-p)!\,(l-q)!\,\langle\!\langle X \vee \Phi, \Psi \rangle\!\rangle \\ &= \langle X \vee \Phi, \Psi \rangle\,.\end{aligned}$$

We thus have the relation

$$\langle \Omega, \Psi \rangle = \langle X \vee \Phi, \Psi \rangle$$

for every tensor $\Psi \in S^{k-p}_{l-q}(E)$ whence $\Omega = X \vee \Phi$.

Finally it will be shown that

$$\begin{gathered}(X_1 \otimes X_2) \vee (\Phi_1 \otimes \Phi_2) = (X_1 \vee \Phi_2) \otimes (X_2 \vee \Phi_1) \\ X_1 \in S^l(E), X_2 \in S_k(E) \\ \Phi_1 \in S^p(E), \Phi_2 \in S_q(E) \\ l \geqq q, k \geqq p\,.\end{gathered} \tag{9.10}$$

We note first that, according to (7.37),

$$\langle\!\langle X_1 \otimes X_2, \Omega_1 \otimes \Omega_2 \rangle\!\rangle = \langle\!\langle X_1, \Omega_2 \rangle\!\rangle \cdot \langle\!\langle X_2, \Omega_1 \rangle\!\rangle \quad \Omega_1 \in S^k(E), \Omega_2 \in S_l(E)\,. \tag{9.11}$$

Now consider two arbitrary tensors $\Psi_1 \in S^{k-p}(E)$ and $\Psi_2 \in S_{l-q}(E)$. Employing the formulas (9.4), (8.49) and (9.11) we obtain

$$\begin{aligned}&\langle\!\langle (X_1 \otimes X_2) \vee (\Phi_1 \otimes \Phi_2), \Psi_1 \otimes \Psi_2 \rangle\!\rangle = \langle\!\langle X_1 \otimes X_2, (\Phi_1 \otimes \Phi_2) \wedge (\Psi_1 \otimes \Psi_2) \rangle\!\rangle \\ &= \langle\!\langle X_1 \otimes X_2, (\Phi_1 \wedge \Psi_1) \otimes (\Phi_2 \wedge \Psi_2) \rangle\!\rangle = \langle\!\langle X_1, \Phi_2 \wedge \Psi_2 \rangle\!\rangle \langle\!\langle X_2, \Phi_1 \wedge \Psi_1 \rangle\!\rangle \\ &= \langle\!\langle X_1 \vee \Phi_2, \Psi_2 \rangle\!\rangle \langle\!\langle X_2 \vee \Phi_1, \Psi_1 \rangle\!\rangle = \langle\!\langle (X_1 \vee \Phi_2) \otimes (X_2 \vee \Phi_1), \Psi_1 \otimes \Psi_2 \rangle\!\rangle\end{aligned}$$

This equation implies (9.10) because the space $S^{k-p}_{l-q}(E)$ is generated by the products $\Psi_1 \otimes \Psi_2$ ($\Psi_1 \in S^{k-p}(E), \Psi_2 \in S_{l-q}(E)$ as it has been shown in sec. 8.11.

9.3. Let a linear mapping $\vartheta: E \to F$ be given and consider the induced mappings

$$(\vartheta)_p : S_p(F) \to S_p(E)$$

and

$$(\vartheta)^p : S^p(E) \to S^p(F)\,.$$

(cf. sec. 8.8). It will be shown that

$$\begin{gathered}(\vartheta)_{k-p}\,(X \vee (\vartheta)^p \Phi) = (\vartheta)_k X \vee \Phi \\ X \in S_k(F), \Phi \in S^p(E), (k \geqq p)\,.\end{gathered} \tag{9.12}$$

We know from sec. 7.16 that the mappings $(\vartheta)^p$ and $(\vartheta)_p$ are dual,

$$\langle\langle(\vartheta)^p\Phi, \Psi\rangle\rangle = \langle\langle\Phi, (\vartheta)_p\Psi\rangle\rangle \quad \Phi \in S^p(E), \Psi \in S_p(F) .$$

Let $\Psi \in S^{k-p}(E)$ be an arbitrary tensor. Then it follows from (7.41) and (9.4) that

$$\begin{aligned}&\langle\langle(\vartheta)_{k-p}(X \vee (\vartheta)^p\Phi), \Psi\rangle\rangle = \langle\langle X \vee (\vartheta)^p\Phi, (\vartheta)^{k-p}\Psi\rangle\rangle\\ &= \langle\langle X, (\vartheta)^p\Phi \wedge (\vartheta)^{k-p}\Psi\rangle\rangle = \langle\langle X, (\vartheta)^k(\Phi \wedge \Psi)\rangle\rangle\\ &= \langle\langle(\vartheta)_k X, \Phi \wedge \Psi\rangle\rangle = \langle\langle(\vartheta)_k X \vee \Phi, \Psi\rangle\rangle\end{aligned}$$

whence (9.12).

In the same way it is proved that

$$(\vartheta)^{k-p}(X \vee (\vartheta)_p\Phi) = (\vartheta)^k X \vee \Phi \tag{9.13}$$

$$X \in S^k(E), \Phi \in S_p(F), (k \geqq p) .$$

Problems: 1. Establish the general *Lagrange Identity*

$$\begin{vmatrix}\sum\limits_{\nu=1}^{n}\xi_1^\nu\eta_\nu^1 & \ldots & \sum\limits_{\nu=1}^{n}\xi_1^\nu\eta_\nu^p\\ \vdots & & \vdots\\ \sum\limits_{\nu=1}^{n}\xi_p^\nu\eta_\nu^1 & \ldots & \sum\limits_{\nu=1}^{n}\xi_p^\nu\eta_\nu^p\end{vmatrix} = \sum_{<}\begin{vmatrix}\xi_1^{\nu_1} & \ldots & \xi_1^{\nu_p}\\ \vdots & & \vdots\\ \xi_p^{\nu_1} & \ldots & \xi_p^{\nu_p}\end{vmatrix}\begin{vmatrix}\eta_1^{\nu_1} & \ldots & \eta_1^{\nu_p}\\ \vdots & & \vdots\\ \eta_p^{\nu_1} & \ldots & \eta_p^{\nu_p}\end{vmatrix} .$$

Hint: Employing a pair of dual bases $\overset{*}{e}{}^\nu$, e_ν $(\nu = 1 \ldots n)$, consider the vectors

$$x_i = \sum_\nu \xi_i^\nu e_\nu \quad \text{and} \quad \overset{*}{y}{}^i = \sum_\nu \eta_\nu^i \overset{*}{e}{}^\nu \quad (i = 1 \ldots p) .$$

Evaluate the scalar-product

$$\langle\langle x_1 \wedge \cdots \wedge x_p, \overset{*}{y}{}^1 \wedge \cdots \wedge \overset{*}{y}{}^p\rangle\rangle$$

in two different ways (cf. sec. 8.15) .

2. Let

$$X = \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^k$$

be a decomposable tensor and $\Phi \in S^p(E)$ $(p \leqq k)$ be an arbitrary tensor. Prove the formula

$$X \vee \Phi = \sum_{\nu_1 < \cdots < \nu_p} (-1)^{\sum\limits_{i=1}^{p}(\nu_i - i)} \Phi(\overset{*}{a}{}^{\nu_1} \ldots \overset{*}{a}{}^{\nu_p})\, \overset{*}{a}{}^{\nu_{p+1}} \wedge \cdots \wedge \overset{*}{a}{}^{\nu_k}$$

where $(\nu_{p+1} \ldots \nu_k)$ denotes the complementary $(n-k)$-tuple of $(\nu_1 \ldots \nu_p)$ in the natural order.

3. Let k linearly independent vectors $\overset{*}{b}{}^\mu \in E^*$ and p linearly independent vectors $a_\nu \in E$ be given $(p \leqq k)$. Denote by V^* and by U the subspaces of E^* and E generated by the vectors $\overset{*}{b}{}^\mu$ $(\mu = 1 \ldots k)$ and

a_ν $(\nu = 1 \ldots p)$ respectively. Prove:

$$(\overset{*}{b}{}^1 \wedge \cdots \wedge \overset{*}{b}{}^k) \vee (a_1 \wedge \cdots \wedge a_p)$$
$$= \begin{cases} 0 & \text{if } \dim (V^* \cap U^\perp) \geqq k - p + 1 \\ \overset{*}{c}{}^{p+1} \wedge \cdots \wedge \overset{*}{c}{}^k & \text{if } \dim (V^* \cap U^\perp) = k - p \end{cases}$$

where the vectors $(\overset{*}{c}{}^{p+1} \ldots \overset{*}{c}{}^k)$ form a basis of $V^* \cap U^\perp$.

Note: It follows from the formula in prob. 5, chap. II, § 2 that

$$\dim (V^* \cap U^\perp) \geqq k - p \,.$$

Hence the two above cases are the only possible ones.

4. Given three tensors $X \in S_p(E)$, $\Phi \in S^p(E)$ and $\Psi \in S_r(E)$ $(k \geqq p, r \geqq p)$ show that the products

$$(X \wedge \Psi) \vee \Phi, \quad X \wedge (\Psi \vee \Phi) \quad \text{and} \quad (X \vee \Phi) \wedge \Psi$$

are generally different.

§ 2. The tensors J^p

9.4. Definition. In the present paragraph we shall apply our general formulas to the tensors

$$J^p = \frac{1}{p!} \underbrace{J \wedge \cdots \wedge J}_{p}$$

where J is the unit-tensor,

$$J(\overset{*}{x}, x) = \langle \overset{*}{x}, x \rangle \,.$$

First of all, it follows from (8.53), that the tensor J^p can be written as the determinant

$$J^p(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; x_1 \ldots x_p) = \begin{vmatrix} \langle \overset{*}{x}{}^1, x_1 \rangle & \cdots & \langle \overset{*}{x}{}^p, x_1 \rangle \\ \vdots & & \vdots \\ \langle \overset{*}{x}{}^1, x_p \rangle & \cdots & \langle \overset{*}{x}{}^p, x_p \rangle \end{vmatrix} \,.$$

Hence, the identity (4.24) can now be written as

$$J^n = \Delta^* \otimes \Delta \tag{9.14}$$

where Δ^* and Δ is a pair of dual determinant-functions.

The tensors J^p satisfy the following relations:

$$J^p \wedge J^q = \binom{p+q}{p} J^{p+q} \tag{9.15}$$

$$C^i_j J^p = (-1)^{i+j} (n - p + 1) J^{p-1} \quad (1 \leqq i \leqq p,\ 1 \leqq j \leqq p) \tag{9.16}$$

$$J^k \vee J^p = \binom{n + p - k}{p} J^{k-p} \quad (0 \leqq p \leqq k) \,. \tag{9.17}$$

The formula (9.15) is an immediate consequence of (8.51). To prove (9.16), let $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$ be a pair of dual bases. Then the formula (7.26) yields

$$
\begin{aligned}
&(C^i_j J^p)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{p-1}) \\
&= \sum_\nu J^p(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{i-1}, \overset{*}{e}{}^\nu, \overset{*}{x}{}^i \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{j-1}, e_\nu, x_j \ldots x_{p-1}) \\
&= (-1)^{i+j} \sum_\nu J^p(\overset{*}{e}{}^\nu, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; e_\nu, x_1 \ldots x_{p-1}) .
\end{aligned}
$$

Hence, it remains to be shown that

$$
\begin{aligned}
&\sum_\nu J^p(\overset{*}{e}{}^\nu, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; e_\nu, x_1 \ldots x_{p-1}) \\
&= (n - p + 1) J^{p-1}(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{p-1}) . \qquad (9.18)
\end{aligned}
$$

We may assume that the vectors x_ν $(\nu = 1 \ldots p - 1)$ are linearly independent, because otherwise both sides of (9.18) are equal to zero. Then we can choose the basis e_ν such that $e_\nu = x_\nu$ $(\nu = 1 \ldots p)$. This implies that the first $(p - 1)$ terms on the left hand-side of (9.18) are zero. Moreover,

$$\langle \overset{*}{e}{}^\nu, x_j \rangle = \langle \overset{*}{e}{}^\nu, e_j \rangle = 0 \quad (p \leqq \nu \leqq n, 1 \leqq j \leqq p - 1)$$

and hence

$$
\begin{aligned}
&\sum_{\nu=p}^n J^p(\overset{*}{e}{}^\nu, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; e_\nu, x_1 \ldots x_{p-1}) \\
&= \sum_{\nu=p}^n \begin{vmatrix} 1 & \langle \overset{*}{x}{}^1, e_\nu \rangle \cdots\cdots & \langle \overset{*}{x}{}^{p-1}, e_\nu \rangle \\ 0 & \langle \overset{*}{x}{}^1, x_1 \rangle \cdots\cdots & \langle \overset{*}{x}{}^{p-1}, x_1 \rangle \\ \vdots & \vdots & \vdots \\ 0 & \langle \overset{*}{x}{}^1, x_{p-1} \rangle \cdots & \langle \overset{*}{x}{}^{p-1}, x_{p-1} \rangle \end{vmatrix} = \sum_{\nu=p}^n \begin{vmatrix} \langle \overset{*}{x}{}^1, x_1 \rangle \cdots\cdots & \langle \overset{*}{x}{}^{p-1}, x_1 \rangle \\ \vdots & \vdots \\ \langle \overset{*}{x}{}^1, x_{p-1} \rangle \cdots & \langle \overset{*}{x}{}^{p-1}, x_{p-1} \rangle \end{vmatrix} \\
&= (n - p + 1) J^{p-1}(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{p-1}; x_1 \ldots x_{p-1})
\end{aligned}
$$

whence (9.18).

Now we are ready to prove (9.17). Assume first that $p = 1$. Then the formulas (9.9) and (9.16) yield

$$J^k \vee J = C^1_1 J^k = (n - k + 1) J^{k-1} . \qquad (9.19)$$

This formula shows that (9.17) is correct if $p = 1$. Now assume, by induction, that the formula (9.17) is correct for the tensors J^ν $(1 \leqq \nu \leqq p - 1)$. To prove (9.17) for $\nu = p$, write J^p in the form

$$J^p = \frac{1}{p} J^{p-1} \wedge J .$$

Then the equations (9.15) and (9.7) yield

$$J^k \vee J^p = \frac{1}{p} J^k \vee (J^{p-1} \wedge J) = \frac{1}{p} (J^k \vee J^{p-1}) \vee J \cdot \qquad (9.20)$$

By induction, we know that

$$J^k \vee J^{p-1} = \binom{n+p-k-1}{p-1} J^{k-p+1} . \tag{9.21}$$

From (9.20) and (9.21) and (9.19) we obtain the relation

$$J^k \vee J^p = \frac{1}{p}\binom{n+p-k-1}{p-1} J^{k-p+1} \vee J = \binom{n+p-k}{p} J^{k-p}$$

showing that (9.17) is correct for $\nu = p$.

The formula (9.17) yields for the cases $k = n$ and $k = p$

$$J^n \vee J^p = J^{n-p} \tag{9.22}$$

and

$$\langle\!\langle J^p, J^p \rangle\!\rangle = \binom{n}{p} . \qquad (0 \leqq p \leqq n)$$

We finally note, that the scalar product of two tensors $\Phi \in S^p(E)$ and $\Psi \in S_p(E)$ can be written in the form

$$\langle\!\langle \Phi, \Psi \rangle\!\rangle = \langle\!\langle J^p, \Phi \otimes \Psi \rangle\!\rangle . \tag{9.23}$$

It is sufficient to prove (9.23) for two decomposable tensors

$$\Phi = a_1 \wedge \ldots \wedge a_p \text{ and } \Psi = \overset{*}{a}{}^1 \wedge \ldots \wedge \overset{*}{a}{}^p .$$

Then, in view of (8.58) and (8.59),

$$\langle\!\langle J^p, \Phi \otimes \Psi \rangle\!\rangle = J^p(\overset{*}{a}{}^1, \ldots, \overset{*}{a}{}^p; a, \ldots, a_p) = \begin{vmatrix} \langle \overset{*}{a}{}^1, a_1 \rangle \cdots \langle \overset{*}{a}{}^p, a_1 \rangle \\ \vdots \qquad\qquad \vdots \\ \langle \overset{*}{a}{}^1, a_p \rangle \cdots \langle \overset{*}{a}{}^p, a_p \rangle \end{vmatrix} = \langle\!\langle \Phi, \Psi \rangle\!\rangle .$$

9.5. The characteristic polynomial. In this section we shall give a new derivation of the characteristic polynomial of an endomorphism. Let ϑ be an endomorphisms of E and define the tensor θ by

$$\theta(\overset{*}{x}, x) = \langle x, \vartheta x \rangle .$$

Then it follows from (8.53) that

$$\theta^p(\overset{*}{x}{}^1, \ldots \overset{*}{x}{}^p; x_1 \ldots x_p) = J^p(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^p; \vartheta x_1 \ldots \vartheta x_p) .$$

This formula yields for $p = n$

$$\theta^n(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n; x_1 \ldots x_n) = J^n(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n; \vartheta x_1 \ldots \vartheta x_n)$$
$$= \det\vartheta \cdot J^n(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n; x_1 \ldots x_n)$$

whence

$$\theta^n = \det\vartheta \cdot J^n .$$

Replacing ϑ by $\vartheta - \lambda\iota$ we obtain the relation

$$(\theta - \lambda J)^n = \det(\vartheta - \lambda\iota) \cdot J^n$$

and using the binomial formula (8.51)

$$\sum_{p=0}^{n} (-1)^{n-p}\,\theta^p \wedge \lambda^{n-p} J^{n-p} = \det(\vartheta - \lambda\,\iota)\,J^n .$$

In this sum every product $\theta^p \wedge J^{n-p}$ is a skew-symmetric tensor of order (n, n). Hence

$$\theta^p \wedge J^{n-p} = (-1)^{n-p}\,\alpha_p\,J^n \quad (0 \leqq p \leqq n) \tag{9.24}$$

where α_p is a scalar. We thus obtain the equation

$$\det(\vartheta - \lambda\,\iota) = \sum_{p=0}^{n} \alpha_p\,\lambda^{n-p}$$

showing that the determinant of $\vartheta - \lambda\,\iota$ is a polynomial in λ of degree n. The coefficients α_p are given by the formula

$$\alpha_p = (-1)^{n-p}\langle\langle J^p, \theta^p\rangle\rangle \quad (0 \leqq p \leqq n) . \tag{9.25}$$

In fact, scalar multiplication of (9.24) by J^n yields

$$\langle\langle J^n, \theta^p \wedge J^{n-p}\rangle\rangle = (-1)^{n-p}\,\alpha_p \langle\langle J^n, J^n\rangle\rangle .$$

Observing that

$$\langle\langle J^n, \theta^p \wedge J^{n-p}\rangle\rangle = \langle\langle J^n, J^{n-p} \wedge \theta^p\rangle\rangle = \langle\langle J^n \vee J^{n-p}, \theta^p\rangle\rangle = \langle\langle J^p, \theta^p\rangle\rangle$$

and

$$\langle\langle J^n, J^n\rangle\rangle = 1$$

we find the formula (9.25).

To obtain another interpretation of the coefficients α_p consider the induced endomorphisms

$$(\vartheta)^p\colon S^p(E) \to S^p(E) .$$

Then, in view of (8.72),

$$\theta^p = (\mathscr{A}\,\vartheta)^p = \mathscr{A}\,(\vartheta)^p$$

and consequently, by (8.66)

$$\langle\langle J^p, \theta^p\rangle\rangle = \langle\langle \mathscr{A}\,(\iota)^p, \mathscr{A}\,(\vartheta)^p\rangle\rangle = \langle \iota^p, (\vartheta)^p\rangle = \operatorname{tr}(\vartheta)^p .$$

We thus obtain the formula

$$\alpha_p = (-1)^{n-p}\operatorname{tr}(\vartheta)^p \qquad (0 \leqq p \leqq n)$$

showing that the coefficient α_p, up to a sign-factor, is equal to the trace of the endomorphism $(\vartheta)^p$. Hence, the characteristic polynomial of ϑ can be written in the form

$$\det(\vartheta - \lambda\,\iota) = \sum_{p=0}^{n} (-1)^{n-p}\operatorname{tr}(\vartheta)^p\,\lambda^{n-p} . \tag{9.26}$$

It follows from (9.26) that the mappings $\vartheta_2 \circ \vartheta_1$ and $\vartheta_1 \circ \vartheta_2$ have the same characteristic polynomial:

$$\det(\vartheta_2 \circ \vartheta_1 - \lambda\,\iota) = \sum_{p=0}^{n} (-1)^{n-p} \operatorname{tr}(\vartheta_2 \circ \vartheta_1)^p \lambda^{n-p}$$

$$= \sum_{p=0}^{n} (-1)^{n-p} \operatorname{tr}((\vartheta_2)^p \circ (\vartheta_1)^p) \lambda^{n-p}$$

$$= \sum_{p=0}^{n} (-1)^{n-p} \operatorname{tr}((\vartheta_1)^p \circ (\vartheta_2)^p) \lambda^{n-p}$$

$$= \sum_{p=0}^{n} (-1)^{n-p} \operatorname{tr}(\vartheta_1 \circ \vartheta_2)^p \lambda^{n-p} = \det(\vartheta_1 \circ \vartheta_2 - \lambda\,\iota) \,.$$

Problems: 1. Show that $\alpha J^p = J^p$ for every automorphism α of E (concerning the definition of αJ^p cf. sec. 7.9).

2. Prove the formulas

$$\langle\!\langle J^{p+q}, J^p \wedge \Phi \rangle\!\rangle = \binom{n-q}{p} \langle\!\langle J^q, \Phi \rangle\!\rangle \qquad \Phi \in S^q_q(E)$$

and

$$J^p \vee \Phi = \Phi \quad \Phi \in S^p(E) \quad \text{or} \quad \Phi \in S_p(E) \,.$$

3. Show that the coefficient of λ^{n-p} in the characteristic polynomial of an $n \times n$-matrix is $(-1)^{n-p}$ times the sum of all principal minors of order p.

4. Given an endomorphism ϑ of E define the endomorphisms $\vartheta^{(p)}$ $(p = 0 \ldots n)$ by

$$\vartheta^{(p)} = \mathscr{A}^{-1}(J^{p+1} \vee (\mathscr{A}\vartheta)^p) \,,$$

where

$$\mathscr{A}: L(E; E) \to T^1_1(E)$$

is the isomorphism defined in sec. 7.2.

a) Establish the recursion-formula

$$\vartheta^{(p)} = \langle\!\langle J^p, (\mathscr{A}\vartheta)^p \rangle\!\rangle\,\iota - \vartheta \circ \vartheta^{(p-1)} \,.$$

Hint: Use the formula (9.9).

b) Derive the formula

$$\vartheta^{(p)} = (-1)^{n-p} \sum_{\nu=0}^{p} \alpha_\nu \vartheta^{p-\nu}$$

where the α_ν are the coefficients in the characteristic polynomial of ϑ.

c) From b) derive the relation

$$\sum_{\nu=0}^{n} \alpha_\nu \vartheta^{n-\nu} = 0$$

(Cavley-Hamilton theorem).

d) Prove the recursion-formula

$$\alpha_p = -\frac{1}{p} \sum_{\nu=0}^{p-1} \alpha_\nu \operatorname{tr} \vartheta^\nu \quad (p = 1 \ldots n).$$

§ 3. The dual isomorphisms

With the help of the dual product we now are ready to define natural isomorphisms

$$S^p(E) \to S_{n-p}(E) \quad \text{and} \quad S_q^p(E) \to S_{n-p}^{n-q}(E)\,.$$

The present chapter is devoted to the investigation of these isomorphisms.

9.6. The operators D^p and D_p. Select two dual determinant-functions Δ and Δ^* in E and E^*. Then the operators D^p and D_p defined by

$$D^p \Phi = \Delta \vee \Phi \quad \Phi \in S^p(E) \tag{9.27}$$

and $\qquad (0 \leqq p \leqq n)$

$$D_p \Psi = \Delta^* \vee \Psi \quad \Psi \in S_p(E) \tag{9.28}$$

determine two linear mappings

$$D^p \colon S^p(E) \to S_{n-p}(E)$$

and $\qquad (0 \leqq p \leqq n)$

$$D^p \colon S_p(E) \to S^{n-p}(E)$$

It will be shown, that these linear mappings have the following fundamental properties:

$$\langle\!\langle D^p \Phi, D_p \Psi \rangle\!\rangle = \langle\!\langle \Phi, \Psi \rangle\!\rangle \qquad \Phi \in S^p(E), \Psi \in S_p(E) \tag{9.29}$$

$$\langle\!\langle D^p \Phi_1, \Phi_2 \rangle\!\rangle = (-1)^{p(n-p)} \langle\!\langle \Phi_1, D^{n-p} \Phi_2 \rangle\!\rangle \quad \Phi_1 \in S^p(E), \Phi_2 \in S^{n-p}(E) \tag{9.30}$$

$$\langle\!\langle D_p \Psi_1, \Psi_2 \rangle\!\rangle = (-1)^{p(n-p)} \langle\!\langle \Psi_1, D_{n-p} \Psi_2 \rangle\!\rangle \quad \Psi_1 \in S_p(E), \Psi_2 \in S_{n-p}(E). \tag{9.31}$$

We start with the proof of formula (9.29). Employing the formulas (9.10) and (9.14) we obtain

$$D^p \Phi \otimes D_p \Psi = (\Delta \vee \Phi) \otimes (\Delta^* \vee \Psi) = (\Delta \otimes \Delta^*) \vee (\Phi \otimes \Psi) = J^n \vee (\Phi \otimes \Psi)\,.$$

Scalar multiplication by J^{n-p} and use of (9.10), (9.8) and (9.22) yields

$$\begin{aligned} \langle\!\langle D^p \Phi \otimes D_p \Psi, J^{n-p} \rangle\!\rangle &= \langle\!\langle J^n \vee (\Phi \otimes \Psi), J^{n-p} \rangle\!\rangle \\ &= \langle\!\langle J^n \vee J^{n-p}, \Phi \otimes \Psi \rangle\!\rangle = \langle\!\langle J^p, \Phi \otimes \Psi \rangle\!\rangle\,. \end{aligned} \tag{9.32}$$

Applying the relation (9.23) on both sides of (9.32) we obtain for the left hand-side

$$\langle\langle D^p \Phi \otimes D_p \Psi, J^{n-p} \rangle\rangle = \langle\langle D^p \Phi, D_p \Psi \rangle\rangle \tag{9.33}$$

and for the right hand-side

$$\langle\langle J^p, \Phi \otimes \Psi \rangle\rangle = \langle\langle \Phi, \Psi \rangle\rangle . \tag{9.34}$$

The equations (9.32), (9.33) and (9.44) imply (9.29).

The formula (9.30) is a consequence of (9.8):

$$\langle\langle D^p \Phi_1, \Phi_2 \rangle\rangle = \langle\langle \Delta \vee \Phi_1, \Phi_2 \rangle\rangle = (-1)^{p(n-p)} \langle\langle \Delta \vee \Phi_2, \Phi_1 \rangle\rangle$$
$$= (-1)^{p(n-p)} \langle\langle D^{n-p} \Phi_2, \Phi_1 \rangle\rangle .$$

The relation (9.31) is proved in the same way.

The formulas (9.30) and (9.31) show that the linear mappings

$$D^p \colon S^p(E) \to S_{n-p}(E)$$

and

$$D^{n-p} \colon S^{n-p}(E) \to S_p(E)$$

are dual to each other up to a sign-factor,

$$D^{n-p} = (-1)^{p(n-p)} (D^p)^\star .$$

Combining (9.29) and (9.31) we obtain

$$\langle\langle \Phi, \Psi \rangle\rangle = \langle\langle D^p \Phi, D_p \Psi \rangle\rangle = (-1)^{p(n-p)} \langle\langle D_{n-p} D^p \Phi, \Psi \rangle\rangle$$
$$\Phi \in S^p(E), \Psi \in S_p(E) .$$

In view of the duality of the spaces $S^p(E)$ and $S_p(E)$ this relation implies that

$$D_{n-p} \circ D^p = (-1)^{p(n-p)} I \tag{9.35}$$

(I identity-operator) *showing that D^p is an isomorphism of S^p onto S_{n-p}* and $(-1)^{p(n-p)} D_{n-p}$ *is the inverse isomorphism**).

9.7. The duals of the basis-tensors. Let $\overset{*}{e}{}^\nu$, e_ν $(\nu = 1 \ldots n)$ be a pair of dual bases such that

$$\Delta (e_1 \ldots e_n) = 1$$

where Δ is the determinant-function used in the definition of D^p. Then

$$\Delta = \overset{*}{e}{}^1 \wedge \cdots \wedge \overset{*}{e}{}^n \quad \text{and} \quad \Delta^* = e_1 \wedge \cdots \wedge e_n .$$

We know from sec. 8.7 that the products

$$e_{\nu_1} \wedge \cdots \wedge e_{\nu_p} \; (\nu_1 < \cdots < \nu_p) \quad \text{and} \quad \overset{*}{e}{}^{\nu_{p+1}} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_n} \; (\nu_{p+1} < \cdots < \nu_n)$$

*) It is to be observed that the isomorphisms D^p and D_p depend on the choice of the determinant-function Δ. If Δ is replaced by $\lambda\Delta$, where $\lambda \neq 0$ is a scalar, then D^p is multiplied by λ and D_p is multiplied by $\frac{1}{\lambda}$.

form a basis of $S^p(E)$ and $S_{n-p}(E)$ respectively. It will be shown that

$$D^p\,(e_{\nu_1}\wedge\cdots\wedge e_{\nu_p}) = (-1)^{\sum_{i=1}^{p}(\nu_i - i)}\,\overset{*}{e}{}^{\nu_{p+1}}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n} \tag{9.36}$$

where $(\nu_{p+1}\dots\nu_n)$ is the complementary $(n-p)$-tuple of $(\nu_1\dots\nu_p)$ in the natural order. To prove the formula (9.36) consider the permutation

$$\sigma\colon (1\dots n)\to(\nu_1\dots\nu_n)\,.$$

Then

$$\overset{*}{e}{}^1\wedge\cdots\wedge\overset{*}{e}{}^n = \varepsilon_\sigma\,\overset{*}{e}{}^{\nu_1}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n}$$

and hence

$$D^p\,(e_{\nu_1}\wedge\cdots\wedge e_{\nu_p}) = (\overset{*}{e}{}^1\wedge\cdots\wedge\overset{*}{e}{}^n)\vee(e_{\nu_1}\wedge\cdots\wedge e_{\nu_p})$$

$$= \varepsilon_\sigma\,(\overset{*}{e}{}^{\nu_1}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n})\vee(e_{\nu_1}\wedge\cdots\wedge e_{\nu_p})\,. \tag{9.37}$$

Now let $(\mu_{p+1}\dots\mu_n)$ be an arbitrary $(n-p)$-tuple in the natural order. Using the formula (9.4) we obtain

$$\langle\!\langle(\overset{*}{e}{}^{\nu_1}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n})\vee(e_{\nu_1}\wedge\cdots\wedge e_{\nu_p}),\,e_{\mu_{p+1}}\wedge\cdots\wedge e_{\mu_n}\rangle\!\rangle$$

$$= \langle\!\langle\overset{*}{e}{}^{\nu_1}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n},\,e_{\nu_1}\wedge\cdots\wedge e_{\nu_p}\wedge e_{\mu_{p+1}}\wedge\cdots\wedge e_{\mu_n}\rangle\!\rangle\,.$$

This scalar-product is equal to one if $\mu_j=\nu_j$ $(j=p+1\dots n)$ and equal to zero in all other cases. This implies that

$$(\overset{*}{e}{}^{\nu_1}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n})\vee(e_{\nu_1}\wedge\cdots\wedge e_{\nu_p}) = \overset{*}{e}{}^{\nu_{p+1}}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n}\,. \tag{9.38}$$

The formulas (9.37) and (9.38) yield

$$D^p\,(e_{\nu_1}\wedge\cdots\wedge e_{\nu_p}) = \varepsilon_\sigma\,\overset{*}{e}{}^{\nu_{p+1}}\wedge\cdots\wedge\overset{*}{e}{}^{\nu_n}\,. \tag{9.39}$$

It follows from the inequalities $\nu_1<\cdots<\nu_p$ and $\nu_{p+1}<\cdots<\nu_n$ that

$$\varepsilon_\sigma = (-1)^{\sum_{i=1}^{n}(\nu_i - i)}$$

Inserting this into (9.39) we obtain the formula (9.36).

9.8. The isomorphism of $S^p_q(E)$ onto $S^{n-q}_{n-p}(E)$. For every skew-symmetric tensor Φ of order (p, q) denote by $*\Phi$ the skew-symmetric tensor of order $(n-q, n-p)$ defined by

$$*\Phi = J^n\vee\Phi\,. \tag{9.40}$$

Then the correspondence $\Phi\to*\Phi$ determines a linear mapping

$$*\colon S^p_q(E)\to S^{n-q}_{n-p}(E)\,.$$

This mapping has the following properties:

$$\langle\!\langle *\Phi, *\Psi\rangle\!\rangle = \langle\!\langle \Phi, \Psi\rangle\!\rangle \qquad \Phi \in S^p_q(E), \Psi \in S^q_p(E) \tag{9.41}$$

$$\langle\!\langle *\Phi_1, \Phi_2\rangle\!\rangle = (-1)^{p(n-p)+q(n-q)} \langle\!\langle \Phi_1, *\Phi_2\rangle\!\rangle$$
$$\Phi_1 \in S^p_q(E), \Phi_2 \in S^{n-p}_{n-q}(E) \tag{9.42}$$

$$**\Phi = (-1)^{p(n-p)+q(n-q)}\Phi \qquad \Phi \in S^p_q(E) \tag{9.43}$$

$$*(\Phi \wedge \Psi) = *\Phi \vee \Psi \qquad \Phi \in S^p_q(E), \Psi \in S^r_s(E) \tag{9.44}$$

$$*J^p = J^{n-p}. \tag{9.45}$$

Before proving the above formulas we note the following relation between the operator $*$ and the two operators D^p and D_p (cf. sec. 9.6):

$$*(\Phi_1 \otimes \Phi_2) = D^p\Phi_1 \otimes D_q\Phi_2 \qquad \Phi_1 \in S^p(E), \Phi_2 \in S_q(E)\,. \tag{9.46}$$

In fact, the formulas (9.40) and (9.10) yield

$$*(\Phi_1 \otimes \Phi_2) = J^n \vee (\Phi_1 \otimes \Phi_2) = (\Delta^* \otimes \Delta) \vee (\Phi_1 \otimes \Phi_2)$$
$$= (\Delta^* \vee \Phi_2) \otimes (\Delta \vee \Phi_1) = D^p\Phi_1 \otimes D_q\Phi_2\,.$$

Next, we observe that it is sufficient to prove the formula (9.41) for two tensors Φ and Ψ of the form

$$\Phi = \Phi_1 \otimes \Phi_2 \qquad \Phi_1 \in S^p(E),\ \Phi_2 \in S_q(E)$$

and

$$\Psi = \Psi_1 \otimes \Psi_2 \qquad \Psi_1 \in S^q(E),\ \Psi_2 \in S_p(E)$$

because the spaces $S^p_q(E)$ and $S^q_p(E)$ are generated by these products. Using (9.46), we obtain

$$*\Phi = J^n \vee \Phi = (\Delta^* \otimes \Delta) \vee (\Phi_1 \otimes \Phi_2)$$
$$= (\Delta^* \vee \Phi_2) \otimes (\Delta \vee \Phi_1) = D_q\Phi_2 \otimes D^p\Phi_1 \tag{9.47}$$

and

$$*\Psi = J^n \vee \Psi = (\Delta^* \otimes \Delta) \vee (\Psi_1 \otimes \Psi_2)$$
$$= (\Delta^* \vee \Psi_2) \otimes (\Delta \vee \Psi_1) = D_p\Psi_2 \otimes D^q\Psi_1\,. \tag{9.48}$$

These relations and the formula (9.10) yield

$$\langle\!\langle *\Phi, *\Psi\rangle\!\rangle = \langle\!\langle D_q\Phi_2 \otimes D^p\Phi_1, D_p\Psi_2 \otimes D^q\Psi_1\rangle\!\rangle$$
$$= \langle\!\langle D_q\Phi_2, D^q\Psi_1\rangle\!\rangle \langle\!\langle D^p\Phi_1, D_p\Psi_2\rangle\!\rangle\,. \tag{9.49}$$

It follows from the formula (9.29) that

$$\langle\!\langle D_q\Phi_2, D^q\Psi_1\rangle\!\rangle = \langle\!\langle \Phi_2, \Psi_1\rangle\!\rangle \tag{9.50}$$

and

$$\langle\!\langle D^p\Phi_1, D_p\Psi_2\rangle\!\rangle = \langle\!\langle \Phi_1, \Psi_2\rangle\!\rangle\,. \tag{9.51}$$

The equations (9.49), (9.50) and (9.51) yield

$$\begin{aligned}\langle\langle *\Phi, *\Psi\rangle\rangle &= \langle\langle \Phi_2, \Psi_1\rangle\rangle\langle\langle \Phi_1, \Psi_2\rangle\rangle \\ &= \langle\langle \Phi_1 \otimes \Phi_2, \Psi_1 \otimes \Psi_2\rangle\rangle = \langle\langle \Phi, \Psi\rangle\rangle,\end{aligned}$$

which proves (9.41). The formula (9.42) follows from (9.8):

$$\begin{aligned}\langle\langle *\Phi_1, \Phi_2\rangle\rangle = \langle\langle J^n \vee \Phi_1, \Phi_2\rangle\rangle &= (-1)^{p(n-p)+q(n-q)} \langle\langle J^n \vee \Phi_2, \Phi_1\rangle\rangle \\ &= (-1)^{p(n-p)+q(n-q)} \langle\langle *\Phi_2, \Phi_1\rangle\rangle .\end{aligned}$$

To prove (9.43), let $\Psi \in S^q_p(E)$ be an arbitrary tensor. Employing the relations (9.42) and (9.41) we obtain

$$\langle\langle **\Phi, \Psi\rangle\rangle = (-1)^{p(n-p)+q(n-q)} \langle\langle *\Phi, *\Psi\rangle\rangle = (-1)^{p(n-p)+q(n-q)} \langle\langle \Phi, \Psi\rangle\rangle$$

whence (9.43). The relation (9.44) is a consequence of (9.7):

$$*(\Phi \wedge \Psi) = J^n \vee (\Phi \wedge \Psi) = (J^n \vee \Phi) \vee \Psi = *\Phi \vee \Psi .$$

The formula (9.45) follows from (9.22).

The relation (9.41) shows that the *mapping* $\Phi \to *\Phi$ *preverses the scalar-product and hence it defines an isomorphism of* $S^p_q(E)$ *onto* $S^{n-q}_{n-p}(E)$. *From* (9.42) *it follows that any two isomorphisms*

$$S^p_q(E) \xrightarrow{*} S^{n-q}_{n-p}(E)$$

and

$$S^{n-p}_{n-q}(E) \xrightarrow{*} S^q_p(E)$$

are dual up to a sign-factor. The relation (9.43) *implies that the composite isomorphism*

$$S^p_q(E) \xrightarrow{*} S^{n-q}_{n-p}(E) \xrightarrow{*} S^p_q(E)$$

is the identity up to the factor $(-1)^{p(n-p)+q(n-q)}$.

Finally we note that

$$\begin{aligned}&*(e_{\nu_1} \wedge \cdots \wedge e_{\nu_p} \otimes \overset{*}{e}{}^{\mu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\mu_q}) \\ &= (-1)^{\sum_{i=1}^{p}(\nu_i - i)} (-1)^{\sum_{j=1}^{q}(\mu_j - j)} e_{\mu_{q+1}} \wedge \cdots \wedge e_{\mu_n} \otimes \overset{*}{e}{}^{\nu_{p+1}} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_n} \qquad (9.52)\end{aligned}$$

for every pair of dual bases. In this formula, $(\mu_{q+1} \ldots \mu_n)$ denotes the complementary $(n-q)$-tuple of $(\mu_1 \ldots \mu_q)$ and $(\nu_{p+1} \ldots \nu_n)$ is the complementary $(n-p)$-tuple of $(\nu_1 \ldots \nu_p)$. The relation (9.52) is a consequence of the formulas (9.46) and (9.36)

Problems: 1. Show that the components of the tensor $*\Phi$ are given by the formula

$$(*\Phi)^{\mu_{q+1}\ldots\mu_n}_{\nu_{p+1}\ldots\nu_n} = (-1)^{\sum_{i=1}^{p}(\nu_i - i) + \sum_{j=1}^{q}(\mu_j - j)} \Phi^{\nu_1\ldots\nu_p}_{\mu_1\ldots\mu_q}$$

$$(\mu_{q+1} < \cdots < \mu_n,\ \nu_{p+1} < \cdots < \nu_n),$$

where $(\nu_1 \ldots \nu_p)$ is the complementary p-tuple of $(\nu_{p+1} \ldots \nu_n)$ and $(\mu_1 \ldots \mu_q)$ is the complementary q-tuple of $(\mu_{q+1} \ldots \mu_n)$ in the natural order.

2. Let Δ be the determinant-function used in the definition of the isomorphism D^p and $e_\nu (\nu = 1 \ldots n)$ be a basis of E, such that $\Delta(e_1 \ldots e_n) = 1$. Prove the formula

$$D^p (x_1 \wedge \cdots \wedge x_p) = \sum_{\nu_1 < \cdots < \nu_p} (-1)^{\sum_{i=1}^{p}(\nu_i - i)}$$
$$\langle \overset{*}{e}{}^{\nu_1}, a_1 \rangle \cdots \langle \overset{*}{e}{}^{\nu_p}, a_p \rangle\, \overset{*}{e}{}^{\nu_{p+1}} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_n}.$$

In the above sum, $(\nu_{p+1} \ldots \nu_n)$ denotes the complementary $(n-p)$-tuple of $(\nu_1 \ldots \nu_p)$ in the natural order.

3. Using the relation in problem 1 prove the *Laplace expansion formula* for a determinant. Given an $n \times n$ matrix A and an integer $p (1 \leqq p \leqq n)$,

$$\det A = \sum_{\nu_1 \ldots \nu_p} (-1)^{\sum_{i=1}^{p}(\nu_i - i)} \det A_{1 \ldots p}^{\nu_1 \ldots \nu_p} \cdot \det A_{p+1 \ldots n}^{\nu_{p+1} \ldots \nu_n}.$$

Hint: Define the vectors x_ν by $x_\nu = \sum_\mu \alpha_\nu^\mu e_\mu$ and the tensors Φ and Ψ by $x_1 \wedge \ldots \wedge x_p$ and $x_{p+1} \wedge \ldots \wedge x_n$ respectively; then apply the following formula:

$$\langle\!\langle D^p \Phi, \Psi \rangle\!\rangle = \langle\!\langle \Delta, \Phi \wedge \Psi \rangle\!\rangle.$$

4. Show that the isomorphism (9.40) reduces to the identity in the spaces S_0^n and S_n^0.

5. Suppose that $n = 2m$ and define the bilinear function (Φ, Ψ) in the space $S_m^m(E)$ by

$$(\Phi, \Psi) = \langle\!\langle * \Phi, \Psi \rangle\!\rangle.$$

a) Show that this bilinear function is symmetric.

b) If φ and ψ are two endomorphisms of a 2-dimensional space prove that

$$(\mathscr{A} \varphi, \mathscr{A} \psi) = \operatorname{tr}(\varphi \circ \psi) - \operatorname{tr} \varphi \cdot \operatorname{tr} \psi.$$

6. Verify the formula

$$*(X \vee \Phi) = (-1)^{(k+l+p+q)(n+1)} * X \vee \Phi$$
$$X \in S_k^l(E), \quad \Phi \in S_q^p(E) \quad (l \geqq s, k \geqq r).$$

7. Let E be a 2-dimensional linear space and Δ be the determinant-function which is used to define the isomorphism $D^1: S^1(E) \to S_1(E)$. Identifying $S^1(E)$ with E and $S_1(E)$ with E^* we obtain an isomorphism $D: E \to E^*$. Prove that

$$\Delta(x, y) = \langle D x, y \rangle \qquad x, y \in E.$$

8. Prove that the isomorphism $*$ is a tensorial mapping but that the isomorphisms D^p and D_p are not tensorial mappings.

9. Consider two decomposable tensors

$$\Phi = a_1 \wedge \cdots \wedge a_p \quad a_\nu \in E \quad (\nu = 1 \ldots p)$$

and

$$\Psi = \overset{*}{a}{}^1 \wedge \cdots \wedge \overset{*}{a}{}^q \quad \overset{*}{a}{}^\mu \in E^* \; (\mu = 1 \ldots q)\, .$$

Show that the tensors $D^p\Phi$, $D_q\Psi$ and $*\,(\Phi \otimes \Psi)$ can be written in the following form

$$(D^p\Phi)\,(x_{p+1} \ldots x_n) = \Delta\,(a_1 \ldots a_p, x_{p+1} \ldots x_n)$$

$$(D_q\Psi)\,(\overset{*}{x}{}^{q+1} \ldots \overset{*}{x}{}^n) = \Delta\,(\overset{*}{a}{}^1 \ldots \overset{*}{a}{}^q, \overset{*}{x}{}^{q+1} \ldots \overset{*}{x}{}^n)$$

and

$$*\,(\Phi \otimes \Psi)\,(\overset{*}{x}{}^{p+1} \ldots \overset{*}{x}{}^n; x_{p+1} \ldots x_n)$$
$$= J^n\,(\overset{*}{a}{}^1 \ldots \overset{*}{a}{}^q, \overset{*}{x}{}^{q+1} \ldots \overset{*}{x}{}^n; a_1 \ldots a_p, x_{p+1} \ldots x_n)\, .$$

10. Consider p linearly independent vectors $a_\nu \in E$ and denote by U the subspace which is generated by these vectors. Prove that the tensor $D^p\,(a_1 \wedge \cdots \wedge a_p)$ can be written as

$$D^p\,(a_1 \wedge \cdots \wedge a_p) = \overset{*}{a}{}^{p+1} \wedge \cdots \wedge \overset{*}{a}{}^n$$

where the vectors $\overset{*}{a}{}^\nu$ $(\nu = p + 1 \ldots n)$ form a basis of $U^\perp$.

11. Let Φ be a skew-symmetric contravariant tensor of order p. Define the subspace $V(\Phi)$ of E as the set of all vectors $v \in E$ for which $\Phi \wedge v = 0$. Prove the relation

$$V(\Phi) = U(D^p\Phi)$$

where $U(D^p\Phi)$ is the subspace of E defined by the covariant tensor $D^p\Phi$ (cf. sec. 8.9.).

§ 4. The adjoint tensor

9.9. The endomorphism $*\alpha$. Let α be an endomorphism of $S^p(E)$. Then an endomorphism

$$*\alpha\colon S_{n-p}(E) \to S_{n-p}(E)$$

is determined by

$$*\alpha = D^p \circ \alpha \circ (D^p)^{-1}. \tag{9.53}$$

It follows immediately from the definition (9.53) that

$$*\,(\beta \circ \alpha) = *\,\beta \circ *\alpha \tag{9.54}$$

and

$$**\alpha = \alpha\, .$$

The connection between α and $*\alpha$ can be expressed by the commutative diagram

$$\begin{array}{ccc} S^p(E) & \xrightarrow{\alpha} & S^p(E) \\ {\scriptstyle D^p}\downarrow & & \downarrow{\scriptstyle D^p} \\ S_{n-p}(E) & \xrightarrow{*\alpha} & S_{n-p}(E) \end{array}$$

Now consider the isomorphisms

$$\mathscr{A}: L\,(S^p(E)\,;\,S^p(E)) \to S^p_p(E)$$

and

$$\mathscr{A}: L\,(S_{n-p}(E)\,;\,S_{n-p}(E)) \to S^{n-p}_{n-p}(E)$$

defined by

$$\langle\!\langle \mathscr{A}\alpha, \Phi \otimes \Phi^* \rangle\!\rangle = \langle\!\langle \alpha\Phi, \Phi^* \rangle\!\rangle \quad \Phi \in S^p(E), \Phi^* \in S_p(E) \tag{9.55}$$

and

$$\langle\!\langle \mathscr{A}\,\beta, \Psi \otimes \Psi^* \rangle\!\rangle = \langle\!\langle \beta\Psi, \Psi^* \rangle\!\rangle \quad \Psi \in S_{n-p}(E), \Psi^* \in S^{n-p}(E) \tag{9.56}$$

respectively (cf. sec. 8.16). It will be shown that

$$\mathscr{A} * \alpha = * \mathscr{A}\alpha\,. \tag{9.57}$$

In other words, the diagram

$$\begin{array}{ccc} L\,(S^p(E)\,;S^p(E)) & \xrightarrow{*} & L\,(S_{n-p}(E)\,;\,S_{n-p}(E)) \\ {\scriptstyle \mathscr{A}}\downarrow & & \downarrow{\scriptstyle \mathscr{A}} \\ S^p_p(E) & \xrightarrow{*} & S^{n-p}_{n-p}(E) \end{array}$$

is commutative.

To prove the relation (9.57) consider two arbitrary tensors $\Phi \in S^{n-p}(E)$ and $\Phi^* \in S_{n-p}(E)$. Then the formulas (8.65), (9.53) and (9.35) yield

$$\begin{aligned} &\langle\!\langle \mathscr{A} * \alpha, \Phi^* \otimes \Phi \rangle\!\rangle = \langle\!\langle * \alpha\Phi^*, \Phi \rangle\!\rangle = \langle\!\langle D^p\,\alpha\,(D^p)^{-1}\Phi^*, \Phi \rangle\!\rangle \\ &= (-1)^{p(n-p)} \langle\!\langle D^p\,\alpha\,D_{n-p}\Phi^*, \Phi \rangle\!\rangle\,. \end{aligned} \tag{9.58}$$

At the same time we obtain from (9.42), (9.46), (8.64) and (9.30)

$$\begin{aligned} &\langle\!\langle * \mathscr{A}\alpha, \Phi^* \otimes \Phi \rangle\!\rangle = \langle\!\langle \mathscr{A}\alpha, * \,(\Phi^* \otimes \Phi) \rangle\!\rangle \\ &= \langle\!\langle \mathscr{A}\alpha, D^{n-p}\Phi \otimes D_{n-p}\Phi^* \rangle\!\rangle = \langle\!\langle (\alpha \circ D_{n-p})\,\Phi^*, D^{n-p}\Phi \rangle\!\rangle \\ &= (-1)^{p(n-p)} \langle\!\langle (D^p \circ \alpha \circ D_{n-p})\,\Phi^*, \Phi \rangle\!\rangle\,. \end{aligned} \tag{9.59}$$

Comparing the relations (9.58) and (9.59) we find that

$$\langle\!\langle \mathscr{A} * \alpha, \Phi^* \otimes \Phi \rangle\!\rangle = \langle\!\langle * \mathscr{A}\alpha, \Phi^* \otimes \Phi \rangle\!\rangle$$

whence (9.57).

9.10. Application to induced endomorphisms. Let us now consider the case that the endomorphism α is induced by an endomorphism ϑ of E,

$$\alpha = (\vartheta)^p .$$

It will be shown that

$$(\vartheta)_{n-p} \circ * (\vartheta)^p = I \cdot \det\vartheta \tag{9.60}$$

and

$$(\vartheta)^{n-p} \circ * (\vartheta)_p = I \cdot \det\vartheta . \tag{9.61}$$

where I is, as usual, the identity-operator.

Applying the formula (9.12) to the tensor $X = \Delta$ and observing that

$$(\vartheta)_n \Delta = \det\vartheta \cdot \Delta$$

we obtain

$$(\vartheta)_{n-p} D^p (\vartheta)^p \Phi = \det\vartheta \cdot D^p\Phi \qquad \Phi \in S^p(E) .$$

Written as an operator-identity this relation reads

$$(\vartheta)_{n-p} \circ D^p \circ (\vartheta)^p = \det\vartheta \cdot D^p .$$

Multiplying from the right by $(D^p)^{-1}$ we obtain

$$(\vartheta)_{n-p} \circ \underbrace{D^p \circ (\vartheta)^p \circ (D^p)^{-1}}_{*(\vartheta)^p} = I \cdot \det\vartheta$$

whence (9.60). The formula (9.61) is proved in the same way.

It will be shown in sec. 9.14 that the relation (9.60) is essentially the Laplace expansion formula for a determinant.

9.11. The adjoint endomorphism. Let ϑ be an endomorphism of E and $(\vartheta)_{n-1}$ the induced endomorphism in the space $S_{n-1}(E)$. Applying the operation $*$ to $(\vartheta)_{n-1}$ we obtain again an endomorphism of E. This endomorphism is called the *classical adjoint* of ϑ and will be denoted by $\mathrm{ad}\vartheta$,

$$\mathrm{ad}\vartheta = * (\vartheta)_{n-1} . \tag{9.62}$$

The adjoint of the product of two endomorphisms is the product of the adjoint mappings with the factors reversed,

$$\mathrm{ad}\,(\vartheta_2 \circ \vartheta_1) = \mathrm{ad}\vartheta_1 \circ \mathrm{ad}\vartheta_2 . \tag{9.63}$$

In fact, the formulas (9.62) and (9.54) yield

$$\begin{aligned} \mathrm{ad}\,(\vartheta_2 \circ \vartheta_1) &= * (\vartheta_2 \circ \vartheta_1)_{n-1} = * ((\vartheta_1)_{n-1} \circ (\vartheta_2)_{n-1}) \\ &= * (\vartheta_1)_{n-1} \circ * (\vartheta_2)_{n-1} = \mathrm{ad}\vartheta_1 \circ \mathrm{ad}\vartheta_2 . \end{aligned}$$

Employing the relation (9.61) for the case $p = n - 1$ we obtain the formula

$$\vartheta \circ \operatorname{ad}\vartheta = \iota \cdot \det\vartheta \tag{9.64}$$

showing that

$$\operatorname{ad}\vartheta = \vartheta^{-1} \cdot \det\vartheta$$

for a regular endomorphism.

It follows from (9.64), that $\operatorname{ad}\vartheta$ is regular whenever ϑ is regular. Conversely, the regularity of $\operatorname{ad}\vartheta$ implies the regularity of ϑ. In fact, it follows from (9.64), that ϑ is either regular or the zero-mapping. But the second case is impossible, because $\operatorname{ad} 0 = 0$.

Passing over to the determinant in (9.64) we find the relation

$$\det \operatorname{ad}\vartheta = (\det\vartheta)^{n-1} \tag{9.65}$$

for every regular endomorphism ϑ. This formula is still correct for a non-regular endomorphism because then both sides are zero.

It should be observed that the correspondence $\vartheta \to \operatorname{ad}\vartheta$ is not linear unless E has the dimension 2.

9.12. The adjoint tensor. In analogy with (9.62) we define the *adjoint* of a tensor $\theta \in S_1^1(E)$ by

$$\operatorname{ad}\theta = *\,\theta^{n-1}\,.$$

Then the operations $\mathscr{A}$ and ad commute,

$$\operatorname{ad}\mathscr{A}\,\vartheta = \mathscr{A}\,\operatorname{ad}\vartheta\,. \tag{9.66}$$

In fact,

$$\mathscr{A}\,\operatorname{ad}\vartheta = \mathscr{A}\,*(\vartheta)_{n-1} = *\,\mathscr{A}\,(\vartheta)_{n-1} = *\,(\mathscr{A}\vartheta)^{n-1} = \operatorname{ad}\mathscr{A}\vartheta\,.$$

It is the purpose of this section to prove the formula

$$*\,(\operatorname{ad}\theta)^p = \theta^{n-p}\,(\det\vartheta)^{p-1} \quad \text{where } \theta = \mathscr{A}\vartheta \quad (1 \leqq p \leqq n)\,. \tag{9.67}$$

This relation will follow from the formula

$$\theta^k \vee (\operatorname{ad}\theta)^p = \binom{n+p-k}{p}\,\theta^{k-p}\,(\det\vartheta)^p \quad (0 \leqq p \leqq k)\,. \tag{9.68}$$

We shall first prove (9.68) for the case $p = 1$,

$$\theta^k \vee \operatorname{ad}\theta = (n - k + 1)\,\theta^{k-1}\det\vartheta\,. \tag{9.69}$$

To prove (9.69) let $\overset{*}{e}{}^\nu, e_\nu$ $(\nu = 1 \ldots n)$ be a pair of dual bases. Then the formula (9.9) yields

$$\begin{aligned}(\theta^k \vee \operatorname{ad}\theta)\,&(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; x_1 \ldots x_{k-1}) \\ &= \sum_{\nu,\mu} \theta^k\,(\overset{*}{e}{}^\nu, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; e_\mu, x_1 \ldots x_{k-1})\,\beta_\nu^\mu\end{aligned} \tag{9.70}$$

where β_ν^μ are the components of the tensor $\operatorname{ad}\theta$,

$$\beta_\nu^\mu = (\operatorname{ad}\theta)\,(\overset{*}{e}{}^\mu, e_\nu) = \langle \overset{*}{e}{}^\mu, \operatorname{ad}\vartheta\, e_\nu\rangle\,. \tag{9.71}$$

Observing that the tensors θ^k and J^k are connected by the equation

$$\theta^k(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^k; x_1 \ldots x_k) = J^k(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^k, \vartheta x_1 \ldots \vartheta x_k)$$

we can write the right hand-side of (9.70) as

$$\sum_{\nu,\mu} J^k(\overset{*}{e}{}^\nu, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; \vartheta e_\mu, \vartheta x_1 \ldots \vartheta x_{k-1})\, \beta^\mu_\nu$$
$$= \sum_\nu J^k(\overset{*}{e}{}^\nu, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; \vartheta \sum_\mu \beta^\mu_\nu e_\mu, \vartheta x_1 \ldots \vartheta x_{k-1})\,. \tag{9.72}$$

Now it follows from (9.64) that

$$\vartheta \sum_\mu \beta^\mu_\nu e_\mu = (\vartheta \circ \operatorname{ad}\vartheta)\, e_\nu = \det\vartheta \cdot e_\nu$$

and hence the sum (9.72) is equal to

$$\det\vartheta \sum_\nu J^k(\overset{*}{e}{}^\nu, \overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; e_\nu, \vartheta x_1 \ldots \vartheta x_{k-1})$$
$$= \det\vartheta\, (C^1_1 J^k)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; \vartheta x_1 \ldots \vartheta x_{k-1})\,.$$

But, in view of (9.16)

$$C^1_1 J^k = (n - k + 1)\, J^{k-1}$$

and we thus obtain altogether

$$(\theta^k \vee \operatorname{ad}\theta)(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; x_1 \ldots x_{k-1})$$
$$= (n - k + 1)\, J^{k-1}(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; \vartheta x_1 \ldots \vartheta x_{k-1}) \det\vartheta$$
$$= (n - k + 1)\, \theta^{k-1}(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{k-1}; x_1 \ldots x_{k-1}) \det\vartheta$$

which proves the relation (9.69).

Now assume, by induction, that (9.68) is correct for the exponent $p - 1$,

$$\theta^k \vee (\operatorname{ad}\theta)^{p-1} = \binom{n-k+p-1}{p-1} \theta^{k-p+1} (\det\vartheta)^{p-1}\,. \tag{9.73}$$

Then the equations (9.73) and (9.69) yield

$$\theta^k \vee (\operatorname{ad}\theta)^p = \frac{1}{p} (\theta^k \vee (\operatorname{ad}\theta)^{p-1}) \vee \operatorname{ad}\theta$$
$$= \frac{1}{p} \binom{n+p-k-1}{p-1} \theta^{k-p+1} \vee \operatorname{ad}\theta\, (\det\vartheta)^{p-1}$$
$$= \frac{n+p-k}{p} \binom{n+p-k-1}{p-1} \theta^{k-p} (\det\vartheta)^p$$
$$= \binom{n+p-k}{p} \theta^{k-p} (\det\vartheta)^p$$

showing that (9.68) is correct for the exponent p. This proves formula (9.68).

Inserting $k = n - 1$ into (9.68) and replacing p by $p - 1$ we obtain

$$\theta^{n-1} \vee (\operatorname{ad}\theta)^{p-1} = p\,\theta^{n-p} (\det\vartheta)^{p-1} \quad (1 \leqq p \leqq n)\,. \tag{9.74}$$

In view of (9.43) and (9.44), the left hand-side of (9.74) can be written as

$$\theta^{n-1} \vee (\mathrm{ad}\,\theta)^{p-1} = * \,\mathrm{ad}\,\theta \vee (\mathrm{ad}\,\theta)^{p-1} =$$
$$= * \,(\mathrm{ad}\,\theta \wedge (\mathrm{ad}\,\theta)^{p-1}) = p * (\mathrm{ad}\,\theta)^p . \tag{9.75}$$

The relations (9.74) and (9.75) imply (9.67).

The formula (9.67) yields for the case $p = n - 1$

$$* \,(\mathrm{ad}\,\theta)^{n-1} = \theta\,(\det\vartheta)^{n-2}. \tag{9.76}$$

Observing that

$$* \,(\mathrm{ad}\,\theta)^{n-1} = \mathrm{ad}\;\mathrm{ad}\,\theta = \mathrm{ad}^2\,\theta$$

we obtain from (9.76) the relation

$$\mathrm{ad}^2\,\theta = \theta\,(\det\vartheta)^{n-2} \tag{9.77}$$

showing that the operation ad^2 consists simply of the multiplication by the factor $(\det\vartheta)^{n-2}$.

Combining the relations (9.66) and (9.77) we find that

$$\mathscr{A}\,\mathrm{ad}^2\vartheta = \mathrm{ad}\,(\mathscr{A}\,\mathrm{ad}\vartheta) = \mathrm{ad}^2\,\mathscr{A}\vartheta = \mathscr{A}\vartheta \cdot (\det\vartheta)^{n-2}$$

whence

$$\mathrm{ad}^2\vartheta = \vartheta \cdot (\det\vartheta)^{n-2\,\star)}. \tag{9.78}$$

9.13. The adjoint matrix. Employing a pair of dual bases we obtain from every relation between the mappings ϑ and $\mathrm{ad}\vartheta$ a corresponding relation between the matrices. First of all, we show that the components

$$\theta^\nu_\mu = \theta(\overset{*}{e}{}^\nu, e_\mu) \quad\text{and}\quad (\mathrm{ad}\,\theta)^\nu_\mu = \mathrm{ad}\,\theta\,(\overset{*}{e}{}^\nu, e_\mu)$$

of the tensors θ and $\mathrm{ad}\,\theta$ are connected by the formula

$$(\mathrm{ad}\,\theta)^\mu_\nu = (-1)^{\mu+\nu} \begin{vmatrix} \theta^1_1 & \dots & \hat\theta^\nu_1 & \dots & \theta^n_1 \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \hat\theta^1_\mu & \dots & \hat\theta^\nu_\mu & \dots & \hat\theta^n_\mu \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \theta^1_n & \dots & \hat\theta^\nu_n & \dots & \theta^n_n \end{vmatrix} . \tag{9.79}$$

In fact, the equations (9.42), (9.52) and (8.53) yield

$$(\mathrm{ad}\,\theta)^\mu_\nu = \langle\!\langle \mathrm{ad}\,\theta,\, e_\nu \otimes \overset{*}{e}{}^\mu \rangle\!\rangle = \langle\!\langle *\,\theta^{n-1},\, e_\nu \otimes \overset{*}{e}{}^\mu \rangle\!\rangle = \langle\!\langle \theta^{n-1},\, *\,(e_\nu \otimes \overset{*}{e}{}^\mu) \rangle\!\rangle$$
$$= (-1)^{\nu+\mu} \langle\!\langle \theta^{n-1},\, e_1 \wedge \cdots \wedge \hat e_\mu \wedge \dots e_n \otimes \overset{*}{e}{}^1 \wedge \cdots \wedge \hat{\overset{*}{e}}{}^\nu \vee \dots \overset{*}{e}{}^n \rangle\!\rangle ,$$

whence (9.79).

Now consider the matrices

$$\alpha^\mu_\nu = \langle \overset{*}{e}{}^\mu, \vartheta e_\nu \rangle \quad\text{and}\quad \beta^\mu_\nu = \langle \overset{*}{e}{}^\mu, (\mathrm{ad}\vartheta)\, e_\nu \rangle$$

$\star$) If ϑ is regular this relation is an immediate consequence of (9.64).

of the endomorphisms ϑ and $\mathrm{ad}\vartheta$. The equations

$$\langle \overset{*}{x}, \vartheta x\rangle = \theta(\overset{*}{x}, x)$$

and

$$\langle \overset{*}{x}, (\mathrm{ad}\vartheta)\, x\rangle = (\mathrm{ad}\theta)\,(\overset{*}{x}, x)$$

imply that

$$\alpha_\nu^\mu = \theta_\nu^\mu \quad \text{and} \quad \beta_\nu^\mu = (\mathrm{ad}\,\theta)_\nu^\mu\,. \tag{9.80}$$

The relations (9.80), and (9.79) yield the formula

$$\beta_\nu^\mu = (-1)^{\nu+\mu} \begin{vmatrix} \alpha_1^1 & \dots & \hat{\alpha}_1^\nu & \dots & \alpha_1^n \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \hat{\alpha}_\mu^1 & \dots & \hat{\alpha}_\mu^\nu & \dots & \hat{\alpha}_\mu^n \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \alpha_n^1 & \dots & \hat{\alpha}_n^\nu & \dots & \alpha_n^n \end{vmatrix}$$

showing that the matrix of $\mathrm{ad}\,\theta$ is the adjoint of the matrix of θ (cf. sec. 4.13). Using the notation introduced in sec. 3.2 we thus can write

$$M\,(\mathrm{ad}\vartheta) = \mathrm{ad}M\,(\vartheta)\,.$$

Now we obtain the following formulas which correspond to the relations (9.63), (9.64), (9.65) and (9.78)*):

$$\begin{aligned} \mathrm{ad}\,(A\,B) &= \mathrm{ad}B \cdot \mathrm{ad}A \\ \mathrm{ad}A \cdot A &= J \cdot \det A \qquad (J \text{ unit-matrix}) \\ \det\mathrm{ad}\,A &= (\det A)^{n-1} \\ \mathrm{ad}^2 A &= A\,(\det A)^{n-2}\,. \end{aligned}$$

9.14. The Laplace expansion formula. In concluding this section we shall derive the Laplace expansion formula for a determinant from the relation (9.60). Let e_ν $(\nu = 1 \dots n)$ be a basis of E and $A = (\alpha_\nu^\mu)$ the corresponding matrix of ϑ. Then it follows from (9.60) that

$$((\vartheta)_p \circ *(\vartheta)^{n-p})\,(\overset{*}{e}^{\nu_1} \wedge \dots \wedge \overset{*}{e}^{\nu_p}) = \overset{*}{e}^{\nu_1} \wedge \dots \wedge \overset{*}{e}^{\nu_p} \cdot \det A \qquad (\nu_1 < \dots < \nu_p)\,.$$

Scalar multiplication by $e_{\nu_1} \wedge \dots \wedge e_{\nu_p}$ and use of (7.41) yields

$$\begin{aligned} \det A &= \langle\langle e_{\nu_1} \wedge \dots \wedge e_{\nu_p}, ((\vartheta)_p \circ *(\vartheta)^{n-p})\,(\overset{*}{e}^{\nu_1} \wedge \dots \wedge \overset{*}{e}^{\nu_p})\rangle\rangle \\ &= \langle\langle(\vartheta)^p\,(e_{\nu_1} \wedge \dots \wedge e_{\nu_p}), *(\vartheta)^{n-p}\,(\overset{*}{e}^{\nu_1} \wedge \dots \wedge \overset{*}{e}^{\nu_p})\rangle\rangle\,. \end{aligned}$$

Now

$$(\vartheta)^p\,(e_{\nu_1} \wedge \dots \wedge e_{\nu_p}) = \sum_{\mu_1 < \dots < \mu_p} \det A_{\nu_1 \dots \nu_p}^{\mu_1 \dots \mu_p}\, e_{\mu_1} \wedge \dots \wedge e_{\mu_p}$$

where $A_{\nu_1 \dots \nu_p}^{\mu_1 \dots \mu_p}$ denotes the submatrix of A consisting of the rows $(\nu_1 \dots \nu_p)$

*) The second one of the above relations has been established in sec. 4.13 in a different way (Cramer's formula).

and the columns $(\mu_1 \ldots \mu_p)$. We thus obtain

$$\det A = \sum_{\mu_1 < \cdots < \mu_p} \det A^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_p} \langle\!\langle e_{\mu_1} \wedge \cdots \wedge e_{\mu_p}, *(\vartheta)^{n-p} (\overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p}) \rangle\!\rangle . \tag{9.81}$$

Applying the formula (9.56) to the scalar product on the right hand-side we find

$$\begin{aligned} &\langle\!\langle e_{\mu_1} \wedge \cdots \wedge e_{\mu_p}, *(\vartheta)^{n-p} (\overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p}) \rangle\!\rangle \\ &= \langle\!\langle \mathscr{A} *(\vartheta)^{n-p}, (e_{\mu_1} \wedge \cdots \wedge e_{\mu_p}) \otimes (\overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p}) \rangle\!\rangle . \end{aligned} \tag{9.82}$$

Now the formulas (9.57) and (8.72) yield

$$\mathscr{A} *(\vartheta)^{n-p} = *\mathscr{A}(\vartheta)^{n-p} = *(\mathscr{A}\vartheta)^{n-p} = *\theta^{n-p}, \quad \theta = \mathscr{A}\vartheta .$$

Hence, the right hand-side of (9.82) can be written as

$$\begin{aligned} &\langle\!\langle *\theta^{n-p}, e_{\mu_1} \wedge \cdots \wedge e_{\mu_p} \otimes \overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p} \rangle\!\rangle \\ &= \langle\!\langle \theta^{n-p}, *(e_{\mu_1} \wedge \cdots \wedge e_{\mu_p} \otimes \overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p}) \rangle\!\rangle \\ &= (-1)^{\sum_{i=1}^{p}(\nu_i + \mu_i)} \langle\!\langle \theta^{n-p}, e_{\nu_{p+1}} \wedge \cdots \wedge e_{\nu_n} \otimes \overset{*}{e}{}^{\mu_{p+1}} \wedge \cdots \wedge \overset{*}{e}{}^{\mu_n} \rangle\!\rangle \\ &= (-1)^{\sum_{i=1}^{p}(\nu_i + \mu_i)} \theta^{n-p} (\overset{*}{e}{}^{\mu_{p+1}} \ldots \overset{*}{e}{}^{\mu_n}; e_{\nu_{p+1}} \ldots e_{\nu_n}) , \end{aligned}$$

where $(\nu_{p+1} \ldots \nu_n)$ and $(\mu_{p+1} \ldots \mu_n)$ are the complementary $(n-p)$-tuples in the natural order (cf. (9.42), (9.52) and (8.58)).

Observing that θ has the components

$$\theta^\nu_\mu = \theta(\overset{*}{e}{}^\nu, e_\mu) = \langle \overset{*}{e}{}^\nu, \vartheta e_\mu \rangle = \alpha^\nu_\mu$$

we see that

$$\theta^{n-p} (\overset{*}{e}{}^{\mu_{p+1}} \ldots \overset{*}{e}{}^{\mu_n}; e_{\nu_{p+1}} \ldots e_{\nu_n}) = \det A^{\mu_{p+1} \ldots \mu_n}_{\nu_{p+1} \ldots \nu_n} .$$

Inserting this into (9.82) we find that

$$\langle\!\langle e_{\mu_1} \wedge \cdots e_{\mu_p}, *\vartheta^{n-p} (\overset{*}{e}{}^{\nu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\nu_p}) \rangle\!\rangle = (-1)^{\sum_{i=1}^{p}(\nu_i + \mu_i)} \det A^{\mu_{p+1} \ldots \mu_n}_{\nu_{p+1} \ldots \nu_n} . \tag{9.83}$$

From (9.81) and (9.83) we finally obtain the *Laplace expansion formula* for the determinant of A:

$$\det A = \sum_{\mu_1 < \cdots < \mu_p} (-1)^{\sum_{i=1}^{p}(\nu_i + \mu_i)} \det A^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_p} \det A^{\mu_{p+1} \ldots \mu_n}_{\nu_{p+1} \ldots \nu_n} \quad (\nu_1 < \cdots < \nu_p) .$$

Problems: 1. Given an endomorphism of a 2-dimensional linear space E show that

$$\operatorname{ad} \vartheta = \iota \operatorname{tr} \vartheta - \vartheta .$$

Verify the relation

$$\operatorname{ad}^2 \vartheta = \vartheta .$$

2. Prove that the ranks of an endomorphism and of the adjoint endomorphism are connected by the relations

$$\begin{aligned} r(\mathrm{ad}\vartheta) &= 0 \quad \text{if} \quad r(\vartheta) \leqq n-2 \\ r(\mathrm{ad}\vartheta) &= 1 \quad \text{if} \quad r(\vartheta) = n-1 \\ r(\mathrm{ad}\vartheta) &= n \quad \text{if} \quad r(\vartheta) = n\,. \end{aligned}$$

Using these relations prove that

$$\mathrm{ad}^2\vartheta = \vartheta\,(\det\vartheta)^{n-2}.$$

3. Derive the formula

$$\langle\!\langle \theta^p, (\mathrm{ad}\,\theta)^p\rangle\!\rangle = \binom{n}{p}(\det\vartheta)^p \quad \theta = \mathscr{A}(\vartheta) \quad (1 \leqq p \leqq n)\,.$$

4. Given an $n \times n$-matrix $A = (\alpha_\nu^\mu)$ define the matrix $(\tilde{\alpha}_\nu^\mu)$ by

$$\tilde{\alpha}_\nu^\mu = \begin{vmatrix} \alpha_1^1 & \dots & \hat{\alpha}_1^\mu & \dots & \alpha_1^n \\ \vdots & & \vdots & & \vdots \\ \hat{\alpha}_\nu^1 & \dots & \hat{\alpha}_\nu^\mu & \dots & \hat{\alpha}_\nu^n \\ \vdots & & \vdots & & \vdots \\ \alpha_n^1 & \dots & \hat{\alpha}_n^\mu & \dots & \alpha_n^n \end{vmatrix} \qquad (\nu, \mu = 1 \dots n)\,.$$

Using the relation (9.67) prove the *Jacobian identities:*

$$\begin{vmatrix} \tilde{\alpha}_{\nu_1}^{\mu_1} & \dots & \tilde{\alpha}_{\nu_p}^{\mu_1} \\ \vdots & & \vdots \\ \tilde{\alpha}_{\nu_1}^{\mu_p} & \dots & \tilde{\alpha}_{\nu_p}^{\mu_p} \end{vmatrix} = \begin{vmatrix} \alpha_{\nu_{p+1}}^{\mu_{p+1}} & \dots & \alpha_{\nu_n}^{\mu_{p+1}} \\ \vdots & & \vdots \\ \alpha_{\nu_{p+1}}^{\mu_n} & \dots & \alpha_{\nu_n}^{\mu_n} \end{vmatrix} (\det A)^{p-1}$$

$$(\nu_1 < \dots < \nu_p,\ \mu_1 < \dots < \mu_p),$$

where $(\mu_{p+1} \dots \mu_n)$ and $(\nu_{p+1} \dots \nu_n)$ are the complementary $(n-p)$-tuples of $(\mu_1 \dots \mu_p)$ and $(\nu_1 \dots \nu_p)$ respectively.

5. Prove that

$$\mathrm{ad}\vartheta = (-1)^{n-1} \sum_{\nu=0}^{n-1} \alpha_\nu \vartheta^{n-\nu-1}$$

where the α_ν are the coefficients in the characteristic polynomial of ϑ. Hint: Use the Cayley-Hamilton theorem (cf. prob. 4, § 2).

6. Show that the determinants of the induced endomorphisms

$$(\vartheta)^p\colon S^p(E) \to S^p(E) \quad \text{and} \quad (\vartheta)^{n-p}\colon S^{n-p}(E) \to S^{n-p}(E)$$

are connected by the relation

$$\det(\vartheta)^p \cdot \det(\vartheta)^{n-p} = (\det\vartheta)^{\binom{n}{p}}.$$

§ 5. The exterior product

Applying the isomorphism

$$D^{n-1}\colon S^{n-1}(E) \to E^*$$

to a skew-symmetric product of $n-1$ vectors $x_\nu \in E$ we obtain an $(n-1)$-linear skew-symmetric mapping of E into E^* which is called the exterior product. However, we shall first give a different definition of the exterior product and establish its relation to the dual product later.

9.15. Definition. Consider a pair of dual spaces E^*, E, and assume that a determinant-function $\Delta \neq 0$ is given in E. Then every system of $(n-1)$ vectors $a_\nu \in E$ defines a linear function f in E by

$$f(x) = \Delta\,(a_1 \ldots a_{n-1}, x) \quad x \in E\,. \tag{9.84}$$

The function f can be written as

$$f(x) = \langle \overset{*}{a}, x\rangle \qquad\qquad \overset{*}{a} \in E^*\,.$$

The vector $\overset{*}{a}$ is uniquely determined by the function f and hence by the vectors $a_\nu\,(\nu = 1 \ldots n-1)$. It is called the *exterior product* of the $(n-1)$ vectors a_ν and will be denoted by $[a_1 \ldots a_{n-1}]$. Now the equation (9.84) can be written as

$$\Delta\,(a_1 \ldots a_{n-1}, x) = \langle [a_1 \ldots a_{n-1}], x\rangle \qquad x \in E\,. \tag{9.85}$$

It follows immediately from the above definition that the exterior product is distributive and skew-symmetric with respect to all factors. Replacing x by one of the vectors $a_\nu (\nu = 1 \ldots n-1)$ we obtain the relation

$$\langle [a_1 \ldots a_{n-1}], a_\nu\rangle = 0 \qquad\qquad (\nu = 1 \ldots n-1)$$

showing that the vector $[a_1 \ldots a_{n-1}]$ is orthogonal to all vectors a_ν $(\nu = 1 \ldots n-1)$. The exterior product of $n-1$ vectors is zero if and only if these vectors are linearly independent. In fact, if the vectors a_ν are linearly dependent, the left hand-side of (9.85) is identically zero, whence $[a_1 \ldots a_{n-1}] = 0$. Conversely, assume that the vectors a_ν $(\nu = 1 \ldots n-1)$ are linearly independent. Then there exists a vector a_n such that $\Delta\,(a_1 \ldots a_n) \neq 0$. According to (9.85), this implies that $[a_1 \ldots a_{n-1}] \neq 0$.

Let $e_\nu\,(\nu = 1 \ldots n)$ be a basis of E such that

$$\Delta\,(e_1 \ldots e_n) = 1 \tag{9.86}$$

and denote by $\overset{*}{e}{}^\nu\,(\nu = 1 \ldots n)$ the dual basis. Then the vector $[e_1 \ldots \hat{e}_\nu \ldots e_n]$ *) is orthogonal to all vectors $e_\mu\,(\mu \neq \nu)$ and hence it must be a multiple of $\overset{*}{e}{}^\nu$,

$$[e_1 \ldots \hat{e}_\nu \ldots e_n] = \lambda_\nu\, \overset{*}{e}{}^\nu \qquad\qquad (\nu = 1 \ldots n)\,.$$

*) The symbol $\hat{e}_\nu$ indicates that the vector e_ν is deleted.

Scalar multiplication with e_ν and use of (9.85) yields

$$\Delta\,(e_1 \ldots \hat{e}_\nu \ldots e_n, e_\nu) = \lambda_\nu\,.$$

It follows from the skew-symmetry of Δ and from (9.86) that

$$\Delta\,(e_1 \ldots \hat{e}_\nu \ldots e_n, e_\nu) = (-1)^{n-\nu}\,\Delta\,(e_1 \ldots e_n) = (-1)^{n-\nu}$$

whence

$$\lambda_\nu = (-1)^{n-\nu}\,.$$

So we obtain the following formulas for the exterior-product of the basis-vectors:

$$[e_1 \ldots \hat{e}_\nu \ldots e_n] = (-1)^{n-\nu}\,\overset{*}{e}{}^\nu \quad (\nu = 1 \ldots n)\,. \tag{9.87}$$

In the case $n = 3$ the formulas (9.87) read

$$[e_1, e_2] = \overset{*}{e}{}^3,\ [e_1, e_3] = -\overset{*}{e}{}^2,\ [e_2, e_3] = \overset{*}{e}{}^1.$$

9.16. The Lagrange identity. Consider a pair of dual determinant-functions Δ and Δ^* in E and E^*. Then an exterior product is defined in E and in E^*. It follows from the definition of these products that

$$\Delta\,(x_1 \ldots x_n) = \langle [x_1 \ldots x_{n-1}], x_n\rangle \quad x_\nu \in E \text{ and} \tag{9.88}$$

$$\Delta\,(\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^n) = \langle \overset{*}{x}{}^n, [\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{n-1}]\rangle \quad \overset{*}{x}{}^\nu \in E^*\,. \tag{9.89}$$

It will be shown that these two exterior products are connected by the *Lagrange identity*

$$\langle [x_1 \ldots x_{n-1}], [\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{n-1}]\rangle = \begin{vmatrix} \langle \overset{*}{x}{}^1, x_1\rangle & \cdots & \langle \overset{*}{x}{}^{n-1}, x_1\rangle \\ \vdots & & \vdots \\ \langle \overset{*}{x}{}^1, x_{n-1}\rangle & \cdots & \langle \overset{*}{x}{}^{n-1}, x_{n-1}\rangle \end{vmatrix}.$$

Multiplying the equations (9.88) and (9.89) and using the formula (4.27) we obtain

$$\langle [x_1 \ldots x_{n-1}], x_n\rangle \langle \overset{*}{x}{}^n, [\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{n-1}]\rangle = \begin{vmatrix} \langle \overset{*}{x}{}^1, x_1\rangle & \cdots & \langle \overset{*}{x}{}^n, x_1\rangle \\ \vdots & & \vdots \\ \langle \overset{*}{x}{}^1, x_n\rangle & \cdots & \langle \overset{*}{x}{}^n, x_n\rangle \end{vmatrix}.$$

Inserting $x_n = [\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{n-1}]$ into this equation we see that all scalar-products in the last row of the determinant become zero except for $\langle \overset{*}{x}{}^n, x_n\rangle$. Expansion of the determinant with respect to the last row yields

$$\langle [x_1 \ldots x_{n-1}], [\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{n-1}]\rangle \langle \overset{*}{x}{}^n, x_n\rangle = \begin{vmatrix} \langle \overset{*}{x}{}^1, x_1\rangle & \cdots & \langle \overset{*}{x}{}^{n-1}, x_1\rangle \\ \vdots & & \vdots \\ \langle \overset{*}{x}{}^1, x_{n-1}\rangle & \cdots & \langle \overset{*}{x}{}^{n-1}, x_{n-1}\rangle \end{vmatrix} \langle \overset{*}{x}{}^n, x_n\rangle. \tag{9.90}$$

We may assume that the vectors $\overset{*}{x}{}^\nu$ $(\nu = 1 \ldots n-1)$ are linearly independent because otherwise both sides of the Lagrange identity are zero. Then $x_n \neq 0$ and hence a vector $\overset{*}{x}{}^n \in E^*$ can be chosen such that $\langle \overset{*}{x}{}^n, x_n\rangle \neq 0$. Dividing (9.90) by $\langle \overset{*}{x}{}^n, x_n\rangle$ we find the Lagrange identity.

9.17. Dual product and exterior product. Employing the isomorphism

$$D^{n-1}: S^{n-1}(E) \to E^*$$

the exterior product we can write as

$$[x_1 \ldots x_{n-1}] = D^{n-1}(x_1 \wedge \cdots \wedge x_{n-1}) \quad x_\nu \in E\,. \tag{9.91}$$

In fact, let x be an arbitrary vector of E. Then it follows from (9.27) and (9.4) that

$$\begin{aligned} &\langle D^{n-1}(x_1 \wedge \cdots \wedge x_{n-1}), x\rangle = \langle\!\langle \Delta \vee (x_1 \wedge \cdots \wedge x_{n-1}), x\rangle\!\rangle \\ &= \langle\!\langle \Delta, x_1 \wedge \cdots \wedge x_{n-1} \wedge x\rangle\!\rangle\,. \end{aligned} \tag{9.92}$$

According to (8.58) and (9.85), the right hand-side of this relation can be written as

$$\langle\!\langle \Delta, x_1 \wedge \cdots \wedge x_{n-1} \wedge x\rangle\!\rangle = \Delta\,(x_1 \ldots x_{n-1}, x) = \langle [x_1 \ldots x_{n-1}], x\rangle\,. \tag{9.93}$$

The equations (9.92) and (9.93) yield

$$\langle D^{n-1}(x_1 \wedge \cdots \wedge x_{n-1}), x\rangle = \langle [x_1 \ldots x_{n-1}], x\rangle$$

whence (9.91). In the same way it is shown that

$$[\overset{*}{x}{}^1 \ldots \overset{*}{x}{}^{n-1}] = D_{n-1}(\overset{*}{x}{}^1 \wedge \cdots \wedge \overset{*}{x}{}^{n-1}) \quad \overset{*}{x}{}^\nu \in E^*.$$

Problems: 1. Show that every tensor $\Phi \in S_{n-1}(E)$ can be written in the form

$$\Phi\,(x_1 \ldots x_{n-1}) = \langle [x_1 \ldots x_{n-1}], a\rangle \quad a \in E$$

and that the vector a is uniquely determined by Φ.

2. Consider an $(n-1)$-linear skew-symmetric mapping φ of E into a linear space F. Prove that there exists exactly one linear mapping $\chi: E^* \to F$ such that

$$\varphi\,(x_1 \ldots x_{n-1}) = \chi\,[x_1 \ldots x_{n-1}] \quad x_\nu \in E \quad (\nu = 1 \ldots n-1)\,.$$

3. If α is an automorphism of E, prove that

$$[\alpha x_1 \ldots \alpha x_{n-1}] = \det\alpha.\,(\alpha^{-1})^*\,[x_1 \ldots x_{n-1}] \quad x_\nu \in E \quad (\nu = 1 \ldots n-1)\,.$$

4. Let e_ν $(\nu = 1 \ldots n)$ be a basis of E such that $\Delta\,(e_1 \ldots e_n) = 1$. Given $(n-1)$ vectors

$$x_i = \sum_\nu \xi_i^\nu e_\nu \quad (i = 1 \ldots n-1)$$

show that the vector $y = [x_1 \ldots x_{n-1}]$ has the components

$$\eta^\nu = (-1)^{n-\nu} \begin{vmatrix} \xi_1^1 & \ldots \hat{\xi}_1^\nu & \ldots \xi_1^n \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \xi_{n-1}^1 & \ldots \hat{\xi}_{n-1}^\nu & \ldots \xi_{n-1}^n \end{vmatrix} \quad (\nu = 1 \ldots n)\,.$$

5. Using the formula in problem 4 derive the relation

$$\begin{vmatrix} \sum\limits_{\nu=1}^{n} \xi_1^\nu \eta_\nu^1 & \cdots & \sum\limits_{\nu=1}^{n} \xi_1^\nu \eta_\nu^{n-1} \\ \vdots & & \vdots \\ \sum\limits_{\nu=1}^{n} \xi_{n-1}^\nu \eta_\nu^1 & \cdots & \sum\limits_{\nu=1}^{n} \xi_{n-1}^\nu \eta_\nu^{n-1} \end{vmatrix}$$

$$= \sum_{\nu=1}^{n} \begin{vmatrix} \xi_1^1 & \cdots & \hat{\xi}_1^\nu & \cdots & \xi_1^n \\ \vdots & & & & \vdots \\ \xi_{n-1}^1 & \cdots & \hat{\xi}_{n-1}^\nu & \cdots & \xi_{n-1}^n \end{vmatrix} \begin{vmatrix} \eta_1^1 & \cdots & \hat{\eta}_1^\nu & \cdots & \eta_1^n \\ \vdots & & & & \vdots \\ \eta_{n-1}^1 & \cdots & \hat{\eta}_{n-1}^\nu & \cdots & \eta_{n-1}^n \end{vmatrix}$$

from the Lagrange identity.

6. Let E^* and E be two dual 3-dimensional spaces and $a \in E$ $(a \neq 0)$ and $\overset{*}{b} \in E^*$ be two given vectors. Prove that there exists a vector $x \in E$ such that $[a, x] = \overset{*}{b}$ if and only if $\langle \overset{*}{b}, a \rangle = 0$.

Chapter X

Inner product spaces

§ 1. The inner product

10.1. Definition. An *inner product* in a real linear space E is a bilinear function (x, y) having the following properties:

1. Symmetry: $(x, y) = (y, x)$.
2. Positive definiteness: $(x, x) \geqq 0$ and $(x, x) = 0$ only for the vector $x = 0$.

A linear space in which an inner product is defined is called an *inner product space.*

The *norm* $|x|$ of a vector $x \in E$ is defined as the positive square-root

$$|x| = \sqrt{(x, x)} \,.$$

A *unit vector* is a vector with the norm 1. The set of all unit vectors is called the *unit-sphere.*

It follows from the bilinearity of the inner product that

$$|x + y|^2 = |x|^2 + 2(x, y) + |y|^2$$

whence

$$(x, y) = \frac{1}{2}\left(|x + y|^2 - |x|^2 - |y|^2\right) .$$

This equation shows that the inner product can be expressed in terms of the norm.

The restriction of the bilinear function (x, y) to a subspace $E_1 \subset E$ has again the properties 1 and 2 and hence every subspace of an inner product space is itself an inner product space.

10.2. Examples. 1. In the real number-space R^n the *standard inner product* is defined by

$$(x, y) = \sum_\nu \xi^\nu \eta^\nu,$$

where

$$x = (\xi^1 \dots \xi^n) \quad \text{and} \quad y = (\eta^1 \dots \eta^n).$$

2. Let E be an n-dimensional real linear space and $x_\nu (\nu = 1 \dots n)$ be a basis of E. Then an inner product can be defined by

$$(x, y) = \sum_\nu \xi^\nu \eta^\nu,$$

where

$$x = \sum_\nu \xi^\nu x_\nu, \quad y = \sum_\nu \eta^\nu x_\nu.$$

3. Consider the space C of all continuous functions f in the interval $0 \leqq t \leqq 1$ and define the inner product of two functions by

$$(f, g) = \int_0^1 f(t)\, g(t)\, dt.$$

10.3. Orthogonality. Two vectors $x \in E$ and $y \in E$ are said to be *orthogonal* if $(x, y) = 0$. The definiteness implies that only the zero-vector is orthogonal to itself. A system of p vectors $x_\nu \neq 0$ in which any two vectors x_ν and x_μ $(\nu \neq \mu)$ are orthogonal, is linearly independent. In fact, the relation

$$\sum_\nu \lambda^\nu x_\nu = 0$$

yields

$$\lambda^\mu (x_\mu, x_\mu) = 0 \qquad (\mu = 1 \dots p)$$

whence

$$\lambda^\mu = 0 \qquad (\mu = 1 \dots p).$$

Two subspaces $E_1 \subset E$ and $E_2 \subset E$ are called *orthogonal*, denoted as $E_1 \perp E_2$, if any two vectors $x_1 \in E_1$ and $x_2 \in E_2$ are orthogonal.

10.4. The Schwarz-inequality. Let x and y be two arbitrary vectors of the inner product space E. Then the *Schwarz-inequality* asserts that

$$(x, y)^2 \leqq |x|^2 \, |y|^2 \tag{10.1}$$

and that the equality holds if and only if the vectors are linearly dependent. To prove this consider the function

$$|x + \lambda y|^2$$

of the real variable λ. The definiteness of the inner product implies that

$$|x+\lambda y|^2 \geqq 0 \qquad (-\infty < \lambda < \infty)\,.$$

Expanding the norm we obtain

$$\lambda^2 |y|^2 + 2\lambda\,(x,y) + |x|^2 \geqq 0\,.$$

Hence the discriminant of the above quadratic expression must be negative or zero,

$$(x,y)^2 \leqq |x|^2\,|y|^2.$$

Now assume that the equality holds in (10.1). Then the discriminant of the quadratic equation*

$$\lambda^2 |y|^2 + 2\lambda\,(x,y) + |x|^2 = 0 \tag{10.2}$$

is zero. Hence the equation (10.2) has a real solution λ_0. It follows from (10.2) that

$$|\lambda_0 y + x|^2 = 0,$$

whence

$$\lambda_0 y + x = 0\,.$$

Thus, the vectors x and y are linearly dependent.

10.5. Angles. Given two vectors $x \neq 0$ and $y \neq 0$, the Schwarz-inequality implies that

$$-1 \leqq \frac{(x,y)}{|x|\,|y|} \leqq 1\,.$$

Consequently, there exists exactly one real number ω $(0 \leqq \omega \leqq \pi)$ such that

$$\cos\omega = \frac{(x,y)}{|x|\,|y|}\,. \tag{10.3}$$

The number ω is called the *angle* between the vectors x and y. The symmetry of the inner product implies that the angle is symmetric with respect to x and y. If the vectors x and y are orthogonal, it follows that $\cos\omega = 0$, whence $\omega = \frac{\pi}{2}$.

Now assume that the vectors x and y are linearly dependent, $y = \lambda x$. Then

$$\cos\omega = \frac{\lambda}{|\lambda|} = \begin{cases} +1 & \text{if } \lambda > 0 \\ -1 & \text{if } \lambda < 0 \end{cases}$$

and hence

$$\omega = \begin{cases} 0 & \text{if } \lambda > 0 \\ \pi & \text{if } \lambda < 0\,. \end{cases}$$

With the help of (10.3) the equation

$$|x-y|^2 = |x|^2 - 2(x,y) + |y|^2$$

*) Without loss of generality we may assume that $y \neq 0$.

can be written in the form

$$|x - y|^2 = |x|^2 + |y|^2 - 2\,|x|\,|y| \cos\omega\,.$$

This formula is known as the *cosine-theorem*. If the vectors x and y are orthogonal, the cosine-theorem reduces to the *Pythagorean theorem*

$$|x - y|^2 = |x|^2 + |y|^2\,.$$

10.6. The triangle-inequality. It follows from the Schwarz-inequality that

$$|x + y|^2 = |x|^2 + 2\,(x, y) + |y|^2 \leqq |x|^2 + 2\,|x|\,|y| + |y|^2 = (|x| + |y|)^2,$$

whence

$$|x + y| \leqq |x| + |y|\,. \tag{10.4}$$

The relation (10.4) is called the *triangle-inequality*. To discuss the equality-sign we may exclude the trivial case $y = 0$. It will be shown that the equality holds in (10.4) if and only if

$$x = \lambda y, \quad \lambda > 0\,.$$

The equation

$$|x + y| = |x| + |y|$$

implies that

$$|x|^2 + 2\,(x, y) + |y|^2 = |x|^2 + 2\,|x|\,|y| + |y|^2\,,$$

whence

$$(x, y) = |x|\,|y|\,. \tag{10.5}$$

Thus, the vectors x and y must be linearly dependent,

$$x = \lambda y\,. \tag{10.6}$$

The equations (10.5) and (10.6) yield $\lambda = |\lambda|$, whence $\lambda \geqq 0$.

Conversely, assume that $x = \lambda y$, where $\lambda \geqq 0$. Then

$$|x + y| = |(\lambda + 1)\,y| = (\lambda + 1)\,|y| = \lambda\,|y| + |y| = |x| + |y|\,.$$

Given three vectors x, y, z, the triangle-inequality can be written in the form

$$|x - y| \leqq |x - z| + |z - y|\,. \tag{10.7}$$

As a generalization of (10.7), we prove the *Ptolemy-inequality*

$$|x - y|\,|z| \leqq |y - z|\,|x| + |z - x|\,|y|\,. \tag{10.8}$$

The relation (10.8) is trivial if one of the three vectors is zero. Hence we may assume that $x \neq 0$, $y \neq 0$ and $z \neq 0$. Define the vectors x', y' and z' by

$$x' = \frac{x}{|x|^2}\,, \quad y' = \frac{y}{|y|^2}\,, \quad z' = \frac{z}{|z|^2}\,.$$

Then

$$|x' - y'|^2 = \frac{1}{|x|^2} - \frac{2\,(x, y)}{|x|^2\,|y|^2} + \frac{1}{|y|^2} = \frac{|x - y|^2}{|x|^2\,|y|^2}\,,$$

Applying the inequality (10.7) to the vectors x', y' and z' we obtain

$$\frac{|x-y|}{|x|\,|y|} \leq \frac{|y-z|}{|y|\,|z|} + \frac{|z-x|}{|z|\,|x|},$$

whence (10.8).

Problems: 1. For $x = (\xi^1, \xi^2)$ and $y = (\eta^1, \eta^2)$ in R^2 show that the bilinear function

$$(x, y) = \xi^1\eta^1 - \xi^2\eta^2 - \xi^1\eta^2 + 4\xi^2\eta^2$$

satisfies the properties listed in sec. 10.1.

2. Consider the space S of all infinite sequences $x = (\xi^1, \xi^2, \ldots)$ such that

$$\sum_\nu \xi_\nu^2 < \infty.$$

Show that $\sum_\nu \xi_\nu \eta_\nu$ converges and that the bilinear function $(x, y) = \sum_\nu \xi_\nu \eta_\nu$ is an inner product.

3. Consider three mutually different vectors $x \neq 0$, $y \neq 0$ and $z \neq 0$. Prove that the equation

$$|x - y|\,|z| = |y - z|\,|x| + |z - x|\,|y|$$

holds if and only if the four points x, y, z, 0 are contained on a circle such that the pairs x, y and z, 0 separate each other.

4. Consider two inner product spaces E_1 and E_2. Prove that an inner product is defined in the Cartesian product $E_1 \times E_2$ by

$$((x_1, x_2), (y_1, y_2)) = (x_1, y_1) + (x_2, y_2) \quad x_1, y_1 \in E_1, \quad x_2, y_2 \in E_2.$$

5. Given a subspace E_1 of a finite dimensional inner product space E, consider the factor-space E/E_1. Prove that every equivalence class contains exactly one vector wich is orthogenal to E_1.

§ 2. Orthonormal bases

10.7. Definition. Let E be an n-dimensional inner product space and $x_\nu (\nu = 1 \ldots n)$ be a basis of E. Then the bilinear function (x, y) determines a symmetric matrix

$$g_{\nu\mu} = (x_\nu, x_\mu) \qquad (\nu, \mu = 1 \ldots n). \quad (10.9)$$

The inner product of two vectors

$$x = \sum_\nu \xi^\nu x_\nu \quad \text{and} \quad y = \sum_\nu \eta^\nu x_\nu$$

can be written as

$$(x, y) = \sum_{\nu,\mu} \xi^\nu \eta^\mu (x_\nu, x_\mu) = \sum_{\nu,\mu} g_{\nu\mu} \xi^\nu \eta^\mu \qquad (10.10)$$

and hence it appears as a bilinear form with the coefficient-matrix $g_{\nu\mu}$.

The basis x_ν $(\nu = 1 \ldots n)$ is called *orthonormal*, if the vectors x_ν $(\nu = 1 \ldots n)$ are mutually orthogonal and have the norm 1,

$$(x_\nu, x_\mu) = \delta_{\nu\mu}. \tag{10.11}$$

Then the formula (10.10) reduces to

$$(x, y) = \sum_\nu \xi^\nu \eta^\nu \tag{10.12}$$

and in the case $y = x$

$$|x|^2 = \sum_\nu \xi^\nu \xi^\nu.$$

The substitution $x = x_\mu$ in (10.12) yields

$$(x, x_\mu) = \xi^\mu \qquad (\mu = 1 \ldots n). \tag{10.13}$$

Now assume that $x \neq 0$, and denote by θ_μ the angle between the vectors x and $x_\mu (\mu = 1 \ldots n)$. The formulas (10.3) and (10.13) imply that

$$\cos\theta_\mu = \frac{\xi^\mu}{|x|} \quad (\mu = 1 \ldots n). \tag{10.14}$$

If x is a unit-vector (10.14) reduces to

$$\cos\theta_\mu = \xi^\mu \quad (\mu = 1 \ldots n). \tag{10.15}$$

These equations show that the components of a unit-vector x relative to an orthonormal basis are equal to the cosines of the angles between x and the basisvectors x_μ.

10.8. The Schmidt-orthogonalization. In this section it will be shown that an orthonormal basis can be constructed in every inner product space of finite dimension. Let $a_\nu (\nu = 1 \ldots n)$ be an arbitrary basis of E. Starting out from this basis a new basis $b_\nu (\nu = 1 \ldots n)$ will be constructed whose vectors are mutually orthogonal. Let

$$b_1 = a_1.$$

Then put

$$b_2 = a_2 + \lambda b_1$$

and determine the scalar λ such that $(b_1, b_2) = 0$. This yields

$$(a_2, b_1) + \lambda (b_1, b_1) = 0.$$

Since $b_1 \neq 0$, this equation can be solved with respect to λ. The vector b_2 thus obtained is different from zero because otherwise a_1 and a_2 would be linearly dependent.

To obtain b_3, set

$$b_3 = a_3 + \mu b_1 + \nu b_2$$

and determine the scalars μ and ν such that

$$(b_1, b_3) = 0 \quad \text{and} \quad (b_2, b_3) = 0.$$

This yields

$$(a_3, b_1) + \mu (b_1, b_1) = 0$$

and

$$(a_3, b_2) + \nu (b_2, b_2) = 0 .$$

Since $b_1 \neq 0$ and $b_2 \neq 0$, these equations can be solved with respect to μ and ν. The linear independence of the vectors a_1, a_2, a_3 implies that $b_3 \neq 0$. Continuing this way we finally obtain a system of n vectors $b_\nu \neq 0$ $(\nu = 1 \ldots n)$ such that

$$(b_\nu, b_\mu) = 0 \quad (\nu \neq \mu) .$$

It follows from the criterion in sec. 10.3, that the vectors b_ν are linearly independent and hence they form a basis of E. Consequently the vectors

$$e_\nu = \frac{b_\nu}{|b_\nu|} \qquad (\nu = 1 \ldots n)$$

form an orthonormal basis.

10.9. Orthogonal transformations. Consider two orthogonal bases x_ν and $\bar{x}_\nu$ $(\nu = 1 \ldots n)$ of E. Denote by α_ν^μ the matrix of the basis-transformation $x_\nu \to \bar{x}_\nu$,

$$\bar{x}_\nu = \sum_\mu \alpha_\nu^\mu x_\mu . \tag{10.16}$$

The relations

$$(x_\nu, x_\mu) = \delta_{\nu\mu} \quad \text{and} \quad (\bar{x}_\nu, \bar{x}_\mu) = \delta_{\nu\mu}$$

imply that

$$\sum_\lambda \alpha_\nu^\lambda \alpha_\mu^\lambda = \delta_{\nu\mu} . \tag{10.17}$$

This equation shows that the product of the matrix (α_ν^μ) and the transposed matrix is equal to the unit-matrix. In other words, the transposed matrix coincides with the inverse matrix. A matrix of this kind is called *orthogonal.*

Hence, two orthonormal bases are related by an orthogonal matrix. Conversely, given an orthonormal basis $x_\nu (\nu = 1 \ldots n)$ and an orthogonal $n \times n$-matrix (α_ν^μ), the basis $\bar{x}_\nu$ defined by (10.16) is again orthonormal.

10.10. Orthogonal complement. Let E be an inner product space (of finite or infinite dimension) and E_1 be a subspace of E. Denote by $E_1^\perp$ the set of all vectors which are orthogonal to E_1. Obviously, $E_1^\perp$ is again a subspace of E and the intersection $E_1 \cap E_1^\perp$ consists of the zero-vector only. $E_1^\perp$ is called the *orthogonal complement* of E_1. Now it will be shown that the spaces E_1 and $E_1^\perp$ form a direct decomposition of E,

$$E = E_1 \oplus E_1^\perp , \tag{10.18}$$

provided that E_1 has finite dimension.

Select an orthonormal basis $y_\mu (\mu = 1 \ldots m)$ of E_1. Given a vector $x \in E$ and a vector

$$y = \sum_\mu \eta^\mu y_\mu \, .$$

of E_1 consider the difference

$$z = x - y \, .$$

Then

$$(z, y_\mu) = (x, y_\mu) - (y, y_\mu) = (x, y_\mu) - \eta^\mu \, .$$

This equation shows that z is contained in $E_1^\perp$ if and only if

$$\eta^\mu = (x, y_\mu) \qquad (\mu = 1 \ldots m) \, .$$

We thus obtain the decomposition

$$x = p + h \tag{10.19}$$

where

$$p = \sum_\mu (x, y_\mu)\, y_\mu \quad \text{and} \quad h = x - p \, .$$

The vector p is called the *orthogonal projection* of x onto E_1.

Passing over to the norm in the decomposition (10.19) we obtain the relation

$$|x|^2 = |p|^2 + |h|^2 \, . \tag{10.20}$$

The formula (10.20) yields the *Bessel's inequality*

$$|x| \geqq |p|$$

showing that the norm of the projection never exceeds the norm of x. The equality holds if and only if $h = 0$, i. e. if and only if $x \in E_1$. The number $|h|$ is called the *distance* of x from the subspace E_1.

If E has finite dimension the decomposition (10.18) implies that

$$\dim E_1^\perp = \dim E - \dim E_1 \, .$$

Problems: 1. Starting from the basis

$$a_1 = (1, 0, 1) \quad a_2 = (2, 1, -3) \quad a_3 = (-1, 1, 0)$$

of the number-space R^3 construct an orthonormal basis by the Schmidt-orthogonalization process.

2. Let E and F be two inner product spaces. Show that the bilinear function

$$(x_1 \otimes y_1, x_2 \otimes y_2) = (x_1, x_2)\,(y_1, y_2) \quad x_1, x_2 \in E, y_1, y_2 \in F$$

defines an inner product in the space $E \otimes F$.

3. Given an inner product space E and a subspace E_1 of finite dimension consider a decomposition

$$x = x_1 + x_2 \qquad x_1 \in E_1$$

and the projection

$$x = p + h \qquad p \in E_1, h \in E_1^{\perp}.$$

Prove that

$$|x_2| \geqq h$$

and that the equality is assumed only if $x_1 = p$ and $x_2 = h$.

4. Let C be the space of all continuous functions in the interval $0 \leqq t \leqq 1$ with the inner product defined as in sec. 10.2. If C^1 denotes the subspace of all continuously differentiable functions, show that $C_1^{\perp} = 0$.

5. Consider a subspace E_1 of E. Assume an orthogonal decomposition

$$E_1 = F_1 \oplus G_1 \qquad F_1 \perp G_1 .$$

Establish the relations

$$F_1^{\perp} = E_1^{\perp} \oplus G_1, E_1 \perp G_1 \text{ and } G_1^{\perp} = E_1^{\perp} \oplus F_1, E_1^{\perp} \perp F_1 .$$

6. Let F^3 be the space of all polynomials of degree $\leqq 2$. Define the inner product of two polynomials as follows:

$$(P, Q) = \int_{-1}^{1} P(t)\, Q(t)\, dt .$$

The vectors 1, t, t^2 form a basis in F^3. Orthogonalize and orthonormalize this basis. Generalize the result for the case of the space F^n of polynomials of degree $\leqq n - 1$.

§ 3. Normed determinant-functions

10.11. Definition. Let E be an n-dimensional inner product space. The bilinear function (x, y) satisfies the conditions (I) and (II) of sec. 2.6 if E^* is identified with E. In fact, assume, that $(x, y) = 0$ for a fixed vector $x \in E$ and all vectors $y \in E$. Then $(x, x) = 0$, whence, by the definiteness, $x = 0$. Thus, an inner product space can be considered as dual to itself. The orthonomal bases of E are exactly those which are dual to themselves.

Now consider a determinant-function $\Delta_0 \neq 0$ in E. Then the formula (4.24) yields

$$\begin{vmatrix} (x_1, y_1) & \dots & (x_1, y_n) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ (x_n, y_1) & \dots & (x_n, y_n) \end{vmatrix} = \alpha \cdot \Delta_0 (x_1 \dots x_n) \cdot \Delta_0 (y_1 \dots y_n) , \qquad (10.21)$$

where α is a constant. Inserting $x_\nu = y_\nu = e_\nu$ $(\nu = 1 \dots n)$ where e_ν is an orthonormal basis we obtain the equation

$$\alpha \cdot [\Delta_0 (e_1 \dots e_n)]^2 = 1$$

showing that $\alpha > 0$. Let us now introduce a new determinant-function Δ by

$$\Delta = \sqrt{\alpha} \cdot \Delta_0 .$$

Then the identity (10.21) reads

$$\begin{vmatrix} (x_1, y_1) & \dots & (x_1, y_n) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ (x_n, y_1) & \dots & (x_n, y_n) \end{vmatrix} = \Delta\,(x_1 \dots x_n)\,\Delta\,(y_1 \dots y_n)\,. \tag{10.22}$$

A determinant-function satisfying the above identity will be called a *normed determinant-function*. It follows from (10.22) that there exist exactly two normed determinant-functions Δ and $-\Delta$ in an inner product space.

Now assume that an orientation is defined in E. Then one of the functions Δ and $-\Delta$ represents the orientation. Consequently, *in an oriented inner product space there exists exactly one normed determinant-function representing the given orientation.*

10.12. Angles in an oriented plane. With the help of a normed determinant-function it is possible to attach a sign to the angle between two vectors of a 2-dimensional oriented inner product space. Consider the normed determinant-function Δ which represents the given orientation. Then the identity (10.22) yields

$$|x|^2\,|y|^2 - (x, y)^2 = \Delta\,(x, y)^2\,. \tag{10.23}$$

Now assume that $x \neq 0$ and $y \neq 0$. Dividing (10.23) by $|x|^2\,|y|^2$ we obtain the relation

$$\frac{(x, y)^2}{|x|^2\,|y|^2} + \frac{\Delta\,(x, y)^2}{|x|^2\,|y|^2} = 1\,.$$

Consequently, there exists exactly one real number θ in the interval $-\pi < \theta \leqq \pi$ such that

$$\cos\theta = \frac{(x, y)}{|x|\,|y|} \quad \text{and} \quad \sin\theta = \frac{\Delta\,(x, y)}{|x|\,|y|}\,. \tag{10.24}$$

This number is called the *oriented angle* between x and y.

If the orientation is changed, Δ has to be replaced by $-\Delta$, and hence θ changes into $-\theta$.

Furthermore it follows from (10.24) that θ changes the sign if the vectors x and y are interchanged and that

$$\theta\,(x, -y) = \theta\,(x, y) + \varepsilon\pi$$

where $\varepsilon = +1$ if $\theta\,(x, y) > 0$ and $\varepsilon = -1$ if $\theta\,(x, y) < 0$.

10.13. The Gram determinant. Given p vectors x_ν $(\nu = 1 \dots p)$ in an inner product space E, the *Gram determinant* $G(x_1 \dots x_p)$ is defined by

$$G(x_1 \dots x_p) = \begin{vmatrix} (x_1, x_1) & \dots & (x_1, x_p) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ (x_p, x_1) & \dots & (x_p, x_p) \end{vmatrix} . \tag{10.25}$$

It will be shown that

$$G(x_1 \dots x_p) \geqq 0 \tag{10.26}$$

and that the equality holds if and only if the vectors $(x_1 \dots x_p)$ are linearly dependent. In the case $p = 2$ (10.26) reduces to the Schwarz-inequality.

To prove (10.26), assume first that the vectors x_ν $(\nu = 1 \dots p)$ are linearly dependent. Then the rows of the matrix (10.25) are also linearly dependent whence

$$G(x_1 \dots x_p) = 0 .$$

If the vectors x_ν $(\nu = 1 \dots p)$ are linearly independent, they generate a p-dimensional subspace E_1 of E. E_1 is again an inner product space. Denote by Δ_1 a normed determinant-function in E_1. Then it follows from (10.22) that

$$G(x_1 \dots x_p) = \Delta_1 (x_1 \dots x_p)^2 .$$

The linear independence of the vectors x_ν $(\nu = 1 \dots p)$ implies that $\Delta_1 (x_1 \dots x_p) \neq 0$, whence

$$G(x_1 \dots x_p) > 0 .$$

10.14. The volume of a parallelepiped. Let p linearly independent vectors a_ν $(\nu = 1 \dots p)$ be given in E. The set

$$x = \sum_\nu \lambda^\nu a_\nu \quad 0 \leqq \lambda^\nu \leqq 1 \quad (\nu = 1 \dots p) \tag{10.27}$$

is called the *p-dimensional parallelepiped* spanned by the vectors a_ν $(\nu = 1 \dots p)$. The volume $V(a_1 \dots a_p)$ of the parallelepiped is defined by

$$V(a_1 \dots a_p) = |\Delta_1 (a_1 \dots a_p)| , \tag{10.28}$$

where Δ_1 is a normed determinant-function in the subspace generated by the vectors a_ν $(\nu = 1 \dots p)$.

In view of the identity (10.22) the formula (10.28) can be written as

$$V(a_1 \dots a_p)^2 = \begin{vmatrix} (a_1, a_1) & \dots & (a_1, a_p) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ (a_p, a_1) & \dots & (a_p, a_p) \end{vmatrix} . \tag{10.29}$$

In the case $p = 2$ the above formula yields

$$V(a_1, a_2)^2 = |a_1|^2 \, |a_2|^2 - (a_1, a_2)^2 = |a_1|^2 \, |a_2|^2 \sin^2\theta , \tag{10.30}$$

where θ denotes the angle between a_1 and a_2. Taking the square-root on both sides of (10.30), we obtain the well-known formula for the area of a parallelogram:

$$V(a_1, a_2) = |a_1|\,|a_2|\,|\sin\theta| .$$

Going back to the general case, select an integer i $(1 \leqq i \leqq p)$ and decompose a_i in the form

$$a_i = \sum_{\nu \neq i} \xi^\nu a_\nu + h_i, \text{ where } (h_i, a_\nu) = 0 \quad (\nu \neq i) . \tag{10.31}$$

Then (10.28) can be written as

$$V(a_1 \dots a_p) = |\Delta_1 (a_1 \dots a_{i-1}, h_i, a_{i+1} \dots a_p)| .$$

Employing the identity (10.22) and observing that $(h_i, a_\nu) \neq 0$ $(\nu \neq i)$ we obtain*)

$$V(a_1 \dots a_p)^2 = \begin{vmatrix} (a_1, a_1) & \dots & (\hat{a}_1, \hat{a}_i) & \dots & (a_1, a_p) \\ \vdots & & & & \vdots \\ (\hat{a}_i, \hat{a}_1) & \dots & (\hat{a}_i, \hat{a}_i) & \dots & (\hat{a}_i, \hat{a}_p) \\ \vdots & & & & \vdots \\ (a_p, a_1) & \dots & (\hat{a}_p, \hat{a}_i) & \dots & (a_p, a_p) \end{vmatrix} (h_i, h_i) . \tag{10.32}$$

The determinant in this equation represents the square of the volume of the $(p-1)$-dimensional parallelepiped generated by the vectors $(a_1 \dots \hat{a}_i \dots a_p)$. We thus obtain the formula

$$V(a_1 \dots a_p) = V(a_1 \dots \hat{a}_i \dots a_p) \cdot |h_i| \quad (1 \leqq i \leqq p)$$

showing that the volume $V(a_1 \dots a_p)$ is the product of the volume $V(a_1 \dots \hat{a}_i \dots a_p)$ of the i^{th} "base" and the corresponding height.

10.15. The exterior product. Let us recall the definition of the exterior product given in sec. 9.15: If E^*, E is a pair of dual spaces of dimension n and $\Delta \neq 0$ is a given determinant-function in E then the exterior product $[a_1, \dots a_{n-1}]$ of $n-1$ vectors $a_\nu \in E$ is defined by

$$\langle [a_1, \dots a_{n-1}], x \rangle = \Delta (a_1 \dots a_{n-1}, x) \qquad x \in E .$$

If E is an inner product space, the dual space E^* can be identified with E and thus the exterior product $[a_1 \dots a_{n-1}]$ becomes a vector of E. Moreover, if an orientation is given in E, there exists exactly one normed determinant-function Δ representing the orientation. Using this determinant-function we now define the exterior product in an oriented inner product space E by the equation

$$([a_1 \dots a_{n-1}], x) = \Delta (a_1 \dots a_{n-1}, x) . \tag{10.33}$$

*) The symbol $\hat{a}_i$ indicates that the vector a_i is deleted.

Inserting $x = a_\nu (1 \leqq \nu \leqq n-1)$ in (10.33) we find the equation

$$([a_1 \dots a_{n-1}], a_\nu) = 0 \qquad (\nu = 1 \dots n-1)$$

showing that the exterior product of the vectors a_ν is orthogonal to every vector a_ν. To obtain a geometric interpretation of the norm of the exterior product write the Lagrange identity (9.89) in the form

$$([a_1 \dots a_{n-1}], [b_1 \dots b_{n-1}]) = \begin{vmatrix} (a_1, b_1) & \dots & (a_1, b_{n-1}) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ (a_{n-1}, b_1) & \dots & (a_{n-1}, b_{n-1}) \end{vmatrix}. \tag{10.34}$$

where a_ν and b_ν $(\nu = 1 \dots n-1)$ are arbitrary vectors of E.

This relation yields for $b_\nu = a_\nu$

$$|[a_1 \dots a_{n-1}]|^2 = \begin{vmatrix} (a_1, a_1) & \dots & (a_1, a_{n-1}) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ (a_{n-1}, a_1) & \dots & (a_{n-1}, a_{n-1}) \end{vmatrix}$$

showing that the norm of the exterior product of $(n-1)$ linearly independent vectors is equal to the volume of the parallelepiped spanned by these vectors.

Substituting $x = [a_1 \dots a_{n-1}]$ in (10.33) we find that

$$|[a_1 \dots a_{n-1}]|^2 = \Delta(a_1, \dots, a_{n-1}, [a_1 \dots a_{n-1}]) .$$

Now assume that the vectors $a_\nu (\nu = 1 \dots n-1)$ are linearly independent. Then $[a_1 \dots a_{n-1}] \neq 0$ and hence it follows from the above equation that

$$\Delta(a_1, \dots, a_{n-1}, [a_1 \dots a_{n-1}]) \neq 0 .$$

This relation shows that the vectors

$$a_1, \dots, a_{n-1}, [a_1 \dots a_{n-1}]$$

form a *positive* basis of E.

Now consider a positive orthonormal basis $e_\nu (\nu = 1 \dots n)$ of E. Then

$$\Delta(e_1 \dots e_n) = +1$$

and hence the formulas (9.87) yield

$$[e_1 \dots \hat{e}_i \dots e_n] = (-1)^{n-i} e_i \qquad (1 \leqq i \leqq n) .$$

In particular, for $n = 3$,

$$[e_1, e_2] = e_3 , \quad [e_1, e_3] = -e_2 , \quad [e_2, e_3] = e_1 .$$

Hence the exterior product of two vectors

$$a = \sum_i \xi^i e_i \quad \text{and} \quad b = \sum_i \eta^i e_i$$

of the 3-space is given by

$$[a, b] = (\xi^2\eta^3 - \xi^3\eta^2)\, e_1 + (\xi^3\eta^1 - \xi^1\eta^3) e_2 + (\xi^1\eta^2 - \xi^2\eta^1)\, e_3 \,.$$

Problems: 1. Given a vector $a \neq 0$ determine the locus of all vectors x such that $x - a$ is orthogonal to $x + a$.

2. Prove that the exterior product defines a Lie-multiplication in a 3-dimensional inner product-space.

3. Let e be a given unit vector of an n-dimensional inner product space E and E_1 be the orthogonal complement of e. Show that the distance of a vector $x \in E$ from the subspace E_1 is given by

$$d = |(x, e)| \,.$$

4. Prove that the area of the parallelogram generated by the vectors x_1 and x_2 is given by

$$A = 2\sqrt{s(s-a)\,(s-b)\,(s-c)} \,,$$

where

$$a = |x_1|\,, \quad b = |x_2|\,, \quad c = |x_2 - x_1|\,, \quad s = \frac{1}{2}(a + b + c) \,.$$

5. Given four vectors a_1, a_2, b_1, b_2 in an oriented 3-space, verify the formula

$$([a_1, a_2], [b_1, b_2]) = (a_1, b_1)\,(a_2, b_2) - (a_1, b_2)\,(a_2, b_1) \,.$$

From this formula derive the identity

$$[a, b]^2 + (a, b)^2 = |a|^2\,|b|^2 \,.$$

6. Prove the formula

$$[[a, b], c] = (a, c)\, b - (b, c)\, a \,.$$

Hint: The product on the left-hand side is orthogonal to $[a, b]$ and hence it can be written as

$$[[a, b], c] = \lambda a + \mu b \,.$$

Determine the factors λ and μ using the formula in problem 5.

7. Let $a \neq 0$ and b be two given vectors of an oriented 3-space. Prove that the equation $[x, a] = b$ has a solution if and only if $(a, b) = 0$. If this condition is satisfied and x_0 is a particular solution, show that the general solution is $x_0 + \lambda a$.

8. Consider an oriented inner product space of dimension 2. Given two positive orthonormal bases (e_1, e_2) and $(\bar{e}_1, \bar{e}_2)$, prove the formulas

$$\bar{e}_1 = e_1 \cos\omega - e_2 \sin\omega$$

$$\bar{e}_2 = e_1 \sin\omega + e_2 \cos\omega \,,$$

where ω is the angle between e_1 and $\bar{e}_1$.

9. Let e_1, e_2, e_3 be three unit-vectors in an oriented plane. Denote by θ_{ij} $(i \neq j)$ the angle between e_i and e_j $(i, j = 1, 2, 3)$. Using the formulas (10.24) prove the relations

$$\sin\theta_{12} = \sin\theta_{13}\cos\theta_{23} - \cos\theta_{13}\sin\theta_{23}$$

and

$$\cos\theta_{12} = \cos\theta_{13}\cos\theta_{23} + \sin\theta_{13}\sin\theta_{23}\,.$$

10. Let a_1 and a_2 be two linearly independent vectors of an oriented Euclidean 3-space and F be the plane generated by a_1 and a_2. Introduce an orientation in F such that the basis a_1, a_2 is positive. Prove that the angle between two vectors

$$x = \xi^1 a_1 + \xi^2 a_2 \quad \text{and} \quad y = \eta^1 a_1 + \eta^2 a_2$$

is determined by the equations

$$\cos\theta = \frac{\sum\limits_{\nu,\mu}(a_\nu, a_\mu)\,\xi^\nu\eta^\mu}{|x|\,|y|} \quad \text{and} \quad \sin\theta = \frac{\xi^1\eta^2 - \xi^2\eta^1}{|x|\,|y|}\,|[a_1, a_2]| \quad (-\pi < \theta \leqq \pi)\,.$$

11. Given an orthonormal basis e_ν $(\nu = 1, 2, 3)$ in the 3-space, define the endomorphisms φ_ν by

$$\varphi_\nu x = [x, e_\nu] \qquad (\nu = 1, 2, 3)\,.$$

Prove that

$$\sum_\nu \varphi_\nu^2 = -2\,\iota\,.$$

12. Let E be an oriented inner product space of dimension 4. Select a unit-vector e and denote by E_1 the orthogonal complement of e. Using the induced orientation in E_1 (see sec. 5.3) define a multiplication in E as follows:

$$x e = x,\; y e = y \qquad x \in E,\; y \in E$$

$$x y = (x, y)\, e + [x, y] \qquad x \in E_1,\; y \in E_1\,.$$

Prove that this multiplication has the following properties:

1. $(xy)\,z = x(yz)$.

2. For every vector $x \neq 0$ there exists a vector x^{-1} such that $xx^{-1} = e$ and $x^{-1}x = e$.

3. $|xy| = |x|\,|y|$.

The algebra which is defined in E by the above multiplication is called the *quaternion-algebra*.

13. Let x, y, z be three vectors of a plane such that x and y are linearly independent and that $x + y + z = 0$.

a) Prove that the ordered pairs x, y; y, z and z, x represent the same orientation. Then show that

$$\theta(x, y) + \theta(y, z) + \theta(z, x) = 2\pi$$

where the angles refer to the above orientation.

b) Prove that

$$\theta(y, -x) + \theta(z, -y) + \theta(x, -z) = \pi .$$

What is the geometric significance of the two above relations?

14. Given p vectors $x_1, \ldots, x_p$ prove the inequality

$$G(x_1, \ldots, x_p) \leqq |x_1|^2 \, |x_2|^2, \ldots, |x_p|^2 .$$

Then derive the *Hadamard's inequality* for a determinant

$$\begin{vmatrix} a_{11} \cdots a_{1n} \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ a_{n1} \cdots a_{nn} \end{vmatrix}^2 \leqq \sum_{k=1}^{n} |a_{1k}|^2 \cdot \sum_{k=1}^{n} |a_{2k}|^2 \cdots \sum_{k=1}^{n} |a_{nk}|^2 .$$

§ 4. The duality in an inner product space

10.16. The Riesz-theorem. Let E be an n-dimensional inner product space and $L(E)$ be the space of all linear functions in E. Assign to every vector $x \in E$ the linear function f_x defined by

$$f_x(y) = (x, y) .$$

The correspondence

$$x \to f_x \tag{10.35}$$

defines a linear mapping $E \to L(E)$. This mapping is regular. In fact, assume that $f_x = 0$ for a vector $x \in E$. Then

$$(x, x) = f_x(x) = 0 ,$$

whence $x = 0$. Since the linear spaces E and $L(E)$ have the same dimension, the mapping (10.35) is an isomorphism of E *onto* $L(E)$. Hence *every linear function f in E can be represented in the form*

$$f(y) = (x, y)$$

and the vector $x \in E$ is uniquely determined by f. (Theorem by F. RIESZ).

10.17. The isomorphism of E and E^*. Now consider an arbitrary linear space E^* which is dual to E. The duality between E and E^* defines an isomorphism of $L(E)$ onto E^* in the following way (see sec. 2.10): Every linear function f can be uniquely written as

$$f(y) = \langle \overset{*}{x}, y \rangle \qquad \overset{*}{x} \in E^* .$$

Combining the isomorphism s $E \to L(E)$ and $L(E) \to E^*$ we obtain an isomorphism

$$\alpha : x \to f_x \to \overset{*}{x}$$

of E onto E^*. The vectors x and x^* are related by the equation

$$(x, y) = \langle \overset{*}{x}, y \rangle . \tag{10.36}$$

The isomorphism $\alpha:E\to E^*$ thus obtained is characterized by the relation

$$(x, y) = \langle \alpha x, y\rangle \qquad x, y \in E .$$

With the help of the isomorphism $\alpha:E\to E^*$ an inner product can be introduced in E^*. In fact, the bilinear function

$$(\overset{*}{x}, \overset{*}{y}) = (\alpha^{-1}\overset{*}{x}, \alpha^{-1}\overset{*}{y}) \quad \overset{*}{x}, \overset{*}{y} \in E^* \tag{10.37}$$

is again symmetric and positive definite and hence it defines an inner product in E^*.

10.18. The metric tensors. The bilinear functions (x, y) and $(\overset{*}{x}, \overset{*}{y})$ determine a symmetric covariant and a symmetric contravariant tensor of second order in E. These tensors are called the *metric tensors* of the inner product-space E.

Now consider a pair of dual bases $\overset{*}{x}{}^{\nu}, x$ $(\nu = 1\ldots n)$ of E^* and E. Then the two metric tensors define two symmetric matrices

$$g_{\nu\lambda} = (x_\nu, x_\lambda) \quad\text{and}\quad g^{\nu\lambda} = (\overset{*}{x}{}^{\nu}, \overset{*}{x}{}^{\lambda}) .$$

Moreover, the isomorphism $\alpha: E\to E^*$ defines a matrix (α_ν^μ) by

$$\alpha x_\nu = \sum_\mu \alpha_{\nu\mu}\overset{*}{x}{}^{\mu} . \tag{10.38}$$

Scalar multiplication of (10.38) by x_λ yields

$$\alpha_{\nu\lambda} = \langle \alpha x_\nu, x_\lambda\rangle = (x_\nu, x_\lambda) = g_{\nu\lambda} .$$

Hence, the equation (10.38) can be written in the form

$$\alpha x_\nu = \sum_\lambda g_{\nu\lambda}\overset{*}{x}{}^{\lambda} \qquad (\nu = 1\ldots n) . \tag{10.39}$$

In the same way it follows that

$$\alpha^{-1}\overset{*}{x}{}^{\nu} = \sum_\lambda g^{\nu\lambda} x_\lambda \qquad (\nu = 1\ldots n) . \tag{10.40}$$

The equations (10.39) and (10.40) show that the matrices $g_{\nu\lambda}$ and $g^{\nu\lambda}$ are inverse to each other,

$$\sum_\lambda g_{\nu\lambda} g^{\lambda\mu} = \delta_\nu^\mu .$$

Now assume that the basis x_ν $(\nu = 1\ldots n)$ is orthonormal, $g_{\nu\lambda} = \delta_{\nu\lambda}$. Then it follows from (10.38) that

$$\alpha x_\nu = \sum_\lambda \delta_{\nu\lambda}\overset{*}{x}{}^{\lambda} = \overset{*}{x}{}^{\nu} \qquad (\nu = 1\ldots n) . \tag{10.41}$$

Thus, the isomorphism α maps every orthonormal basis x_ν $(\nu = 1\ldots n)$ of E into its dual basis. The formula (10.37) shows in addition that the dual basis of an orthonormal basis of E is an orthonormal basis of E^*.

10.19. Covariant components. Let x_ν $(\nu = 1 \dots n)$ be an arbitrary basis of E. Given a vector $x \in E$ the numbers

$$\xi_\nu = (x, x_\nu) \tag{10.42}$$

are called the *covariant components* of x relative to the basis x_ν. The covariant components ξ_ν and the "contravariant" components ξ^ν are related by the formula

$$\xi_\mu = \sum_\nu g_{\mu\nu} \xi^\nu . \tag{10.43}$$

In fact, inner multiplication of $x = \sum_\nu \xi^\nu x_\nu$ by x_μ yields (10.43). Multiplying (10.43) by the inverse matrix $g^{\lambda\mu}$ we find that

$$\xi^\lambda = \sum_\mu g^{\lambda\mu} \xi_\mu . \tag{10.44}$$

If the basis $x_\nu (\nu = 1 \dots n)$ is orthonormal the formulas (10.43) and (10.44) reduce to

$$\xi^\nu = \xi_\nu \tag{10.45}$$

showing that the contravariant components and the covariant components coincide if the underlying basis is orthonormal.

To justify the notation "covariant components" consider a dual space E^* of E and the isomorphism $\alpha: E \to E^*$ induced by the inner product. Denote by $\overset{*}{\xi}_\nu$ the components of the vector αx relative to the dual basis $\overset{*}{x}{}^\nu$ $(\nu = 1 \dots n)$,

$$\alpha x = \sum_\nu \overset{*}{\xi}_\nu \overset{*}{x}{}^\nu .$$

Then

$$\overset{*}{\xi}_\nu = \langle \alpha x, x_\nu \rangle = (x, x_\nu) = \xi_\nu .$$

This equation shows that the covariant components of x are equal to the components of the "covariant" vector αx relative to the dual basis.

10.20. Tensors in an inner product space. Consider the space $T_q^p(E)$ of all tensor of order (p, q) in the n-dimensional inner product space E. The isomorphism $\alpha: E \to E^*$ defined by the inner product permits to associate with every tensor Φ of order (p, q) $(p \geqq 1)$ a tensor of order $(p-1, q+1)$ in the following way: Fix an integer i $(1 \leqq i \leqq p)$ and define the tensor $G^i\Phi$ by the equation

$$\begin{aligned}&(G^i\Phi)(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^{p-1}; x_1 \dots x_{q+1})\\ &= \Phi(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^{i-1}, \alpha x_{q+1}, \overset{*}{x}{}^i \dots \overset{*}{x}{}^{p-1}; x_1 \dots x_q) .\end{aligned} \tag{10.46}$$

The operation $\Phi \to G^i\Phi$ obviously defines an isomorphism of $T_q^p(E)$

onto $T^{p-1}_{q+1}(E)$. In the same way we define the operator $G_j\colon T^p_q(E) \to T^{p+1}_{q-1}(E)$ $(1 \leqq j \leqq q)$ by

$$(G_j\Phi)(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^{p+1}; x_1 \dots x_{q-1})$$
$$= \Phi(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^p; x_1 \dots x_{j-1}, \alpha^{-1}\overset{*}{x}{}^{p+1}, x_j \dots x_{q-1}) .$$

To obtain a relation between the components of Φ and $G^i\Phi$ let $\overset{*}{e}{}^\nu$, e_ν $(\nu = 1 \dots n)$ a pair of dual bases. Substituting $\overset{*}{x}{}^i = \overset{*}{e}{}^{\nu_i}$ $(\nu = 1 \dots p-1)$ and $x_j = e_{\mu_j}$ $(j = 1 \dots q+1)$ in (10.46) we obtain

$$(G^i\Phi)(\overset{*}{e}{}^{\nu_1} \dots \overset{*}{e}{}^{\nu_{p-1}}; e_{\mu_1} \dots e_{\mu_{q+1}})$$
$$= \Phi(\overset{*}{e}{}^{\nu_1} \dots \overset{*}{e}{}^{\nu_{i-1}}, \alpha e_{\mu_{q+1}}, \overset{*}{e}{}^{\nu_i} \dots \overset{*}{e}{}^{\nu_{p-1}}; e_{\mu_1} \dots e_{\mu_q})$$
$$= \sum_\lambda g_{\mu_{q+1}\lambda}\Phi(\overset{*}{e}{}^{\nu_1} \dots \overset{*}{e}{}^{\nu_{i-1}}, \overset{*}{e}{}^{\lambda}, \overset{*}{e}{}^{\nu_i} \dots \overset{*}{e}{}^{\nu_{p-1}}; e_{\mu_1} \dots e_{\mu_p}) ,$$

whence

$$(G^i\Phi)^{\nu_1 \dots \nu_{p-1}}_{\mu_1 \dots \mu_{q+1}} = \sum_\lambda \Phi^{\nu_1 \dots \nu_{i-1}\lambda\nu_i \dots \nu_{p-1}}_{\mu_1 \dots \mu_q} g_{\lambda\mu_{q+1}} .$$

This equation shows that the components of $G^i\Phi$ are obtained from the components of Φ by "lowering" the i-the superscript. In the same way it is shown that the components of $G_j\Phi$ are obtained by "raising" the j-th subscript of Φ,

$$(G_j\Phi)^{\nu_1 \dots \nu_{p+1}}_{\mu_1 \dots \mu_{q-1}} = \sum_\lambda \Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_{j-1}\lambda\mu_j \dots \mu_{q-1}} g^{\lambda\nu_{p+1}} .$$

We finally note that the operations

$$G^1\colon T^1(E) \to T_1(E) \quad \text{and} \quad G_1\colon T_1(E) \to T^1(E)$$

coincide with the isomorphisms

$$\alpha\colon E \to E^* \quad \text{and} \quad \alpha^{-1}\colon E^* \to E$$

respectively, if $T^1(E)$ is identified with E and $T_1(E)$ is identified with E^*.

10.21. The isomorphism G. Let $\Phi \in T^p_q(E)$ be a given tensor. Raising all the subscripts and lowering all the superscripts we obtain a tensor of order (q, p) which will be denoted by $G\Phi$. The tensor $G\Phi$ is given by the explicit formula

$$G\Phi(\overset{*}{x}{}^1 \dots \overset{*}{x}{}^q; x_1 \dots x_p) = \Phi(\alpha x_1 \dots \alpha x_p, \alpha^{-1}\overset{*}{x}{}^1 \dots \alpha^{-1}\overset{*}{x}{}^q) . \quad (10.47)$$

The operation $\Phi \to G\Phi$ defines obviously an isomorphism

$$G\colon T^p_q(E) \to T^q_p(E) .$$

This isomorphism has the following properties:

$$G(\Phi \otimes \Psi) = G\Phi \otimes G\Psi \qquad \Phi \in T^p_q(E), \Psi \in T^r_s(E) \qquad (10.48)$$
$$G^2\Phi = \Phi \qquad \Phi \in T^p_q(E) \qquad (10.49)$$
$$\langle G\Phi, \Psi\rangle = \langle\Phi, G\Psi\rangle \qquad \Phi \in T^p_q(E), \Psi \in T^p_q(E) \qquad (10.50)$$
$$\langle G\Phi, G\Psi\rangle = \langle\Phi, \Psi\rangle \qquad \Phi \in T^p_q(E), \Psi \in T^q_p(E) \qquad (10.51)$$

The formulas (10.48) and (10.49) follow immediately from (10.47). The equation (10.50) is a consequence of (10.37), if Φ and Ψ are tensors of order 1. Now it follows from (10.48) and from the linearity of G that (10.50) is generally valid. The relation (10.51) is a consequence of (10.50) and (10.49):

$$\langle G\Phi, G\Psi\rangle = \langle \Phi, G^2\Psi\rangle = \langle \Phi, \Psi\rangle .$$

10.22. The inner product in $T_q^p(E)$. We now introduce a bilinear function (Φ, Ψ) in the space $T_q^p(E)$ by the equation

$$(\Phi, \Psi) = \langle G\Phi, \Psi\rangle \qquad \Phi, \Psi \in T_q^p(E) . \tag{10.52}$$

It will be shown that this bilinear function defines an inner product in $T_q^p(E)$. The symmetry of the above bilinear function follows from (10.50). To prove the definiteness observe that the bilinear function (10.52) can be expressed as

$$(\Phi, \Psi) = \sum_{(\nu)(\mu)} (G\Phi)^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_p} \Psi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} \tag{10.53}$$

in terms of the components of $G\Phi$ and Ψ. Using an orthonormal basis we find

$$(G\Phi)^{\mu_1 \dots \mu_q}_{\nu_1 \dots \nu_p} = \Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q}$$

and hence we obtain from (10.53)

$$(\Phi, \Psi) = \sum_{(\nu)(\mu)} \Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} \Psi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} .$$

Inserting $\Psi = \Phi$ we find the formula

$$(\Phi, \Phi) = \sum_{(\nu)(\mu)} \Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q} \Phi^{\nu_1 \dots \nu_p}_{\mu_1 \dots \mu_q}$$

showing that $(\Phi, \Phi) > 0$ for every tensor $\Phi \neq 0$.

In this way, an inner product is induced in every tensor-space $T_q^p(E)$ by the inner product of E. This inner product has the following additional property:

$$|\Phi \otimes \Psi| = |\Phi|\,|\Psi| \quad \Phi \in T_q^p(E), \Psi \in T_s^r(E) .$$

This formula follows from the relations (10.52), (10.48) and (7.36):

$$(\Phi \otimes \Psi, \Phi \otimes \Psi) = \langle G(\Phi \otimes \Psi), \Phi \otimes \Psi\rangle = \langle G\Phi \otimes G\Psi, \Phi \otimes \Psi\rangle$$
$$= \langle G\Phi, \Phi\rangle \langle G\Psi, \Psi\rangle = (\Phi, \Phi)\,(\Psi, \Psi) .$$

We finally notice the isomorphism G preserves the inner product. In fact,

$$(G\Phi, G\Psi) = \langle G^2\Phi, G\Psi\rangle = \langle \Phi, G\Psi\rangle = (\Phi, \Psi) .$$

10.23. The relations between G and the dual isomorphisms. Let σ be a permutation of $(1 \ldots p)$ and τ be a permutation of $(1 \ldots q)$. Then it follows from (10.47) that

$$(\tau, \sigma)\, G\Phi = G\,(\sigma, \tau)\,\Phi \qquad\qquad \Phi \in T_q^p(E).$$

Multiplication by $\varepsilon_\sigma\,\varepsilon_\tau$ and summation over all permutations σ and τ yields the formula

$$A \circ G = G \circ A \tag{10.54}$$

showing that the operator G commutes with the antisymmetry-operator. In particular, this implies that the isomorphism G sends skew-symmetric tensors into skew-symmetric tensors and thus it induces an isomorphism

$$G: S_q^p(E) \to S_p^q(E).$$

This isomorphism commutes with the skew-symmetric product and with the dual product,

$$G\,(\Phi \wedge \Psi) = G\Phi \wedge G\Psi \quad \Phi \in S_q^p(E), \Psi \in S_s^r(E) \tag{10.55}$$

$$G\,(X \vee \Phi) = GX \vee G\Phi \quad X \in S_l^k(E), \Phi \in S_q^p(E). \tag{10.56}$$

The relation (10.55) follows from (10.48) and (10.54). To prove (10.56), let $\Psi \in S_{k-q}^{l-p}(E)$ be an arbitrary tensor. Then the formulas (10.50), (9.4), (10.51) and (10.55) yield

$$\begin{aligned}\langle\!\langle G\,(X \vee \Phi), \Psi\rangle\!\rangle &= \langle\!\langle X \vee \Phi, G\Psi\rangle\!\rangle = \langle\!\langle X, \Phi \wedge G\Psi\rangle\!\rangle = \langle\!\langle GX, G\,(\Phi \wedge G\Psi)\rangle\!\rangle \\ &= \langle\!\langle GX, G\Phi \wedge \Psi\rangle\!\rangle = \langle\!\langle GX \vee G\Phi, \Psi\rangle\!\rangle,\end{aligned}$$

whence (10.56).

Now consider the isomorphism $D^p: S^p(E) \to S_{n-p}(E)$ defined by

$$D^p\Phi = \Delta \vee \Phi$$

where Δ is a normed determinant-function in E. Employing the relation (10.56) for $X = \Delta$ and observing that $G\Delta = \Delta^*$ is the dual determinant-function we obtain the formula

$$G \circ D^p = D_p \circ G. \tag{10.57}$$

In the same way it is shown that the isomorphism

$$*: S_q^p(E) \to S_{n-p}^{n-q}(E)$$

defined by

$$*\,\Phi = J^n \vee \Phi$$

commutes with the operator G: Inserting $X = J^n$ in (10.56) and observing that $G\,J^n = J^n$ we obtain

$$G\,(*\Phi) = G\,J^n \vee G\Phi = J^n \vee G\Phi = *\,(G\Phi) \qquad \Phi \in S_q^p(E),$$

whence

$$G\,* = *\,G. \tag{10.58}$$

Similarly as in sec. (8.14) we now introduce a new inner product $((\Phi,\Psi))$ in the space $S_q^p(E)$ which differs from the product (Φ,Ψ) by the factor $\frac{1}{p!\,q!}$.

$$((\Phi,\Psi)) = \frac{1}{p!\,q!}(\Phi,\Psi) = \langle\langle G\Phi,\Psi\rangle\rangle. \tag{10.59}$$

Then the isomorphisms D^p, and $*$ preserve the inner product as follows from (10.57) and (10.58):

$$\begin{aligned}((D^p\Phi, D^p\Psi)) &= \langle\langle G\, D^p\Phi, D^p\Psi\rangle\rangle \\ &= \langle\langle D_p\, G\Phi, D^p\Psi\rangle\rangle = \langle\langle G\Phi,\Psi\rangle\rangle = ((\Phi,\Psi))\end{aligned}$$

and

$$\begin{aligned}((*\Phi, *\Psi)) &= \langle\langle G *\Phi, *\Psi\rangle\rangle \\ &= \langle\langle * G\Phi, *\Psi\rangle\rangle = \langle\langle G\Phi,\Psi\rangle\rangle = ((\Phi,\Psi)).\end{aligned}$$

Problems: 1. Let e_ν $(\nu = 1 \dots n)$ be a basis of E which consists of unit vectors. Show that the i-th covariant component of a vector relative to this basis is equal to its distance from the plane defined by $(x, e_i) = 0$.

2. Given an orthonormal basis e_ν $(\nu = 1 \dots n)$ of E prove that the products

$$e_{\nu_1} \otimes \cdots \otimes e_{\nu_p} \otimes \overset{*}{e}{}^{\mu_1} \otimes \cdots \otimes \overset{*}{e}{}^{\mu_p}$$

and

$$(e_{\nu_1} \wedge \cdots \wedge e_{\nu_p}) \otimes (\overset{*}{e}{}^{\mu_1} \wedge \cdots \wedge \overset{*}{e}{}^{\mu_p}) \qquad \nu_1 < \cdots < \nu_p,\ \mu_1 < \cdots < \mu_p$$

form an orthonormal basis of $T_q^p(E)$ and $S_q^p(E)$ respectively.

3. Show that the bilinear function defined in chapter IX § 3 problem 5 is not positive definite.

4. Prove the formula

$$((a_1 \wedge \cdots \wedge a_p, b_1 \wedge \cdots \wedge b_p)) = \begin{vmatrix} (a_1, b_1) & \dots & (a_1, b_p) \\ \vdots & & \vdots \\ (a_p, b_1) & \dots & (a_p, b_p) \end{vmatrix} \qquad a_\nu \in E, \quad b_\nu \in E.$$

5. Let E be a Euclidean plane and Δ be a normed determinant-function in E. Show that the mapping

$$\psi x = G\, D^1 x \qquad x \in E$$

is a rotation through the angle $+\frac{\pi}{2}$ provided that E is oriented by means of Δ.

6. Prove that the exterior product of two vectors in an oriented Euclidean 3-space can be written as

$$[a, b] = G\, D^2\, (a \wedge b).$$

Using the above definition prove the formula

$$([a, b], [c, d]) = (a, c)\,(b, d) - (a, d)\,(b, c)\,.$$

7. Consider two p-dimensional subspaces E_1 and E_2 of E and assume that an orientation is defined in E_1 and in E_2. Select two positive bases $a_\nu (\nu = 1 \ldots p)$ and $b_\mu (\mu = 1 \ldots p)$ of E_1 and E_2 respectively and define the tensors Φ and Ψ by

$$\Phi = a_1 \wedge \cdots \wedge a_p \quad \text{and} \quad \Psi = b_1 \wedge \cdots \wedge b_p\,.$$

Prove that the quotient

$$\frac{(\Phi, \Psi)}{(\Phi, \Phi)\,(\Psi, \Psi)}$$

does not depend on the choice of the bases a_ν and b_ν. The number θ defined by

$$\cos\theta = \frac{(\Phi, \Psi)}{(\Phi, \Phi)\,(\Psi, \Psi)} \qquad (0 \leqq \theta \leqq \pi)$$

is called the angle between the subspaces E_1 and E_2. Prove: $\theta = 0$ if and only if $E_1 = E_2$ and the two orientations coincide; $\theta = \pi$ if and only if $E_1 = E_2$ and the two orientations are opposite.

Note: It should be observed that the equation $\theta = \frac{\pi}{2}$ does not imply in general that the subspaces E_1 and E_2 are orthogonal.

§ 5. Normed linear spaces

10.24. Norm-functions. Let E be a real linear space of finite or infinite dimension. A *norm-function* in E is a real-valued function $\|x\|$ having the following properties:

N_1: $\|x\| \geqq 0$ for every $x \in E$ and $\|x\| = 0$ only if $x = 0$.

N_2: $\|x + y\| \leqq \|x\| + \|y\|$.

N_3: $\|\lambda x\| = |\lambda| \cdot \|x\|$.

A linear space in which a norm-function is defined is called a *normed linear space*. The distance of two vectors x and y of a normed linear space is defined by

$$\varrho\,(x, y) = \|x - y\|\,.$$

The property N_2 implies the *triangle-inequality*

$$\varrho\,(x, y) \leqq \varrho\,(x, z) + \varrho\,(z, y)\,.$$

10.25. Examples. 1. Every inner product space is a normed linear space with the norm defined by

$$\|x\| = \sqrt{(x, x)}\,. \tag{10.60}$$

2. Let C be the linear space of all continuous functions f in the interval $0 \leqq t \leqq 1$. Then a norm is defined in C by

$$\|f\| = \max_{0 \leqq t \leqq 1} |f(t)| .$$

The conditions N_1 and N_3 are obviously satisfied. To prove N_2 observe that

$$|f(t) + g(t)| \leqq |f(t)| + |g(t)| \leqq \|f\| + \|g\|, \quad (0 \leqq t \leqq 1)$$

whence

$$\|f + g\| \leqq \|f\| + \|g\| .$$

3. Consider an n-dimensional (real or complex) linear space E and let x_ν $(\nu = 1 \ldots n)$ be a basis of E. Define the norm of a vector

$$x = \sum_\nu \xi^\nu x_\nu$$

by the equation

$$\|x\| = \sum_\nu |\xi^\nu| . \tag{10.61}$$

10.26. The norm-topology. A norm in a linear space defines a topology by the following system of neighborhoods: Given a vector $a \in E$ consider the "spheres"

$$U_a\colon \|x - a\| < \delta$$

for all possible positive numbers δ. In this way a system of subsets is distinguished in E. It is obvious that this system satisfies the conditions H_1 and H_3 of sec. 5.5. To prove H_2 let b be a point of the neighborhood U_a. Then the neighborhood

$$U_b\colon \|x - b\| < \delta_1, \quad \delta_1 = \delta - \|b - a\|$$

is contained in U_a. In fact, if $x \in U_b$, then

$$\|x - a\| \leqq \|x - b\| + \|b - a\| < \delta_1 + \|b - a\| = \delta$$

whence $x \in U_a$. It remains to be shown that every two different vectors a and b have disjoint neighborhoods. It follows from N_1 that

$$\varrho\,(a, b) > 0 .$$

Define the neighborhoods U_a and U_b by

$$U_a\colon \|x - a\| < \frac{1}{2}\,\varrho\,(a, b), \quad U_b\colon \|x - b\| < \frac{1}{2}\,\varrho\,(a, b) .$$

Assume that there exists a vector $z \in U_a \cap U_b$. Then

$$\|z - a\| < \frac{1}{2}\,\varrho\,(a, b) \quad \text{and} \quad \|z - b\| < \frac{1}{2}\,\varrho\,(a, b)$$

whence, by the triangle-inequality,

$$\varrho\,(a, b) = \|b - a\| \leqq \|b - z\| + \|z - a\| < \varrho\,(a, b) .$$

This is a contradiction and hence the intersection $U_a \cap U_b$ is empty. In this way every normed linear space becomes a Hausdorff-space.

10.27. Equivalent norms. Two norms $\|x\|$ and $|x|$ in a linear space are called *equivalent* if there exist positive numbers m and M such that

$$m\,\|x\| \leqq |x| \leqq M\,\|x\| \quad \text{for every vector } x \in E.$$

Obviously, equivalent norms induce the same topology in E (cf. sec. 5.6).

Now assume that E has a finite dimension. Then every two norms are equivalent. To prove this it is sufficient to show that an arbitrary norm $|x|$ is equivalent to the norm defined by (10.61). The triangle-inequality yields

$$|x| = |\sum_\nu \xi^\nu a_\nu| \leqq \sum_\nu |\xi^\nu|\,|a_\nu| \leqq M \sum_\nu |\xi^\nu| = M\,\|x\| \tag{10.62}$$

where

$$M = \max_{1 \leqq \nu \leqq n} |a_\nu|\,.$$

To obtain an estimation in the other direction consider $|x|$ as a function of the n components ξ^ν. The relation

$$|x - y| \leqq M\,\|x - y\| = M \sum_\nu |\xi^\nu - \eta^\nu|$$

shows that this function is continuous. Moreover, the function $|x|$ is positive on the compact set $\|x\| = 1$. Consequently, there exists a number $m > 0$ such that

$$|x| \geqq m \quad \text{for} \quad \|x\| = 1\,.$$

Now it follows from N_3 that

$$|x| \geqq m\,\|x\|\,. \tag{10.63}$$

The inequalities (10.62) and (10.63) imply that the norms $\|x\|$ and $|x|$ are equivalent.

Our result shows especially that the topology which is induced in an n-dimensional linear space by an inner product coincides with the topology defined in sec. 5.5.

10.28. The norm of an endomorphism. An endomorphism φ of a normed linear space E is called *bounded* if there exists a number M such that

$$\|\varphi x\| \leqq M\,\|x\| \qquad x \in E\,. \tag{10.64}$$

A linear combination of two bounded endomorphisms is again bounded. Hence, the set $B\,(E; E)$ of all bounded endomorphisms is a linear space.

Given a bounded endomorphism φ denote by $\|\varphi\|$ the greatest lower bound of all the numbers M in (10.64). Then

$$\|\varphi x\| \leqq \|\varphi\|\,\|x\|\,.$$

Now it will be shown that the function $\|\varphi\|$ thus obtained is indeed a norm-function in $B(E;E)$. The conditions N_1 and N_3 are obviously satisfied. To prove N_2 let φ and ψ be two bounded endomorphisms. Then

$$\|(\varphi+\psi)\,x\| = \|\varphi x + \psi x\| \leqq \|\varphi x\| + \|\psi x\| \leqq (\|\varphi\| + \|\psi\|) \cdot \|x\| \quad x \in E$$

and consequently,

$$\|\varphi + \psi\| \leqq \|\varphi\| + \|\psi\| .$$

The norm-function $\|\varphi\|$ has the following additional property:

$$\|\psi \circ \varphi\| \leqq \|\psi\| \cdot \|\varphi\| . \tag{10.65}$$

In fact,

$$\|(\psi \circ \varphi)\,x\| \leqq \|\psi\| \cdot \|\varphi x\| \leqq \|\psi\| \cdot \|\varphi\| \cdot \|x\| \quad x \in E$$

whence (10.65).

Problems: 1. Let E be a normed linear space and E_1 be a subspace of E. Show that a norm-function is defined in the factor-space E/E_1 by

$$\|\bar{x}\| = \inf_{x \in \bar{x}} \|x\| \qquad \bar{x} \in E/E_1.$$

2. An inifinite sequence of vectors x_ν $(\nu = 1, 2 \ldots)$ of a normed linear space E is called *convergent* towards x if the following condition holds: To every positive number ε there exists an integer N such that

$$\|x_n - x\| < \varepsilon \quad \text{if} \quad n > N .$$

a) Prove that every convergent sequence satisfies the following *Cauchy-criterion:* To every positive number ε there exists an integer N such that

$$|x_n - x_m| < \varepsilon \quad \text{if} \quad n > N \quad \text{and} \quad m > N .$$

b) Prove that every Cauchy-sequence*) in a normed linear space of finite dimension is convergent.

c) Give an example showing that the assertion b) is not necessarily correct if the dimension of E is infinite.

3. A normed linear space is called *complete* if every Cauchy-sequence is convergent. Let E be a complete normed linear space and φ be an endomorphism of E such that $\|\varphi\| < 1$. Prove that the series $\sum\limits_{\nu=0}^{\infty} \varphi^\nu$ is convergent and that the endomorphism

$$\psi = \sum_{\nu=0}^{\infty} \varphi^\nu$$

has the following properties:

a) $(\iota - \varphi) \circ \psi = \psi \circ (\iota - \varphi) = \iota.$

b) $\|\psi\| \leqq \dfrac{1}{1 - \|\varphi\|}$.

*) i. e. a sequence satisfying the Cauchy-criterion.

Chapter XI

Linear mappings of inner product spaces

In this chapter all the linear spaces are assumed to have finite dimension

§ 1. The adjoint mapping

11.1. Definition. Consider two inner product spaces E and F and assume that a linear mapping $\varphi: E \to F$ is given. If E^* and F^* are two linear spaces dual to E and F respectively, the mapping φ induces a dual mapping $\varphi^*: F^* \to E^*$. The mappings φ and φ^* are related by

$$\langle \overset{*}{y}, \varphi x \rangle = \langle \varphi^* \overset{*}{y}, x \rangle \qquad x \in E, \overset{*}{y} \in F^*. \quad (11.1)$$

Since inner products are defined in E and in F, these linear spaces can be considered as dual to themselves. Then the dual mapping is a linear mapping of F into E. This mapping is called the *adjoint mapping* of φ and will be denoted by $\tilde{\varphi}$. Replacing the scalar-product by the inner product in (11.1) we obtain the relation

$$(\varphi x, y) = (x, \tilde{\varphi} y) \qquad x \in E, y \in F. \quad (11.2)$$

In this way every linear mapping φ of an inner product space E into an inner product space F determines a linear mapping $\tilde{\varphi}$ of F into E.

The adjoint mapping $\tilde{\tilde{\varphi}}$ of $\tilde{\varphi}$ is again φ. In fact, the mappings $\tilde{\varphi}$ and $\tilde{\tilde{\varphi}}$ are related by

$$(\tilde{\varphi} y, x) = (y, \tilde{\tilde{\varphi}} x). \quad (11.3)$$

The equations (11.2) and (11.3) yield

$$(\varphi x, y) = (\tilde{\tilde{\varphi}} x, y) \qquad x \in E, y \in F$$

whence $\tilde{\tilde{\varphi}} = \varphi$. Hence, the relation between a linear mapping and the adjoint mapping is symmetric.

As it has been shown in see 2.12, the subspaces $\varphi(E)$ and $K(\tilde{\varphi})$ are orthogonal complements. We thus obtain the orthogonal decomposition

$$F = \varphi(E) \oplus K(\tilde{\varphi}). \quad (11.4)$$

11.2. The relation between the matrices. Employing two bases x_ν ($\nu = 1 \ldots n$) and y_μ ($\mu = 1 \ldots m$) of E and of F, we obtain from the mappings φ and $\tilde{\varphi}$ two matrices α_ν^μ and $\tilde{\alpha}_\mu^\nu$ *) defined by the equations

$$\varphi x_\nu = \sum_\varkappa \alpha_\nu^\varkappa y_\varkappa$$

and

$$\tilde{\varphi} y_\mu = \sum_\lambda \tilde{\alpha}_\mu^\lambda x_\lambda.$$

*) The subscript indicates the row.

Substituting $x = x_\nu$ and $y = y_\mu$ in (11.2) we obtain the relation

$$\sum_\varkappa \alpha_\nu^\varkappa (y_\varkappa, y_\mu) = \sum_\lambda \tilde{\alpha}_\mu^\lambda (x_\nu, x_\lambda) . \tag{11.5}$$

Introducing the components

$$g_{\nu\lambda} = (x_\nu, x_\lambda) \quad \text{and} \quad h_{\mu\varkappa} = (y_\mu, y_\varkappa)$$

of the metric tensors we can write the relation (11.5) as

$$\sum_\varkappa \alpha_\nu^\varkappa h_{\varkappa\mu} = \sum_\nu \tilde{\alpha}_\mu^\lambda g_{\nu\lambda} .$$

Multiplication by the inverse matrix $g^{\nu\varrho}$ yields the formula

$$\tilde{\alpha}_\mu^\varrho = \sum_{\varkappa, \nu} \alpha_\nu^\varkappa h_{\varkappa\mu} g^{\nu\varrho} . \tag{11.6}$$

Now assume that the bases x_ν $(\nu = 1 \ldots n)$ and y_μ $(\mu = 1 \ldots m)$ are orthonormal,

$$g_{\nu\lambda} = \delta_{\nu\lambda}, \quad h_{\varkappa\mu} = \delta_{\varkappa\mu} .$$

Then the formula (11.6) reduces to

$$\tilde{\alpha}_\mu^\varrho = \alpha_\varrho^\mu .$$

This relation shows that with respect to orthonormal bases, the matrices of adjoint mappings are transposed to each other.

11.3. The adjoint endomorphism. Let us now consider the case that $F = E$. Then to every endomorphism φ of E corresponds an adjoint endomorphism $\tilde{\varphi}$. Since $\tilde{\varphi}$ can be considered as the dual mapping of φ relative to the inner product, it follows that

$$\det \tilde{\varphi} = \det \varphi \text{ and } \operatorname{tr} \tilde{\varphi} = \operatorname{tr} \varphi .$$

(see sec. 4.8). The adjoint mapping of the product $\psi \circ \varphi$ is given by

$$\widetilde{\psi \circ \varphi} = \tilde{\varphi} \circ \tilde{\psi} .$$

The matrices of $\tilde{\varphi}$ and φ relative to an orthonormal basis are transposed to each other.

11.4. Normal mappings. An endomorphism φ of E is called *normal* if the mappings φ and $\tilde{\varphi}$ commute,

$$\tilde{\varphi} \circ \varphi = \varphi \circ \tilde{\varphi} . \tag{11.7}$$

The above relation is equivalent to the identity

$$(\varphi x, \varphi y) = (\tilde{\varphi} x, \tilde{\varphi} y) \qquad x \in E, y \in E . \tag{11.8}$$

In fact, assume that φ is normal. Then

$$(\varphi x, \varphi y) = (x, \tilde{\varphi}\, \varphi y) = (x, \varphi\, \tilde{\varphi} y) = (\tilde{\varphi} x, \tilde{\varphi} y) .$$

Conversely, the relation (11.8) implies that

$$(y, \tilde{\varphi}\,\varphi x) = (\varphi y, \varphi x) = (\tilde{\varphi} y, \tilde{\varphi} x) = (y, \varphi\,\tilde{\varphi} x)\,,$$

whence (11.7).

Substituting $y = x$ in (11.8) we obtain the equation

$$|\varphi x|^2 = |\tilde{\varphi} x|^2 \qquad x \in E$$

showing that the kernel of $\tilde{\varphi}$ coincides with the kernel of $\varphi, K(\tilde{\varphi}) = K(\varphi)$. Thus, the orthogonal decomposition (11.4) can be written as

$$E = \varphi(E) \oplus K(\varphi) \qquad \varphi(E) \perp K(\varphi)\,. \qquad (11.9)$$

It follows from (11.9) that the restriction of a normal mapping to the subspace $\varphi(E)$ is regular. Consequently, φ^2 has the same rank as φ. The same argument shows that all the mappings φ^ν ($\nu = 2, 3 \ldots$) have the same rank as φ.

11.5. The relation between endomorphisms and bilinear functions. Given an endomorphism $\varphi: E \to E$ consider the bilinear function

$$\Phi(x, y) = (\varphi x, y)\,. \qquad (11.10)$$

The correspondence $\varphi \to \Phi$ defines a linear mapping

$$L\,(E; E) \to B\,(E, E)\,. \qquad (11.11)$$

It will be shown that this linear mapping is an isomorphism of $L\,(E; E)$ onto $B\,(E, E)$. To prove that the mapping (11.11) is regular, assume that a certain endomorphism φ determines the zero-function. Then $(\varphi x, y) = 0$ for every $x \in E$ and every $y \in E$, whence $\varphi = 0$.

It remains to be shown that (11.11) is a mapping onto $B\,(E, E)$. Given a bilinear function Φ, choose a fixed vector $x \in E$ and consider the linear function f_x defined by

$$f_x(y) = \Phi\,(x, y)\,.$$

By the Riesz-theorem (cf. sec. 10.16) this function can be written in the form

$$f_x(y) = (x', y)$$

where the vector $x' \in E$ is uniquely determined by x.

Define the endomorphism φ: $E \to E$ by

$$\varphi x = x'\,.$$

Then

$$\Phi\,(x, y) = (\varphi x, y) \qquad x \in E, y \in E\,.$$

Thus, there is a one-to-one correspondence between the endomorphisms of E and the bilinear functions in E. In particular, the identity-map corresponds to the bilinear function defined by the inner product.

Let $\tilde{\Phi}$ be the bilinear function which corresponds to the adjoint endomorphism. Then

$$\tilde{\Phi}(x, y) = (\tilde{\varphi} x, y) = (x, \varphi y) = (\varphi y, x) = \Phi(y, x) .$$

This equation shows that the bilinear functions $\tilde{\Phi}$ and Φ are obtained from each other by interchanging the arguments.

Problems: 1. Consider two inner product spaces E and F. Prove that an inner product is defined in the space $L(E; F)$ by

$$(\varphi, \psi) = \operatorname{tr}(\tilde{\psi} \circ \varphi) \qquad \varphi, \psi \in L(E; F) .$$

2. If φ is normal, prove that the mappings φ and $\tilde{\varphi}$ have the same eigenvectors and the same eigenvalues.

3. Let φ and $\tilde{\varphi}$ be two adjoint endomorphisms of E. Assume that e is an eigenvector of φ and that $\tilde{e}$ is an eigenvector of $\tilde{\varphi}$. Prove that $(e, \tilde{e}) = 0$ if the corresponding eigenvalues are different.

4. Let φ be a normal endomorphism of a plane. Prove that the matrix of φ relative to an orthonormal basis has the form

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} .$$

§ 2. Selfadjoint mappings

11.6. Eigenvalue problem. An endomorphism $\varphi: E \to E$ is called *selfadjoint* if $\tilde{\varphi} = \varphi$. Written in a less condensed form this condition reads

$$(\varphi x, y) = (x, \varphi y) \qquad x, y \in E .$$

The above equation implies that the matrix of a selfadjoint endomorphism relative to an orthonormal basis is symmetric.

It is the aim of this paragraph to show that a selfadjoint endomorphism of an n-dimensional inner product space E has n eigenvectors which are mutually orthogonal.

Define the function F by

$$F(x) = \frac{(x, \varphi x)}{(x, x)} \qquad x \neq 0 . \quad (11.12)$$

This function is defined for all vectors $x \neq 0$. As a quotient of continuous functions, F is also continuous. Moreover, F is homogenous of degree zero, i. e.

$$F(\lambda x) = F(x) \qquad (\lambda \neq 0) . \quad (11.13)$$

Consider the function F on the unit sphere $|x| = 1$. Since the unit sphere is a bounded and closed subset of E, F assumes a minimum on the sphere $|x| = 1$. Let e_1 be a unit vector such that

$$F(e_1) \leqq F(x) \quad (11.14)$$

for all vectors $|x| = 1$. The relations (11.13) and (11.14) imply that

$$F(e_1) \leqq F(x) \tag{11.15}$$

for all vectors $x \neq 0$. In fact, if $x \neq 0$ is an arbitrary vector, consider the corresponding unit-vector e. Then $x = |x|\, e$, whence

$$F(x) = F(e) \geqq F(e_1)\,.$$

Now it will be shown that e_1 is an eigenvector of φ. Let y be an arbitrary vector and define the function f by

$$f(t) = F\,(e_1 + t y)\,. \tag{11.16}$$

Then it follows from (11.15) that f assumes a minimum at $t = 0$, whence $\dot{f}(0) = 0$. Inserting the expression (11.12) into (11.16) we can write

$$f\,(t) = \frac{(e_1 + t y, \varphi e_1 + t\varphi y)}{(e_1 + t y, e_1 + t y)}\,.$$

Differentiating this function at the point $t = 0$ we obtain

$$\dot{f}(0) = (e_1, \varphi y) + (y, \varphi e_1) - 2\,(e_1, \varphi e_1)\,(e_1, y)\,. \tag{11.17}$$

Since φ is selfadjoint,

$$(e_1, \varphi y) = (\varphi e_1, y)$$

and hence the equation (11.17) can be written as

$$\dot{f}(0) = 2\,(\varphi e_1, y) - 2\,(e_1, \varphi e_1)\,(e_1, y)\,. \tag{11.18}$$

We thus obtain

$$(\varphi e_1 - (e_1, \varphi e_1)\, e_1, y) = 0 \tag{11.19}$$

for every vector $y \in E$. This implies that

$$\varphi e_1 = (e_1, \varphi e_1)\, e_1\,,$$

i. e. e_1 is an eigenvector of φ and the corresponding eigenvalue is

$$\lambda_1 = (e_1, \varphi e_1)\,.$$

11.7. Representation in diagonal form. Once an eigenvector of φ has been constructed it is easy to find a system of n orthogonal eigenvectors. The construction of further eigenvectors is based upon the following property of a selfadjoint mapping: Let J be an invariant subspace of E, i. e. a subspace such that $\varphi(J) \subset J$. Then the orthogonal complement $J^{\perp}$ is again invariant. In fact, consider a vector $y \in J^{\perp}$. Then

$$(\varphi y, x) = (y, \varphi x) = 0$$

for all vectors $x \in J$, whence $\varphi y \in J^{\perp}$.

It follows from this remark that the orthogonal complement E_1 of e_1 is invariant under φ. The induced endomorphism is obviously again selfadjoint and hence the above construction can be applied to E_1. Hence, there exists an eigenvector e_2 such that $(e_1, e_2) = 0$.

Continuing this way we finally obtain a system of n eigenvectors e_ν $(\nu = 1 \ldots n)$ such that

$$(e_\nu, e_\mu) = \delta_{\nu\mu} .$$

The eigenvectors e_ν form an orthonormal basis of E. In this basis the mapping φ has the form

$$\varphi e_\nu = \lambda_\nu e_\nu \tag{11.20}$$

where λ_ν denotes the eigenvalue of e_ν. These equations show that the matrix of a selfadjoint mapping has diagonal form if the eigenvectors are used as a basis.

11.8. The eigenvector-spaces. If λ is an eigenvalue of φ, the corresponding *eigen-space* $E(\lambda)$ is the set of all vectors x satisfying the equation $\varphi x = \lambda x$ including the zero vector. Two eigen-spaces $E(\lambda)$ and $E(\lambda')$ corresponding to different eigenvalues are orthogonal. In fact, assume that

$$\varphi e = \lambda e \quad \text{and} \quad \varphi e' = \lambda' e' .$$

Then

$$(e', \varphi e) = \lambda (e, e') \quad \text{and} \quad (e, \varphi e') = \lambda' (e, e') .$$

Subtracting these equations we obtain

$$(\lambda' - \lambda)(e, e') = 0 ,$$

whence $(e, e') = 0$ if $\lambda' \neq \lambda$.

Denote by λ_ν $(\nu = 1 \ldots r)$ the different eigenvalues of φ. Then every two eigenspaces $E(\lambda_i)$ and $E(\lambda_j)$ $(i \neq j)$ are orthogonal. Since every vector $x \in E$ can be written as a linear combination of eigenvectors it follows that the direct sum of the spaces $E(\lambda_i)$ is E. We thus obtain the orthogonal decomposition

$$E = E(\lambda_1) \oplus \cdots \oplus E(\lambda_r) . \tag{11.21}$$

Let φ_i be the endomorphism induced by φ in $E(\lambda_i)$. Then

$$\varphi_i x = \lambda_i x \qquad\qquad x \in E(\lambda_i) .$$

This implies that the characteristic polynomial of φ_i is given by

$$\det(\varphi_i - \lambda\,\iota) = (\lambda - \lambda_i)^{k_i} \qquad (i = 1 \ldots r) \tag{11.22}$$

where k_i is the dimension of $E(\lambda_i)$. It follows from (11.21) and (11.22) that the characteristic polynomial of φ is equal to the product

$$\det(\varphi - \lambda\,\iota) = (\lambda_1 - \lambda)^{k_1} \ldots (\lambda_r - \lambda)^{k_r}. \tag{11.23}$$

The representation (11.23) shows that the characteristic polynomial of a selfadjoint endomorphism has n real zeros, if every zero is counted with this multiplicity. As another consequence of (11.23) we note that the dimension of the eigen-space $E(\lambda_i)$ is equal to the multiplicity of the zero λ_i in the characteristic polynomial.

11.9. The characteristic polynomial of a symmetric matrix. The above result implies that a symmetric $n \times n$-matrix $A = (\alpha_\nu^\mu)$ has n real eigenvalues. In fact, consider the endomorphism

$$\varphi x_\nu = \sum_\mu \alpha_\nu^\mu x_\mu \qquad (\nu = 1 \ldots n)$$

where x_ν $(\nu = 1 \ldots n)$ is an orthonormal basis of E. Then φ is selfadjoint and hence the characteristic polynomial of φ has the form (11.23). At the same time we know that

$$\det(\varphi - \lambda\,\iota) = \det(A - \lambda J)\,. \tag{11.24}$$

The equations (11.23) and (11.24) yield

$$\det(A - \lambda J) = (\lambda_1 - \lambda)^{k_1} \ldots (\lambda_r - \lambda)^{k_r}\,.$$

11.10. Eigenvectors of bilinear functions. In sec. 11.5 a one-to-one correspondence between all the bilinear functions Φ in E and all the endomorphisms $\varphi: E \to E$ has been defined. A bilinear function Φ and the corresponding endomorphism φ are related by the equation

$$\Phi(x, y) = (\varphi x, y) \qquad x, y \in E\,.$$

Using this relation, we define eigenvectors and eigenvalues of a bilinear function as the eigenvectors and eigenvalues of the corresponding endomorphism. Let e be an eigenvector of Φ and λ be the corresponding eigenvalue. Then

$$\Phi(e, y) = (\varphi e, y) = \lambda(e, y) \tag{11.25}$$

for every vector $y \in E$.

Now assume that the bilinear function Φ is symmetric,

$$\Phi(x, y) = \Phi(y, x)\,.$$

Then the corresponding endomorphism φ is selfadjoint. Consequently, there exists an orthonormal system of n eigenvectors e_ν

$$\varphi e_\nu = \lambda_\nu e_\nu \qquad (\nu = 1 \ldots n)\,. \tag{11.26}$$

This implies that

$$\Phi(e_\nu, e_\mu) = \lambda_\nu (e_\nu, e_\mu) = \lambda_\nu \delta_{\nu\mu}\,.$$

Hence, to every symmetric bilinear function Φ in E there exists an orthonormal basis of E in which the matrix of Φ has diagonal-form.

Problems: 1. Prove by direct computation that a symmetric 2×2-matrix has two real eigenvalues.

2. Compute the eigenvalues of the matrix

$$\begin{pmatrix} 4 & -1 & 2 \\ -1 & -2 & -\frac{5}{2} \\ 2 & -\frac{5}{2} & 1 \end{pmatrix}$$

3. Find the eigenvalues of the bilinear function

$$\Phi(x, y) = \sum_{\nu \neq \mu} \xi^\nu \eta^\mu .$$

4. Prove that the product of two selfadjoint endomorphisms φ and ψ is selfadjoint if and only if $\psi \circ \varphi = \varphi \circ \psi$.

5. A selfadjoint endomorphism φ is called *positive,* if

$$(x, \varphi x) \geqq 0$$

for every $x \in E$. Given a positive selfadjoint endomorphism φ, prove that there exists exactly one positive selfadjoint endomorphism ψ such that $\psi^2 = \varphi$.

6. Given a selfadjoint mapping φ, consider a vector $b \in K(\varphi)^\perp$. Prove that there exists exactly one vector $a \in K(\varphi)^\perp$ such that $\varphi a = b$.

7. Let φ be a selfadjoint mapping and e_ν $(\nu = 1 \ldots n)$ be a system of n orthonormal eigenvectors. Define the mapping φ_λ by

$$\varphi_\lambda = \varphi - \lambda\, \iota$$

where λ is a real parameter. Prove that

$$\varphi_\lambda^{-1} x = \sum_\nu \frac{(x, e_\nu)}{\lambda_\nu - \lambda} e_\nu \qquad x \in E,$$

provided that λ is not an eigenvalue of φ.

8. Let φ be an endomorphism of a real n-dimensional linear space E. Show that an inner product can be introduced in E such that φ becomes a selfadjoint mapping if and only φ has n linearly independent eigenvectors.

9. Let φ be an endomorphism of E and $\tilde{\varphi}$ the adjoint endomorphism. Denote by $|\varphi|$ the norm of φ which is induced by the Euclidean norm of E (cf. sec. 10.28). Prove that

$$|\varphi|^2 = \lambda$$

where λ is the largest eigenvalue of the selfadjoint mapping $\tilde{\varphi} \circ \varphi$.

§ 3. Orthogonal projections

11.11. Definition. Consider a subspace E_1 of E. Then every vector $x \in E$ can be uniquely decomposed as

$$x = p + h \qquad p \in E_1,\ h \in E_1^\perp .$$

Consequently, an endomorphism π of E is defined by

$$\pi x = p\,.$$

This endomorphism is called the *orthogonal projection* of E onto E_1. The mapping π reduces to the identity-map in E_1 and to the zero-map in $E_1^\perp$. Since $\pi x \in E_1$ for every vector $x \in E$, it follows that

$$\pi^2 x = \pi(\pi x) = \pi x \qquad x \in E$$

whence $\pi^2 = \pi$.

An orthogonal projection is selfadjoint. In fact,

$$(\pi x, y) = (\pi x, \pi y + h) = (\pi x, \pi y)\,.$$

Interchanging x and y in this relation, we obtain

$$(\pi y, x) = (\pi y, \pi x)\,,$$

whence

$$(\pi x, y) = (x, \pi y) \qquad x, y \in E\,.$$

Conversely, it will be shown that every selfadjoint endomorphism φ satisfying the condition $\varphi^2 = \varphi$ is an orthogonal projection onto a certain subspace. Since φ is selfadjoint, the space E is the direct sum of kernel and image-space,

$$E = \varphi(E) \oplus K\,, \qquad \varphi(E) \perp K\,.$$

The equation $\varphi^2 = \varphi$ implies that φ reduces to the identity in the subspace $\varphi(E)$. In fact, let $y = \varphi x$ be a vector of $\varphi(E)$. Then

$$\varphi y = \varphi^2 x = \varphi x = y\,.$$

Now the above decomposition shows that φ is the orthogonal projection of E onto $\varphi(E)$.

Consider two subspaces E_1 and E_2 of E and the corresponding projections $\pi_1: E \to E_1$ and $\pi_2: E \to E_2$. It will be shown that $\pi_2 \circ \pi_1 = 0$ if and only if E_1 and E_2 are orthogonal to each other. Assume first that $E_1 \perp E_2$. Then $\pi_1 x \in E_2^\perp$ for every vector $x \in E$, whence $\pi_2 \circ \pi_1 = 0$. Conversely, the equation $\pi_2 \circ \pi_1 = 0$ implies that $\pi_1 x \in E_2^\perp$ for every vector $x \in E$, whence $E_1 \in E_2^\perp$.

11.12. Sum of two projections. The sum of two projections $\pi_1: E \to E_1$ and $\pi_2: E \to E_2$ is again a projection if and only if the subspaces E_1 and E_2 are orthogonal. Assume first that $E_1 \perp E_2$ and consider the endomorphism $\pi = \pi_1 + \pi_2$. Then

$$\pi(x_1 + x_2) = \pi x_1 + \pi x_2 = x_1 + x_2 \qquad x_1 \in E_1, x_2 \in E_2\,.$$

Hence, π reduces to the identity-map in the sum $E_1 \oplus E_2$. On the other hand,

$$\pi x = 0 \quad \text{if} \quad x \in E_1^\perp \cap E_2^\perp\,.$$

But $E_1^\perp \cap E_2^\perp$ is the orthogonal complement of the sum $E_1 + E_2$ and hence π is the projection of E onto $E_1 + E_2$.

Conversely, assume that $\pi_1 + \pi_2$ is a projection. Then

$$(\pi_1 + \pi_2)^2 = \pi_1 + \pi_2\,,$$

whence

$$\pi_1 \circ \pi_2 + \pi_2 \circ \pi_1 = 0\,. \tag{11.27}$$

This equation implies that

$$\pi_1 \circ \pi_2 \circ \pi_1 + \pi_2 \circ \pi_1 = 0 \tag{11.28}$$

and

$$\pi_1 \circ \pi_2 + \pi_1 \circ \pi_2 \circ \pi_1 = 0\,. \tag{11.29}$$

Adding (11.28) and (11.29) and using (11.27) we obtain

$$\pi_1 \circ \pi_2 \circ \pi_1 = 0\,. \tag{11.30}$$

The equations (11.30) and (11.28) yield

$$\pi_2 \circ \pi_1 = 0\,.$$

This implies that $E_1 \perp E_2$, as it has been shown at the end of sec. 11.11.

11.13. Difference of two projections. The difference $\pi_1 - \pi_2$ of two projections $\pi_1 \colon E \to E_1$ and $\pi_2 \colon E \to E_2$ is a projection if and only if E_2 is a subspace of E_1. To prove this, consider the endomorphism

$$\varphi = \iota - (\pi_1 - \pi_2) = (\iota - \pi_1) + \pi_2\,.$$

Since $\iota - \pi_1$ is the projection $E \to E_1^\perp$, it follows that φ is a projection if and only if $E_1^\perp \subset E_2^\perp$, i. e., if and only if $E_1 \supset E_2$. If this condition is fulfilled, φ is the projection onto the subspace $E_1^\perp + E_2$. This implies that $\pi_1 - \pi_2 = \iota - \varphi$ is the projection onto the subspace

$$(E_1^\perp + E_2)^\perp = E_1 \cap E_2^\perp\,.$$

This subspace is the orthogonal complement of E_2 relative to E_1.

11.14. Product of two projections. The product of two projections $\pi_1 \colon E \to E_1$ and $\pi_2 \colon E \to E_2$ is an orthogonal projection if and only if the projections commute. Assume first that $\pi_2 \circ \pi_1 = \pi_1 \circ \pi_2$. Then

$$\pi_2\, \pi_1 x = \pi_2 x = x \quad \text{for every vector} \quad x \in E_1 \cap E_2\,. \tag{11.31}$$

On the other hand, $\pi_2 \circ \pi_1$ reduces to the zero-map in the subspace $(E_1 \cap E_2)^\perp = E_1^\perp + E_2^\perp$. In fact, consider a vector

$$x = x_1^\perp + x_2^\perp \qquad x_1^\perp \in E_1^\perp,\, x_2^\perp \in E_2^\perp\,.$$

Then

$$\pi_2\, \pi_1 x = \pi_2\, \pi_1 x_1^\perp + \pi_2\, \pi_1 x_2^\perp = \pi_2\, \pi_1 x_1^\perp + \pi_1\, \pi_2 x_2^\perp = 0. \tag{11.32}$$

The equations (11.31) and (11.32) show that $\pi_2 \circ \pi_1$ is the projection $E \to E_1 \cap E_2$.

Conversely, if $\pi_2 \circ \pi_1$ is a projection, it follows that

$$\pi_2 \circ \pi_1 = (\pi_2 \circ \pi_1)^* = \pi_1^* \circ \pi_2^* = \pi_1 \circ \pi_2 .$$

Problems: 1. Prove that a subspace $J \subset E$ is invariant under the projection $\pi: E \to E_1$ (i. e. $\pi(J) \subset J$) if and only if

$$J = J \cap E_1 \oplus J \cap E_1^{\perp} .$$

2. Prove that two projections $\pi_1: E \to E_1$ and $\pi_2: E \to E_2$ commute if and only if

$$E_1 + E_2 = E_1 \cap E_2 + E_1 \cap E_2^{\perp} + E_1^{\perp} \cap E_2 .$$

3. The *reflection* ϱ of E at a subspace E_1 is defined by

$$\varrho x = p - h$$

where $x = p + h$ ($p \in E_1, h \in E_1^{\perp}$). Show that the reflection ϱ and the projection $\pi: E \to E_1$ are related by

$$\varrho = 2\pi - \iota .$$

4. Consider an endomorphism φ of a real linear space E. Prove that an inner product can be introduced in E such that becomes an orthogonal projection if and only if $\varphi^2 = \varphi$.

5. Given two projections π_1 and π_2 prove that

$$\pi = \pi_1 + \pi_2 - (\pi_1 \circ \pi_2 + \pi_2 \circ \pi_1)$$

is again a projection.

6. Given a selfadjoint endomorphism φ of E, consider the distinct eigenvalues λ_i and the corresponding eigenspaces E_i ($i = 1 \ldots r$). If π_i denotes the orthogonal projection $E \to E_i$ prove the relations:

a) $\pi_i \circ \pi_j = 0 \quad (i \neq j)$.

b) $\sum_i \pi_i = \iota$.

c) $\sum_i \lambda_i \pi_i = \varphi$.

§ 4. Skew mappings

11.15. Definition. An endomorphism ψ of E is called *skew* if $\tilde{\psi} = -\psi$. The above condition is equivalent to the relation

$$(\psi x, y) + (x, \psi y) = 0 \qquad x, y \in E . \tag{11.33}$$

It follows from (11.33) that the matrix of a skew mapping relative to an orthonormal basis is skew-symmetric.

Substitution of $y = x$ in (11.33) yields the equation

$$(x, \psi x) = 0 \qquad x \in E \qquad (11.34)$$

showing that every vector is orthogonal to its image-vector. Conversely, an endomorphism ψ having this property is skew. In fact, replacing x by $x + y$ in (11.34) we obtain

$$(x + y, \psi x + \psi y) = 0 ,$$

whence

$$(y, \psi x) + (x, \psi y) = 0 .$$

It follows from (11.34) that a skew mapping can only have the eigenvalue $\lambda = 0$.

The relation $\tilde{\psi} = -\psi$ implies that

$$\operatorname{tr} \psi = 0$$

and

$$\det \psi = (-1)^n \det \psi .$$

The last equation shows that

$$\det \psi = 0$$

if the dimension of E is odd. More general, it will now be shown that the rank of a skew endomorphism is always even. Since every skew mapping is normal (see sec. 11.4) the image space $\psi(E)$ is the orthogonal complement of the kernel. Consequently, the induced endomorphism $\psi_1 : \psi(E) \rightarrow \psi(E)$ is regular. At the same time ψ_1 is again skew and thus the dimension of $\psi(E)$ must be even.

It follows from this result that the rank of a skew-symmetric matrix is always even.

11.16. The normal-form of a skew-symmetric matrix. In this section it will be shown that to every skew mapping ψ an orthonormal basis a_ν $(\nu = 1 \ldots n)$ can be constructed in which the matrix of ψ has the form

$$\begin{pmatrix} 0 & \varkappa_1 & & & & & & & \\ -\varkappa_1 & 0 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 0 & \varkappa_p & & & & \\ & & & -\varkappa_p & 0 & & & & \\ & & & & & 0 & & & \\ & & & & & & \ddots & & \\ & & & & & & & & 0 \end{pmatrix} \qquad (11.35)$$

Consider the mapping $\varphi = \psi^2$. Then $\tilde{\varphi} = \varphi$. According to the result of sec. 11.7, there exists an orthonormal basis $e_\nu (\nu = 1 \dots n)$ in which φ has the form

$$\varphi e_\nu = \lambda_\nu e_\nu \qquad (\nu = 1 \dots n) .$$

All the eigenvalues λ_ν are negative or zero. In fact, the equation

$$\varphi e = \lambda e$$

implies that

$$\lambda = (e, \varphi e) = (e, \psi^2 e) = -(\psi e, \psi e) \leqq 0 .$$

Since the rank of ψ is even and ψ^2 has the same rank as ψ, the rank of φ must be even. Consequently, the number of the negative eigenvalues is even and we can enumerate the vectors e_ν $(\nu = 1 \dots n)$ such that

$$\lambda_\nu < 0 \quad (\nu = 1 \dots 2p) \quad \text{and} \quad \lambda_\nu = 0 \quad (\nu = 2p + 1 \dots n) .$$

Define the orthonormal basis a_ν $(\nu = 1 \dots n)$ by

$$a_{2\nu-1} = e_\nu, \quad a_{2\nu} = \frac{1}{\varkappa_\nu} \psi e_\nu \quad \varkappa_\nu = \sqrt{-\lambda_\nu} \quad (\nu = 1 \dots p)$$

and

$$a_\nu = e_\nu \quad (\nu = 2p + 1 \dots n) .$$

In this basis the matrix of ψ has the form (11.35).

11.17. Skew endomorphisms of an even-dimensional space. It is a classical theorem in matrix-theory, that the determinant of a skew-symmetric matrix of even order is the square of the rational function of the entries. In the present section a prove of this fact will be supplied which is based upon the skew-symmetric product.

Let ψ be a skew endomorphism of the even-dimensional space E. Then a skew-symmetric tensor of second order is defined by

$$\Psi(x, y) = (\psi x, y)$$

and hence the product

$$\Psi^k = \frac{1}{k!} \Psi \wedge \dots \wedge \Psi \qquad (n = 2k)$$

is a skew-symmetric tensor of order n.

Now introduce an orientation in E by a normed determinant-function Δ. Then Ψ^k is a multiple of Δ,

$$\Psi^k (x_1 \dots x_n) = \lambda(\psi)\, \Delta (x_1 \dots x_n) \quad x_\nu \in E . \tag{11.36}$$

In this way, a scalar $\lambda(\psi)$ is assigned to every skew endomorphism of E. Note that the scalar $\lambda(\psi)$ depends on the orientation. If the orientation is reversed, $\lambda(\psi)$ changes into $-\lambda(\psi)$.

It will be shown, that

$$\det \psi = \lambda(\psi)^2. \tag{11.37}$$

As it has been proved in sec. 11.16 there exists an orthonormal basis $a_\nu (\nu = 1 \dots n)$ of E, in which the matrix of ψ has the form (11.35). The components of the tensor Ψ relative to this basis are given by

$$(\Psi_{\nu\mu}) = \begin{pmatrix} 0 & \varkappa_1 & & & & \\ -\varkappa_1 & 0 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & 0 & \varkappa_k \\ & & & & -\varkappa_k & 0 \end{pmatrix} \quad n = 2k, \; \varkappa_\nu \geqq 0 \quad (11.38)$$

To find $\Psi^k (a_1 \dots a_n)$ we employ the formula (8.11) for the skew-symmetric product. This yields

$$\Psi^k (a_1 \dots a_n) = \frac{1}{2^k k!} \sum_\sigma \varepsilon_\sigma \Psi_{\sigma(1)\sigma(2)} \dots \Psi_{\sigma(n-1)\sigma(n)} .$$

In view of the skew symmetry of Ψ this can be written as

$$\Psi^k (a_1 \dots a_n) = \frac{1}{k!} \sum_\sigma \varepsilon_\sigma \Psi_{\sigma(1)\sigma(2)} \dots \Psi_{\sigma(n-1)\sigma(n)}$$

where the sum is extended over all those permutations σ of $(1 \dots n)$ for which $\sigma(2\nu - 1) < \sigma(2\nu)$ $(\nu = 1 \dots k)$.

Now it follows from (11.38) that

$$\Psi_{\nu\mu} = \begin{cases} 0 \text{ for } \nu < \mu \,, \text{ if } \nu \text{ is even} \\ \delta_{\nu\,\nu+1} \varkappa_\nu, \text{ if } \nu \text{ is odd} . \end{cases} \quad (11.39)$$

The first relation (11.39) implies that we can restrict the summation to all those permutations σ which permute the even and the odd numbers separately. Every such permutation is even and hence we can write

$$\Psi^k (a_1 \dots a_n) = \frac{1}{k!} \sum_\sigma \Psi_{\sigma(1)\sigma(2)} \dots \Psi_{\sigma(n-1)\sigma(n)}$$

where σ runs over all permutations transforming the even numbers into numbers. Now it follows from the second relation (11.39) that there remain only those terms for which $\sigma(2\nu) = \sigma(2\nu - 1) + 1$. We thus obtain

$$\Psi^k (a_1 \dots a_n) = \frac{1}{k!} \sum_\varrho \Psi_{\varrho(1)\,\varrho(1)+1} \dots \Psi_{\varrho(n-1)\,\varrho(n-1)+1}$$

where the sum is taken over all permutations ϱ of the numbers $(1, 3 \dots n - 1)$. Rearranging the factors we see that every term of this sum is equal to $\Psi_{12} \Psi_{34} \dots \Psi_{n-1\,n}$. Hence,

$$\Psi^k (a_1 \dots a_n) = \Psi_{12} \Psi_{34} \dots \Psi_{n-1\,n} = \varkappa_1 \varkappa_2 \dots \varkappa_k .$$

Inserting $x_\nu = a_\nu (\nu = 1 \dots n)$ into (11.36) we thus find that

$$\varkappa_1 \dots \varkappa_k = \lambda(\psi) \Delta (a_1 \dots a_n) = \varepsilon \lambda(\psi)$$

where $\varepsilon = +1$ or $\varepsilon = -1$ depending on whether the basis a_ν is positive or negative with respect to the given orientation. At the same time it follows from (11.38) that

$$\det \psi = \varkappa_1^2 \dots \varkappa_k^2 .$$

Combining these equations we obtain the formula (11.37).

The theorem which has been mentioned in the beginning is an immediate consequence of the relation (11.37). In fact, let $e_\nu (\nu = 1 \dots n)$ be an orthonormal basis of E and $A = (\alpha_\nu^\mu)$ the corresponding matrix of ψ. Then the equation (11.36) yields

$$\lambda(\psi) = \varepsilon \, \Psi^k (e_1, \dots e_n) = \frac{1}{2^k k!} \varepsilon \sum_\sigma \varepsilon_\sigma \Psi_{\sigma(1)\,\sigma(2)} \dots \Psi_{\sigma(n-1)\,\sigma(n)}$$

$$= \frac{1}{2^k k!} \varepsilon \sum_\sigma \varepsilon_\sigma \alpha_{\sigma(1)\,\sigma(2)} \dots \alpha_{\sigma(n-1)\,\sigma(n)} .$$

We thus obtain from (11.37) the formula

$$\det A = \left\{ \frac{1}{2^k k!} \sum_\sigma \varepsilon_\sigma \alpha_{\sigma(1)\,\sigma(2)} \dots \alpha_{\sigma(n-1)\,\sigma(n)} \right\}^2$$

showing that the determinant of A is the square of a rational function of the entries $\alpha_{\nu\mu}$.

Problems: 1. Show that every skew endomorphism φ of a 2-dimensional inner product space satisfies the relation

$$(\varphi x, \varphi y) = \det \varphi \cdot (x, y) .$$

2. Show that every skew endomorphism φ of an oriented 3-space can be written as

$$\varphi x = [b, x]$$

and that the vector b is uniquely determined by φ. Prove that the components of b relative to a positive orthonormal basis are obtained from the matrix $(\alpha_{\nu\mu})$ of φ by the formulas

$$\beta^1 = \alpha_{23}, \ \beta^2 = \alpha_{31}, \ \beta^3 = \alpha_{12} .$$

3. Assume that $\varphi \neq 0$ and ψ are two skew endomorphisms of the 3-space having the same kernel. Prove that $\psi = \lambda \varphi$ where φ is a scalar.

4. Applying the result of problem 3 to the mappings

$$\varphi x = [[a_1, a_2], x]$$

and

$$\psi x = a_2 (a_1, x) - a_1 (a_2, x)$$

prove the formula

$$[[a_1, a_2], a_3] = a_2 (a_1, a_3) - a_1 (a_2, a_3) .$$

5. Show that every skew-symmetric bilinear function Φ in an oriented 3-space E can be represented in the form

$$\Phi(x, y) = ([x, y], a)$$

and that the vector a is uniquely determined by Φ.

6. Prove that the product of a selfadjoint mapping and a skew mapping has the trace zero.

7. Prove that the characteristic polynomial of a skew mapping satisfies the equation

$$\chi(-\lambda) = (-1)^n \chi(\lambda).$$

From this relation derive that the coefficient of $\lambda^{n-\nu}$ is zero for every odd ν.

8. Let φ be an endomorphism of a real linear space E. Prove that an inner product can be defined in E such that φ becomes a skew mapping if and only if the following conditions are satisfied: 1. The space E can be decomposed into invariant planes. 2. The endomorphisms which are induced in these planes have positive determinant and the trace zero.

9. Given a skew-symmetric 4×4-matrix $A = (\alpha_{\nu\mu})$ verify the identity

$$\det A = (\alpha_{12}\alpha_{34} + \alpha_{13}\alpha_{42} + \alpha_{14}\alpha_{23})^2.$$

§ 5. Isometric mappings

11.18. Definition. Consider two inner product spaces E and F. A linear mapping $\varphi: E \to F$ is called *isometric* if the inner product is preserved under φ,

$$(\varphi x_1, \varphi x_2) = (x_1, x_2) \qquad x_1, x_2 \in E.$$

Inserting $x_1 = x_2 = x$ in we find

$$|\varphi x| = |x| \qquad x \in E.$$

Conversely, the above relation implies that φ is isometric. In fact,

$$\begin{aligned} 2(\varphi x_1, \varphi x_2) &= |\varphi(x_1 + x_2)|^2 - |\varphi x_1|^2 - |\varphi x_2|^2 \\ &= |x_1 + x_2|^2 - |x_1|^2 - |x_2|^2 = 2(x_1, x_2). \end{aligned}$$

Since an isometric mapping preserves the norm it is always regular.

We assume in the following that the spaces E and F have the same dimension. Then every isometric mapping $\varphi: E \to F$ is an isomorphism of E onto F and hence there exists an inverse isomorphism $\varphi^{-1}: F \to E$. The isometry of φ implies that

$$(\varphi x, y) = (x, \varphi^{-1} y) \qquad x \in E, y \in F,$$

whence

$$\tilde{\varphi} = \varphi^{-1}. \tag{11.40}$$

Conversely every isomorphism φ satisfying the equation (11.40) is isometric. In fact,

$$(\varphi x_1, \varphi x_2) = (x_1, \tilde{\varphi}\, \varphi x_2) = (x_1, \varphi^{-1} \varphi x_2) = (x_1, x_2) .$$

The image of an orthonormal basis a_ν $(\nu = 1 \dots n)$ of E under an isometric mapping is an orthonormal basis of F. Conversely, a linear mapping which sends an orthonormal basis of E into an orthonormal basis b_ν $(\nu = 1 \dots n)$ of F is isometric. To prove this, consider two vectors

$$x_1 = \sum_\nu \xi_1^\nu a_\nu \quad \text{and} \quad x_2 = \sum_\nu \xi_2^\nu a_\nu ;$$

then

$$\varphi x_1 = \sum_\nu \xi_1^\nu b_\nu \quad \text{and} \quad \varphi x_2 = \sum_\nu \xi_2^\nu b_\nu ,$$

whence

$$(\varphi x_1, \varphi x_2) = \sum_{\nu,\mu} \xi_1^\nu \xi_2^\mu (y_\nu, y_\mu) = \sum_{\nu,\mu} \xi_1^\nu \xi_2^\mu \delta_{\nu\mu} = \sum_\nu \xi_1^\nu \xi_2^\nu = (x_1, x_2) .$$

It follows from this remark that an isometric mapping can be defined between any two inner product spaces E and F of the same dimension: Select orthonormal bases a_ν and $b_\nu (\nu = 1 \dots n)$ in E and in F respectively and define φ by $\varphi a_\nu = b_\nu (\nu = 1 \dots n)$.

11.19. The condition for the matrix. Assume that an isometric mapping $\varphi : E \to F$ is given. Employing two bases a_ν and b_ν $(\nu = 1 \dots n)$ we obtain from φ an $n \times n$-matrix α_ν^μ by the equations

$$\varphi a_\nu = \sum_\mu \alpha_\nu^\mu b_\mu .$$

Then the equations

$$(\varphi a_\nu, \varphi a_\mu) = (a_\nu, a_\mu)$$

can be written as

$$\sum_{\lambda,\varkappa} \alpha_\nu^\lambda \alpha_\mu^\varkappa (b_\lambda, b_\varkappa) = (a_\nu, a_\mu) .$$

Introducing the components

$$g_{\nu\mu} = (a_\nu, a_\mu) \quad \text{and} \quad h_{\lambda\varkappa} = (b_\lambda, b_\varkappa)$$

of the metric tensors in E and in F we obtain the relation

$$\sum_{\lambda,\varkappa} \alpha_\nu^\lambda \alpha_\mu^\varkappa h_{\lambda\varkappa} = g_{\nu\mu} . \tag{11.41}$$

Conversely, (11.41) implies that the inner products of the basis-vectors are preserved under φ and hence that φ is an isometric mapping.

If the bases a_ν and b_ν are orthonormal,

$$g_{\nu\mu} = \delta_{\nu\mu} , \quad h_{\lambda\varkappa} = \delta_{\lambda\varkappa} ,$$

the relation (11.41) reduces to

$$\sum_{\lambda} \alpha_{\nu}^{\lambda} \alpha_{\mu}^{\lambda} = \delta_{\nu\mu}$$

showing that the matrix of an isometric mapping relative to orthonormal bases is orthogonal.

11.20. Rotations. A *rotation* of an inner product space E is an isometric mapping of E into itself. The formula (11.40) implies that

$$(\det \varphi)^2 = 1$$

showing that the determinant of a rotation is ± 1.

A rotation is called *proper* if $\det \varphi = + 1$ and *improper* if $\det \varphi = - 1$.

Every eigenvalue of a rotation is ± 1. In fact, the equation $\varphi e = \lambda e$ implies that $|e| = |\lambda| \, |e|$, whence $\lambda = \pm 1$. A rotation need not have eigenvectors as can already be seen in the plane. Now assume that the dimension of E is odd. Then every endomorphism with positive determinant has at least one positive eigenvalue (cf. 4.21). This implies, that *a proper rotation of an odd-dimensional space has at least one fix-vector.* In the same way it is shown, that to every improper rotation of an odd-dimensional space there exists always a vector $e \neq 0$ such that $\varphi e = - e$.

If the dimension of E is even, nothing can be said about proper rotations. However, to every improper rotation, there is a fix-vector and a vector e such that $\varphi e = - e$ (cf. sec. 4.21).

The product of two rotations is obviously again a rotation and the inverse of a rotation is also a rotation. In other words, the set of all rotations of an n-dimensional inner product space forms a group, called the *general orthogonal group*. The relation

$$\det (\varphi_2 \circ \varphi_1) = \det \varphi_2 \cdot \det \varphi_1$$

implies that the set of all proper rotations forms a subgroup, the *special orthogonal group*.

11.21. Decomposition into invariant planes and lines. Using the results of § 2 it will now be shown, that, for every rotation φ there exists an orthogonal decomposition of E into invariant subspaces of dimensions 1 and 2. The construction below is based on the following property of a rotation: The orthogonal complement of an invariant subspace is again invariant. In fact, let J be an invariant subspace and y a vector of $J^{\perp}$. Then

$$(x, \varphi y) = (\varphi^{-1} x, y) = 0$$

for all vectors $x \in J$, whence $\varphi y \in J^{\perp}$.

Denote by E_1 and by E_2 the eigenspaces which correspond to the eigenvectors $\lambda = 1$ and $\lambda = -1$ respectively. Then E_1 is orthogonal to E_2:

Let $x_1 \in E_1$ and $x_2 \in E_2$ be two arbitrary vectors. Then

$$\varphi x_1 = x_1 \quad \text{and} \quad \varphi x_2 = -x_2 .$$

These equations yield

$$(x_1, x_2) = -(x_1, x_2) ,$$

whence $(x_1, x_2) = 0$.

It follows from the above remark that the subspace $F = (E_1 \oplus E_2)^{\perp}$ is again invariant under φ. Moreover, F does not contain an eigenvector of φ and hence F has even dimension. Now consider the selfadjoint endomorphism

$$\psi = \varphi + \tilde{\varphi} = \varphi + \varphi^{-1}$$

of F. The result of sec. 11.6 assures that there exists an eigenvector e of ψ. If λ denotes the corresponding eigenvalue we have the relation

$$\varphi e + \varphi^{-1} e = \lambda e .$$

Applying φ we obtain

$$\varphi^2 e = \lambda \varphi e - e . \tag{11.42}$$

Since there are no eigenvectors of φ in F the vectors e and φe are linearly independent and hence they generate a plane F_1. The equation (11.42) shows that this plane is invariant under φ. The induced endomorphism is a *proper* rotation (otherwise there would be eigenvectors in F_1).

The orthogonal complement $F_1^{\perp}$ of F_1 with respect to F is again invariant under φ and hence the same construction can be applied to $F_1^{\perp}$. Continuing in this way we finally obtain an orthogonal decomposition of F into orthogonal invariant planes.

Now select orthonormal bases in E_1, E_2 and in every invariant plane. These bases determine together an orthonormal basis of E. In this basis the matrix of φ has the form

$$\begin{pmatrix} \varepsilon_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \varepsilon_p & & & & & \\ & & & \cos\theta_1 & \sin\theta_1 & & & \\ & & & -\sin\theta_1 & \cos\theta_1 & & & \\ & & & & & \ddots & & \\ & & & & & & \cos\theta_k & \sin\theta_k \\ & & & & & & -\sin\theta_k & \cos\theta_k \end{pmatrix} \quad \begin{array}{l} \varepsilon_\nu = \pm 1 \quad (\nu = 1 \ldots p) \\ 2k = n - p \end{array}$$

Problems: 1. Given a skew endomorphism ψ of E, prove that

$$\varphi = (\psi + \iota) \circ (\psi - \iota)^{-1}$$

is a rotation and that -1 is not an eigenvalue of φ. Conversely, if φ is a rotation, not having the eigenvalue -1 prove that

$$\psi = (\varphi - \iota) \circ (\varphi + \iota)^{-1}$$

is a skew endomorphism.

2. An endomorphism φ of E is called *homothetic* if $\varphi = \lambda\tau$ where λ is a real number and τ is a proper rotation. Show that a homothetic mapping $\varphi \neq 0$ preserves the angles. Conversely, prove that every automorphism of E which preserves the angles is homothetic.

3. Assume two inner products Φ and Ψ in E such that all angles with respect to Φ and Ψ coincide. Prove that $\Psi(x, y) = \lambda\Phi(x, y)$ where $\lambda > 0$ is a constant.

4. Prove that every normal endomorphism of a plane is homothetic.

5. Let φ be a mapping of the inner product space E into itself such that

$$|\varphi x - \varphi y| = |x - y| \qquad x, y \in E .$$

Prove that φ is then linear.

6. Prove that to every proper rotation φ there exists a continuous family φ_t $(0 \leqq t \leqq 1)$ of rotations such that $\varphi_0 = \varphi$ and $\varphi_1 = \iota$.

7. Let φ be an automorphism of an n-dimensional real linear space E. Show that an inner product can be defined in E such that φ becomes a rotation if and only if the following conditions are fulfilled:

1. The space E can be decomposed into ineducible invariant planes*) and invariant straight lines.

2. Every invariant straight line remains pointwise fixed or is reflected at the point 0.

3. In every irreducible invariant plane an automorphism ψ is induced such that

$$\det\psi = 1 \quad \text{and} \quad |\operatorname{tr}\psi| < 2 .$$

8. If φ is a rotation of an n-dimensional Euclidean space, show that $|\operatorname{tr}\varphi| \leqq n$.

§ 6. Rotations of the plane and of the 3-space

11.22. Proper rotations of the plane. Let E be a Euclidean plane and φ be a proper rotation of E. Employing an orthonormal basis e_1, e_2 of E we can write

$$\varphi e_1 = \alpha e_1 + \beta e_2 \qquad \alpha^2 + \beta^2 = 1 . \tag{11.43}$$

Since φe_2 is orthogonal to φe_1, it follows that

$$\varphi e_2 = \pm(-\beta e_1 + \alpha e_2) . \tag{11.44}$$

*) i. e. a plane which can not be decomposed into invariant straight lines.

Computing the determinant of φ from (11.43) and (11.44) we obtain

$$\det\varphi = \pm(\alpha^2 + \beta^2) = \pm 1 .$$

Since φ is a proper rotation the determinant must be $+1$ and hence the $+$ sign stands in (11.44),

$$\varphi e_2 = -\beta e_1 + \alpha e_2 . \tag{11.45}$$

Given an arbitrary vector

$$x = \xi e_1 + \eta e_2$$

we obtain from (11.43) and (11.45)

$$\varphi x = (\alpha\xi - \beta\eta)\, e_1 + (\beta\xi + \alpha\eta)\, e_2$$

whence

$$(x, \varphi x) = \alpha(\xi^2 + \eta^2) = \alpha(x, x) . \tag{11.46}$$

This equation shows that the inner product $(x, \varphi x)$ only depends on the norm of x.

Let us now introduce an orientation in E by a normed determinant-function Δ. Then

$$\begin{aligned}\Delta(x, \varphi x) &= \eta(\alpha\xi - \beta\eta)\,\Delta(e_2, e_1) + \xi(\beta\xi + \alpha\eta)\,\Delta(e_1, e_2)\\ &= \beta\,|x|^2\,\Delta(e_1, e_2) = \varepsilon\beta\,|x|^2\end{aligned} \tag{11.47}$$

where $\varepsilon = \pm 1$ depending on whether the basis e_1, e_2 is positive or negative. Denote by θ the oriented angle between the vectors x and φx. The equations (10.24), (11.46) and (11.47) yield

$$\begin{aligned}\cos\theta &= \alpha\\ \sin\theta &= \varepsilon\,\beta\end{aligned} \tag{11.48}$$

showing that the angle θ does not depend on x. Hence, it makes sense to call θ the *rotation-angle of* φ. If the orientation of E is reversed, the rotation-angle changes the sign.

Now the equations (11.43) and (11.45) can be written in the form

$$\begin{aligned}\varphi e_1 &= e_1\cos\theta + \varepsilon e_2\sin\theta\\ \varphi e_2 &= -\varepsilon e_1\sin\theta + e_2\cos\theta\end{aligned} \tag{11.49}$$

where $\varepsilon = 1$ if the basis (e_1, e_2) is positive, and $\varepsilon = -1$ if the basis (e_1, e_2) is negative. The above equations show that

$$\cos\theta = \frac{1}{2}\operatorname{tr}\varphi . \tag{11.50}$$

11.23. Proper rotations of the 3-space. Consider a proper rotation φ of a 3-dimensional inner product space E. As it has been shown in sec. 11.20, there exists a 1-dimensional subspace E_1 of E whose vectors remain fixed. If φ is different from the identity-map there are no other fix-vectors. In fact, assume that a and b are two linearly independent fix-vectors. Let c $(c \neq 0)$ be a vector which is orthogonal to a and to b. Then $\varphi c = \lambda c$ where $\lambda = \pm 1$. Now the equation $\det \varphi = 1$ implies that $\lambda = +1$ showing that φ is the identity.

In the following discussion it is assumed that $\varphi \neq \iota$. Then the fix-vectors generate a 1-dimensional subspace E_1 called the *fix-axis*.

To determine the axis of a given rotation φ consider the skew mapping

$$\psi = \frac{1}{2}(\varphi - \tilde{\varphi}) \tag{11.51}$$

and introduce an orientation in E. Then ψ can be written in the form

$$\psi x = [u, x] \qquad u \in E. \tag{11.52}$$

The vector u which is uniquely determined by φ is called the *rotation-vector*. The rotation-vector is contained in the fix-axis. In fact, let $a \neq 0$ be a vector of the fix-axis. Then the equations (11.52) and (11.51) yield

$$[u, a] = \psi a = \frac{1}{2}(\varphi a - \tilde{\varphi} a) = \frac{1}{2}(\varphi a - \varphi^{-1} a) = 0 \tag{11.53}$$

showing that u is a multiple of a. Hence (11.52) can be used to find the rotation axis provided that the rotation-vector is different from zero.

This exceptional case occurs if and only if $\varphi = \tilde{\varphi}$ i. e. if and only if $\varphi = \varphi^{-1}$. Then φ has the eigenvalues $1, -1$ and -1. In other words, φ is a reflection at the fix-axis.

11.24. The rotation-angle. Consider the plane F which is orthogonal to E_1. Then φ transforms F into itself and the induced rotation φ_1 is again proper. Denote by θ the rotation-angle of φ_1. Then, in view of (11.50)

$$\cos\theta = \frac{1}{2} \operatorname{tr} \varphi_1 .$$

Observing that

$$\operatorname{tr} \varphi = \operatorname{tr} \varphi_1 + 1$$

we obtain the formula

$$\cos\theta = \frac{1}{2}(\operatorname{tr} \varphi - 1) .$$

To find a formula for $\sin\theta$ consider the orientation of F which is induced by E and the vector u (cf. sec. 5.3) *). This orientation is represented by the normed determinant-function

$$\Delta_1(y, z) = \frac{1}{|u|} \Delta(u, y, z)$$

where Δ is the normed determinant-function representing the orientation

*) It is assumed that $u \neq 0$.

of E. Now the formula (10.24) yields

$$\sin\theta = \Delta_1(y, \varphi y) = \frac{1}{|u|}\Delta(u, y, \varphi y) \tag{11.54}$$

where y is an arbitrary unit vector of F. Now

$$\Delta(u, y, \varphi y) = \det\varphi\,\Delta(\varphi^{-1}u, \varphi^{-1}y, y) = \Delta(u, \varphi^{-1}y, y) = -\Delta(u, y, \varphi^{-1}y)$$

and hence the equation (11.54) can be written as

$$\sin\theta = -\frac{1}{|u|}\Delta(u, y, \varphi^{-1}y). \tag{11.55}$$

By adding the formulas (11.54) and (11.55) we obtain

$$\sin\theta = \frac{1}{2|u|}\Delta(u, y, \varphi y - \varphi^{-1}y) = \frac{1}{|u|}\Delta(u, y, \psi y). \tag{11.56}$$

Inserting the expression (11.52) in (11.56), we thus obtain

$$\sin\theta = \frac{1}{|u|}\Delta(u, y, [u, y]) = \frac{1}{|u|}|[u, y]|^2. \tag{11.57}$$

Since y is a unit-vector orthogonal to u, it follows that

$$|[u, y]| = |u|\,|y| = |u|$$

and hence (11.57) yields the formula

$$\sin\theta = |u|.$$

This equation shows that $\sin\theta$ is positive and hence that $0 < \theta < \pi$ if the above orientation of F is used.

Altogether we thus obtain the following geometric significance of the rotation-vector u:

1. u is contained in the fix-axis.
2. The norm of u is equal to $\sin\theta$.
3. If the orientation induced by u is used in F, then θ is contained in the interval $0 < \theta < \pi$.

Let us now compare the rotations φ and φ^{-1}. φ^{-1} has obviously the same fix-axis as φ. Observing that $\varphi^{-1} = \tilde{\varphi}$ we see that the rotation-vector of φ^{-1} is $-u$. This implies that the inverse rotation induces the inverse orientation in the plane F.

To obtain an explicit expression for u select a positive orthonormal basis e_1, e_2, e_3 in E and let α_ν^μ be the corresponding matrix of φ. Then ψ has the matrix

$$\beta_\nu^\mu = \frac{1}{2}(\alpha_\nu^\mu - \alpha_\mu^\nu)$$

and the components of u are given by

$$u^1 = \frac{1}{2}(\alpha_3^2 - \alpha_2^3) \quad u^2 = \frac{1}{2}(\alpha_1^3 - \alpha_3^1) \quad u^3 = \frac{1}{2}(\alpha_2^1 - \alpha_1^2).$$

It should be observed that the rotation-vector u does not determine the mapping φ completely. In fact, two rotations about the same axis through the angles θ and $\pi - \theta$ have the same rotation-vector. To characterize the mapping φ completely, we need both, the rotation-vector and the cosine of the rotation-angle.

Problems: 1. Prove that any two proper rotations of the plane commute.

2. Given a proper rotation φ of the plane denote by $\theta(\varphi)$ the corresponding rotation-angle. Show that

$$\theta(\varphi_2 \circ \varphi_1) = \theta(\varphi_1) + \theta(\varphi_2) .$$

3. Let φ be an automorphism of a real 2-dimensional linear space E. Prove that an inner product can be introduced in E such that φ becomes a proper rotation if and only if

$$\det \varphi = 1 \quad \text{and} \quad |\operatorname{tr} \varphi| \leqq 2 .$$

4. Consider the set H of all homothetic*) endomorphisms φ of the plane such that $\det \varphi \geqq 0$. Prove:

a) If $\varphi_1 \in H$ and $\varphi_2 \in H$, then $\lambda \varphi_1 + \mu \varphi_2 \in H$.

b) If the multiplication is defined in H in the natural way, the set H becomes a commutative field.

c) Choose a fixed unit-vector e. Then, to every vector $x \in E$ there exists exactly one homothetic mapping φ_x such that $\varphi_x e = x$. Define a multiplication in E by $xy = \varphi_x y$. Prove that E becomes a field under this multiplication and that the mapping $x \to \varphi_x$ defines an isomorphism of E onto H.

d) Prove that E is isomorphic to the field of complex numbers.

5. Given an improper rotation φ of the plane construct an orthonormal basis e_1, e_2 such that $\varphi e_1 = e_1$ and $\varphi e_2 = - e_2$.

6. Show that every skew endomorphism ψ of the plane is homothetic. If $\psi \neq 0$, prove that the angle of the corresponding rotation is equal to $+\frac{\pi}{2}$ if the orientation is defined by the determinant-function

$$\Delta(x, y) = (\psi x, y) \qquad\qquad x, y \in E .$$

7. Find the axis and the angle of the rotation defined by

$$\varphi e_1 = \frac{1}{3}(-e_1 + 2e_2 - 2e_3)$$

$$\varphi e_2 = \frac{1}{3}(2e_1 + 2e_2 + e_3)$$

$$\varphi e_3 = \frac{1}{3}(2e_1 - e_2 - 2e_3)$$

where e_ν $(\nu = 1, 2, 3)$ is a positive orthonormal basis.

*) cf. § 5, prob. 2.

8. If φ is a proper rotation of the 3-space, prove the relation

$$\det(\varphi + \iota) = 4\,(1 + \cos\theta)$$

where θ is the corresponding angle.

9. Consider an orthogonal 3×3-matrix (α_ν^μ) whose determinant is $+1$. Prove the relation

$$\left(\sum_\nu \alpha_\nu^\nu - 1\right)^2 + \sum_{\nu<\mu} (\alpha_\mu^\nu - \alpha_\nu^\mu)^2 = 4\,.$$

10. Let e be a unit-vector of an oriented 3-space and θ $(-\pi < \theta \leqq \pi)$ be a given angle. Denote by F the plane orthogonal to e. Consider the proper rotation φ whose axis is generated by e and whose angle is θ if the orientation induced by e is used in F. Prove the formula

$$\varphi x = x\cos\theta + e\,(e, x)\,(1 - \cos\theta) + [e, x]\sin\theta\,.$$

11. Prove that two proper rotations of the 3-space commute if and only if they have the same axis.

12. Let φ be a proper rotation not having the eigenvalue -1.

a) Prove that $\chi = (\varphi - \iota) \circ (\varphi + \iota)^{-1}$ is a skew mapping.

b) Prove that the mappings χ and $\psi = \frac{1}{2}(\varphi - \tilde{\varphi})$ are connected by the equation

$$\chi = \frac{1}{1 + \cos\theta}\,\psi$$

where θ denotes the rotation-angle of φ.

13. Assume that an improper rotation $\varphi \neq -\iota$ of the Euclidean 3-space is given.

a) Prove that the vectors x for which $\varphi x = -x$, form a 1-dimensional subspace E_1.

b) Prove that a proper rotation φ_1 is induced in the plane F orthogonal to E_1. Defining the rotation-vector u as in sec. 11.23, prove that φ_1 is the identity if and only if $u = 0$.

c) Show that the rotation-angle of φ_1 is given by

$$\cos\theta = \frac{1}{2}(\operatorname{tr}\varphi + 1)$$

and that $0 < \theta < \pi$ if the induced orientation is used in F.

14. Prove that an endomorphism φ of the Euclidean 3-space is homothetic if and only if

$$\varphi[x, y] = [\varphi x, \varphi y]\,.$$

§ 7. Differentiable families of automorphisms

11.25. Differentiation formulas. Let E be an n-dimensional inner product space and $L(E; E)$ the space of all endomorphisms of E. It has been shown in sec. 10.28 that a norm is defined in the space $L(E;E)$ by the equation

$$|\varphi| = \max_{|x|=1} |\varphi x| .$$

A continuous mapping $t \to \varphi(t)$ of a closed interval $t_0 \leqq t \leqq t_1$ into the space $L(E;E)$ will be called a *continuous family of endomorphisms* or a *continuous curve in* $L(E;E)$. A continuous curve $\varphi(t)$ is called *differentiable* if the limit

$$\lim_{\Delta t \to 0} \frac{\varphi(t+\Delta t) - \varphi(t)}{\Delta t} = \dot{\varphi}(t)$$

exists for every t ($t_0 \leqq t \leqq t_1$). The mapping $\dot{\varphi}(t)$ is obviously again linear for every fixed t.

The following formulas are immediate consequences of the above definition:

1. $(\lambda\varphi + \mu\psi)^{\cdot} = \lambda\dot{\varphi} + \mu\dot{\psi}$ (λ, μ constants)
2. $(\psi \circ \varphi)^{\cdot} = \dot{\psi} \circ \varphi + \psi \circ \dot{\varphi}$
3. $\dot{\tilde{\varphi}} = \tilde{\dot{\varphi}}$

4. If $\varphi_\nu(t)$ ($\nu = 1 \dots p$) are p differentiable curves in $L(E;E)$ and Φ is a p-linear function in $L(E;E)$, then

$$\frac{d}{dt} \Phi(\varphi_1(t) \dots \varphi_p(t)) = \sum_{\nu=1}^{p} \Phi(\varphi_1(t) \dots \dot{\varphi}_\nu(t) \dots \varphi_p(t)).$$

A curve $\varphi(t)$ ($t_0 \leqq t \leqq t_1$) is called *continuously differentiable* if the mapping $t \to \dot{\varphi}(t)$ is again continuous. Throughout this paragraph all differentiable curves are assumed to be continuously differentiable.

11.26. Differentiable families of automorphisms. Our first aim is to establish a one-to-one correspondence between all differentiable families of automorphisms on the one hand and all continuous families of endomorphisms on the other hand. Let a differentiable family $\varphi(t)$ ($t_0 \leqq t \leqq t_1$) of automorphisms be given such that $\varphi(t_0) = \iota$. Then a continuous family $\psi(t)$ of endomorphisms is defined by

$$\psi(t) = \dot{\varphi}(t) \circ \varphi(t)^{-1}.$$

Interpreting t as the time we obtain the following physical significance of the mappings $\psi(t)$: Let x be a fixed vector of E and

$$x(t) = \varphi(t)\, x$$

the corresponding orbit. Then the velocity-vector $\dot{x}(t)$ is given by

$$\dot{x}(t) = \dot{\varphi}(t)\, x = \dot{\varphi}(t)\, \varphi(t)^{-1} x(t) = \psi(t)\, x(t) .$$

Hence, the mapping $\psi(t)$ associates with every vector $x(t)$ its velocity at the instant t.

Now it will be shown that, conversely, to every continuous curve $\psi(t)$ in $L(E; E)$ there exists exactly one differentiable family $\varphi(t)$ $(t_0 \leqq t \leqq t_1)$ of automorphisms satisfying the differential equation

$$\dot{\varphi}(t) = \psi(t) \circ \varphi(t) \tag{11.58}$$

and the initial-contition $\varphi(t_0) = \iota$. First of all we notice that the differential equation (11.58) together with the above initial-condition is equivalent to the integral equation

$$\varphi(t) = \iota + \int_{t_0}^{t} \psi(t) \circ \varphi(t)\, dt \quad (t_0 \leqq t \leqq t_1)\,. \tag{11.59}$$

In the next section the solution of the integral equation (11.59) will be constructed by the method of successive approximations.

11.27. The Picard iteration process. Define the curves $\varphi_n(t)$ $(n=0,1\ldots)$ by the equations

$$\left.\begin{aligned} &\varphi_0(t) = \iota \\ &\text{and} \\ &\varphi_{n+1}(t) = \iota + \int_{t_0}^{t} \psi(t) \circ \varphi_n(t)\, dt \quad (n = 0, 1\ldots) \end{aligned}\right\} t_0 \leqq t \leqq t_1\,. \tag{11.60}$$

Introducing the differences

$$\Delta_n(t) = \varphi_n(t) - \varphi_{n-1}(t) \quad (n = 1, 2\ldots) \tag{11.61}$$

we obtain from (11.60) the relations

$$\Delta_n(t) = \int_{t_0}^{t} \psi(t) \circ \Delta_{n-1}(t)\, dt \quad (n = 2, 3, \ldots)\,. \tag{11.62}$$

The equation (11.61) yields for $n = 1$

$$\Delta_1(t) = \varphi_1(t) - \varphi_0(t) = \int_{t_0}^{t} \psi(t)\, dt\,.$$

Define the number M by

$$M = \max_{t_0 \leqq t \leqq t_1} |\psi(t)|\,.$$

Then

$$|\Delta_1(t)| \leqq M\,(t - t_0)\,. \tag{11.63}$$

Employing the equation (11.62) for $n = 2$ we obtain in view of (11.63)

$$|\Delta_2(t)| \leqq M^2 \int_{t_0}^{t} (t - t_0)\, dt \leqq \frac{M^2}{2}(t - t_0)^2$$

and in general

$$|\Delta_n(t)| \leqq \frac{M^n}{n!}(t - t_0)^n \qquad (n = 1, 2 \ldots).$$

Now the relations (11.61) imply that

$$\varphi_{n+p}(t) - \varphi_n(t) = \sum_{\nu = n+1}^{n+p} \Delta_\nu(t),$$

whence

$$\begin{aligned} |\varphi_{n+p}(t) - \varphi_n(t)| \leqq \sum_{\nu=n+1}^{n+p} |\Delta_\nu(t)| &\leqq \sum_{\nu=n+1}^{n+p} \frac{M^\nu}{\nu!}(t - t_0)^\nu \leqq \\ &\leqq \sum_{\nu=n+1}^{n+p} \frac{M^\nu}{\nu!}(t_1 - t_0)^\nu . \end{aligned} \tag{11.64}$$

Let $\varepsilon > 0$ be an arbitrary number. It follows from the convergence of the series $\sum_\nu \frac{M^\nu}{\nu!}(t_1 - t_0)^\nu$ that there exists an integer N such that

$$\sum_{\nu=n+1}^{n+p} \frac{M^\nu}{\nu!}(t_1 - t_0)^\nu < \varepsilon \quad \text{for } n > N \text{ and } p \geqq 1. \tag{11.65}$$

The inequalities (11.64) and (11.65) yield

$$|\varphi_{n+p}(t) - \varphi_n(t)| < \varepsilon \quad \text{for } n > N \text{ and } p \geqq 1 .$$

These relations show that the sequence $\varphi_n(t)$ is uniformly convergent in the interval $t_0 \leqq t \leqq t_1$,

$$\lim_{n \to \infty} \varphi_n(t) = \varphi(t) .$$

In view of the uniform convergence, the equation (11.60) implies that

$$\varphi(t) = \iota + \int_{t_0}^{t} \psi(t) \circ \varphi(t)\, dt \qquad (t_0 \leqq t \leqq t_1) . \tag{11.66}$$

As a uniform limit of continuous curves the curve $\varphi(t)$ is itself continuous. Hence, the right hand-side of (11.66) is differentiable and so $\varphi(t)$ must be differentiable. Differentiating (11.66) we obtain the relation

$$\dot{\varphi}(t) = \psi(t) \circ \varphi(t)$$

showing that $\varphi(t)$ satisfies the differential equation (11.58). The equation $\varphi(t_0) = \iota$ is an immediate consequence of the relations $\varphi_n(t_0) = \iota$ $(n = 0, 1 \ldots)$.

11.28. The determinant of $\varphi(t)$. It remains to be shown that the mappings $\varphi(t)$ are automorphisms. This will be done by proving the formula

$$\det \varphi(t) = e^{\int_{t_0}^{t} tr\, \psi(t)\, dt} . \tag{11.67}$$

Let $\Delta \neq 0$ be a determinant-function in E. Then

$$\Delta(\varphi(t)\, x_1 \ldots \varphi(t)\, x_n) = \det \varphi(t)\, \Delta\,(x_1 \ldots x_n) \quad x_\nu \in E\,.$$

Differentiating this equation and using the differential equation (11.58) we obtain

$$\begin{aligned}&\sum_\nu \Delta\,(\varphi(t)\, x_1 \ldots \psi(t)\; \varphi(t)\, x_\nu \ldots \varphi(t)\, x_n)\\ &= \frac{d}{dt} \det \varphi(t) \cdot \Delta\,(x_1 \ldots x_n)\,.\end{aligned} \tag{11.68}$$

Observing that

$$\begin{aligned}&\sum_\nu \Delta\,(\varphi(t)\, x_1 \ldots \psi(t)\; \varphi(t)\, x_\nu \ldots \varphi(t)\, x_n)\\ &= \operatorname{tr} \psi(t)\, \Delta\,(\varphi(t)\, x_1 \ldots \varphi(t)\, x_n)\\ &= \operatorname{tr} \psi(t) \det \varphi(t)\, \Delta\,(x_1 \ldots x_n)\,.\end{aligned}$$

We obtain from (11.68) the differential equation

$$\frac{d}{dt} \det \varphi(t) = \operatorname{tr} \psi(t) \cdot \det \varphi(t) \tag{11.69}$$

for the function $\det \varphi(t)$. Integrating this differential equation and observing the initial-condition

$$\det \varphi(t_0) = \det \iota = 1$$

we find the formula (11.67).

11.29. Uniqueness of the solution. Assume that $\varphi_1(t)$ and $\varphi_2(t)$ are two solutions of the differential equation (11.58) and the initial condition $\varphi(t_0) = \iota$. Consider the difference

$$\varphi(t) = \varphi_2(t) - \varphi_1(t)\,.$$

The curve $\varphi(t)$ is again a solution of the differential equation (11.58) and it satisfies the initial condition $\varphi(t_0) = 0$. This implies the inequality

$$|\varphi(t)| = \left|\int_{t_0}^{t} \dot\varphi(t)\, dt\right| \leqq \int_{t_0}^{t} |\dot\varphi(t)|\, dt \leqq M \int_{t_0}^{t} |\varphi(t)|\, dt\,. \tag{11.70}$$

Now define the function F by

$$F(t) = \int_{t_0}^{t} |\varphi(t)|\, dt\,. \tag{11.71}$$

Then (11.70) implies the relation

$$\dot F(t) \leqq M\, F(t)\,.$$

Multiplying by e^{-tM} we obtain

$$\dot F(t)\, e^{-tM} - M\, e^{-tM}\, F(t) \leqq 0\,,$$

whence

$$\frac{d}{dt}(F(t)\, e^{-tM}) \leqq 0\,.$$

Integrating this inequality and observing that $F(t_0) = 0$, we obtain

$$F(t)\, e^{-(t-t_0)M} \leqq 0$$

and consequently

$$F(t) \leqq 0 \qquad (t_0 \leqq t \leqq t_1)\,. \tag{11.72}$$

On the other hand it follows from (11.71) that

$$F(t) \geqq 0 \qquad (t_0 \leqq t \leqq t_1)\,. \tag{11.73}$$

The relations (11.72) and (11.73) imply that $F(t) \equiv 0$ whence $\varphi(t) \equiv 0$. Consequently, the two solutions $\varphi_1(t)$ and $\varphi_2(t)$ coincide.

11.30. 1-parameter groups of automorphisms. A differentiable family of automorphisms $\varphi(t)$ $(-\infty < t < \infty)$ is called a *1-parameter group*, if

$$\varphi(t+\tau) = \varphi(t) \circ \varphi(\tau)\,. \tag{11.74}$$

The equation (11.74) implies indeed that the automorphisms $\varphi(t)$ form a group. Inserting $t = 0$ we find $\varphi(0) = \iota$. Now the equation (11.74) yields

$$\varphi(t) \circ \varphi(-t) = \iota$$

showing that with every automorphism $\varphi(t)$ the inverse automorphism $\varphi(t)^{-1}$ is contained in the family $\varphi(t)$ $(-\infty < t < \infty)$. In addition it follows from (11.74) that the group $\varphi(t)$ is commutative.

Differentiation of (11.74) with respect to t yields

$$\dot{\varphi}(t+\tau) = \dot{\varphi}(t) \circ \varphi(\tau)\,.$$

Inserting $t = 0$ we obtain the differential equation

$$\dot{\varphi}(\tau) = \psi \circ \varphi(\tau) \qquad (-\infty < \tau < \infty) \tag{11.75}$$

where $\psi = \dot{\varphi}(0)$. Conversely, consider the differential equation (11.75) where ψ is a given endomorphism of E which does not depend on τ. It will be shown that the solution $\varphi(\tau)$ of this differential equation to the initial condition $\varphi(0) = \iota$ is a 1-parameter group of automorphisms. To prove this, let τ be fixed and consider the curves

$$\varphi_1(t) = \varphi(t+\tau) \tag{11.76}$$

and

$$\varphi_2(t) = \varphi(t) \circ \varphi(\tau)\,. \tag{11.77}$$

Differentiating the equations (11.76) and (11.77) we obtain

$$\dot{\varphi}_1(t) = \dot{\varphi}(t+\tau) = \psi \circ \varphi(t+\tau) = \psi \circ \varphi_1(t) \tag{11.78}$$

and

$$\dot{\varphi}_2(t) = \dot{\varphi}(t) \circ \varphi(\tau) = \psi \circ \varphi(t) \circ \varphi(\tau) = \psi \circ \varphi_2(t)\,. \tag{11.79}$$

The relations (11.78) and (11.79) show that the two curves $\varphi_1(t)$ and $\varphi_2(t)$ satisfy the same differential equation. Moreover,

$$\varphi_1(0) = \varphi_2(0) = \varphi(\tau) .$$

Thus, it follows from the uniqueness-theorem of sec. 11.29 that $\varphi_1(t) \equiv \varphi_2(t)$ whence (11.74).

11.31. Differentiable families of rotations. Let $\varphi(t)$ $(t_0 \leqq t \leqq t_1)$ be a differentiable family of *rotations* such that $\varphi(t_0) = \iota$. Since $\det \varphi(t) = \pm 1$ for every t and $\det \varphi(0) = +1$ it follows from the continuity that $\det \varphi(t) = +1$, i. e. all rotations $\varphi(t)$ are proper.

Now it will be shown that the endomorphisms

$$\psi(t) = \dot{\varphi}(t) \circ \varphi(t)^{-1}$$

are skew. Differentiating the identity

$$\tilde{\varphi}(t) \circ \varphi(t) = \iota$$

we obtain

$$\dot{\tilde{\varphi}}(t) \circ \varphi(t) + \tilde{\varphi}(t) \circ \dot{\varphi}(t) = 0 .$$

Inserting

$$\dot{\varphi}(t) = \psi(t) \circ \varphi(t)$$

and

$$\dot{\tilde{\varphi}}(t) = \tilde{\dot{\varphi}}(t) = \tilde{\varphi}(t) \circ \tilde{\psi}(t)$$

into this equation we find

$$\tilde{\varphi}(t) \circ (\tilde{\psi}(t) + \psi(t)) \circ \varphi(t) = 0 ,$$

whence

$$\tilde{\psi}(t) + \psi(t) = 0 .$$

Conversely, let the family of automorphisms $\varphi(t)$ be defined as the differential equation

$$\dot{\varphi}(t) = \psi(t) \circ \varphi(t) , \quad \varphi(t_0) = \iota$$

where $\psi(t)$ is a continuous family of *skew* mappings. Then every automorphism $\varphi(t)$ is a proper rotation. To prove this, define the family $\chi(t)$ by

$$\chi(t) = \tilde{\varphi}(t) \circ \varphi(t) .$$

Then

$$\begin{aligned} \dot{\chi}(t) &= \dot{\tilde{\varphi}}(t) \circ \varphi(t) + \tilde{\varphi}(t) \circ \dot{\varphi}(t) \\ &= -\tilde{\varphi}(t) \circ \psi(t) \circ \varphi(t) + \tilde{\varphi}(t) \circ \psi(t) \circ \varphi(t) = 0 \end{aligned}$$

and

$$\chi(t_0) = \iota .$$

Now the uniqueness theorem implies that $\chi(t) \equiv \iota$, whence

$$\tilde{\varphi}(t) \circ \varphi(t) = \iota .$$

This equation shows that the mappings $\varphi(t)$ are rotations.

11.32. Angular velocity. As an example, let $\varphi(t)$ be a differentiable family of rotations of the 3-space such that $\varphi(0) = \iota$. If t is interpreted as the time, the family $\varphi(t)$ can be considered as a rigid motion of the space E. Given a vector x, the curve

$$x(t) = \varphi(t)\, x$$

describes its orbit. The corresponding velocity-vector is determined by

$$\dot{x}(t) = \dot{\varphi}(t)\, x = \psi(t)\, \varphi(t)\, x = \psi(t)\, x(t) . \tag{11.80}$$

Now assume that an orientation is defined in E. Then every mapping $\psi(t)$ can be written as

$$\psi(t)\, y = [y, u(t)] . \tag{11.81}$$

The vector $u(t)$ is uniquely determined by $\psi(t)$ and hence by t. The equations (11.80) and (11.81) yield

$$\dot{x}(t) = [x(t), u(t)] . \tag{11.82}$$

The vector $u(t)$ is called the *angular velocity* at the time t. To obtain a physical interpretation of the angular velocity fix a certain instant t and assume that $u(t) \neq 0$. Then the equation (11.82) shows that $\dot{x}(t) = 0$ if and only if $x(t)$ is a multiple of $u(t)$. In other words, the straight line generated by $u(t)$ consists of all vectors having the velocity zero at the instant t. This straight line is called the *instantaneous axis*. The equation (11.82) implies that the velocity-vector $\dot{x}(t)$ is orthogonal to the instantaneous axis.

Passing over to the norm the equation (11.82) we find that

$$|\dot{x}(t)| = |u(t)|\, |h(t)|$$

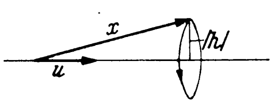


Fig. 1

where $|h(t)|$ is the distance of the vector $x(t)$ from the instantaneous axis (fig. 1). Consequently, the norm of $u(t)$ is equal to the magnitude of the velocity of a vector having the distance 1 from the instantaneous axis.

The uniqueness-theorem in sec. **11.29** implies that the rigid motion $\varphi(t)$ is uniquely determined if the angular velocity is a given function of t.

11.33. The trigonometric functions. In this concluding section we shall apply our general results about families of rotations to the Euclidean plane and show that this leads to the trigonometric function $\cos t$ and $\sin t$. This definition has the advantage that the addition-theorems can be proved in a simple fashion, without making use of the geometric intuition.

Let E be an oriented Euclidean plane and Δ be the normed determinant function representing the given orientation. Consider the skew mapping ψ which is defined by the equation

$$(\psi x, y) = \Delta(x, y) . \tag{11.83}$$

First of all we notice that ψ is a proper rotation. In fact, the identity (10.23) yields

$$(\psi x, y)^2 = \Delta(x, y)^2 = (x, x)(y, y) - (x, y)^2 .$$

Inserting $y = \psi x$ we find

$$(\psi x, \psi x)^2 = (x, x)(\psi x, \psi x) .$$

Now ψ is regular as follows from (11.83). Hence the above equation implies that

$$(\psi x, \psi x) = (x, x) .$$

Replacing x and y by ψx and ψy respectively in (11.83) we obtain the relation

$$\Delta(\psi x, \psi y) = (\psi^2 x, \psi y) = (\psi x, y) = \Delta(x, y)$$

showing that

$$\det \psi = +1 .$$

Let $\varphi(t)$ $(-\infty < t < \infty)$ be the family of rotations defined by the differential equation

$$\dot{\varphi}(t) = \psi \circ \varphi(t) \tag{11.84}$$

and the initial condition

$$\varphi(0) = \iota .$$

Then it follows from the result of sec. 11.30 that

$$\varphi(t + \tau) = \varphi(t) \circ \varphi(\tau) . \tag{11.85}$$

We now define the functions $c(t)$ and $s(t)$ by

$$\left.\begin{aligned} c(t) &= \frac{1}{2}\operatorname{tr}\varphi(t) \\ &\text{and} \\ s(t) &= -\frac{1}{2}\operatorname{tr}(\psi \circ \varphi(t)) \end{aligned}\right\} \quad -\infty < t < \infty . \tag{11.86}$$

These functions are the well-know functions $\cos t$ and $\sin t$. In fact, all the properties of the trigonometric functions can easily be derived from (11.86). Select an arbitrary unit-vector e. Then the vectors e and ψe form an orthonormal basis of E. Consequently,

$$\operatorname{tr}\varphi(t) = (\varphi(t)e, e) + (\varphi(t)\psi e, \psi e) . \tag{11.87}$$

Since ψ is itself a proper rotation, the mappings $\varphi(t)$ and ψ commute.

Hence, the second term in (11.87) can be written as

$$(\varphi(t)\,\psi e, \psi e) = (\psi\,\varphi(t)\,e, \psi e) = (\varphi(t)\,e, e)\,.$$

We thus obtain

$$c(t) = (\varphi(t)\,e, e)\,. \tag{11.88}$$

In the same way it is shown that

$$s(t) = (\varphi(t)\,e, \psi e)\,. \tag{11.89}$$

The equations (11.88) and (11.89) imply that

$$\varphi(t)\,e = c(t)\,e + s(t)\,\psi e\,. \tag{11.90}$$

Replacing t by $t+\tau$ in (11.88) and using the formulas (11.85) and (11.90) we obtain

$$\begin{aligned} c\,(t+\tau) &= (\varphi\,(t+\tau)\,e, e) = (\varphi(t)\,\varphi(\tau)\,e, e) \\ &= c(t)\,(\varphi(\tau)e, e) - s(t)\,(\varphi(\tau)e, \psi e)\,. \end{aligned} \tag{11.91}$$

The equations (11.91), (11.88) and (11.89) yield the addition-theorem of the function $c(t)$:

$$c\,(t+\tau) = c(t)\,c(\tau) - s(t)\,s(\tau)\,.$$

In the same way it is shown that

$$s\,(t+\tau) = s(t)\,c(\tau) + c(t)\,s(\tau)\,.$$

Problems: 1. Let ψ be an endomorphism of the inner product space E. Define the automorphism $\exp\psi$ by

$$\exp\psi = \varphi(1)$$

where $\varphi(t)$ is the family of automorphisms defined by

$$\dot\varphi(t) = \psi\circ\varphi(t)\,,\ \varphi(0) = \iota\,.$$

Prove that

$$\varphi(t) = \exp\,(t\psi) \qquad\qquad (-\infty < t < \infty)\,.$$

2. Show that the mapping $\psi\to\exp\psi$ defined in problem 1 has the following properties:

1. $\exp\,(\psi_1+\psi_2) = \exp\psi_1\circ\exp\psi_2$ if $\psi_2\circ\psi_1 = \psi_1\circ\psi_2$.
2. $\exp\,(-\psi) = (\exp\psi)^{-1}$.
3. $\exp 0 = 1$.
4. $\widetilde{\exp\psi} = \exp\tilde\psi$.
5. $\det\exp\psi = e^{\operatorname{tr}\psi}$.

From these formulas derive that $\exp\psi$ is selfadjoint if ψ is selfadjoint and that $\exp\psi$ is a proper rotation if ψ is skew.

3. Consider the family of rotations $\varphi(t)$ defined by (11.84).

a) Assuming that there is a real number $p \neq 0$ such that $\varphi(p) = \iota$, prove that $\varphi(t+p) = \varphi(t)$ $(-\infty < t < \infty)$.

b) Prove that $\varphi(t_0) = \iota$ if and only if

$$t_0 = 4k \int_0^1 \frac{d\tau}{\sqrt{1-\tau^2}} \qquad (k = 0, \pm 1, \pm 2, \ldots)$$

c) Show that the family $\varphi(t)$ has derivatives of every order and that

$$\varphi^{(\nu+2)}(t) = -\varphi^{(\nu)}(t) \qquad (\nu = 0, 1 \ldots).$$

d) Define the curve $x(t)$ by

$$x(t) = \varphi(t)\, e$$

where e is a fixed unit-vector. Show that

$$\int_0^t |\dot{x}(t)|\, dt = t.$$

4. Derive from the formulas (11.86) that the function $c(t)$ is even and $s(t)$ is odd.

5. Let ψ be the skew mapping defined by (11.83). Prove the *De Moivre's formula*

$$\exp(t\psi) = c(t)\,\iota + s(t)\,\psi.$$

6. Let ψ a skew endomorphism of an n-dimensional inner product space and $\varphi(t)$ the corresponding family of rotations. Consider the normal form (11.35) of the matrix of ψ. Prove that the function $\varphi(t)$ $(-\infty < t < \infty)$ is periodic if and only if all the ratios $\varkappa_\nu : \varkappa_\mu$ are rational.

7. Let E be a linear space and $\Phi(t)$ be a skew-symmetric tensor which is a differentiable function of t.

Prove that

$$\frac{d}{dt}\Phi(t)^k = \Phi(t)^{k-1} \wedge \frac{d\Phi(t)}{dt}$$

(For the definition of $\Phi(t)^k$ cf. sec. 8.13).

Chapter XII

Symmetric bilinear functions

All the properties of an inner product space discussed in Chapter X are based upon the bilinearity, the symmetry and the definiteness of the inner product. The question arises which of these properties do not depend on the definiteness and hence can be carried over to a linear space with an indefinite inner product. Linear spaces of this type will be discussed in § 4. First of all, the general properties of a symmetric bilinear function will be investigated.

§ 1. Bilinear and quadratic functions

12.1. Definition. Let Φ be a symmetric bilinear function in the real linear space E, i. e. a bilinear function such that

$$\Phi(x, y) = \Phi(y, x) .$$

Replacing y by x we obtain a function Ψ of one vector,

$$\Psi(x) = \Phi(x, x) . \tag{12.1}$$

Conversely, the bilinear function Φ can be expressed in terms of the function Ψ. In fact, replacing x by $x + y$ in (12.1) we obtain

$$\Psi(x + y) = \Phi(x + y, x + y) = \Psi(x) + 2\Phi(x, y) + \Psi(y) , \tag{12.2}$$

whence

$$\Phi(x, y) = \frac{1}{2}\{\Psi(x + y) - \Psi(x) - \Psi(y)\} . \tag{12.3}$$

The equation (12.3) shows that different symmetric bilinear functions Φ lead to different functions Ψ.

Replacing y by $-y$ in (12.2) we find

$$\Psi(x - y) = \Psi(x) - 2\Phi(x, y) + \Psi(y) . \tag{12.4}$$

Adding the equations (12.3) and (12.4) we obtain the so-called *parallelogram-identity*

$$\Psi(x + y) + \Psi(x - y) = 2(\Psi(x) + \Psi(y)) . \tag{12.5}$$

12.2. Quadratic functions. A continuous function Ψ of one vector which satisfies the parallelogram-identity will be called a *quadratic function.* Every symmetric bilinear function yields a quadratic function by setting $x = y$. We shall now prove that, conversely, every quadratic function can be obtained in this way.

Substituting $x = y = 0$ in the parallelogram-identity we find that

$$\Psi(0) = 0 . \tag{12.6}$$

Now the same identity yields for $x = 0$

$$\Psi(-y) = \Psi(y)$$

showing that a quadratic function is an even function.

If there exists at all a symmetric bilinear function Φ such that

$$\Phi(x, x) = \Psi(x)$$

this function is given by the equation

$$\Phi(x, y) = \frac{1}{2}\{\Psi(x+y) - \Psi(x) - \Psi(y)\}. \tag{12.7}$$

Therefore it remains to be shown that the function Φ defined by (12.7) is indeed bilinear and symmetric. The symmetry is an immediate consequence of (12.7). Next, we prove the relation

$$\Phi(x_1 + x_2, y) = \Phi(x_1, y) + \Phi(x_2, y). \tag{12.8}$$

The equation (12.7) yields

$$\begin{aligned} 2\Phi(x_1 + x_2, y) &= \Psi(x_1 + x_2 + y) - \Psi(x_1 + x_2) - \Psi(y) \\ 2\Phi(x_1, y) &= \Psi(x_1 + y) - \Psi(x_1) - \Psi(y) \\ 2\Phi(x_2, y) &= \Psi(x_2 + y) - \Psi(x_2) - \Psi(y), \end{aligned}$$

whence

$$\begin{aligned} 2\{\Phi(x_1 + x_2, y) - \Phi(x_1, y) - \Phi(x_2, y)\} &= \{\Psi(x_1 + x_2 + y) + \Psi(y)\} - \\ - \{\Psi(x_1 + y) + \Psi(x_2 + y)\} &- \{\Psi(x_1 + x_2) - \Psi(x_1) - \Psi(x_2)\}. \end{aligned} \tag{12.9}$$

It follows from (12.5) that

$$\Psi(x_1 + x_2 + y) + \Psi(y) = \frac{1}{2}\{\Psi(x_1 + x_2 + 2y) + \Psi(x_1 + x_2)\} \tag{12.10}$$

and

$$\Psi(x_1 + y) + \Psi(x_2 + y) = \frac{1}{2}\{\Psi(x_1 + x_2 + 2y) + \Psi(x_1 - x_2)\}. \tag{12.11}$$

Subtracting (12.11) from (12.10) and using the parallelogram-identity again we find that

$$\begin{aligned} &\{\Psi(x_1 + x_2 + y) + \Psi(y)\} - \{\Psi(x_1 + y) + \Psi(x_2 + y)\} \\ &= \frac{1}{2}\{\Psi(x_1 + x_2) - \Psi(x_1 - x_2)\} = -\Psi(x_1) - \Psi(x_2) + \Psi(x_1 + x_2). \end{aligned} \tag{12.12}$$

Now the equations (12.9) and (12.12) imply (12.8). Inserting $x_1 = x$ and $x_2 = -x$ into (12.8) we obtain

$$\Phi(-x, y) = -\Phi(x, y). \tag{12.13}$$

It remains to be shown that

$$\Phi(\lambda x, y) = \lambda\Phi(x, y) \tag{12.14}$$

for every real number λ. First of all it follows from (12.8) that

$$\Phi(kx, y) = k\Phi(x, y)$$

for a positive integer k. The equation (12.13) shows that (12.14) is also correct for negative integers. Now consider a rational number

$$\lambda = \frac{p}{q} \quad (p, q \text{ integers}).$$

Then

$$q\Phi\left(\frac{p}{q}x, y\right) = \Phi(px, y) = p\Phi(x, y),$$

whence

$$\Phi\left(\frac{p}{q}x, y\right) = \frac{p}{q}\Phi(x, y).$$

To prove (12.13) for an irrational factor λ we note first that Φ is a continuous function of x and y, as follows from the continuity of Ψ. Now select a sequence of rational numbers λ_n such that

$$\lim_{n\to\infty} \lambda_n = \lambda.$$

Then we have for every n the relation

$$\Phi(\lambda_n x, y) = \lambda_n \Phi(x, y). \tag{12.15}$$

For $n \to \infty$ we obtain from (12.15) the relation (12.14).

Our result shows that the relations (12.1) and (12.7) define a one-to-one correspondence between all symmetric bilinear functions and all quadratic functions. If no ambiguity is possible we shall designate a symmetric bilinear function and the corresponding quadratic function by the same symbol, i. e., we shall simply write

$$\Phi(x, x) = \Phi(x).$$

12.3. Bilinear and quadratic forms. Now assume that E has the dimension n and let x_ν $(\nu = 1 \ldots n)$ be a basis of E. Then a symmetric bilinear function Φ can be expressed as a bilinear form

$$\Phi(x, y) = \sum_{\nu, \mu} \alpha_{\nu\mu}\, \xi^\nu \eta^\mu \tag{12.16}$$

where

$$x = \sum_\nu \xi^\nu x_\nu \quad \text{and} \quad y = \sum_\nu \eta^\nu x_\nu$$

and the matrix $\alpha_{\nu\mu}$ is defined by

$$\alpha_{\nu\mu} = \Phi(x_\nu, x_\mu)\,^{*)}.$$

It follows from the symmetry of Φ that the matrix $\alpha_{\nu\mu}$ is symmetric:

$$\alpha_{\nu\mu} = \alpha_{\mu\nu}.$$

Replacing y by x in (12.16) we obtain the corresponding *quadratic form*

$$\Phi(x) = \sum_{\nu, \mu} \alpha_{\nu\mu}\, \xi^\nu \xi^\mu.$$

---

*) The first index counts the row.

Problems: 1. Assume that Ψ is a function in E satisfying the relation

$$\Psi(x + \lambda y) + \lambda \Psi(x - y) = (1 + \lambda)(\Psi(x) + \lambda \Psi(y))$$

for every real number λ. Without using continuity-arguments, show that the function Φ defined by (12.7) is bilinear.

2. Prove that a symmetric bilinear function in E defines a quadratic function in the Cartesian product $E \times E$.

3. Denote by A and by $\bar{A}$ the matrices of the bilinear function Φ with respect to two bases x_ν and $\bar{x}_\nu$ $(\nu = 1 \dots n)$. Show that

$$\bar{A} = T A T^*$$

where T is the matrix of the basis-transformation $x_\nu \to \bar{x}_\nu$.

§ 2. The decomposition of E

12.4. Nullspace. Let Φ be a symmetric bilinear function in the n-dimensional space E. The set of all vectors x_0 with the property that

$$\Phi(x_0, y) = 0 \quad \text{for all vectors} \quad y \in E \tag{12.17}$$

is obviously a subspace of E. This subspace is called the *null-space* of Φ and will be denoted by E_0. The difference of the dimensions of E and E_0 is called the *rank* of the bilinear function Φ.

Introducing a dual space E^* of E we obtain another interpretation of the null-space. Consider the linear mapping $\varphi: E \to E^*$ which is defined by

$$\Phi(x, y) = \langle \varphi x, y \rangle \qquad x, y \in E\,. \tag{12.18}$$

Then the null-space of Φ obviously coincides with the kernel of φ,

$$E_0 = K(\varphi)\,.$$

Consequently, the rank of Φ is equal to the rank of the mapping φ. Let $(\alpha_{\nu\mu})$ be the matrix of Φ relative to a basis x_ν $(\nu = 1 \dots n)$ of E. Then the relation (12.18) yields

$$\langle \varphi x_\nu, x_\mu \rangle = \Phi(x_\nu, x_\mu) = \alpha_{\nu\mu}$$

showing that $\alpha_{\nu\mu}$ is the matrix of the mapping φ. This implies that the rank of the matrix $(\alpha_{\nu\mu})$ is equal to the rank of φ and hence equal to the rank of Φ.

12.5. Non-degenerate bilinear functions. A symmetric bilinear function Φ is called *non-degenerate,* if the null-space reduces to the zero-vector. In other words, the function Φ is non-degenerate if the equation

$$\Phi(x_0, y) = 0$$

for a fixed vector $x_0 \in E$ and all vectors $y \in E$ implies that $x_0 = 0$. It follows from there that a non-degenerate bilinear function defines a scalar-product in the space E (cf. sec. 2.6). With respect to this scalar-product the space E is dual to itself.

The rank of a non-degenerate bilinear function is obviously equal to the dimension of E. It follows from the last remark in sec. 12.4, that a bilinear function is non-degenerate if and only if the determinant of the matrix $(\alpha_{\nu\mu})$ is different from zero.

12.6. Definiteness. A symmetric bilinear function Φ is called *positive definite* if

$$\Phi(x) > 0$$

for all vectors $x \neq 0$. As it has been shown in sec. 10.4, a positive definite bilinear function satisfies the Schwarz-inequality

$$\Phi(x, y)^2 \leqq \Phi(x)\,\Phi(y) \qquad x, y \in E .$$

Equality holds if and only if the vectors x and y are linearly dependent. A positive definite function Φ is non-degenerate. In fact, assume that $\Phi(x_0, y) = 0$ for a fixed vector x_0 and all vectors y. Then $\Phi(x_0) = 0$, whence $x_0 = 0$.

If $\Phi(x) \geqq 0$ for all vectors $x \in E$, but $\Phi(x) = 0$ for some vectors $x \neq 0$, the function Φ is called *positive semidefinite.* The Schwarz-inequality is still valid for a semidefinite function. But now the equality may hold without the vectors x and y being linearly dependent. A semidefinite function is always degenerate. In fact, consider a vector $x_0 \neq 0$ such that $\Phi(x_0) = 0$. Then the Schwarz-inequality implies that

$$\Phi(x_0, y)^2 \leqq \Phi(x_0)\,\Phi(y) = 0$$

whence $\Phi(x_0, y) = 0$ for all vectors y.

In the same way negative definite and negative semidefinite bilinear functions are defined.

The bilinear function Φ is called *indefinite* if the function $\Phi(x)$ assumes positive and negative values. An indefinite function may be degenerate or non-degenerate.

12.7. The decomposition of E. Let a non-degenerate indefinite bilinear function Φ be given in the n-dimensional space E. It will be shown that the space E can be decomposed into two subspaces E^+ and E^- such that Φ is positive definite in E^+ and is negative definite in E^-.

Since Φ is indefinite, there are subspaces of E in which Φ is positive definite. For instance, every vector a for which $\Phi(a) > 0$ generates such a subspace.

Let E^+ be a subspace of maximal dimension such that Φ is positive definite in E^+. Consider the orthogonal complement E^- of E^+ with respect to the scalar-product defined by Φ. Since Φ is positive definite in E^+, the intersection $E^+ \cap E^-$ consists only of the zero-vector. At the same time we have the relation (cf. sec. **2.8**)

$$\dim E^+ + \dim E^- = \dim E .$$

This yields the direct decomposition

$$E = E^+ \oplus E^- .$$

Now it can be shown that Φ is negative definite in E^-. Given a vector $z \neq 0$ of E^-, consider the subspace E_1 generated by E^+ and z. Every vector of this subspace can be written as

$$x = y + \lambda z \qquad y \in E^+.$$

This implies that

$$\Phi(x) = \Phi(y) + \lambda^2 \Phi(z) . \tag{12.19}$$

Now assume that $\Phi(z) > 0$. Then the equation (12.19) shows that Φ is positive definite in the subspace E_1 which is a contradiction to the maximum-property of E^+. Consequently,

$$\Phi(z) \leqq 0 \quad \text{for all vectors} \quad z \in E^-$$

i. e., Φ is negative semidefinite in E^-. Using the Schwarz-inequality

$$\Phi(z_1, z)^2 \leqq \Phi(z_1)\,\Phi(z) \qquad z_1 \in E^-, z \in E^- \tag{12.20}$$

we can prove that Φ is even negative definite in E^-. Assume that $\Phi(z_1) = 0$ for a vector $z_1 \in E^-$. Then the Schwarz-inequality (12.20) yields

$$\Phi(z_1, z) = 0$$

for all vectors $z \in E^-$. At the same time we know that

$$\Phi(z_1, y) = 0$$

for all vectors $y \in E^+$. These two equations imply that

$$\Phi(z_1, x) = 0$$

for all vectors $x \in E$, whence $z_1 = 0$.

12.8. The decomposition in the degenerate case. If the bilinear function Φ is degenerate, select a subspace E_1 complementary to the null-space E_0,

$$E = E_0 \oplus E_1 .$$

Then Φ is non-degenerate in E_1. In fact, assume that

$$\Phi(x_1, y_1) = 0$$

for a fixed vector $x_1 \in E_1$ and all vectors $y_1 \in E_1$. Consider an arbitrary vector $y \in E$. This vector can be written as

$$y = y_0 + y_1 \qquad y_0 \in E_0, y_1 \in E_1$$

whence

$$\Phi(x_1, y) = \Phi(x_1, y_0) + \Phi(x_1, y_1) = 0 . \tag{12.21}$$

This equation shows that x_1 is contained in E_1 and hence it is contained in the intersection $E_0 \cap E_1$. This implies that $x_1 = 0$.

Now the construction of sec. 12.7 can be applied to the subspace E_1. We thus obtain altogether a direct decomposition

$$E = E^+ \oplus E^- \oplus E_0 \tag{12.22}$$

of E such that Φ is positive definite in E^+ and negative definite in E^-.

12.9. Diagonalization of the matrix. Let $(x_1 \dots x_s)$ be a basis of E^+, which is orthonormal with respect to Φ, $(x_{s+1} \dots x_r)$ be a basis of E^- which is orthonormal with respect to $-\Phi$, and $(x_{r+1} \dots x_n)$ be an arbitrary basis of E_0. Then

$$\Phi(x_\nu, x_\mu) = \varepsilon_\nu \, \delta_{\nu\mu} \qquad \text{where } \varepsilon_\nu = \begin{cases} +1 \; (\nu = 1 \dots s) \\ -1 \; (\nu = s+1 \dots r) \\ 0 \; (\nu = r+1 \dots n) \end{cases}$$

The vectors $(x_1 \dots x_n)$ then form a basis of E in which the matrix of Φ has the following diagonal-form:

$$\begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{s} & & \\ & \underbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}_{r-s} & \\ & & \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{n-r} \end{pmatrix}.$$

12.10. The index. It is clear from the above construction that there are infinitely many different decompositions of the form (12.22). However, the dimensions of E^+ and E^- are uniquely determined by the bilinear function Φ. To prove this, consider two decompositions

$$E = E_1^+ \oplus E_1^- \oplus E_0 \tag{12.23}$$

and

$$E = E_2^+ \oplus E_2^- \oplus E_0 \tag{12.24}$$

such that Φ is positive definite in E_1^+ and E_2^+ and negative definite in E_1^- and E_2^-. This implies that

$$E_2^+ \cap (E_0 \oplus E_1^-) = 0$$

whence

$$\dim E_2^+ + \dim E_1^- + \dim E_0 \leqq n \,. \tag{12.25}$$

Comparing the dimensions in (12.23) we find

$$\dim E_1^+ + \dim E_1^- + \dim E_0 = n\,. \tag{12.26}$$

The equations (12.25) and (12.26) yield

$$\dim E_2^+ \leqq \dim E_1^+\,.$$

Interchanging E_1^+ and E_2^+ we obtain

$$\dim E_1^+ \leqq \dim E_2^+\,,$$

whence

$$\dim E_1^+ = \dim E_2^+\,.$$

Consequently, the dimension of E^+ is uniquely determined by Φ. This number is called the *index* of the bilinear function Φ and the number $\dim E^+ - \dim E^- = 2s - r$ is called the *signature* of Φ.

Now suppose that x_ν $(\nu = 1 \dots n)$ is a basis of E in which the quadratic function Φ has diagonal form

$$\Phi(x) = \sum_\nu \lambda_\nu \xi^\nu \xi^\nu$$

and assume that

$$\lambda_\nu > 0 \; (\nu = 1 \dots p) \quad \text{and} \quad \lambda_\nu \leqq 0 \; (\nu = p + 1 \dots n)\,.$$

Then p is the index of Φ. In fact, the vectors $x_\nu (\nu = 1 \dots p)$ generate a subspace of maximal dimension in which Φ is positive definite.

From the above result we obtain *Sylvester's law of inertia* which asserts that the number of positive coefficients is the same for every diagonal-form.

12.11. The rank and the index of a symmetric bilinear function can be determined explicitly from the corresponding quadratic form

$$\Phi(x) = \sum_{\nu,\mu} \alpha_{\nu\mu} \xi^\nu \xi^\mu\,.$$

We can exclude the case $\Phi = 0$. Then at least one coefficient α_{ij} is different from zero. If $i \neq j$, apply the substitution

$$\xi^i = \bar{\xi}^i + \bar{\xi}^j\,,\; \xi^j = \bar{\xi}^i - \bar{\xi}^j\,.$$

Then

$$\Phi(x) = \sum_{\nu,\mu} \bar{\alpha}_{\nu\mu} \bar{\xi}^\nu \bar{\xi}^\mu$$

where $\bar{\alpha}_{ii} \neq 0$ and $\bar{\alpha}_{jj} \neq 0$. Thus, we may assume that at least one coefficient α_{ii}, say α_{11}, is different from zero. Then $\Phi(x)$ can be written as

$$\Phi(x) = \alpha_{11} \left\{ (\xi^1)^2 + \frac{1}{\alpha_{11}} \sum_{\mu=2}^n \alpha_{1\mu} \xi^1 \xi^\mu \right\} + \sum_{\nu,\mu=2}^n \alpha_{\nu\mu} \xi^\nu \xi^\mu\,.$$

The substitution

$$\eta^1 = \xi^1 + \frac{1}{\alpha_{11}} \sum_{\mu=2}^{n} \alpha_{1\mu} \xi^\mu$$

$$\eta^\nu = \xi^\nu \qquad (\nu = 2 \ldots n)$$

yields

$$\Phi(x) = \alpha_{11} (\eta^1)^2 + \sum_{\nu, \mu=2}^{n} \beta_{\nu\mu} \eta^\nu \eta^\mu . \tag{12.27}$$

The sum in (12.27) is a symmetric bilinear form in $(n-1)$ variables and hence the same reduction can be applied to this sum. Continuing this way we finally obtain an expression of the form

$$\Phi(x) = \sum_n \lambda^\nu \zeta^\nu \zeta^\nu .$$

Rearranging the variables we can achieve that

$$\begin{aligned} \lambda^\nu &> 0 \quad (\nu = 1 \ldots s) \\ \lambda^\nu &< 0 \quad (\nu = s+1 \ldots r) \\ \lambda^\nu &= 0 \quad (\nu = r+1 \ldots n) . \end{aligned}$$

Then r is the rank and s is the index of Φ.

Problems: 1. Let $\Phi \neq 0$ be a given quadratic form. Prove that Φ can be written in the form

$$\Phi(x) = \varepsilon f(x)^2 , \varepsilon = \pm 1$$

where f is a linear function, if and only if the corresponding bilinear function has the rank 1.

2. Consider two linear spaces E and F and a bilinear function Φ in E and F. Define the null-spaces $E_0 \subset E$ and $F_0 \subset F$ as follows: A vector $x_0 \in E$ is contained in E_0 if $\Phi(x_0, y) = 0$ for all vectors $y \in F$. A vector $y_0 \in F$ is contained in F_0 if $\Phi(x, y_0) = 0$ for all vectors $x \in E$. Prove the relation

$$\dim E - \dim E_0 = \dim F - \dim F_0 .$$

3. Given a non degenerate symmetric bilinear form Φ in E, let J be a subspace of maximal dimension such that $\Phi(x, x) = 0$ for every $x \in J$.

Prove that

$$\dim J = \min(s, n-s) .$$

Hint: Introduce two dual spaces E^* and F^* and the linear mappings

$$\varphi_1 : E \to E^* \quad \text{and} \quad \varphi_2 : F \to F^*$$

defined by

$$\Phi(x, y) = \langle \varphi_1 x, y \rangle \quad \text{and} \quad \Phi(x, y) = \langle x, \varphi_2 y \rangle .$$

4. Define the bilinear function Φ in the space $L(E; E)$ of all endomorphisms by

$$\Phi(\varphi, \psi) = \operatorname{tr}(\psi \circ \varphi).$$

Let $S(E; E)$ be the space of all selfadjoint mappings and $A(E; E)$ be the space of all skew mappings with respect to a positive definite inner product. Prove:

a) $\Phi(\varphi, \varphi) > 0$ for every $\varphi \neq 0$ in $S(E; E)$,

b) $\Phi(\varphi, \varphi) < 0$ for every $\varphi \neq 0$ in $A(E; E)$,

c) $\Phi(\varphi, \psi) = 0$ if $\varphi \in S(E; E)$ and $\psi \in A(E; E)$,

d) The index of Φ is $\frac{n(n+1)}{2}$, where $n = \dim E$.

5. Find the index of the quadratic form

$$\Phi(x) = \sum_{i<j} \xi^i \xi^j.$$

6. Let Φ be a bilinear function in E. Assume that E_1 is a subspace of E such that Φ is non-degenerate in E_1. Define the subspace E_2 as follows: A vector $x_2 \in E$ is contained in E_2 if

$$\Phi(x_1, x_2) = 0 \quad \text{for all vectors } x_1 \in E_1.$$

Prove that

$$E = E_1 \oplus E_2.$$

7. Consider a bilinear function Φ such that $\Phi(x,x) > 0$ for all vectors $x \neq 0$. Construct a basis of E in which the matrix of Φ has the form

$$\begin{pmatrix} 1 & \varkappa_1 & & & & & & \\ -\varkappa_1 & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & \varkappa_p & & & \\ & & & -\varkappa_p & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

Hint: Decompose Φ in the form

$$\Phi = \Phi_1 + \Phi_2,$$

where

$$\Phi_1(x, y) = \frac{1}{2}(\Phi(x, y) + \Phi(y, x))$$

and

$$\Phi_2(x, y) = \frac{1}{2}(\Phi(x, y) - \Phi(y, x)).$$

§ 3. Simultaneous diagonalization of two bilinear forms

12.12. It has been shown in sec. 11.10 that a pair of symmetric bilinear forms Φ and Ψ can be simultaneously reduced to diagonal-form provided that one of the two functions is positive definite. In the present paragraph it will be shown that in a space of dimension $n \geqq 3$ the simultaneous reduction of Φ and Ψ to diagonal form is still possible under the following less restrictive hypothesis: The quadratic functions $\Phi(x)$ and $\Psi(x)$ do not both assume the value zero for the same vector $x \neq 0$. More precisely:

Let E be a linear space of dimension $n \geqq 3$ and let Φ and Ψ be two symmetric bilinear functions in E such that

$$\Phi(x)^2 + \Psi(x)^2 > 0 \quad \text{for all vectors} \quad x \neq 0 .$$

Then there exists a basis e_ν $(\nu = 1 \ldots n)$ of E in which Φ and Ψ have diagonal form.

The above theorem is not correct for $n = 2$, as it can be seen from the following example:

$$\Phi(x) = (\xi^1)^2 - (\xi^2)^2, \quad \Psi(x) = 2\xi^1 \xi^2 .$$

These functions satisfy the above hypothesis. On the other hand, it is impossible to find a basis in which Φ and Ψ have diagonal form. To prove this observe that the hyperbolas

$$\Phi(x) = 1 \quad \text{and} \quad \Psi(x) = 1$$

have exactly two points in common. Now assume that there exists a basis such that

$$\Phi(x) = \alpha_1(\eta^1)^2 + \alpha_2(\eta^2)^2 \quad \text{and} \quad \Psi(x) = \beta_1(\eta^1)^2 + \beta_2(\eta^2)^2 .$$

Then the intersection of the above hyperbolas would consist either of four points or no point. This gives a contradiction.

12.13. To prove the above theorem we employ a similar method as in sec. 11.6. If one of the functions Φ and Ψ, say Ψ, is positive definite the desired basis-vectors are those for which the function

$$F(x) = \frac{\Phi(x)}{\Psi(x)} \qquad x \neq 0 . \tag{12.28}$$

assumes a relative minimum. However, if Ψ is indefinite, the denominator in (12.28) assumes the value zero for certain vectors $x \neq 0$ and hence the function F is no longer defined in the entire space $x \neq 0$. The method of sec. 11.6 can still be carried over to the present case if the function F is replaced by the function

$$\operatorname{arc\,tan} F(x) . \tag{12.29}$$

To avoid difficulties arising from the fact that the function arc tan is not single-valued, we shall write the function as a line-integral. At this point the hypothesis $n \geqq 3$ will be essential*).

Let $\dot{E}$ be the deleted space $x \neq 0$ and $x = x(t)$ $(0 \leqq t \leqq 1)$ be a differentiable curve in $\dot{E}$. Consider the line-integral

$$J = \int_0^1 \frac{\Phi(x)\,\Psi(x,\dot{x}) - \Phi(x,\dot{x})\,\Psi(x)}{\Phi(x)^2 + \Psi(x)^2}\,dt \tag{12.30}$$

taken along the curve $x(t)$. First of all it will be shown that the integral J depends only on the initial point $x_0 = x(0)$ and the endpoint $x = x(1)$ of the curve $x(t)$. For this purpose define the following mapping of E into the complex w-plane

$$\omega(x) = \Phi(x) + i\,\Psi(x)\,.$$

The image of the curve $x(t)$ under this mapping is the curve

$$\omega(t) = \Phi(x(t)) + i\,\Psi(x(t)) \qquad (0 \leqq t \leqq 1) \tag{12.31}$$

in the w-plane. The hypothesis $\Phi(x)^2 + \Psi(x)^2 \neq 0$ implies that the curve $\omega(t)$ $(0 \leqq t \leqq 1)$ does not go through the point $\omega = 0$. The integral (12.30) can now be written as

$$J = \int_0^1 \frac{u\dot{v} - \dot{u}v}{u^2 + v^2}\,dt$$

where the integration is taken along the curve (12.31).

Now let $\theta(t)$ be an angle-function for the curve $\omega(t)$ i.e. a continuous function of t such that

$$\cos\theta(t) = \frac{u(t)}{|\omega(t)|} \quad \text{and} \quad \sin\theta(t) = \frac{v(t)}{|\omega(t)|} \tag{12.32}$$

(cf. fig. 2)**). It follows from the differentiability of the curve $\omega(t)$ that the angle-function θ is also differentiable and we thus obtain from (12.32)

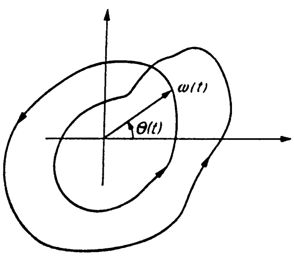


Fig. 2

$$-\sin\theta\cdot\dot{\theta} = \frac{-\dot{u}\,|\omega| + u\,\dfrac{d|\omega|}{dt}}{|\omega|^2} \tag{12.33}$$

and

$$\cos\theta\cdot\dot{\theta} = \frac{\dot{v}\,|\omega| + v\,\dfrac{d|\omega|}{dt}}{|\omega|^2}\,. \tag{12.34}$$

Multiplying (12.33) by $-\sin\theta$ and (11.34) by $\cos\theta$ and adding these equations we find that

$$\dot{\theta} = \frac{u\dot{v} - \dot{u}v}{u^2 + v^2}\,.$$

*) The given proof is due to JOHN MILNOR.

**) For more details cf. ALEXANDROV. Combinatorial Topology, Vol. I, chapter II, § 2.

Integration from $t = 0$ to $t = 1$ gives

$$\int_0^1 \frac{u\dot{v} - \dot{u}v}{u^2 + v^2}\, dt = \theta(1) - \theta(0)$$

showing that the integral J is equal to the change of the angle-function θ along the curve $\omega(t)$,

$$J = \theta(1) - \theta(0)\,. \tag{12.35}$$

Now consider another differentiable curve $x = \bar{x}(t)$ $(0 \leqq t \leqq 1)$ in $\dot{E}$ with the initial point x_0 and the endpoint x and denote by $\bar{J}$ the integral (12.30) taken along the curve $\bar{x}(t)$. Then the formula (12.35) yields

$$\bar{J} = \bar{\theta}(1) - \bar{\theta}(0) \tag{12.36}$$

where $\bar{\theta}$ is an angle-function for the curve

$$\bar{\omega}(t) = \Phi(\bar{x}(t)) + i\Psi(\bar{x}(t)) \qquad (0 \leqq t \leqq 1)\,.$$

Since the curves $\omega(t)$ and $\bar{\omega}(t)$ $(0 \leqq t \leqq 1)$ have the same initial point and the same endpoint it follows that

$$\bar{\omega}(0) - \omega(0) = 2k_0\pi \quad \text{and} \quad \bar{\omega}(1) - \omega(1) = 2\,k_1\pi \tag{12.37}$$

where k_0 and k_1 are integers. The equations (12.35), (12.36) and (12.37) show that the difference $\bar{J} - J$ is a multiple of $2\,\pi$,

$$\bar{J} - J = 2\,k\pi\,.$$

It remains to be shown that $k = 0$. The hypothesis $n \geqq 3$ implies that the deleted space $\dot{E}$ is simply connected. In other words, there exists a continuous mapping $x = x(t, \tau)$ of the square $0 \leqq t \leqq 1$, $0 \leqq \tau \leqq 1$ into $\dot{E}$ such that

$$x(t, 0) = x(t), \quad x(t, 1) = \bar{x}(t) \quad 0 \leqq t \leqq 1$$

and

$$x(0, \tau) = x_0, \quad x(1, \tau) = x \quad 0 \leqq \tau \leqq 1\,.$$

The mapping $x(t, \tau)$ can even be assumed to be differentiable. Then, for every fixed τ, we can form the integral (12.30) along the curve

$$x(t, \tau) \qquad (0 \leqq t \leqq 1)\,.$$

This integral is a continuous function $J(\tau)$ of τ. At the same time we know that the difference $J(\tau) - J$ is a multiple of $2\,\pi$,

$$J(\tau) - J = 2\,\pi\,k(\tau)\,. \tag{12.38}$$

Hence, $k(\tau)$ is a continuous integer-valued function in the interval $0 \leqq \tau \leqq 1$ and thus $k(\tau)$ must be a constant. Since $k(0) = 0$ it follows that $k(\tau) = 0$ $(0 \leqq \tau \leqq 1)$. Now the equation (12.38) yields

$$J(\tau) = J \qquad (0 \leqq \tau \leqq 1).$$

Inserting $\tau = 1$ we obtain the relation

$$\bar{J} = J$$

showing that the integral (12.30) is indeed independent of the curve $x(t)$.

12.14. The function *F*. We now can define a single-valued function F in the deleted space $\dot{E}$ by

$$F(x) = \int_{x_0}^{x} \frac{\Phi(x)\,\Psi(x,\dot{x}) - \Phi(x,\dot{x})\,\Psi(x)}{\Phi(x)^2 + \Psi(x)^2}\,dt \tag{12.39}$$

where the integration is taken along an arbitrary differentiable curve $x(t)$ leading from x_0 to x. The function F is homogeneous of degree zero,

$$F(\lambda x) = F(x), \qquad \lambda > 0. \tag{12.40}$$

To prove this, observe that

$$F(\lambda x) - F(x) = \int_{x}^{\lambda x} \frac{\Phi(x)\,\Psi(x,\dot{x}) - \Phi(x,\dot{x})\Psi(x)}{\Phi(x)^2 + \Psi(x)^2}\,dt.$$

Choosing the straight segment

$$x(t) = (1-t)\,\lambda x + t x \quad (0 \leqq t \leqq 1)$$

as path of integration we find that

$$\dot{x} = (1-\lambda)\,x$$

whence

$$\Phi(x)\,\Psi(x,\dot{x}) - \Phi(x,\dot{x})\,\Psi(x) = 0.$$

This implies the equation (12.40).

12.15. The construction of eigenvectors. From now on our proof will follow the same line as in sec. 11.6. We consider first the case that Ψ is non-degenerate. Introduce a positive definite inner product in E. Then the continuous function F assumes a minimum on the sphere $|x| = 1$. Let e_1 be a vector on $|x| = 1$ such that

$$F(e_1) \leqq F(x)$$

for all vectors $|x| = 1$. Then the homogenuity of F implies that

$$F(e_1) \leqq F(x)$$

for all vectors $x \neq 0$.

Consequently, the function

$$f(t) = F(e_1 + t y),$$

where y is an arbitrary vector, assumes a minimum at $t = 0$, whence

$$\dot{f}(0) = 0\,. \tag{12.41}$$

Carrying out the differentiation we find that

$$\dot{f}(0) = \frac{\Phi(e_1, y)\,\Psi(e_1) - \Phi(e_1)\,\Psi(e_1, y)}{\Phi(e_1)^2 + \Psi(e_1)^2}\,, \tag{12.42}$$

The equations (12.41) and (12.42) imply that

$$\Phi(e_1, y)\,\Psi(e_1) - \Phi(e_1)\,\Psi(e_1, y) = 0 \tag{12.43}$$

for all vectors $y \in E$. In this equation $\Psi(e_1) \neq 0$. In fact, assume that $\Psi(e_1) = 0$. Then $\Phi(e_1) \neq 0$ and hence the equation (12.43) yields $\Psi(e_1, y) = 0$ for all vectors $y \in E$. This is a contradiction to our assumption that Ψ is non-degenerate.

Define the number λ_1 by

$$\lambda_1 = \frac{\Phi(e_1)}{\Psi(e_1)}\,;$$

then the equation (12.43) can be written as

$$\Phi(e_1, y) = \lambda_1 \Psi(e_1, y) \qquad\qquad y \in E.$$

12.16. Now consider the subspace E_1 defined by the equation

$$\Psi(e_1, z) = 0\,.$$

Since Ψ is non-degenerate, E_1 has the dimension $n - 1$. Moreover, the restriction of Ψ to E_1 is again non-degenerate: Assume that z_1 is a vector of E_1 such that

$$\Psi(z_1, z) = 0 \tag{12.44}$$

for all vectors $z \in E_1$. The equation (12.44) implies that

$$\Psi(z_1, x) = 0 \tag{12.45}$$

for every vector $x \in E$ because x can be decomposed in the form

$$x = \xi\, e_1 + z \qquad\qquad z \in E_1\,.$$

Now it follows from (12.45) that $z_1 = 0$, showing that Ψ is non-degenerate in E_1. Therefore, the construction of sec. (12.15) can be applied to E_1. We thus obtain a vector $e_2 \in \mathrm{E}_1$ with the property that

$$\Phi(e_2, z) = \lambda_2 \Psi(e_2, z) \quad \text{for every vector } z \in E_1 \tag{12.46}$$

where

$$\lambda_2 = \frac{\Phi(e_2)}{\Psi(e_2)}\,.$$

The equation (12.46) implies that

$$\Phi(e_2, y) = \lambda_2 \Psi(e_2\, y) \tag{12.47}$$

for every vector $y \in E$;

In fact, y can be decomposed in the form

$$y = \xi e_1 + z \qquad z \in E_1$$

and we thus obtain

$$\begin{aligned}\Phi(e_2, y) &= \xi\Phi(e_2, e_1) + \Phi(e_2, z) = \xi\Phi(e_1, e_2) + \Phi(e_2, z) \\ &= \xi\lambda_1\Psi(e_1, e_2) + \Phi(e_2, z) = \Phi(e_2, z)\end{aligned} \tag{12.48}$$

and

$$\Psi(e_2, y) = \xi\Psi(e_2, e_1) + \Psi(e_2, z) = \Psi(e_2, z)\,. \tag{12.49}$$

The equations (12.46), (12.48) and (12.49) yield (12.47).

Continuing this construction we obtain after n steps a system of n vectors e_ν subject to the following conditions:

$$\begin{aligned}&\Phi(e_\nu, y) = \lambda_\nu\Psi(e_\nu, y) \qquad y \in E \\ &\Psi(e_\nu, e_\nu) \neq 0 \\ &\Psi(e_\nu, e_\mu) = 0 \quad (\nu \neq \mu)\,.\end{aligned} \tag{12.50}$$

Rearranging the vectors e_ν and multiplying them with appropriate scalars we can achieve that

$$\Psi(e_\nu, e_\mu) = \varepsilon_\nu\,\delta_{\nu\mu} \quad \varepsilon_\nu = \begin{cases} +1 & (\nu = 1 \ldots s) \\ -1 & (\nu = s+1 \ldots n) \end{cases} \tag{12.51}$$

where s denotes the signature of Ψ. It follows from the above relations that the vectors e_ν form a basis of E.

Inserting $y = e_\mu$ in the first equation (12.50) we find

$$\Phi(e_\nu, e_\mu) = \lambda_\nu\,\varepsilon_\nu\,\delta_{\nu\mu}\,. \tag{12.52}$$

The equations (12.51) and (12.52) show that the bilinear functions Φ and Ψ have diagonal form in the basis $e_\nu (\nu = 1 \ldots n)$.

12.17. The degenerate case. The degenerate case now remains to be considered. We may assume that $\Psi \neq 0$. Then it will be shown that there exists a scalar λ_0 such that the bilinear function $\Phi + \lambda_0\Psi$ is non-degenerate.

Let E^* be a dual space of E and consider the linear mappings

$$\varphi: E \to E^* \quad \text{and} \quad \psi: E \to E^*$$

defined by the equations

$$\Phi(x, y) = \langle\varphi x, y\rangle \quad \text{and} \quad \Psi(x, y) = \langle\psi x, y\rangle\,.$$

Then

$$\psi(E) \cap \varphi(K(\psi)) = 0\,. \tag{12.53}$$

To prove the relation (12.53) let y be a vector of $\psi(E) \cap \varphi(K(\psi))$. Then $y = \varphi x$, where $x \in K(\psi)$. Hence

$$\Psi(x) = \langle x, \psi x\rangle = 0 \tag{12.54}$$

and

$$\Phi(x) = \langle\varphi x, x\rangle = \langle y, x\rangle = 0\,, \tag{12.55}$$

because $y \in \psi(E)$ and $x \in K(\psi)$.

The equations (12.54) and (12.55) imply that $x = 0$ and hence that $y = \varphi x = 0$.

Now let x_ν $(\nu = 1 \ldots n)$ be a basis of E such that the vectors $(x_{r+1} \ldots x_n)$ form a basis of $K(\psi)$. Employing a determinant-function $\Delta \neq 0$ in E we obtain

$$\Delta(\varphi x_1 + \lambda \psi x_1 \ldots \varphi x_n + \lambda \psi x_n)$$
$$= \Delta(\varphi x_1 + \lambda \psi x_1 \ldots \varphi x_r + \lambda \psi x_r, \varphi x_{r+1} \ldots \varphi x_n) .$$

The expansion of this expression yields a polynomial $f(\lambda)$ starting with the term

$$\lambda^r \Delta(\psi x_1 \ldots \psi x_r, \varphi x_{r+1} \ldots \varphi x_n) .$$

The coefficient of λ^r is not identically zero. This follows from the relation (12.53) and the fact that the r vectors $\psi x_\varrho \in \psi(E)$ $(\varrho = 1 \ldots r)$ and the $(n - r)$ vectors $\varphi x_\sigma \in \varphi(K(\psi))$ $(\sigma = r + 1 \ldots n)$ are linearly independent.

Hence, f is a polynomial of degree r. Our assumption $\Psi \neq 0$ implies that $r \geqq 1$. Consequently, a number λ_0 can be chosen such that $f(\lambda_0) \neq 0$. Then $\Phi + \lambda_0 \Psi$ is non-degenerate.

By the previous theorem there exists a basis e_ν $(\nu = 1 \ldots n)$ of E in which the bilinear functions Φ and $\Phi + \lambda \Psi$ both have diagonal form. Then the functions Φ and Ψ have also diagonal form in this basis.

Problem: 1. Let $A = (\alpha_{\nu\mu})$ and $B = (\beta_{\nu\mu})$ be two symmetric $n \times n$-matrices and assume that the equations

$$\sum_{\nu,\mu} \alpha_{\nu\mu} \xi^\nu \xi^\mu = 0 \quad \text{and} \quad \sum_{\nu,\mu} \beta_{\nu\mu} \xi^\nu \xi^\mu = 0$$

together imply that $\xi^\nu = 0$ $(\nu = 1 \ldots n)$. Prove that the polynomial

$$f(\lambda) = \det(A + \lambda B)$$

has r real roots where r is the rank of B.

§ 4. Pseudo-Euclidean spaces

12.18. Definition. A *pseudo-Euclidean* space is a real linear space in which a non-degenerate indefinite bilinear function is given. As in the positive definite case, this bilinear function is called the *inner product* and is denoted by (x, y). The index of the inner product is called the *index of the pseudo-Euclidean space.*

Since the inner product is indefinite, the number (x, x) can be positive, negative or zero, depending on the vector x. A vector $x \neq 0$ is called

space-like, if $(x, x) > 0$

time-like, if $(x, x) < 0$

a *light-vector,* if $(x, x) = 0$

The *light-cone* is the set of all light-vectors.

As in the definite case two vectors x and y are called *orthogonal* if $(x, y) = 0$. The light-cone consists of all vectors which are orthogonal to themselves.

A basis $e_\nu (\nu = 1 \ldots n)$ is called *orthonormal* if

$$(e_\nu, e_\mu) = \varepsilon_\nu \, \delta_{\nu\mu}$$

where

$$\varepsilon_\nu = \begin{cases} +1 & (\nu = 1 \ldots s) \\ -1 & (\nu = s+1 \ldots n) \end{cases}.$$

In sec. 12.9 we have shown that an orthonormal basis can always be constructed.

If an orthonormal basis $e_\nu (\nu = 1 \ldots n)$ is chosen, the inner product of two vectors

$$x = \sum_\nu \xi^\nu e_\nu \quad \text{and} \quad y = \sum_\nu \eta^\nu e_\nu$$

is given by the bilinear form

$$(x, y) = \sum_{\nu=1}^{n} \varepsilon_\nu \xi^\nu \eta^\nu = \sum_{\nu=1}^{s} \xi^\nu \eta^\nu - \sum_{\nu=s+1}^{n} \xi^\nu \eta^\nu \tag{12.56}$$

and the equation of the light-cone reads

$$\sum_{\nu=1}^{s} \xi^\nu \xi^\nu - \sum_{\nu=s+1}^{n} \xi^\nu \xi^\nu = 0.$$

12.19. The duality between E and itself. In sec. 12.5 we mentioned that every non-degenerate bilinear function in an n-dimensional linear space E defines a scalar-product. Hence, a pseudo-Euclidean space can be regarded as dual to itself relative to the indefinite inner product.

In particular, every subspace E_1 of E determines an orthogonal complement $E_1^\perp$ which is again a subspace of E. The subspaces E_1 and $E_1^\perp$ have complementary dimensions (cf. sec. 2.8),

$$\dim E_1^\perp = \dim E - \dim E_1. \tag{12.57}$$

However, the intersection $E_1 \cap E_1^\perp$ does not necessarily consist of the zero-vector alone, as in the positive definite case. Assume, for instance, that E_1 is the 1-dimensional subspace generated by a light-vector l. Then E_1 is contained in $E_1^\perp$.

It will be shown that $E_1 \cap E_1^\perp = 0$ if and only if the inner product is non-degenerate in E_1. Assume first that this condition is fulfilled. Let x_1 be a vector of $E_1 \cap E_1^\perp$. Then

$$(x_1, y_1) = 0 \quad \text{for all vectors } y_1 \in E_1, \tag{12.58}$$

whence $x_1 \in E_1 \cap E_1^\perp$ and thus $x_1 = 0$. Conversely, assume that $E_1 \cap E_1^\perp = 0$. Then the equation (12.57) implies that

$$E = E_1 \oplus E_1^\perp. \tag{12.59}$$

Now let x_1 be a vector of E_1 such that

$$(x_1, y_1) = 0 \quad \text{for all vectors} \quad y_1 \in E_1 .$$

It follows from (12.59) that every vector y of E can be written as

$$y = y_1 + y_1^{\perp} \qquad y_1 \in E_1, \, y_1^{\perp} \in E_1^{\perp},$$

whence

$$(x_1, y) = (x_1, y_1) + (x_1, y_1^{\perp}) = 0 \quad \text{for all vectors} \quad y \in E .$$

This equation implies that $x_1 = 0$. Consequently, the inner product is non-degenerate in E_1.

12.20. Normed determinant-functions. Let Δ_0 be a determinant-function in E. Since E is dual to itself, the identity (4.24) applies to E yielding

$$\begin{vmatrix} (x_1, y_1) & \dots & (x_1, y_n) \\ \vdots & & \vdots \\ (x_n, y_1) & \dots & (x_n, y_n) \end{vmatrix} = \alpha \, \Delta_0 (x_1 \dots x_n) \, \Delta_0 (y_1 \dots y_n), \; \alpha \neq 0 . \tag{12.60}$$

Substituting $x_\nu = y_\nu = e_\nu$ in (12.60), where $e_\nu (\nu = 1 \dots n)$ is an orthonormal basis, we obtain

$$\alpha \, \Delta_0 (e_1 \dots e_n)^2 = (-1)^{n-s} .$$

This equation shows that

$$\alpha (-1)^{n-s} > 0 .$$

Consequently, another determinant-function Δ can be defined by

$$\Delta = \pm \sqrt{(-1)^{n-s} \alpha} \cdot \Delta_0 \tag{12.61}$$

Then the identity (12.60) assumes the form

$$\Delta (x_1 \dots x_n) \, \Delta (y_1 \dots y_n) = (-1)^{n-s} \begin{vmatrix} (x_1, y_1) & \dots & (x_1, y_n) \\ \vdots & & \vdots \\ (x_n, y_1) & \dots & (x_n, y_n) \end{vmatrix} \tag{12.62}$$

A determinant-function satisfying the relation (12.62) is called a *normed determinant-function.* The equation (12.61) shows that there exist exactly two normed determinant-functions Δ and $-\Delta$ in E.

12.21. The pseudo-Euclidean plane. The simplest example of a pseudo-Euclidean space is a 2-dimensional linear space with an inner product of index 1. Then the light-cone consists of two straight lines. Selecting two vectors l_1 and l_2 which generate these lines we have the equations

$$(l_1, l_1) = 0 \quad \text{and} \quad (l_2, l_2) = 0. \tag{12.63}$$

But

$$(l_1, l_2) \neq 0$$

because otherwise the inner product would be identically zero. We can therefore assume that

$$(l_1, l_2) = -1\,. \tag{12.64}$$

Given a vector

$$x = \xi^1 l_1 + \xi^2 l_2$$

of E the equations (12.63) and (12.64) yield

$$(x, x) = -2\,\xi^1\,\xi^2$$

showing that x is space-like if $\xi^1\,\xi^2 < 0$ and x is time-like if $\xi^1\,\xi^2 > 0$. In other words, the space-like vectors are contained in the two sectors S_1 and S_2 of fig. 3 and the time-like vectors are contained in T^+ and T^-.

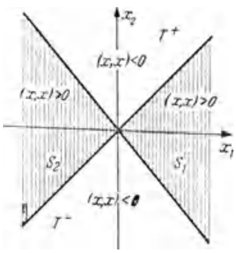


Fig. 3

The inner product of two vectors

$$x = \xi^1\,l_1 + \xi^2\,l_2 \quad \text{and} \quad y = \eta^1\,l_1 + \eta^2\,l_2$$

is given by

$$(x, y) = -\,(\xi^1\eta^2 + \xi^2\eta^1)\,.$$

This formula shows that the inner product of two space-like vectors is positive if and only if these vectors are contained in the same one of the sectors S_1 and S_2.

Let an orientation be defined in E by the normed determinant-function Δ. Then the identity (12.62) yields ($n = 2$, $s = 1$)

$$(x, y)^2 - \Delta(x, y)^2 = (x, x)\,(y, y)\,. \tag{12.65}$$

If x and y are not light-vectors the equation (12.65) may be written in the form

$$\frac{(x, y)^2}{(x, x)\,(y, y)} - \frac{\Delta\,(x, y)^2}{(x, x)\,(y, y)} = 1\,. \tag{12.66}$$

Now assume in addition, that the vectors x and y are space-like and are contained in the same one of the sectors S_1 and S_2. Then

$$(x, y) > 0\,. \tag{12.67}$$

The relations (12.66) and (12.67) imply that there exists exactly one real number θ ($-\infty < \theta < \infty$) such that

$$\cosh\theta = \frac{(x, y)}{|x|\,|y|} \quad \text{and} \quad \sinh\theta = \frac{\Delta(x, y)}{|x|\,|y|}\,. \tag{12.68}$$

This number is called the pseudo-Euclidean angle between the space-like vectors x and y.

We finally note that the vectors

$$e_1 = \frac{1}{\sqrt{2}}\,(l_1 - l_2) \quad \text{and} \quad e_2 = \frac{1}{\sqrt{2}}\,(l_1 + l_2)$$

form an orthonormal basis of E. Relative to this basis the equation of the light-cone assumes the form

$$(\xi^1)^2 - (\xi^2)^2 = 0\,.$$

12.22. Pseudo-Euclidean spaces of index *n—1*. More generally let us consider an n-dimensional pseudo-Euclidean space with the index $n-1$. Then every fixed time-like unit vector z determines an orthogonal decomposition of E into an $(n-1)$-dimensional subspace consisting of space-like vectors and the 1-dimensional subspace generated by z. In fact, every vector $x \in E$ can be uniquely decomposed in the form

$$x = \lambda z + y \qquad (z, y) = 0$$

where the scalar λ is given by

$$\lambda = -(x, z)\,.$$

Passing over to the norm we obtain the equation

$$(x, x) = -\lambda^2 + (y, y)$$

showing that

$$\begin{aligned} &\lambda^2 < (y, y) \text{ if } x \text{ is space-like} \\ &\lambda^2 > (y, y) \text{ if } x \text{ is time-like} \\ &\lambda^2 = (y, y) \text{ if } x \text{ is a light-vector.} \end{aligned} \tag{12.69}$$

From this decomposition we shall now derive the following properties:

(1) Two time-like vectors are never orthogonal.

(2) A time-like vector is never orthogonal to a light-vector.

(3) Two light-vectors are orthogonal if and only if they are linearly dependent.

(4) The orthogonal complement of a light-vector is an $(n-1)$-dimensional subspace of E in which the inner product is positive semidefinite and has the rank $n-2$.

To prove (1), consider another time-like vector z_1. This vector z_1 can be written as

$$z_1 = \lambda z + y_1 \qquad (z, y_1) = 0\,. \tag{12.70}$$

Then

$$\lambda^2 > (y_1, y_1),$$

whence $\lambda \neq 0$. Inner multiplication of (12.70) by z yields

$$(z, z_1) = \lambda\,(z, z) \neq 0\,.$$

Next, consider a light-vector l. Then

$$l = \lambda z + y \qquad (z, y) = 0$$

and

$$\lambda^2 = (y, y) > 0\,.$$

These two relations imply that

$$(l, z) = \lambda (z, z) \neq 0$$

which proves (2).

Now let l_1 and l_2 be two orthogonal light-vectors. Then we have the decompositions

$$l_1 = \lambda_1 z + y_1 \quad \text{and} \quad l_2 = \lambda_2 z + y_2 ,$$

whence

$$-\lambda_1 \lambda_2 + (y_1, y_2) = 0 . \tag{12.71}$$

Observing that

$$\lambda_1^2 = (y_1, y_1) \quad \text{and} \quad \lambda_2^2 = (y_2, y_2)$$

we obtain from (12.71) the equation

$$(y_1, y_1)\,(y_2, y_2) = (y_1, y_2)^2 . \tag{12.72}$$

The vectors y_1 and y_2 are contained in the orthogonal complement of z. In this space the inner product is positive definite and hence the equation (12.72) implies that y_1 and y_2 are linearly dependent , $y_2 = \lambda y_1$. Inserting this into (12.71) we find $\lambda_2 = \lambda \lambda_1$, whence $l_2 = \lambda l_1$.

Finally, let l be a light-vector and E_1 be the orthogonal complement of l. It follows from property (2) that E_1 does not contain time-like vectors. In other words, the inner product is positive semidefinite in E_1. To find the null-space of the inner product, assume that y_1 is a vector of E_1 such that

$$(y_1, y) = 0 \quad \text{for all vectors } y \in E_1 .$$

This implies that $(y_1, y_1) = 0$ showing that y_1 is a light-vector. Now it follows from property (3) that y_1 is a multiple of l. Consequently, the null-space of the inner product in E_1 is generated by l.

12.23. Fore-cone and past-cone. As another consequence of the properties established in the last section it will now be shown that the set of all time-like vectors consists of two disjoint sectors T^+ and T^- (cf. fig. 4.) To this purpose we define an equivalence relation in the set T of all time-like vectors in the following way:

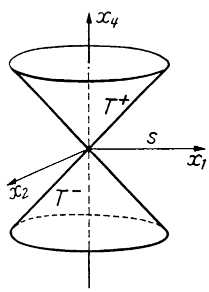


Fig. 4

$$z_1 \sim z_2 \quad \text{if} \quad (z_1, z_2) < 0 . \tag{12.73}$$

The relation (12.73) is obviously symmetric and reflexive. To prove the transitivity, consider three time-like vectors z_i $(i = 1, 2, 3)$ and assume that

$$(z_1, z_3) < 0 \quad \text{and} \quad (z_2, z_3) < 0 .$$

We have to show that

$$(z_1, z_2) < 0 .$$

We may assume that z_3 is a time-like *unit-vector*. Then the vectors z_1 and z_2 can be decomposed in the form

$$z_i = \lambda_i z_3 + y_i , \quad \lambda_i = (z_i, z_3) \quad (i = 1, 2) \tag{12.74}$$

where the vectors y_1 and y_2 are contained in the orthogonal complement

F of z_3. The equations (12.74) yield

$$(z_i, z_i) = -\lambda_i^2 + (y_i, y_i) \quad (i = 1, 2) \tag{12.75}$$

and

$$(z_1, z_2) = -\lambda_1 \lambda_2 + (y_1, y_2) . \tag{12.76}$$

It follows from (12.69) that

$$(y_i, y_i) < \lambda_i^2 \quad (i = 1, 2) .$$

Now observe that the inner product is positive definite in the subspace F. Consequently, the Schwarz-inequality applies to the vectors y_1 and y_2, yielding

$$(y_1, y_2)^2 \leqq (y_1, y_1)\,(y_2, y_2) \leqq \lambda_1^2 \lambda_2^2 .$$

This inequality shows that the first term determines the sign on the right-hand side of (12.76). But this term is negative because $\lambda_i = (z_i, z_3) < 0$ $(i = 1, 2)$ and we thus obtain

$$(z_1, z_2) < 0 .$$

The equivalence-relation (12.73) induces a decomposition of the set T into two classes T^+ and T^- which are obtained from each other by the reflection $x \to -x$.

12.24. The two subsets T^+ and T^- are *convex*, i. e., they contain with any two vectors z_1 and z_2 the straight segment

$$z(t) = (1 - t)\, z_1 + t z_2 \quad (0 \leqq t \leqq 1) .$$

In fact, assume that $z_1 \in T^+$ and $z_2 \in T^+$. Then

$$(z_1, z_1) < 0 \,, \; (z_2, z_2) < 0 \quad \text{and} \quad (z_1, z_2) < 0 \,,$$

whence

$$(z(t), z(t)) = (1 - t)^2\, (z_1, z_1) + 2t\,(1 - t)\,(z_1, z_2) + t^2\, (z_2, z_2) < 0 \,, \qquad (0 \leqq t \leqq 1) .$$

In the special theory of relativity the sectors T^+ and T^- are called the *fore-cone* and the *past-cone*.

The set S of the space-like vectors is not convex as fig. 4 shows.

Problems: 1. Let E be a pseudo-Euclidean plane and g_1, g_2 be the two straight lines generated by the light-vectors. Introduce a Euclidean metric in E such that g_1 and g_2 are orthogonal. Prove that two vectors $x \neq 0$ and $y \neq 0$ are orthogonal with respect to the pseudo-Euclidean metric if and only if they generate the Euclidean bisectors of g_1 and g_2.

2. Consider a pseudo-Euclidean space of dimension 3 and index 2. Assume that an orientation is defined in E by the normed determinant-function Δ. As in a Euclidean space define the exterior product of two vectors x_1 and x_2 by the relation

$$([x_1, x_2], x_3) = \Delta\,(x_1, x_2, x_3) .$$

Prove: a) $[x_1, x_2] = 0$ if and only if the vectors x_1 and x_2 are linearly dependent.

b) $([x_1, x_2], [x_1, x_2]) = (x_1, x_2)^2 - (x_1, x_1)(x_2, x_2)$

c) If e_1, e_2, e_3 is a positive orthonormal basis of E then

$$[e_1, e_2] = -e_3, \ [e_1, e_3] = -e_2, \ [e_2, e_3] = e_1 .$$

3. Let E be an n-dimensional pseudo-Euclidean space of index $n-1$. Given two time-like vectors z_1 and z_2 prove: a) The vector $z_1 + z_2$ is time-like or space-like depending on whether z_1 and z_2 are contained in the same cone or in different cones. b) The Schwarz-inequality holds in the reversed form

$$(z_1, z_2)^2 \geqq (z_1, z_1)(z_2, z_2) .$$

The equality-sign stands if and only if z_1 and z_2 are linearly dependent.

4. Denote by S the set of all space-like vectors. Prove that the set S is connected if $n \geqq 3$. More precisely: Given two vectors $x_0 \in S$ and $x_1 \in S$ there exists a continuous curve $x = x(t)$ $(0 \leqq t \leqq 1)$ in S such that $x(0) = x_0$ and $x(1) = x_1$.

§ 5. Linear mappings of pseudo-Euclidean spaces

12.25. The adjoint mapping. Let φ an endomorphism of the n-dimensional pseudo-Euclidean space E. Since E is dual to itself with respect to the inner product the adjoint endomorphism $\tilde{\varphi}$ can be defined as in sec. 11.1. The mappings φ and $\tilde{\varphi}$ are connected by the relation

$$(\varphi x, y) = (x, \tilde{\varphi} y) \qquad x, y \in E . \qquad (12.77)$$

The duality of the mappings φ and $\tilde{\varphi}$ implies that

$$\det \tilde{\varphi} = \det \varphi \quad \text{and} \quad \operatorname{tr} \tilde{\varphi} = \operatorname{tr} \varphi .$$

Let (α_ν^μ) and $(\tilde{\alpha}_\nu^\mu)$ $(\nu, \mu = 1 \dots n)$ be the matrices of φ and $\tilde{\varphi}$ relative to an orthonormal basis e_ν. Inserting $x = e_\nu$ and $y = e_\mu$ into (12.77) we find that

$$\varepsilon_\mu \tilde{\alpha}_\nu^\mu = \varepsilon_\nu \alpha_\mu^\nu \qquad (\nu, \mu = 1 \dots n)$$

where

$$\varepsilon_\nu = \begin{cases} +1 \ (\nu = 1 \dots s) \\ -1 \ (\nu = s + 1 \dots n) . \end{cases}$$

Now assume that the endomorphism φ is selfadjoint, $\tilde{\varphi} = \varphi$. In the positive definite case we have shown that there exists a system of n orthonormal eigenvectors. This result can be carried over to pseudo-Euclidean spaces of dimension $n \geqq 3$ if we add the hypothesis that $(x, \varphi x) \neq 0$ for all light-vectors. To prove this, consider the symmetric bilinear functions

$$\Phi(x, y) = (\varphi x, y) \quad \text{and} \quad \Psi(x, y) = (x, y) .$$

It follows from the above assumption that

$$\Phi(x)^2 + \Psi(x)^2 > 0 \quad \text{for all vectors} \quad x \neq 0\,.$$

Hence the theorem of sec. 12.12 applies to Φ and Ψ, showing that there exists an orthonormal basis e_ν $(\nu = 1 \ldots n)$ such that

$$(\varphi e_\nu, e_\mu) = \lambda_\nu \varepsilon_\nu \delta_{\nu\mu} \qquad (\nu, \mu = 1 \ldots n)\,. \tag{12.78}$$

The equations (12.78) imply that

$$\varphi e_\nu = \lambda_\nu \varepsilon_\nu e_\nu \qquad (\nu = 1 \ldots n)$$

showing that the e_ν are eigenvectors of φ.

12.26. Pseudo-Euclidean rotations. An endomorphism φ of the pseudo-Euclidean space E which preserves the inner product,

$$(\varphi x, \varphi y) = (x, y) \tag{12.79}$$

is called a *pseudo-Euclidean rotation*. Replacing y by x in (12.79) we obtain the equation

$$(\varphi x, \varphi x) = (x, x) \qquad x \in E$$

showing that a pseudo-Euclidean rotation sends space-like vectors into space-like vectors, time-like vectors into time-like vectors and light-vectors into light-vectors. A rotation is always regular. In fact, assume that $\varphi x = 0$ for a vector $x \in E$. Then it follows from (12.79) that

$$(x, y) = (\varphi x, \varphi y) = 0$$

for all vectors $y \in E$, whence $x = 0$.

Comparing the relations (12.77) and (12.79) we see that the adjoint and the inverse of a pseudo-Euclidean rotation coincide,

$$\tilde{\varphi} = \varphi^{-1}. \tag{12.80}$$

The equation (12.80) shows that the determinant of φ must be ± 1, as in the Euclidean case.

Now let e be an eigenvector of φ and λ be the corresponding eigenvalue,

$$\varphi e = \lambda e\,.$$

Passing over to the norms we obtain

$$(\varphi e, \varphi e) = \lambda^2 (e, e)\,.$$

This equation shows that $\lambda = \pm 1$ provided that e is not a light-vector. An eigenvector which is contained in the light-cone may have an eigenvalue $\lambda \neq \pm 1$ as can be seen from examples.

If an orthonormal basis is chosen in E the matrix of φ satisfies the relations

$$\sum_\lambda \varepsilon_\lambda \alpha_\nu^\lambda \alpha_\mu^\lambda = \varepsilon_\nu \delta_{\nu\mu}\,.$$

A matrix of this kind is called *pseudo-orthogonal*.

12.27. Pseudo-Euclidean rotations of the plane. In particular, consider a pseudo-Euclidean rotation φ of a 2-dimensional space with the index 1. Then the light-cone consists of two straight lines. Since the light-cone is preserved under the rotation φ, if follows that these straight lines are either transformed into themselves or they are interchanged. Now assume that φ is a *proper* rotation i.e. det $\varphi = +1$. Then the second case is impossible because the inner product is preserved under φ. Thus we can write

$$\varphi l_1 = \lambda l_1 \quad \text{and} \quad \varphi l_2 = \frac{1}{\lambda} l_2, \tag{12.81}$$

where l_1, l_2 is the basis of E defined in sec. 12.19. The number λ is positive or negative depending on whether the sectors T^+ and T^- are mapped onto themselves or interchanged.

Now consider an arbitrary vector

$$x = \xi^1 l_1 + \xi^2 l_2 . \tag{12.82}$$

Then the equations (12.81) and (12.82) yield

$$\varphi x = \lambda \xi^1 l_1 + \frac{1}{\lambda} \xi^2 l_2 ,$$

whence

$$(x, \varphi x) = \left(\lambda + \frac{1}{\lambda}\right) \xi^1 \xi^2 = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)(x, x) . \tag{12.83}$$

This equation shows, that the inner product of x and φx depends only on the norm of x as in the case of a Euclidean rotation (cf. sec. 11.22).

To find the "rotation-angle" of φ, introduce an orientation in E such that the basis l_1, l_2 is positive. Let Δ be a normed determinant-function which represents this rotation. Then the identity (12.65) yields

$$\Delta\,(l_1, l_2)^2 = (l_1, l_2)^2 = 1 ,$$

whence

$$\Delta\,(l_1, l_2) = 1 .$$

Inserting the vectors x and φx into Δ we find that

$$\Delta\,(x, \varphi x) = \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)(x, x)\,\Delta\,(l_1, l_2) = \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)(x, x) .$$

Now assume in addition that φ transforms the sectors T^+ and T^- into themselves (i. e. that φ does not interchange T^+ and T^-). Then $\lambda > 0$ and the equation (12.83) shows that $(x, \varphi x) > 0$ for every space-like vector x. Using the formulas (12.68) we obtain the equations

$$\cosh\theta = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) \quad \text{and} \quad \sinh\theta = \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right), \tag{12.84}$$

where θ denotes the pseudo-Euclidean angle between the vectors x and φx.

Now consider the orthonormal basis of E which is determined by the vectors

$$e_1 = \frac{1}{\sqrt{2}}(l_1 - l_2) \quad \text{and} \quad e_2 = \frac{1}{\sqrt{2}}(l_1 + l_2) .$$

Then the equations (12.81) and (12.84) yield

$$\varphi e_1 = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) e_1 + \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right) e_2$$
$$\varphi e_2 = \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right) e_1 + \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) e_2 .$$

We thus obtain the following representation of φ, which corresponds to the representation (11.49) of a Euclidean rotation:

$$\varphi e_1 = e_1 \cosh \theta + e_2 \sinh \theta$$
$$\varphi e_2 = e_1 \sinh \theta + e_2 \cosh \theta .$$

12.28. Lorentz-transformations. A 4-dimensional pseudo-Euclidean space with the index 3 is called *Minkowski-space*. A *Lorentz-transformation* is a rotation of the Minkowski-space. The purpose of this section is to show that *a proper Lorentz-transformation φ possesses always at least one eigenvector on the light-cone**). We may restrict ourselves to Lorentz-transformations which do not interchange fore-cone and past-cone because this can be achieved by multiplication with -1. These transformations are called *orthochroneous*. First of all we observe that a light-vector l is an eigenvector of φ if and only if $(l, \varphi l) = 0$. In fact, the equation $\varphi l = \lambda l$ yields

$$(l, \varphi l) = \lambda (l, l) = 0 .$$

Conversely, assume that l is a light-vector with the property that $(l, \varphi l) = 0$. Then it follows from sec. 11.22 property (3) that the vectors l and φl are linearly dependent. In other words, l is an eigenvector of φ.

Now consider the selfadjoint mapping

$$\psi = \frac{1}{2}(\varphi + \tilde{\varphi}) = \frac{1}{2}(\varphi + \varphi^{-1}) . \qquad (12.85)$$

Then

$$(x, \psi x) = \frac{1}{2}(x, \varphi x) + \frac{1}{2}(x, \tilde{\varphi} x) = (x, \varphi x) \quad x \in E . \qquad (12.86)$$

It follows from the above remark and from (12.86) that a light-vector l is an eigenvector of φ if and only if $(l, \psi l) = 0$. We now preceed indirectly and assume that φ does not have an eigenvector on the light-cone. Then $(x, \psi x) = (x, \varphi x) \neq 0$ for all light-vectors and hence we can apply the result of sec. 12.25 to the mapping ψ: There exist four eigenvectors $e_\nu (\nu = 1 \ldots 4)$ such that

$$(e_\nu, e_\mu) = \varepsilon_\nu \, \delta_{\nu\mu} \quad \varepsilon_\nu = \begin{cases} +1 \ (\nu = 1, 2, 3) \\ -1 \ (\nu = 4) . \end{cases}$$

Let us denote the time-like eigenvector e_4 by e and the corresponding eigenvalue by λ. Then $\psi e = \lambda e$ and hence it follows from (12.85) that

$$\varphi^2 e = 2\lambda \, \psi e - e .$$

*) Observe that a proper *Euclidean* rotation of a 4-dimensional space need not have eigenvectors.

Next, we wish to construct a time-like eigenvector of the mapping φ. If φe is a multiple of e, e is such a vector. We thus may assume that the vectors e and φe are linearly independent. Then these two vectors generate a plane F which is invariant under φ. This plane intersects the light-cone in two straight lines. Since the plane F and the light-cone are both invariant under φ, these two lines are either interchanged or transformed into themselves. In the first case we have two eigenvectors of φ on the light-cone, in contradiction to our assumption. In the second case select two generating vectors l_1 and l_2 on these lines such that $(l_1, l_2) = 1$. Then

$$\varphi l_1 = \alpha l_2 \quad \text{and} \quad \varphi l_2 = \beta l_1 .$$

The condition

$$(\varphi l_1, \varphi l_2) = (l_1, l_2)$$

implies that $\alpha\,\beta = 1$. Now consider the vector z

$$z = l_1 + \alpha l_2 .$$

Then

$$\varphi z = \alpha l_2 + \alpha\,\beta\, l_1 = \alpha l_2 + l_1 = z$$

i. e. z is a fix-vector of φ. To show that z is timelike observe that

$$(z, z) = 2\alpha\,(l_1, l_2) = 2\alpha$$

and that the vector $t = l_1 - l_2$ is time-like. Moreover,

$$(t, \varphi t) = \alpha + \frac{1}{\alpha} . \tag{12.87}$$

Now $(t, \varphi t) < 0$ because φ leaves the fore-cone and the past-cone invariant (cf. sec. 12.23). Hence the equation (12.87) implies that $\alpha < 0$ showing that z is time-like.

Using the time-like fix-vector z we shall now construct an eigenvector on the light-cone which will give us a contradiction. Let E_1 be the orthogonal complement of z. E_1 is a 3-dimensional Euclidean subspace of E which is invariant under φ. Since φ is a proper Lorentz-transformation it induces a proper Euclidean rotation in E_1. Consequently, there exists a fix-axis in E_1 (cf. sec. 11.20). Let y be a vector of the fix-axis such that $(y, y) = -2\,\alpha$. Then $l = y + z$ is a light-vector and

$$\varphi l = \varphi y + \varphi z = y + z = l$$

i. e. l is an eigenvector of φ.

Hence, the assumption that there are no eigenvectors on the light-cone leads to a contradiction and the assertion in the beginning of this section is proved.

We finally note that every eigenvalue λ of φ whose eigenvector l lies on the light-cone is positive. In fact, select a space-like unit-vector y such that $(l, y) = 1$ and consider the vector $z = l + \tau y$ where τ is a real parameter. Then we have the relation

$$(z, z) = 2\tau + \tau^2$$

showing that z is time-like for $-2 < \tau < 0$. Since φ preserves fore-cone and past-cone it follows that

$$(z, \varphi z) < 0 \qquad\qquad (-2 < \tau < 0)\,.$$

But

$$(z, \varphi z) = (l + \tau y, \lambda l + \tau \varphi y) = \tau\left(\lambda + \frac{1}{\lambda}\right) + \tau^2 (y, \varphi y)$$

and we thus obtain

$$\tau\left(\lambda + \frac{1}{\lambda}\right) + \tau^2 (y, \varphi y) < 0 \qquad\qquad (-2 < \tau < 0)\,.$$

Letting $\tau \to 0$ we see that λ must be positive.

Problems: 1. Let φ be an automorphism of the plane E. Prove that an inner product of index 1 can be introduced in E such that φ becomes a proper pseudo-Euclidean rotation if and only if the following conditions are satisfied:

1. There are two linearly independent eigenvectors.
2. $\det\varphi = 1$.
3. $|\mathrm{tr}\,\varphi| \geqq 2$.

2. Find the eigenvectors of the Lorentz-transformation defined by the matrix

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

Verify that there exists an eigenvector on the light-cone.

3. Let a and b be two linearly independent light-vectors in the pseudo-Euclidean plane E. Then an endomorphism ψ of E is defined by

$$\psi a = a\,,\ \psi b = -b\,.$$

Consider the family of automorphisms $\varphi(t)$ which is defined by the differential equation

$$\dot{\varphi}(t) = \psi \circ \varphi(t)$$

and the initial condition

$$\varphi(0) = \iota\,.$$

a) Prove that $\varphi(t)$ is a family of proper rotations carrying fore-cone and past-cone into itself.

b) Define the functions $C(t)$ and $S(t)$ by

$$C(t) = \frac{1}{2}\,\mathrm{tr}\,\varphi(t) \quad\text{and}\quad S(t) = \frac{1}{2}\,\mathrm{tr}\,(\psi \circ \varphi(t))\,.$$

Prove the functional-equations

$$C(t_1+t_2) = C(t_1)\,C(t_2) + S(t_1)\,S(t_2)$$

and

$$S(t_1+t_2) = S(t_1)\,C(t_2) + S(t_2)\,C(t_1)\,.$$

c) Prove that

$$\varphi(t)\,a = e^{-t}a \quad \text{and} \quad \varphi(t)\,b = e^{-t}\,b\,.$$

Chapter XIII

Quadrics

In the present Chapter the theory of the bilinear functions developed in Chapter XII will be applied to the discussion of quadrics. In this context we shall have to deal with *affine spaces.*

§ 1. Affine spaces

13.1. Points and vectors. Let E be a real n-dimensional linear space and A a set of elements $P, Q \dots$ which will be called *points*. Assume that a relation between points and vectors is defined in the following way:

1. To every ordered pair P, Q of A there is assigned a vector of E, called the *difference vector* and denoted by $\overrightarrow{PQ}$.

2. To every point $P \in A$ and every vector $x \in E$ there exists exactly one point $Q \in A$ such that $\overrightarrow{PQ} = x$.

3. If P, Q, R are three arbitrary points, then

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}\,. \tag{13.1}$$

A is called an *n-dimensional affine space* with the *difference space E.*

Insertion of $Q = P$ in (13.1) yields $\overrightarrow{PP} + \overrightarrow{PR} = \overrightarrow{PR}$, whence $\overrightarrow{PP} = 0$ for every point $P \in A$. Using this relation we obtain from (13.1)

$$\overrightarrow{QP} = -\overrightarrow{PQ}\,.$$

The equation $\overrightarrow{P_1Q_1} = \overrightarrow{P_2Q_2}$ implies that $\overrightarrow{P_1P_2} = \overrightarrow{Q_1Q_2}$ (parallelogram-law). In fact,

$$\overrightarrow{P_1P_2} = \overrightarrow{P_1Q_2} - \overrightarrow{P_2Q_2}$$

$$\overrightarrow{Q_1Q_2} = \overrightarrow{P_1Q_2} - \overrightarrow{P_1Q_1}\,.$$

Subtraction of these equations yields $\overrightarrow{P_1P_2} = \overrightarrow{Q_1Q_2}$.

For any given linear space E, an affine space can be constructed which possesses E as difference space:

Define the points as the vectors of E and the difference-vector of two points x and y as the vector $y - x$. Then the above conditions are obviously satisfied.

Let A be a given affine space. If a fixed point O is distinguished as origin, every point P is uniquely determined by the vector $x = \overrightarrow{OP}$. x is called the *position-vector* of P and every point P can be identified with the corresponding position-vector x. The difference-vector of two points x and y is simply the vector $y - x$.

13.2. Affine coordinate systems. An *affine coordinate-system* $(O; x_1 \dots x_n)$ consists of a fixed point $O \in A$, the *origin*, and a basis x_ν $(\nu = 1 \dots n)$ of the difference-space E. Then every point $P \in A$ determines a system of n numbers ξ^ν $(\nu = 1 \dots n)$ by

$$\overrightarrow{OP} = \sum_\nu \xi^\nu x_\nu .$$

The numbers ξ^ν $(\nu = 1 \dots n)$ are called the *affine coordinates* of P relative to the given system. The origin O has the coordinates $\xi^\nu = 0$.

Now consider two affine coordinate-systems

$$(O; x_1 \dots x_n) \quad \text{and} \quad (O'; y_1 \dots y_n) .$$

Denote by α_ν^μ the matrix of the basis-transformation $x_\nu \to y_\nu$ and by β^ν the affine coordinates of O' relative to the system $(O; x_1 \dots x_n)$,

$$y_\nu = \sum_\mu \alpha_\nu^\mu x_\mu , \qquad \overrightarrow{OO'} = \sum_\nu \beta^\nu x_\nu .$$

The affine coordinates ξ^ν and η^ν of a point P, corresponding to the systems $(O; x_1 \dots x_n)$ and $(O'; y_1 \dots y_n)$ respectively, are given by

$$\overrightarrow{OP} = \sum_\nu \xi^\nu x_\nu \quad \text{and} \quad \overrightarrow{O'P} = \sum_\nu \eta^\nu y_\nu . \tag{13.2}$$

Inserting $\overrightarrow{O'P} = \overrightarrow{OP} - \overrightarrow{OO'}$ in the second equation (13.2) we obtain

$$\sum_\nu \eta^\nu y_\nu = \sum_\nu (\xi^\nu - \beta^\nu) x_\nu ,$$

whence

$$\sum_\nu \alpha_\nu^\mu \eta^\nu = \xi^\mu - \beta^\mu \qquad (\mu = 1 \dots n) .$$

Multiplication by the inverse matrix yields the transformation-formula for the affine coordinates:

$$\eta^\nu = \sum_\mu \check{\alpha}_\mu^\nu (\xi^\mu - \beta^\mu) \qquad (\nu = 1 \dots n) .$$

13.3. Affine subspaces. An affine subspace of A is a subset A_1 of A such that the vectors $\overrightarrow{PQ}$ $(P \in A_1, Q \in A_1)$ form a subspace of E. If O is the origin of A and $(O_1; x_1 \dots x_p)$ is an affine coordinate-system of A_1, the points of A_1 can be represented as

$$\overrightarrow{OP} = \overrightarrow{OO_1} + \sum_{\nu=1}^{p} \xi^\nu x_\nu \, . \tag{13.3}$$

For $p = 1$ we obtain a straight line through O_1 with the "direction-vector" x,

$$\overrightarrow{OP} = \overrightarrow{OO_1} + \xi x \, .$$

In the case $p = 2$ the equation (13.3) reads

$$\overrightarrow{OP} = \overrightarrow{OO_1} + \xi^1 x_1 + \xi^2 x_2 \, .$$

It then represents the plane through O_1 generated by the vectors x_1 and x_2. An affine subspace of dimension $n - 1$ is called a *hyperplane.*

Two affine subspaces A_1 and A_2 of A are called *parallel* if the difference-space E_1 of A_1 is contained in the difference-space E_2 of A_2, or conversely. Parallel subspaces are either disjoint or contained in each other. Assume, for instance, that E_2 is contained in E_1. Let Q be a point of the intersection $A_1 \cap A_2$ and P_2 be an arbitrary point of A_2. Then $\overrightarrow{QA_2}$ is contained in E_2 and hence is contained in E_1. This implies that $P_2 \in A_1$, whence $A_2 \subset A_1$.

13.4. Affine mappings. Let $P \to P'$ be a mapping of A into itself subject to the following conditions:

1. $\overrightarrow{P_1Q_1} = \overrightarrow{P_2Q_2}$ implies that $\overrightarrow{P_1'Q_1'} = \overrightarrow{P_2'Q_2'}$.
2. The mapping $\varphi: E \to E$ defined by $\varphi(\overrightarrow{PQ}) = \overrightarrow{P'Q'}$ is linear.

Then $P \to P'$ is called an *affine mapping*. Given two points O and O' and a linear mapping $\varphi: E \to E$, there exists exactly one affine mapping which sends O into O' and induces φ. This mapping is defined by

$$\overrightarrow{OP'} = \overrightarrow{OO'} + \varphi(\overrightarrow{OP}) \, .$$

If a fixed origin is used in A, every affine mapping $x \to x'$ can be written in the form

$$x' = \varphi x + b \, ,$$

where φ is the induced linear mapping and $b = \overrightarrow{OO'}$.

A *translation* is an affine mapping which induces the identity in E,

$$\overrightarrow{P'Q'} = \overrightarrow{PQ} \, .$$

For two arbitrary points P and P' there obviously exists exactly one translation which sends P into P'.

13.5. Euclidean space. Let A be an n-dimensional affine space and assume that a positive definite inner-product is defined in the difference-space E. Then A is called a *Euclidean space.* The distance of two points P and Q is defined by

$$\varrho(P, Q) = |\overrightarrow{PQ}|.$$

It follows from this definition that the distance has the following properties:

1. $\varrho(P, Q) \geqq 0$ and $\varrho(P, Q) = 0$ if and only if $P = Q$.
2. $\varrho(P, Q) = \varrho(Q, P)$.
3. $\varrho(PQ) \leqq \varrho(P, R) + \varrho(R, Q)$.

All the metric concepts defined for an inner-product space (cf. Chap. X) can be applied to a Euclidean space. Given a point $x_1 \in A$ of A and a vector $p \neq 0$ there exists exactly one hyperplane through x_1 whose difference-space is orthogonal to p. This hyperplane is represented by the equation

$$(x - x_1, p) = 0.$$

A *rigid motion* of a Euclidean space is an affine mapping $P \to P'$ which preserves the distance,

$$\varrho(P', Q') = \varrho(P, Q). \tag{13.4}$$

The condition (13.4) implies that the linear mapping, which is induced in the difference-space by a rigid motion, is a rotation. Conversely, given a rotation φ and two points $O \in A$ and $O' \in A$, there exists exactly one rigid motion which induces φ and sends O into O'.

Problems: 1. $(p+1)$ points P_ν $(\nu = 0 \ldots p)$ in an affine space are said to be *in general position,* if the points P_ν are not contained in a $(p-1)$-dimensional subspace. Prove that the points P_ν $(\nu = 0 \ldots p)$ are in general position if and only if the vectors $\overrightarrow{P_0 P_\nu}$ are linearly independent.

2. Given $(p+1)$ points $P_\nu (\nu = 0 \cdots p)$ in general position, the set of all points P defined by

$$\overrightarrow{P_0 P} = \sum_{\nu=1}^{p} \xi^\nu \overrightarrow{P_0 P_\nu} \qquad \xi^\nu \geqq 0, \qquad \sum_{\nu=1}^{p} \xi^\nu \leqq 1$$

is called the p-simplex spanned by the points $P_\nu (\nu = 0 \ldots p)$. If O is the origin of A, prove that a point P of the above simplex can be uniquely represented as

$$\overrightarrow{OP} = \sum_{\nu=0}^{p} \xi^\nu \overrightarrow{OP_\nu}, \qquad \xi^\nu \geqq 0, \qquad \sum_{\nu=0}^{p} \xi^\nu = 1.$$

The numbers ξ^ν $(\nu = 0 \ldots p)$ are called the *barycentric coordinates* of P. The point B with the barycentric coordinates $\xi^\nu = \dfrac{1}{p+1}$ $(\nu = 0 \ldots p)$ is called the *center* of S.

3. Given a p-simplex $(P_0 \dots P_p)$ $(p \geqq 2)$ consider the $(p-1)$ simplex S_i defined by the points $(P_0 \dots \hat{P}_i \dots P_p)$ $(0 \leqq i \leqq p)$ and denote by B_i the center of $S_i (0 \leqq i \leqq p)$. Show that the straight lines (P_i, B_i) and (P_j, B_j) $(i \neq j)$ intersect each other at the center of S and that

$$\overrightarrow{B_i S} = \frac{1}{p+1} \overrightarrow{B_i P_i} .$$

4. An *equilateral simplex* of length a in *a* Euclidean space is a simplex $(P_0 \dots P_p)$ with the property that $|\overrightarrow{P_\nu P_\mu}| = a\,(\nu \neq \mu)$. Find the angle between the vectors $\overrightarrow{P_\nu P_\mu}$ and $\overrightarrow{P_\nu P_\lambda} (\mu \neq \nu,\ \lambda \neq \nu)$ and between the vectors $\overrightarrow{BP_\nu}$ and $\overrightarrow{BP_\mu}$ where B is the center of $(P_0 \dots P_p)$.

5. Assume that an orientation is defined in the difference-space E by the determinant-function Δ. An ordered system of $(n+1)$ points $(P_0 \dots P_n)$ in general position is called *positive* with respect to the given orientation, if

$$\Delta\,(\overrightarrow{P_0 P_1} \dots \overrightarrow{P_0 P_n}) > 0 .$$

a) If the system $(P_0 \dots P_n)$ is positive and σ is a permutation of the numbers $(0, 1 \dots n)$, show that the system $(P_{\sigma(0)} \dots P_{\sigma(n)})$ is again positive if and only if the permutation σ is even.

b) Let A_i be the $(n-1)$-dimensional subspace spanned by the points $P_0 \dots \hat{P}_i \dots P_n$. Introduce an orientation in the difference-space of A_i with the help of the determinant-function

$$\Delta_i(x_1 \dots x_{n-1}) = \Delta\,(\overrightarrow{P_i Q}, x_1 \dots x_{n-1}) ,$$

where Q is an arbitrary point of A_i. Prove that the ordered n-tuple $(P_0 \dots \hat{P}_i \dots P_n)$ is positive with respect to the determinant-function $(-1)^i \Delta_i$.

6. Let $(P_0 \dots P_n)$ be a n-simplex and S be its center. Denote by $S_{i_1 \dots i_k}$ the center of the $(n-k)$-simplex obtained by deleting the vertices $P_{i_1}, \dots, P_{i_k}$. Now select an ordered system of n integers $i_1, \dots, i_n$ $(0 \leqq i_\nu \leqq n)$ and define the affine mapping α by

$$\alpha : S \to P_0,\ S_{i_1} \to P_1,\ S_{i_1 i_2} \to P_2, \dots, S_{i_1 \dots i_n} \to P_n .$$

Prove that $\det \alpha = \frac{1}{(n+1)!} \varepsilon_\sigma$. In this equation σ denotes the permutation $\sigma(\nu) = i_\nu (\nu = 1 \dots n)$, $\sigma(0) = k$ where k is the integer not appearing among the numbers $(i_1 \dots i_n)$.

7. Let g_1 and g_2 be two straight lines in a Euclidean space which are not parallel and do not intersect. Prove that there exists exactly one point P_i on $g_i (i = 1, 2)$ such that $\overrightarrow{P_1 P_2}$ is orthogonal to g_1 and to g_2.

8. Let A_1 and A_2 be two subspaces of the affine space such that the difference-spaces E_1 and E_2 form a direct decomposition of E. Prove that the intersection $A_1 \cap A_2$ consists of exactly one point.

9. Prove that a rigid motion $x' = \tau x + a$ has a fix-point if and only if the vector a is orthogonal to all fix-vectors of τ.

A rigid motion is called *proper*, if $\det \tau = +1$. Prove that every proper rigid motion of the Euclidean plane without fix-points is a translation.

10. Consider a proper rigid motion $x' = \tau x + a$ $(\tau \neq \iota)$ of the Euclidean plane. Prove that there is exactly one fix-point x_0 and that

$$x_0 = \frac{1}{2} |a| \left(a + b \cot \frac{\theta}{2}\right).$$

In this equation, b is a vector of the same length as a and orthogonal to a. θ designates the rotation-angle relative to the orientation defined by the basis (a, b).

11. Prove that two proper rigid motions $\alpha \neq \iota$ and $\beta \neq \iota$ of the plane commute, if and only if one of the two conditions holds:

1. α and β are both translations
2. α and β have the same fix-point.

§ 2. Quadrics in the affine space

13.6. Definition. From elementary analytic geometry it is well known that every conic section can be represented by an equation of the form

$$\sum_{\nu,\mu=1}^{2} \alpha_{\nu\mu} \xi^\nu \xi^\mu + 2 \sum_{\nu=1}^{2} \beta_\nu \xi^\nu = \alpha ,$$

where $\alpha_{\nu\mu}$, β_ν and α are constants. Generalizing this to higher dimensions we define a *quadric* Q in an n-dimensional affine space A as the set of all points satisfying an equation of the form

$$\Phi(x) + 2 f(x) = \alpha , \tag{13.5}$$

where $\Phi \neq 0$ is a quadratic function, f a linear function and α a constant.

For the following discussion it will be convenient to introduce a dual space E^* of the difference-space E. Then the bilinear function Φ can be written in the form

$$\Phi(x, y) = \langle \varphi x, y \rangle \qquad x, y \in E ,$$

where φ is a linear mapping of E into E^* which is dual to itself: $\varphi^* = \varphi$. At the same time the linear function f determines a vector $\overset{*}{a} \in E^*$ such that

$$f(x) = \langle \overset{*}{a}, x \rangle \qquad x \in E .$$

Hence, the equation (13.5) can be written in the form

$$\langle \varphi x, x \rangle + 2 \langle \overset{*}{a}, x \rangle = \alpha . \tag{13.6}$$

We recall that the null-space of the bilinear function Φ coincides with the kernel of the linear mapping φ.

13.7. Cones. Let us assume that there exists a point $x_0 \in Q$ such that $\varphi x_0 + \overset{*}{a} = 0$. Then (13.6) can be written as

$$\langle \varphi x, x \rangle - 2 \langle \varphi x_0, x \rangle = \alpha \tag{13.7}$$

and the substitution $x = x_0$ gives

$$\alpha = - \langle \varphi x_0, x_0 \rangle .$$

Inserting this into (13.7) we obtain

$$\langle \varphi x, x \rangle - 2 \langle \varphi x_0, x \rangle + \langle \varphi x_0, x_0 \rangle = 0 .$$

Hence, the equation of Q assumes the form

$$\Phi (x - x_0) = 0 .$$

A quadric of this kind is called a *cone with the vertex* x_0. For the sake of simplicity, cones will be excluded in the following discussion. In other words, it will be assumed that

$$\varphi x + \overset{*}{a} \neq 0 \quad \text{for all points } x \in Q . \tag{13.8}$$

13.8. Tangent-space. Consider a fixed point $x_0 \in Q$. It follows from the condition (13.8) that the orthogonal complement of the vector $\varphi x_0 + \overset{*}{a}$ is an $(n-1)$-dimensional subspace T_{x_0} of E. This subspace is called the *tangent-space* of Q at the point x_0. A vector $y \in E$ is contained in T_{x_0} if and only if

$$\langle \overset{*}{a} + \varphi x_0, y \rangle = 0 . \tag{13.9}$$

In terms of the functions Φ and f the equation (13.9) can be written as

$$\Phi (x_0, y) + f(y) = 0 . \tag{13.10}$$

The $(n-1)$-dimensional affine subspace which is determined by the point x_0 and the tangent-space T_{x_0} is called the *tangent-hyperplane* of Q at x_0. It consists of all points

$$x = x_0 + y \qquad\qquad y \in T_{x_0} .$$

Inserting $y = x - x_0$ into the equation (13.10) we obtain

$$\Phi (x_0, x - x_0) + f (x - x_0) = 0 . \tag{13.11}$$

Observing that

$$\Phi (x_0) + 2 f (x_0) = \alpha$$

we can write the equation (13.11) of the tangent-hyperplane in the form

$$\Phi (x_0, x) + f (x_0 + x) = \alpha . \tag{13.12}$$

To obtain a geometric picture of the tangent-space, consider a 2-dimensional plane

$$F: x = x_0 + \xi a + \eta b \tag{13.13}$$

through x_0 where a and b are two linearly independent vectors. Inserting (13.13) into the equation (13.5) we obtain the relation

$$\begin{aligned} &\xi^2 \Phi(a) + 2\,\xi\eta\,\Phi(a, b) + \eta^2 \Phi(b) + \\ &\quad + 2\xi\,(\Phi(x_0, a) + f(a)) + 2\eta\,(\Phi(x_0, b) + f(b)) = 0 \end{aligned} \tag{13.14}$$

showing that the plane F intersects Q in a conic γ. Upon introduction of the linear function

$$g(x) = 2\,(\Phi(x_0, x) + f(x)) \tag{13.15}$$

the equation of the conic γ can be written in the form

$$\xi^2 \Phi(a) + 2\xi\eta\,\Phi(a, b) + \eta^2 \Phi(b) + \xi g(a) + \eta\, g(b) = 0\,. \tag{13.16}$$

Now assume that the vectors a and b are chosen such that $g(a)$ and $g(b)$ are not both equal to zero. Then the conic has a unique tangent at the point $\xi = \eta = 0$ and this tangent is generated by the vector

$$t = -g(b)\, a + g(a)\, b\,. \tag{13.17}$$

The vector t is contained in the tangent-space T_{x_0}; this follows from the equation

$$g(t) = -\,g(b)\, g(a) + g(a)\, g(b) = 0\,.$$

Every vector $y \neq 0$ of the tangent-space T_{x_0} can be obtained in this way. In fact, let a be a vector such that $g(a) = 1$ and consider the plane through x_0 spanned by a and y. Then the equation (13.17) yields

$$t = -\,g(y)\, a + g(a)\, y = y \tag{13.18}$$

showing that y is the tangent-vector of the intersection $Q \cap F$ at the point $\xi = \eta = 0$.

Note: If $g(a) = 0$ and $g(b) = 0$ the equation (13.16) reduces to

$$\xi^2 \Phi(a) + 2\xi\eta\,\Phi(a, b) + \eta^2 \Phi(b) = 0\,.$$

Then the intersection of Q and F consists of

a) two straight lines intersecting at x_0, if

$$\Phi(a, b)^2 - \Phi(a)\,\Phi(b) > 0\,,$$

b) the point x_0 only, if

$$\Phi(a, b)^2 - \Phi(a)\,\Phi(b) < 0\,,$$

c) one straight line through x_0, if

$$\Phi(a, b)^2 - \Phi(a)\Phi(b) = 0,$$

but not all three coefficients $\Phi(a)$, $\Phi(b)$ and $\Phi(a, b)$ are zero

d) The entire plane F, if

$$\Phi(a) = \Phi(b) = \Phi(a, b) = 0 .$$

13.9. Uniqueness of the representation. Assume that a quadric Q is represented in two ways

$$\Phi_1(x) + 2f_1(x) = \alpha_1 \tag{13.19}$$

and

$$\Phi_2(x) + 2f_2(x) = \alpha_2 . \tag{13.20}$$

It will be shown that

$$\Phi_2 = \lambda\Phi_1 ,\ f_2 = \lambda f_1 ,\ \alpha_2 = \lambda\alpha_1$$

where $\lambda \neq 0$ is a real number. Let x_0 be a fixed point of Q. It follows from the hypothesis (13.8) that the linear functions g_1 and g_2, defined by

$$g_1(x) = \Phi_1(x, x_0) + f_1(x) \quad \text{and} \quad g_2(x) = \Phi_2(x, x_0) + f_2(x) \tag{13.21}$$

are not identically zero.

Choose a vector a such that $g_1(a) \neq 0$ and $g_2(a) \neq 0$, and a vector $b \neq 0$ such that $g_1(b) = 0$. Obviously a and b are linearly independent. The plane

$$x = x_0 + \xi a + \eta b$$

then intersects the quadric Q in a conic γ whose equation is given by each one of the equations

$$\xi^2\Phi_1(a) + 2\xi\eta\Phi_1(a, b) + \eta^2\Phi_1(b) + \xi g_1(a) = 0 \tag{13.22}$$

and

$$\xi^2\Phi_2(a) + 2\xi\eta\Phi_2(a, b) + \eta^2\Phi_2(b) + \xi g_2(a) + \eta g_2(b) = 0 . \tag{13.23}$$

The tangent of this curve at the point $\xi = \eta = 0$ is generated by the vector

$$t_1 = g_1(a)\, b$$

and also by

$$t_2 = -\, g_2(b)\, a + g_2(a)\, b .$$

This implies that

$$t_2 = \lambda t_1 \quad (\lambda \neq 0)$$

whence

$$g_2(a) = \lambda g_1(a) \quad \text{and} \quad g_2(b) = 0 . \tag{13.24}$$

But $b \neq 0$ was an arbitrary vector of the kernel of g_1. Hence the second equation (13.24) shows that $g_2(b) = 0$ whenever $g_1(b) = 0$. In other words, the linear functions g_1 and g_2 have the same kernel. Consequently g_2 is a constant multiple of g_1,

$$g_2 = \lambda g_1 \quad \lambda \neq 0 . \tag{13.25}$$

Multiplying the equation (13.22) by λ and substracting it from (13.23) we obtain in view of (13.24)

$$\xi^2 (\Phi_2 - \lambda \Phi_1)(a) + 2\xi\eta(\Phi_2 - \lambda\Phi_1)(a, b) + \eta^2(\Phi_2 - \lambda\Phi_1)(b) = 0. \tag{13.26}$$

In this equation all three coefficients must be zero. In fact, if at least one coefficient is different from zero, the equation (13.26) implies that the conic γ consists of two straight lines, one straight line or the point x_0 only. But this is impossible because $g_1(a) \neq 0$. We thus obtain from (13.26) the relations

$$\Phi_2(a) = \lambda\Phi_1(a) , \Phi_2(a, b) = \lambda\Phi_1(a, b) , \Phi_2(b) = \lambda\Phi_1(b) . \tag{13.27}$$

These equations show that

$$\Phi_2(x) = \lambda\Phi_1(x) \tag{13.28}$$

for all vectors $x \in E$: If $g_1(x) = 0$, (13.28) follows from the third equation (13.27); if $g_1(x) \neq 0$, then $g_2(x) \neq 0$ [in view of (13.25)] and (13.28) follows from the first equation (13.27).

Altogether we thus obtain the identities

$$\Phi_2 = \lambda\Phi_1 \quad \text{and} \quad g_2 = \lambda g_1 \qquad \lambda \neq 0 .$$

Now the relations (13.21) imply that

$$f_2 = \lambda f_1$$

and the equation (13.20) can be written as

$$\lambda(\Phi_1(x) + 2f_1(x)) = \alpha_2 . \tag{13.29}$$

Comparing the equations (13.19) and (13.29) we finally obtain $\alpha_2 = \lambda\alpha_1$. This completes the proof of the uniqueness-theorem.

13.10. Centers. Let

$$Q : \Phi(x) + 2f(x) = \alpha$$

be a given quadric and c be an arbitrary point of the space A. If we introduce c as a new origin,

$$x = c + x' ,$$

the equation of Q is transformed into

$$\Phi(x') + 2(\Phi(c, x') + f(x')) = \alpha - \Phi(c) - 2f(c) . \tag{13.30}$$

Here the question arises whether the point c can be chosen such that the linear terms in (13.30) disappear, i. e. that

$$\Phi(c, x') + f(x') = 0 \tag{13.31}$$

for all vectors $x' \in E$. If this is possible, c is called a *center* of Q. Writing the equation (13.31) in the form

$$\langle \varphi c + \overset{*}{a}, x' \rangle = 0 \qquad x' \in E$$

we see that c is a center of Q if and only if

$$\varphi c = - \overset{*}{a}. \tag{13.32}$$

This implies that the quadric Q has a center if and only if the vector $\overset{*}{a}$ is contained in the image-space $\varphi(E)$. Observing that the image-space $\varphi(E)$ is the orthogonal complement of the kernel K of φ we obtain the following criterion: *A quadric Q has a center if and only if the vector $\overset{*}{a}$ is orthogonal to the kernel of φ.*

If this condition is satisfied, the center is determined up to a vector of K. In other words, the set of all centers is an affine subspace of A with the kernel $K(\varphi)$ as difference-space.

Now assume that the bilinear function Φ is non-degenerate. Then φ is a regular mapping of E onto itself and hence the equation (13.32) has exactly one solution. Thus, it follows from the above criterion that *a non-degenerate quadric has exactly one center.*

13.11. Normal-form of a quadric with center. Suppose that Q is a quadric with centers. If a center is used as origin the equation of Q assumes the form

$$\Phi(x) = \beta \quad \beta \neq 0\,. \tag{13.33}$$

Then the tangent-vectors y at a point $x_0 \in Q$ are characterized by the equation $\langle \varphi x_0, y \rangle = 0$. Observing that $\langle \varphi x_0, y \rangle = \langle x_0, \varphi y \rangle$ we see that every tangent-space T_{x_0} contains the null-space of Φ.

The equation of the tangent-hyperplane of Q at x_0 is given by

$$\Phi(x_0, y) = \beta\,. \tag{13.34}$$

It follows from (13.34) that a center of Q is never contained in a tangent-hyperplane.

Dividing (13.33) by β and replacing the quadratic function Φ by $\frac{1}{\beta}\Phi$ we can write the equation of Q in the *normal-form*

$$\Phi(x) = 1\,. \tag{13.35}$$

Now select a basis x_ν $(\nu = 1 \ldots n)$ of E such that

$$\Phi(x_\nu, x_\mu) = \varepsilon_\nu \delta_{\nu\mu} \qquad \varepsilon_\nu = \begin{cases} +1 \; (\nu = 1 \ldots s) \\ -1 \; (\nu = s+1 \ldots r) \\ \;\;\,0 \; (\nu = r+1 \ldots n) \end{cases} \tag{13.36}$$

where r denotes the rank and s denotes the index of Φ. Then the normal-form (13.36) can be written as

$$\sum_{\nu=1}^{r} \varepsilon_\nu \xi^\nu \xi^\nu = 1 . \tag{13.37}$$

13.12. Normal-form of a quadric without center. Now consider a quadric Q without a center. If a point of Q is chosen as origin the constant α in (13.5) becomes zero and the equation of Q reads

$$\Phi(x) + 2\langle \overset{*}{a}, x\rangle = 0 . \tag{13.38}$$

By multiplying the equation (13.38) with -1 if necessary we can achieve that $2s \geqq r$. In other words, we can assume that the signature of Φ is not negative.

To reduce the equation (13.38) to a normal-form consider the tangent-space T_0 at the origin. The equation (13.9) shows that T_0 is the orthogonal complement of $\overset{*}{a}$. Hence, $\overset{*}{a}$ is contained in the orthogonal complement $T_0^\perp$. On the other hand, $\overset{*}{a}$ is not contained in the orthogonal complement $K^\perp$ because otherwise Q would have a center (cf. sec. 13.10). The relations $\overset{*}{a} \in T_0^\perp$ and $\overset{*}{a} \notin K^\perp$ show that $T_0^\perp \not\subset K^\perp$. Taking the orthogonal complement we obtain the relation $T_0 \not\supset K$ showing that there exists a vector $a \in K$ which is not contained in T_0 (cf. fig. 5). Then $\langle \overset{*}{a}, a\rangle \neq 0$ and hence we may assume that $\langle \overset{*}{a}, a\rangle = 1$.

Now T_0 has the dimension $n-1$ and hence every vector $x \in E$ can be written in the form

$$x = y + \xi a \qquad y \in T_0 . \tag{13.39}$$

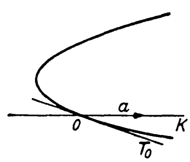


Fig. 5

Inserting (13.39) into the equation (13.38) we obtain

$$\Phi(y) + 2\xi\Phi(y, a) + \xi^2\Phi(a) + 2\langle \overset{*}{a}, y + \xi a\rangle = 0 . \tag{13.40}$$

Now

$$\Phi(y, a) = 0 \quad \text{and} \quad \Phi(a) = 0 ,$$

because $a \in K$, and

$$\langle \overset{*}{a}, y\rangle = 0 ,$$

because $y \in T_0$. Hence, the equation (13.40) reduces to the following *normal-form:*

$$\Phi(y) + 2\xi = 0 . \tag{13.41}$$

Since a is contained in the nullspace of Φ it follows from the decomposition (13.39) that the restriction of Φ to the tangent-space T_0 has again the rank r and the index s. Therefore we can select a basis x_ν $(\nu = 1 \ldots n-1)$ of T_0 such that

$$\Phi(x_\nu, x_\mu) = \varepsilon_\nu \delta_{\nu\mu} \quad (\nu, \mu = 1 \ldots n-1) .$$

Then the vectors x_ν $(\nu = 1 \ldots n-1)$ and a form a basis of E in which the normal form (13.41) can be written as

$$\sum_{\nu=1}^{r} \varepsilon_\nu \xi^\nu \xi^\nu + 2\xi = 0 . \tag{13.42}$$

Problems: 1. Let E be a 3-dimensional pseudo-Euclidean space with the index 2. Given an orientation in E define the exterior product $[x, y]$ by

$$([x, y], z) = \Delta(x, y, z) \qquad x, y, z \in E$$

where Δ is a normed determinant-function (cf. sec. 12.20) which represents the orientation. Consider a point $x_0 \neq 0$ of the light-cone $(x, x) = 0$ and a plane

$$F: x = x_0 + \xi a + \eta b$$

which does not contain the point O. Prove that the intersection of the plane F and the light-cone is

an *ellipse* if $[a, b]$ is time-like

a *hyperbola* if $[a, b]$ is space-like

a *parabola* if $[a, b]$ is a light-vector.

2. Consider the quadric $Q: \Phi(x) = 1$ where Φ is a non-degenerate quadratic function. Then every point $x_1 \neq 0$ defines an $(n-1)$-dimensional subspace $P(x_1)$ by the equation

$$\Phi(x, x_1) = 1 .$$

This subspace is called the *polar* of x_1. It follows from the above equation that the polar $P(x_1)$ does not contain the center O.

a) Prove that $x_2 \in P(x_1)$ if and only if $x_1 \in P(x_2)$.

b) Given an $(n-1)$-dimensional affine subspace A_1 of A which does not contain O, show that there exists exactly one point x_1 such that $A_1 \in P(x_1)$.

c) Show that $P(x_1)$ is a tangent-plane of Q if and only if $x_1 \in Q$.

3. Let x_1 be a point of the quadric $\Phi(x) = 1$. Prove that the restriction of the bilinear function Φ to the tangent-space T_{x_1} has the rank $r-1$ and the index $s-1$.

4. Let x_0 be a point of the quadric

$$\Phi(x) + 2f(x) + \alpha = 0$$

and consider the skew bilinear mapping $\omega: (E, E) \to E$ defined by

$$\omega(x, y) = g(x)\, y - g(y)\, x \qquad x, y \in E$$

where the linear function g is defined by (13.15). Show that the linear closure of the set $\omega(x, y)$ under this mapping is the tangent-space T_{x_0}.

§ 3. Affine equivalence of quadrics

13.13. Definition. Let an affine mapping $x' = \tau x + b$ of A onto itself be given. Then the image of a quadric

$$Q: \Phi(x) + 2 f(x) = \alpha$$

is the quadric Q' defined by the equation

$$Q': \Psi(x) + 2 g(x) = \beta,$$

where

$$\Psi(x) = \Phi(\tau^{-1} x), \tag{13.43}$$

$$g(x) = -\Phi(\tau^{-1} x, \tau^{-1} b) + f(\tau^{-1} x) \tag{13.44}$$

and

$$\beta = -\Phi(\tau^{-1} b) + 2 f(\tau^{-1} b) + \alpha. \tag{13.45}$$

In fact, the relations (13.43), (13.44) and (13.45) yield

$$\Psi(\tau x + b) + 2g(\tau x + b) - \beta = \Phi(x) + 2 f(x) - \alpha$$

showing that a point $x \in A$ is contained in Q if and only if the point $x' = \tau x + b$ is contained in Q'.

Two quadrics Q_1 and Q_2 are called *affine equivalent* if there exists a one-to-one affine mapping of A onto itself which carries Q_1 into Q_2. The affine equivalence induces a decomposition of all possible quadrics into affine equivalence classes. It is the purpose of this paragraph to construct a complete system of representatives of these equivalence classes.

13.14. The affine classification of quadrics is based upon the following theorem: Let E and F be two n-dimensional linear spaces and Φ and Ψ two symmetric bilinear functions in E and in F. Then there exists an isomorphism $\tau: E \to F$ with the property that

$$\Phi(x, y) = \Psi(\tau x, \tau y) \qquad x, y \in E \tag{13.46}$$

if and only if Φ and Ψ have the same rank and the same index.

To prove this assume first that the relation (13.46) holds. Select a basis $a_\nu (\nu = 1 \dots n)$ of E such that

$$\Phi(a_\nu, a_\mu) = \varepsilon_\nu \delta_{\nu\mu} \quad \varepsilon_\nu = \begin{cases} +1 \ (\nu = 1 \dots s) \\ -1 \ (\nu = s + 1 \dots r) \\ \ \ 0 \ (\nu = r + 1 \dots n). \end{cases} \tag{13.47}$$

Then the equations (13.46) and (13.47) yield

$$\Psi(\tau a_\nu, \tau a_\mu) = \Phi(a_\nu, a_\mu) = \varepsilon_\nu \delta_{\nu\mu}.$$

showing that Ψ has the rank r and the index s.

Conversely, assume that this condition is satisfied. Then there exist bases a_ν and b_ν $(\nu = 1 \dots n)$ of E and of F such that

$$\Phi(a_\nu, a_\mu) = \varepsilon_\nu \delta_{\nu\mu} \quad \text{and} \quad \Psi(b_\nu, b_\mu) = \varepsilon_\nu \delta_{\nu\mu}.$$

Define the isomorphism $\tau: E \to F$ by the equations

$$\tau a_\nu = b_\nu \qquad (\nu = 1 \dots n).$$

Then

$$\Phi(a_\nu, a_\mu) = \Psi(\tau a_\nu, \tau a_\mu) \qquad (\nu, \mu = 1 \dots n)$$

and consequently

$$\Phi(x, y) = \Psi(\tau x, \tau y) \qquad x, y \in E.$$

13.15. Affine classification. First of all it will be shown that the centers are invariant under an affine mapping. In fact, let

$$Q: \Phi(x - c) = \beta$$

be a quadric with c as center and $x' = \tau x + b$ an affine mapping of A onto itself. Then the image Q' of Q is given by the equation

$$Q': \Psi(x - c') = \beta,$$

where

$$\Psi(x) = \Phi(\tau^{-1} x)$$

and

$$c' = b + \tau c.$$

This equation shows that c' is a center of Q'.

Now consider to quadrics with center

$$Q_1: \Phi_1(x - c_1) = 1 \tag{13.48}$$

and

$$Q_2: \Phi_2(x - c_2) = 1 \tag{13.49}$$

and assume that $x \to x'$ is an affine mapping carrying Q_1 into Q_2. Since centers are transformed into centers we may assume the mapping $x \to x'$ sends c_1 into c_2 and hence it has the form

$$x' = \tau(x - c_1) + c_2.$$

By hypothesis, Q_1 is mapped onto Q_2 and hence the equation

$$\Phi_2(\tau(x - c_1)) = 1 \tag{13.50}$$

must represent the quadric Q_1. Comparing (13.48) and (13.50) and applying the uniqueness-theorem of sec. 13.9 we find that

$$\Phi_1(x) = \Phi_2(\tau x).$$

This relation implies that

$$r_1 = r_2 \quad \text{and} \quad s_1 = s_2. \tag{13.51}$$

Conversely, the relations (13.51) imply that there exists an automorphism τ of E such that

$$\Phi_1(x) = \Phi_2(\tau x).$$

Then the affine mapping $x \to x'$ defined by

$$x' = \tau(x - c_1) + c_2$$

transforms Q_1 into Q_2. We thus obtain the following criterion: *The two normal forms* (13.48) *and* (13.49) *represent affine equivalent quadrics if and only if the bilinear functions Φ_1 and Φ_2 have the same rank and the same index.*

13.16. Next, let

$$Q_1 : \Phi_1(x - q_1) + 2\langle \overset{*}{a}_1, x - q_1 \rangle = 0 \quad q_1 \in Q_1 \tag{13.52}$$

and

$$Q_2 : \Phi_2(x - q_2) + 2\langle \overset{*}{a}_2, x - q_2 \rangle = 0 \quad q_2 \in Q_2 \tag{13.53}$$

be two quadrics without a center. It is assumed that the equations (13.52) and (13.53) are written in such a way that $2s_1 \geqq r_1$ and $2s_2 \geqq r_2$. If $x' = \tau(x - q_1) + q_2$ is an affine mapping transforming Q_1 into Q_2, the equation of Q_1 can be written in the form

$$\Phi_2(\tau(x - q_1)) + 2(\overset{*}{a}_2, \tau(x - q_1)) = 0.$$

Now the uniqueness-theorem yields

$$\Phi_1(x) = \lambda \Phi_2(\tau x)$$

where $\lambda \neq 0$ is a constant. This relation implies that the bilinear functions Φ_1 and Φ_2 have the same rank r and that $s_2 = s_1$ or $s_2 = r - s_1$ depending on whether $\lambda > 0$ or $\lambda < 0$. But the equation $s_2 = r - s_1$ is only compatible with the inequalities $2s_1 \geqq r_1$ and $2s_2 \geqq r_2$ if $s_1 = s_2 = \frac{r}{2}$ and hence we see that $s_1 = s_2$ in either case.

Conversely, assume that $r_1 = r_2 = r$ and $s_1 = s_2 = s$. To find an affine mapping which transforms Q_1 into Q_2 consider the tangent-spaces $T_{q_1}(Q_1)$ and $T_{q_2}(Q_2)$. As it has been mentioned in sec. 13.12 the restriction of Φ_i to the subspace $T_{q_i}(Q_i)$ $(i = 1, 2)$ has the same rank and the same index as Φ_i. Consequently, there exists an isomorphism $\varrho : T_{q_1}(Q_1) \to T_{q_2}(Q_2)$ such that

$$\Phi_1(y) = \Phi_2(\varrho y) \qquad y \in T_{q_1}(Q_1).$$

Now select a vector a_i in the nullspace of Φ_i $(i = 1, 2)$ such that

$$\langle \overset{*}{a}_i, a_i \rangle = 1 \qquad (i = 1, 2)$$

and define the automorphism τ of E by the equations

$$\tau y = \varrho y \qquad y \in T_{q_1}(Q_1)$$

and

$$\tau a_1 = a_2. \tag{13.54}$$

Then

$$\Phi_2(\tau x) + \langle \overset{*}{a}_2, \tau x \rangle = \Phi_1(x) + \langle \overset{*}{a}_1, x \rangle \qquad x \in E\,. \tag{13.55}$$

In fact, every vector $x \in E$ can be decomposed in the form

$$x = y + \xi a_1 \qquad y \in T_{q_1}(Q_1)\,. \tag{13.56}$$

The equations (13.54), (13.55) and (13.56) imply that

$$\Phi_2(\tau x) = \Phi_2(\tau y + \xi a_2) = \Phi_2(\tau y) = \Phi_1(y) = \Phi_1(y + \xi a_1) = \Phi_1(x) \tag{13.57}$$

and

$$\langle \overset{*}{a}_2, \tau x \rangle = \langle \overset{*}{a}_2, \varrho y + \xi a_2 \rangle = \xi \langle \overset{*}{a}_2, a_2 \rangle = \xi = \langle \overset{*}{a}_1, x \rangle\,. \tag{13.58}$$

Adding (13.57) and (13.58) we obtain (13.55). The relation (13.55) shows that the affine mapping $x' = \tau(x - q_1) + q_2$ sends Q_1 into Q_2 and we have the following result: *The normal-forms* (13.52) *and* (13.53) *represent affine equivalent quadrics if and only if the bilinear functions Φ_1 and Φ_2 have the same rank and the same index.*

13.17. The affine classes. It follows from the two criteria in sec. 13.15, and 13.16 that the normal forms

$$\xi^1\xi^1 + \cdots + \xi^s\xi^s - \xi^{s+1}\xi^{s+1} - \cdots - \xi^r\xi^r = 1 \quad (1 \leqq s \leqq r)$$

and

$$\xi^1\xi^1 + \cdots + \xi^s\xi^s - \xi^{s+1}\xi^{s+1} - \cdots - \xi^r\xi^r + 2\xi = 0 \quad (0 \leqq 2s \leqq r)$$

form a complete system of representatives of the affine classes. Denote by $N_1(r)$ and by $N_2(r)$ the total number of affine classes with center and without center respectively of a given rank r. Then the above equations show that

$$N_1(r) = r \quad \text{and} \quad N_2(r) = \begin{cases} \dfrac{r+1}{2} \text{ if } r \text{ is odd} & \\ & 1 \leqq r \leqq n-1 \\ \dfrac{r+2}{2} \text{ if } r \text{ is even} & \\ 0 & r = n\,. \end{cases}$$

The following list contains a system of representatives of the affine classes in the plane and in the 3-space*):

Plane:

I. Quadrics with center:

1. $r = 2$: a) $s = 2$: $\xi^2 + \eta^2 = 1$ ellipse,

   b) $s = 1$: $\xi^2 - \eta^2 = 1$ hyperbola.

2. $r = 1$: $s = 1$: $\xi = \pm 1$ two parallel lines.

---

*) In the following equations the coordinates are denoted by ξ, η, ζ and the superscripts indicate exponents.

II. Quadrics without center:

$r = 1, s = 1$: $\quad \xi^2 - 2\eta = 0 \quad$ parabola.

3-space:

I. Quadrics with center:

1. $r = 3$: a) $s = 3$: $\xi^2 + \eta^2 + \zeta^2 = 1$ ellipsoid,
   b) $s = 2$: $\xi^2 + \eta^2 - \zeta^2 = 1$ hyperboloid with one shell,
   c) $s = 1$: $\xi^2 - \eta^2 - \zeta^2 = 1$ hyperboloid with two shells.
2. $r = 2$: a) $s = 2$: $\xi^2 + \eta^2 = 1$ elliptic cylinder,
   b) $s = 1$: $\xi^2 - \eta^2 = 1$ hyperbolic cylinder.
3. $r = 1$: $s = 1$: $\xi = \pm 1$ two parallel planes.

II. Quadrics without center:

1. $r = 2$: a) $s = 2$: $\xi^2 + \eta^2 - 2\zeta = 0$ elliptic paraboloid,
   b) $s = 1$: $\xi^2 - \eta^2 - 2\zeta = 0$ hyperbolic paraboloid.
2. $r = 1, s = 1$: $\xi^2 - 2\zeta = 0$ parabolic cylinder.

Problems: 1. Let Q be a given quadric and C be a given point. Show that C is a center of Q if and only if the affine mapping $P \to P'$ defined by $\overrightarrow{CP'} = -\overrightarrow{CP}$ transforms Q into itself.

2. If Φ is an indefinite quadratic function, show that the quadrics

$$\Phi(x) = 1 \quad \text{and} \quad \Phi(x) = -1$$

are equivalent if and only if the signature of Φ is zero.

3. Denote by N_1 and by N_2 the total number of affine classes with center and without center respectively. Prove that

$$N_1 = \frac{n(n+1)}{2}$$

$$N_2 = \begin{cases} k^2 + k + 1 & \text{if} \quad n = 2k \\ k^2 + 2k & \text{if} \quad n = 2k + 1 \,. \end{cases}$$

4. Let x_1 and x_2 be two points of the quadric

$$Q : \Phi(x) = 1 \,.$$

Assume that an isomorphism τ: $T_{x_1} \to T_{x_2}$ is given such that

$$\Phi(\tau y, \tau z) = \Phi(y, z) \qquad y, z \in T_{x_1} \,.$$

Construct an affine mapping $A \to A$ which transforms Q into itself and which induces the isomorphism τ in the tangent-space T_{x_1}.

5. Prove the assertion of problem 4 for the quadric

$$\Phi(x) + 2\langle \overset{*}{a}, x\rangle = 0 \,.$$

§ 4. Quadrics in the Euclidean space

13.18. Normal-vector. Let A be an n-dimensional Euclidean space and

$$Q: \Phi(x) + 2f(x) = \alpha$$

be a quadric in A. The bilinear function Φ determines a selfadjoint endomorphism φ of E by the equation

$$\Phi(x, y) = (\varphi x, y) .$$

The linear function f can be written as

$$f(x) = (a, x)$$

where a is a fixed vector of E. Cones will again be excluded, i. e. we shall assume that

$$\Phi x \neq -a$$

for all points $x \in Q$. Let x_0 be a fixed point of Q. Then the equation (13.10) shows that the tangent-space T_{x_0} consists of all vectors y satisfying the relation

$$(\varphi x_0 + a, y) = 0 .$$

In other words, the tangent-space T_{x_0} is the orthogonal complement of the *normal-vector*

$$p(x_0) = \varphi x_0 + a .$$

The straight line determined by the point x_0 and the vector $p(x_0)$ is called the *normal* of Q at x_0.

13.19. Quadrics with center. Now consider a quadric with center

$$Q: \Phi(x) = 1 . \tag{13.59}$$

Then the normal-vector $p(x_0)$ is simply given by

$$p(x_0) = \varphi x_0 .$$

This equation shows that the linear mapping φ associates with every point $x_0 \in Q$ the corresponding normal-vector. In particular, let x_0 be a point of Q whose position-vector is an eigenvector of φ. Then we have the relation

$$\varphi x_0 = \lambda x_0$$

showing that the normal-vector is a multiple of the position-vector x_0. Inserting this into the equation (13.59) we see that the corresponding eigenvalue is equal to

$$\lambda = \frac{1}{|x_0|^2} .$$

As it has been shown in sec. 11.7 there exists an orthonormal system of n eigenvectors e_ν $(\nu = 1 \ldots n)$. Then

$$\varphi e_\nu = \lambda_\nu e_\nu \qquad (\nu = 1 \ldots n) \tag{13.60}$$

whence

$$\Phi(e_\nu, e_\mu) = \lambda_\nu \delta_{\nu\mu} .$$

Let us enumerate the eigenvectors e_ν such that

$$\begin{gathered} 0 < \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_s \\ 0 > \lambda_{s+1} \geqq \lambda_{s+2} \geqq \cdots \geqq \lambda_r \\ \lambda_{r+1} = \cdots = \lambda_n = 0 \end{gathered} \tag{13.61}$$

where r is the rank and s is the index of Φ. Then the equation (13.59) can be written as

$$\sum_{\nu=1}^{r} \lambda_\nu \xi^\nu \xi^\nu = 1 . \tag{13.62}$$

The vectors

$$a_\nu = \frac{e_\nu}{\sqrt{\lambda_\nu}} \quad (\nu = 1 \ldots s) \quad \text{and} \quad a_\nu = \frac{e_\nu}{\sqrt{-\lambda_\nu}} \quad (\nu = s + 1 \ldots r)$$

are called the *principal axes* and the *conjugate principal axes* of Q. Inserting

$$\lambda_\nu = \frac{e_\nu}{|a_\nu|^2} \quad (\nu = 1 \ldots s) \quad \text{and} \quad \lambda_\nu = -\frac{e_\nu}{|a_\nu|^2} \quad (\nu = s + 1 \ldots r)$$

into (13.62) we obtain the *metric normal-form* of Q:

$$\sum_{\nu=1}^{s} \frac{\xi^\nu \xi^\nu}{|a_\nu|^2} - \sum_{\nu=s+1}^{r} \frac{\xi^\nu \xi^\nu}{|a_\nu|^2} = 1 . \tag{13.63}$$

Every principal axis a_ν generates a straight line which intersects the quadric Q in the points a_ν and $-a_\nu$. The straight lines generated by the conjugate axes have no points in common with Q but they intersect the conjugate quadric

$$Q' : \Phi(x) = -1$$

at the points a_ν and $-a_\nu$ $(\nu = s + 1 \ldots r)$.

13.20. Quadrics without center. Now consider a quadric Q without centers. Using an arbitrary point of Q as origin, we can write the equation of Q in the form

$$\Phi(x) + 2(a, x) = 0 \tag{13.64}$$

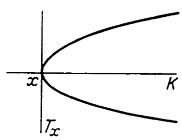


Fig. 6

where a is a normal vector of Q at the point $x = 0$. For every point $x \in Q$ the vector

$$p(x) = \varphi x + a \tag{13.65}$$

is contained in the normal of Q. A point $x \in Q$ is called a *vertex* if the corresponding normal is contained in the null-space K of Φ (cf. fig. 6).

It will be shown that every quadric without centers has at least one vertex.

Applying φ to the equation (13.65) we obtain

$$\varphi p(x) = \varphi^2 x + \varphi a$$

showing that a point $x \in Q$ is a vertex if and only if

$$\varphi^2 x = -\varphi a . \tag{13.66}$$

To find all vertices of Q we thus have to determine all the solutions of the equations (13.64) and (13.66). The self-adjointness of φ implies that the mappings φ and φ^2 have the same image-space and the same kernel (cf. sec. 11.4). Consequently, the equation (13.66) has at least one solution. The general solution of (13.66) can be written in the form

$$x = x_0 + z$$

where z is an arbitrary vector of the nullspace K. Inserting this into the equation (13.64) we obtain

$$\tfrac{1}{2}\Phi(x_0) + (a, x_0) + (a, z) = 0 . \tag{13.67}$$

Now $a \notin K^{\perp}$ (otherwise Q would have a center) and consequently (13.67) has a solution $z \in K$. This solution is determined up to an arbitrary vector of the intersection $K \cap T_0$. In other words, the set of all vertices of Q forms an affine subspace with the difference-space $K \cap T_0$. This subspace has the dimension $(n - r - 1)$.

Now we are ready to construct the normal form of the quadric (13.64). First of all we select a vertex of Q as origin. Then the vector a in (13.64) is contained in the kernel K. Multiplying the equation (13.64) by an appropriate scalar we can achieve that $|a| = 1$ and that $2s \geqq r$. Now let e_ν $(\nu = 1 \ldots n-1)$ be a basis of T_0 consisting of eigenvectors of Φ. Then the vectors e_ν $(\nu = 1 \ldots n-1)$ and a form an orthonormal basis of E such that

$$\Phi(e_\nu, e_\mu) = \lambda_\nu \delta_{\nu\mu} \qquad (\nu, \mu = 1 \ldots n-1)$$

and

$$\Phi(e_\nu, a) = (e_\nu, \varphi a) = 0 \qquad (\nu = 1 \ldots n-1) .$$

In this basis the equation of Q assumes the *metric normal-form*

$$\sum_{\nu=1}^{r} \lambda_\nu \xi^\nu \xi^\nu + 2\xi = 0 . \tag{13.68}$$

Upon introduction of the principal axes and the principal conjugate axes

$$a_\nu = \frac{e_\nu}{\sqrt{\lambda_\nu}} \quad (\nu = 1 \ldots s) \quad \text{and} \quad a_\nu = \frac{e_\nu}{\sqrt{-\lambda_\nu}} \quad (\nu = s+1 \ldots r)$$

the normal-form (13.68) can be also written as

$$\sum_{\nu=1}^{s} \frac{\xi^\nu \xi^\nu}{|a_\nu|^2} - \sum_{\nu=s+1}^{r} \frac{\xi^\nu \xi^\nu}{|a_\nu|^2} + 2\xi = 0 . \tag{13.69}$$

13.21. Metric classification of bilinear forms. Two quadrics Q and Q' in the Euclidean space A are called *metrically equivalent*, if there exists a rigid motion $x \to x'$ which transforms Q into Q'. Two metrically equivalent quadrics are a fortiori affine equivalent. Hence, the metric classification of quadrics consists in the construction of the metric subclasses within every affine equivalence-class.

It will be shown that the lengths of the principal axes form a complete system of metric invariants. In other words, *two affine equivalent quadrics Q and Q' are metrically equivalent if and only if the principal axes of Q and Q' respectively have the same length.*

We prove first the following criterion: Let E and F be two n-dimensional Euclidean spaces and consider two symmetric bilinear functions Φ and Ψ having the same rank and the same index. Then there exists an isometric mapping $\tau: E \to F$ such that

$$\Phi(x, y) = \Psi(\tau x, \tau y) \qquad x, y \in E \tag{13.70}$$

if and only if Φ and Ψ have the same eigenvalues.

Define the endomorphisms $\varphi: E \to E$ and $\psi: F \to F$ by the equations

$$\Phi(x, y) = (\varphi x, y) \quad x, y \in E \quad \text{and} \quad \Psi(x, y) = (\psi x, y) \quad x, y \in F .$$

Then the eigenvalues of Φ and Ψ are equal to the eigenvalues of φ and ψ respectively (cf. sec. 11.10).

Now assume that τ is an isometric mapping of E onto F such that the relation (13.70) holds. Then

$$(\varphi x, y) = (\psi \tau x, \tau y) \tag{13.71}$$

whence

$$\varphi = \tau^{-1} \circ \psi \circ \tau .$$

This relation implies that the endomorphisms φ and ψ have the same eigenvalues.

Conversely, assume that φ and ψ have the same eigenvalues. Then there is an orthonormal basis a_ν in E and an orthonormal basis b_ν $(\nu = 1 \dots n)$ in F such that

$$\varphi a_\nu = \lambda_\nu a_\nu \quad \text{and} \quad \psi b_\nu = \lambda_\nu b_\nu \qquad (\nu = 1 \dots n) . \tag{13.72}$$

Hence, an isometric mapping $\tau: E \to F$ is defined by the equations

$$\tau a_\nu = b_\nu \qquad (\nu = 1 \dots n) . \tag{13.73}$$

The equations (13.72) and (13.73) imply that

$$(\varphi a_\nu, a_\mu) = (\psi \tau a_\nu, \tau a_\mu) \qquad (\nu, \mu = 1 \dots n)$$

whence (13.71).

13.22. Metric classification of quadrics. Consider first two quadrics Q and Q' with center. Since a translation does not change the principal axis we may assume that Q and Q' have the common center O. Then the equations of Q and Q' read

$$Q:\Phi(x) = 1$$

and

$$Q':\Phi'(x) = 1 .$$

Now assume that there exists a rotation of E carrying Q into Q'. Then

$$\Phi(x) = \Phi'(\tau x) \qquad x \in E . \tag{13.74}$$

It follows from the criterion in sec. 13.21 that the bilinear functions Φ and Φ' have the same eigenvalues. This implies that the principal axes of Q and Q' have the same length,

$$|a_\nu| = |a'_\nu| \qquad (\nu = 1 \ldots r) . \tag{13.75}$$

Conversely, assume the relations (13.75). Then

$$|\lambda_\nu| = |\lambda'_\nu| \qquad (\nu = 1 \ldots n) .$$

Observing the conditions (13.61) we see that $\lambda_\nu = \lambda'_\nu$ $(\nu = 1 \ldots n)$. According to the criterion in sec. 13.21 there exists a rotation τ of E such that

$$\Phi(x) = \Phi'(\tau x) .$$

This rotation obviously transforms Q into Q'.

Now let Q and Q' be two quadrics without center. Without loss of generality we may assume that Q and Q' have the common vertex O. Then the equations of Q and Q' are

$$Q:\Phi(x) + 2(a, x) = 0 \quad a \in K \quad |a| = 1 , \tag{13.76}$$

and

$$Q':\Phi'(x) + 2(a', x) = 0 \quad a' \in K \quad |a'| = 1 , \tag{13.77}$$

If Q and Q' are metrically equivalent there exists a rotation τ such that

$$\Phi(x) = \Phi'(\tau x) .$$

Then

$$|a_\nu| = |a'_\nu| \qquad (\nu = 1 \ldots r) . \tag{13.78}$$

Conversely, the equations (13.78) imply that the bilinear functions Φ and Φ' have the same eigenvalues,

$$\lambda_\nu = \lambda'_\nu \qquad (\nu = 1 \ldots n) . \tag{13.79}$$

Now consider the restriction Ψ of Φ to the subspace $T_0(Q)$. Then every eigenvalue of Ψ is also an eigenvalue of Φ. In fact, assume that

$$\Psi(e, y) = \lambda(e, y)$$

for a fixed vector $e \in T_0(Q)$ and all vectors $y \in T_0(Q)$. Then

$$\Phi(e, x) = \Phi(e, \xi a + y) = \xi\Phi(e, a) + \Psi(e, y) = \xi\Phi(e, a) + \lambda(e, y) \tag{13.80}$$

for an arbitrary vector $x \in E$. Since the point O is a vertex of Q we have the equation $\Phi(e, a) = 0$. We thus obtain from (13.80) the relation

$$\Phi(e, x) = \lambda(e, y) = \lambda(e, \xi a + y) = \lambda(e, x)$$

showing that λ is an eigenvalue of Φ. Hence we see that the bilinear function Ψ has the eigenvalues $\lambda_1 \ldots \lambda_{n-1}$. In the same way we see that the restriction Ψ' of Φ' to the subspace $T_0(Q')$ has the eigenvalues $\lambda'_1 \ldots \lambda'_{n-1}$. Now it follows from (13.79) and the criterion in sec. 13.21 that there exists an isometric mapping

$$\varrho: T_0(Q) \to T_0(Q')$$

with the property that

$$\Phi'(\varrho y) = \Phi(y) \qquad y \in T_0(Q)\,.$$

Define the rotation τ of E by the equations

$$\tau y = \varrho y \qquad y \in T_0(Q)$$

$$\tau a = a'\,.$$

Then

$$\Phi'(\tau x) + 2(a', \tau x) = \Phi(x) + 2(a, x) \qquad x \in E$$

and consequently, τ transforms Q into Q'.

13.23. The metric normal-forms in the plane and in 3-space. The equations (13.63) and (13.69) yield the following metric normal forms in the dimensions $n = 2$ and $n = 3$:

Plane:

I. Quadrics with center:

1. $\frac{\xi^2}{a^2} + \frac{\eta^2}{b^2} = 1\,, a \geqq b$ — ellipse with the axes a and b.
2. $\frac{\xi^2}{a^2} - \frac{\eta^2}{b^2} = 1$ — hyperbola with the axes a and b.
3. $\xi = \pm a$ — two parallel lines with the distance $2a$.

II. Quadrics without center:

$\frac{\xi^2}{a^2} = 2\eta$ — parabola with latus rectum of length a.

3-space:

I. Quadrics with center:

1. $\frac{\xi^2}{a^2}+\frac{\eta^2}{b^2}+\frac{\zeta^2}{c^2}=1\,,\ a \geqq b \geqq c$ — ellipsoid with axes a, b, c.

2. $\frac{\xi^2}{a^2}+\frac{\eta^2}{b^2}-\frac{\zeta^2}{c^2}=1\,,\ a \geqq b$ — hyperboloid with one shell and axes a, b, c.

3. $\frac{\xi^2}{a^2}-\frac{\eta^2}{b^2}-\frac{\zeta^2}{c^2}=1\,,\ b \geqq c$ — hyperboloid with two shells and axes a, b, c.

4. $\frac{\xi^2}{a^2}+\frac{\eta^2}{b^2}=1,\ a \geqq b$ — elliptic cylinder with the axes a and b.

5. $\frac{\xi^2}{a^2}-\frac{\eta^2}{b^2}=1$ — hyperbolic cylinder with the axes a and b.

6. $\xi = \pm a$ — two parallel planes with the distance $2a$.

II. Quadrics without center:

1. $\frac{\xi^2}{a^2}-\frac{\eta^2}{b^2}=2\zeta\,,\ a \geqq b$ — elliptic paraboloid with axes a and b.

2. $\frac{\xi^2}{a^2}-\frac{\eta^2}{b^2}=2\zeta$ — hyperbolic paraboloid with axes a and b.

3. $\frac{\xi^2}{a^2}=2\zeta$ — parabolic cylinder with latus rectum of length a.

Problems: 1. Give the center or vertex, the type and the axes of the following quadrics in the 3-space:

a) $2\xi^2+2\eta^2-\zeta^2+8\,\xi\eta-4\,\xi\zeta-4\,\eta\zeta=2\,.$

b) $4\,\xi^2+3\,\eta^2-\zeta^2-12\,\xi\eta+4\,\xi\zeta-8\,\eta\zeta=1\,.$

c) $\xi^2+\eta^2+7\,\zeta^2-16\,\xi\eta-8\,\xi\zeta-8\,\eta\zeta=9\,.$

d) $3\,\xi^2+3\,\eta^2+\zeta^2-2\,\xi\eta+6\,\xi-2\,\eta-2\,\zeta+3=0\,.$

2. Given a non-degenerate quadratic function Φ, consider the family (Q_α) of quadrics defined by

$$\Phi(x)=\alpha \qquad\qquad (\alpha \neq 0)\,.$$

Show that every point $x \neq 0$ is contained in exactly one quadric Q_α. Prove that the endomorphism φ of E defined by

$$\Phi\,(x, y) = (\varphi x, y)$$

associates with every point $x \neq 0$ the normal vector of the quadric passing through x.

3. Consider the quadric

$$Q: \Phi(x) = 1\,,$$

where Φ is a non-degenerate bilinear function. Denote by Q' the image of Q under the mapping φ which corresponds to Φ. Prove that the principal axes of Q' and Q are connected by the relation

$$a'_\nu = \frac{a_\nu}{|a_\nu|^2} \qquad (\nu = 1 \ldots n)\,.$$

4. Given two points p, q and a number 2α ($\alpha > |p - q|$), consider the locus Q of all points x such that

$$|x - p| + |x - q| = 2\alpha\,.$$

Prove that Q is a quadric of index n whose principal axes have the length

$$|a_1| = \alpha, \qquad |a_\nu| = \sqrt{\alpha^2 - \frac{1}{4}|p - q|^2} \quad (\nu = 2 \ldots n)\,.$$

5. Let $\Phi(x) = 1$ be the equation of a non-degenerate quadric Q with the property that x is a normal vector at every point of Q. Prove that Q is a sphere.

Chapter XIV

Unitary Spaces

§ 1. Hermitian functions

14.1. Sesquilinear functions in a complex space. Let E be an n-dimensional complex linear space and Φ be a function of two vectors which is linear in the first argument and "conjugate linear" in the second argument, i. e.

$$\begin{aligned} \Phi(\lambda x_1 + \mu x_2, y) &= \lambda\Phi(x_1, y) + \mu\Phi(x_2, y) \\ \Phi(x, \lambda y_1 + \mu y_2) &= \bar\lambda\Phi(x, y_1) + \bar\mu\Phi(x, y_2) \end{aligned} \tag{14.1}$$

where $\bar\lambda$ and $\bar\mu$ are the complex conjugate coefficients. Then Φ will be called a *sesquilinear function*. Replacing y by x we obtain from Φ the corresponding quadratic function

$$\Psi(x) = \Phi(x, x)\,. \tag{14.2}$$

It follows from (14.1) that Ψ satisfies the relations

$$\Psi(x + y) + \Psi(x - y) = 2(\Psi(x) + \Psi(y)) \tag{14.3}$$

and

$$\Psi(\lambda x) = |\lambda|^2\,\Psi(x)\,.$$

Conversely, the function Φ can be expressed in terms of Ψ. In fact, the equation (14.2) yields

$$\Psi(x+y) = \Psi(x) + \Psi(y) + \Phi(x, y) + \Phi(y, x)\,. \tag{14.4}$$

Replacing y by iy we obtain

$$\Psi(x+iy) = \Psi(x) + \Psi(y) - i\Phi(x, y) + i\Phi(y, x)\,. \tag{14.5}$$

Multiplying (14.5) by i and adding it to (14.4) we find

$$\Psi(x+y) + i\Psi(x+iy) = (1+i)(\Psi(x) + \Psi(y)) + 2\Phi(x, y)\,,$$

whence

$$2\Phi(x,y) = \{\Psi(x+y) - \Psi(x) - \Psi(y)\} + i\{\Psi(x+iy) - \Psi(x) - \Psi(y)\}\,. \tag{14.6}$$

Note: The fact that Φ is uniquely determined by the function Ψ is due to the sesquilinearity. We recall that a bilinear function has to by *symmetric* in order to be uniquely determined by the corresponding quadratic function.

14.2. Hermitian functions. With every sesquilinear function Φ we can associate another sesquilinear function $\tilde{\Phi}$ by the equation

$$\tilde{\Phi}(x, y) = \overline{\Phi(y, x)}\,.$$

A sesquilinear function Φ is called *Hermitian* if $\tilde{\Phi} = \Phi$, i. e.

$$\Phi(x, y) = \overline{\Phi(y, x)}\,. \tag{14.7}$$

Inserting $y = x$ in the relation (14.7) we find that

$$\Psi(x) = \overline{\Psi(x)} \tag{14.8}$$

Hence the quadratic function Ψ is real valued. Conversely, a sesquilinear function Φ whose quadratic function is real valued is Hermitian. In fact, if Ψ is real valued, both parantheses in (14.6) are real. Interchange of x and y yields

$$2\Phi(y, x) = \{\Psi(x+y) - \Psi(x) - \Psi(y)\} + i\{\Psi(y+ix) - \Psi(x) - \Psi(y)\}. \tag{14.9}$$

Comparison of (14.6) and (14.9) shows that the real parts coincide. The sum of the imaginary parts is equal to

$$\Psi(x+iy) + \Psi(y+ix) - 2\Psi(x) - 2\Psi(y)\,.$$

Replacing y by iy in the relation (14.3) we see that this is equal to zero, whence

$$\Phi(y, x) = \overline{\Phi(x, y)}\,.$$

A Hermitian function Φ is called *positive definite,* if $\Psi(x) > 0$ for all vectors $x \neq 0$.

14.3. Hermitian matrices. Let x_ν $(\nu = 1 \ldots n)$ be a basis of E. Then every sesquilinear function Φ defines a complex $n \times n$-matrix

$$\alpha_{\nu\mu} = \Phi(x_\nu, x_\mu)\,.$$

The function Φ is uniquely determined by the matrix $(\alpha_{\nu\mu})$. In fact, if

$$x = \sum_\nu \xi^\nu x_\nu \quad \text{and} \quad y = \sum_\nu \eta^\nu x_\nu$$

are two arbitrary vectors, we see that

$$\Phi(x, y) = \sum_{\nu,\mu} \alpha_{\nu\mu}\, \xi^\nu \bar{\eta}^\mu\,.$$

The matrices $(\alpha_{\nu\mu})$ and $(\tilde{\alpha}_{\nu\mu})$ of Φ and $\tilde{\Phi}$ are obviously connected by the relation

$$\tilde{\alpha}_{\nu\mu} = \bar{\alpha}_{\mu\nu}\,.$$

If Φ is a Hermitian function it follows that

$$\alpha_{\nu\mu} = \bar{\alpha}_{\mu\nu}\,.$$

A complex $n \times n$-matrix satisfying this relation is called a *Hermitian matrix.*

Problems: 1. Prove that a skew-symmetric sesquilinear function is identically zero.

2. Show that the decomposition constructed in sec. 12.7. can be carried over to Hermitian functions.

§ 2. Unitary spaces

14.4. Definition. *A unitary space* is a complex linear space E in which a positive definite Hermitian function, denoted by (x, y), is distinguished. The number (x, y) is called the *inner product* of the vectors x and y. It has the following properties:

1. $(\lambda x_1 + \mu x_2, y) = \lambda(x_1, y) + \mu(x_2, y)$

   $(x, \lambda y_1 + \mu y_2) = \bar{\lambda}(x, y_1) + \bar{\mu}(x, y_2)$.

2. $(x, y) = \overline{(y, x)}$.

3. $(x, x) > 0$ for all vectors $x \neq 0$.

In a similar way as for real linear spaces, the *standard-inner product* in the complex number-space C^n is defined by

$$(x, y) = \sum_\nu \xi^\nu \bar{\eta}^\nu \quad \text{where } x = (\xi^1 \dots \xi^n) \quad \text{and} \quad y = (\eta^1 \dots \eta^n)\,.$$

The *norm* of a vector x of a unitary space is defined as the positive square-root

$$|x| = \sqrt{(x, x)}\,.$$

The Schwarz-inequality

$$|(x, y)| \leqq |x|\,|y| \tag{14.10}$$

is proved in the same way as for real inner product spaces. Equality holds if and only if the vectors x and y are linearly dependent.

From (14.10) we obtain the triangle-inequality

$$|x + y| \leqq |x| + |y|\,.$$

The equality sign holds if and only if $y = \lambda x$ where λ is real and non-negative. In fact, assume that

$$|x + y| = |x| + |y|\,. \tag{14.11}$$

Squaring this equation we obtain

$$(x, y) + \overline{(x, y)} = 2\,|x|\,|y|\,. \tag{14.12}$$

This can be written as

$$\mathrm{Re}\,(x, y) = |x|\,|y|$$

where Re denotes the real part. The above relation yields

$$|(x, y)| = |x|\,|y|$$

and hence it implies that the vectors x and y are linearly dependent, $y = \lambda x$. Inserting this into (14.11) we obtain

$$\lambda + \bar{\lambda} = 2\,|\lambda|$$

whence

$$\mathrm{Re}\,\lambda = |\lambda|\,.$$

Hence, λ is real and non-negative. Conversely, it is clear that

$$|(1 + \lambda)\,x| = |x| + \lambda\,|x|$$

for every real, non-negative number λ.

Two vectors $x \in E$ and $y \in E$ are called *orthogonal*, if

$$(x, y) = 0\,.$$

Every subspace $E_1 \subset E$ determines an orthogonal complement $E_1^\perp$ consisting of all vectors which are orthogonal to E_1. The spaces E_1 and $E_1^\perp$ form a direct decomposition of E:

$$E = E_1 \oplus E_1^\perp\,.$$

14.5. Orthonormal bases. A basis x_ν $(\nu = 1 \ldots n)$ of E is called *orthonormal*, if

$$(x_\nu, x_\mu) = \delta_{\nu\mu} .$$

Then the inner product of two vectors

$$x = \sum_\nu \xi^\nu x_\nu \quad \text{and} \quad y = \sum_\nu \eta^\nu x_\nu$$

is then given by

$$(x, y) = \sum_\nu \xi^\nu \bar{\eta}^\nu .$$

Replacing y by x we obtain

$$|x|^2 = \sum_\nu \xi^\nu \bar{\xi}^\nu .$$

Orthogonal bases can be constructed in the same way as in a real inner product space by the Schmidt-orthogonalization process.

Consider two orthonormal bases x_ν and $\bar{x}_\nu$ $(\nu = 1 \ldots n)$. Then the matrix (α_ν^μ) of the basis-transformation $x_\nu \to \bar{x}_\nu$ satisfies the relations

$$\sum_\mu \alpha_\nu^\mu \bar{\alpha}_\lambda^\mu = \delta_{\nu\lambda} .$$

A complex matrix of this kind is called a *unitary matrix*. Conversely, if an orthonormal basis x_ν and a unitary matrix (α_ν^μ) is given, the basis

$$\bar{x}_\nu = \sum_\mu \alpha_\nu^\mu x_\mu$$

is again orthonormal.

14.6. The duality in a unitary space. A *conjugation* in a complex linear space is a mapping $x \to \bar{x}$ of E into itself having the following properties:

1. $\overline{x_1 + x_2} = \bar{x}_1 + \bar{x}_2$.
2. $\overline{\lambda x} = \bar{\lambda} \bar{x}$.
3. $\bar{\bar{x}} = x$.

If an inner product is defined in E we require in addition that

$$(\bar{x}, \bar{y}) = (x, y) .$$

A conjugation can always be defined in an n-dimensional complex space. In fact, select an orthonormal basis x_ν $(\nu = 1 \ldots n)$ and define the mapping $x \to \bar{x}$ by

$$\bar{x} = \sum_\nu \bar{\xi}^\nu x_\nu$$

where

$$x = \sum_\nu \xi^\nu x_\nu .$$

Then the above conditions are obviously satisfied.

Assume that a conjugation is given in E. Then the function $\langle x, y\rangle$ defined by

$$\langle x, y\rangle = (x, \bar{y}) \tag{14.13}$$

is linear with respect to both arguments. The definiteness of the inner product implies that the function (14.13) is non-degenerate. Hence, the space E may be considered as dual to itself, relative to the scalar-product (14.13).

Now all the properties arising from the duality can be carried over to unitary spaces. The Riesz-theorem asserts that every linear function f in E can be represented in the form

$$f(x) = (x, a)$$

where a is a vector of E which is uniquely determined by f. In fact, there exists a unique vector $b \in E$ such that

$$f(x) = \langle x, b\rangle .$$

Then $a = \bar{b}$.

14.7. Normed determinant-functions. Assuming that a conjugation is defined in E, let $\Delta_0 \neq 0$ be a determinant-function in E. Then the function $\bar{\Delta}_0$ defined by

$$\bar{\Delta}_0 (x_1 \ldots x_n) = \overline{\Delta_0 (\bar{x}_1 \ldots \bar{x}_n)}$$

is obviously again a determinant-function. It will be called the *conjugate determinant-function*. Application of the identity (4.24) to the spaces E and $E^* = E$ yields

$$\begin{vmatrix} \langle x_1, y_1\rangle & \ldots & \langle x_1, y_n\rangle \\ \vdots & & \vdots \\ \langle x_n, y_1\rangle & \ldots & \langle x_n, y_n\rangle \end{vmatrix} = \alpha \Delta_0 (x_1 \ldots x_n)\, \bar{\Delta}_0 (y_1 \ldots y_n)$$

where α is a constant. Replacing the vectors y_ν by $\bar{y}_\nu$ and observing that

$$\langle x_\nu, \bar{y}_\mu\rangle = (x_\nu, y_\mu)$$

we find the relation

$$\begin{vmatrix} (x_1, y_1) & \ldots & (x_1, y_n) \\ \vdots & & \vdots \\ (x_n, y_1) & \ldots & (x_n, y_n) \end{vmatrix} = \alpha \Delta_0 (x_1 \ldots x_n)\, \bar{\Delta}_0 (\bar{y}_1 \ldots \bar{y}_n) \tag{14.14}$$

$$= \alpha \Delta_0 (x_1 \ldots x_n)\, \overline{\Delta_0 (y_1 \ldots y_n)} .$$

Now let e_ν $(\nu = 1 \ldots n)$ be an orthonormal basis of E. Inserting $x_\nu = y_\nu = e_\nu$ in (14.14) we obtain the relation

$$\alpha\, |\Delta_0 (e_1 \ldots e_n)|^2 = 1$$

showing that α is real and positive. Let λ be a complex number such that

$|\lambda|^2 = \alpha$ and define a new determinant-function Δ by $\Delta = \lambda \Delta_0$. Then the identity (14.14) reads

$$\begin{vmatrix} (x_1, y_1) & \dots & (x_1, y_n) \\ \vdots & & \vdots \\ (x_n, y_1) & \dots & (x_n, y_n) \end{vmatrix} = \Delta(x_1 \dots x_n)\,\overline{\Delta(y_1 \dots y_n)}\,. \tag{14.15}$$

Every determinant-function satisfying the relation (14.15) is called a *normed determinant-function.* The normed determinant-functions in E are uniquely determined up to a complex factor of absolute value one.

The equation (14.15) shows that a normed determinant-function assumes the absolute value 1 on every orthonormal basis.

Problems: 1. Prove that the *Gram determinant*

$$G(x_1 \dots x_p) = \begin{vmatrix} (x_1, x_1) & \dots & (x_1, x_p) \\ \vdots & & \vdots \\ (x_p, x_1) & \dots & (x_p, x_p) \end{vmatrix}$$

of p vectors of a unitary space is real and non negative. Show that $G(x_1 \dots x_p) = 0$ if and only if the vectors x_ν are linearly dependent.

2. Assume that a conjugation $x \to \bar{x}$ is defined in the n-dimensional complex linear space E. A vector $z \in E$ is called *real* relative to the given conjugation, if $\bar{z} = z$.

a) Prove that the real vectors form a real n-dimensional linear space $R(E)$.

b) Show that every vector $z \in E$ can be written in exactly one way as $z = x + iy$ where the vectors x and y are real.

c) If E is a unitary space, prove that a real positive definite inner product is induced in $R(E)$ by the inner product in E.

3. Let E be a complex linear space and $z \to \bar{z}$ be a conjugation. Define the space of real vectors as in problem 2. If an inner product is defined in $R(E)$, prove that an inner product is defined in E by

$$(z_1, z_2) = (x_1, x_2) + (y_1, y_2) + i\,((x_1, y_2) - (x_2, y_1))$$

and that

$$\overline{(z_1, z_2)} = (\bar{z}_1, \bar{z}_2)\,.$$

§ 3. Linear mappings of unitary spaces

14.8. The adjoint mapping. Let E and F be two unitary spaces and $\varphi: E \to F$ a linear mapping of E into F. Similarly as in the real case we can associate with φ an adjoint mapping $\tilde{\varphi}$ of F into E. Let $x \to \bar{x}$ and $y \to \bar{y}$ be a conjugation in E and in F, respectively. Then E and F are dual to themselves and hence φ determines a dual mapping $\varphi^*: F \to E$

by the relation

$$\langle \varphi x, y\rangle = \langle x, \varphi^* y\rangle . \tag{14.16}$$

Replacing y by $\bar{y}$ in (14.16) we obtain

$$\langle \varphi x, \bar{y}\rangle = \langle x, \varphi^* \bar{y}\rangle . \tag{14.17}$$

Observing the relation (14.13) between the inner product and the scalar product we can write (14.17) in the form

$$(\varphi x, y) = (x, \overline{\varphi^* \bar{y}}) . \tag{14.18}$$

Now define the mapping $\tilde{\varphi}: F \to E$ by

$$\tilde{\varphi} y = \overline{\varphi^* \bar{y}} . \tag{14.19}$$

Then the relation (14.18) reads

$$(\varphi x, y) = (x, \tilde{\varphi} y) \qquad x \in E , \quad y \in F . \tag{14.20}$$

The mapping $\tilde{\varphi}$ does not depend on the conjugations in E and F. In fact, assume that $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$ are two linear mappings of F into E satisfying the relation (14.20). Then

$$(x, (\tilde{\varphi}_2 - \tilde{\varphi}_1)\, y) = 0 .$$

This equation holds for every fixed $y \in F$ and all vectors $x \in E$ and hence it implies that $\tilde{\varphi}_2 = \tilde{\varphi}_1$. The mapping $\tilde{\varphi}$ is called the *adjoint* of the mapping φ.

It follows from the equation (14.18) that the relations

$$\widetilde{\varphi + \psi} = \tilde{\varphi} + \tilde{\psi} \quad \text{and} \quad \widetilde{\lambda \varphi} = \bar{\lambda} \tilde{\varphi}$$

hold for any two linear mappings and for every complex coefficient λ.

The equation (14.20) implies that the matrices of φ and $\tilde{\varphi}$ relative to two orthonormal bases of E and F are connected by the relation

$$\tilde{\alpha}^{\nu}_{\mu} = \bar{\alpha}^{\mu}_{\nu} \qquad (\nu = 1 \ldots n, \lambda = 1 \ldots m) .$$

Now consider the case $E = F$. Then the determinants of φ and $\tilde{\varphi}$ are complex conjugates. To prove this, let $\Delta \neq 0$ be a determinant-function in E and $\bar{\Delta}$ be the conjugate determinant-function. Then it follows from the definition of $\bar{\Delta}$ that

$$\begin{aligned}\bar{\Delta}\,(\tilde{\varphi} x_1 \ldots \tilde{\varphi} x_n) &= \overline{\Delta\,(\overline{\tilde{\varphi} x_1} \ldots \overline{\tilde{\varphi} x_n})} = \overline{\Delta\,(\varphi^* \bar{x}_1 \ldots \varphi^* \bar{x}_n)} \\ &= \overline{\det \varphi}\, \overline{\Delta\,(\bar{x}_1 \ldots \bar{x}_p)} = \overline{\det \varphi}\, \bar{\Delta}\,(x_1 \ldots x_n) .\end{aligned}$$

This equation implies that

$$\det \tilde{\varphi} = \overline{\det \varphi}\,. \tag{14.21}$$

If φ is replaced by $\varphi - \lambda\,\iota$, where λ is a complex parameter, the relation (14.21) yields

$$\det(\tilde{\varphi} - \lambda\,\iota) = \overline{\det(\varphi - \bar{\lambda}\,\iota)}\,.$$

Expanding both sides with respect to λ we obtain

$$\sum_\nu \tilde{c}_\nu\,\lambda^{n-\nu} = \sum_\nu \bar{c}_\nu\,\lambda^{n-\nu}\,.$$

This equation shows that corresponding coefficients in the characteristic polynomials of φ and $\tilde{\varphi}$ are complex conjugates. In particular,

$$\operatorname{tr}\tilde{\varphi} = \overline{\operatorname{tr}\varphi}\,.$$

14.9. The inner product in the space of all endomorphisms. Now consider the space $L(E; E)$ of all endomorphisms of E. Then an inner product can be introduced in $L(E; E)$ by the equation

$$(\varphi, \psi) = \frac{1}{n}\operatorname{tr}(\varphi \circ \tilde{\psi})\,. \tag{14.22}$$

It follows immediately form (14.22) that the function (φ, ψ) is sesquilinear. Interchange of φ and ψ yields

$$(\psi, \varphi) = \frac{1}{n}\operatorname{tr}(\psi \circ \tilde{\varphi}) = \frac{1}{n}\overline{\operatorname{tr}(\varphi \circ \tilde{\psi})} = \overline{(\varphi, \psi)}\,.$$

To prove that the Hermitian function (14.22) is positive definite let e_ν $(\nu = 1 \ldots n)$ be an orthonormal basis. Then

$$\varphi e_\nu = \sum_\mu \alpha_\nu^\mu\, e_\mu \quad \text{and} \quad \tilde{\varphi} e_\nu = \sum_\mu \tilde{\alpha}_\nu^\mu\, e_\mu \tag{14.23}$$

where $\tilde{\alpha}_\nu^\mu = \bar{\alpha}_\mu^\nu$. The equations (14.23) yield

$$(\varphi \circ \tilde{\varphi})\, e_\nu = \sum_{\mu,\nu} \tilde{\alpha}_\nu^\mu\, \alpha_\mu^\lambda\, e_\lambda$$

whence

$$\operatorname{tr}(\varphi \circ \tilde{\varphi}) = \sum_{\nu,\mu} \tilde{\alpha}_\nu^\mu\, \alpha_\mu^\nu = \sum_{\nu,\mu} \bar{\alpha}_\mu^\nu\, \alpha_\mu^\nu = \sum_{\nu,\mu} |\alpha_\mu^\nu|^2\,.$$

This formula shows that $(\varphi, \varphi) > 0$ for every endomorphism $\varphi \neq 0$.

14.10. Normal mappings. An endomorphism $\varphi\colon E \to E$ is called *normal*, if the mappings φ and $\tilde{\varphi}$ commute,

$$\tilde{\varphi} \circ \varphi = \varphi \circ \tilde{\varphi}\,. \tag{14.24}$$

In the same way as for a real inner product (cf. sec. 11.4) it is shown that the condition (14.24) is equivalent to

$$|\varphi x|^2 = |\tilde{\varphi} x|^2 \qquad x \in E . \quad (14.25)$$

It follows from (14.25) that the kernels of φ, and $\tilde{\varphi}$ coincide, $K(\varphi) = K(\tilde{\varphi})$. We thus obtain the direct decomposition

$$E = K(\varphi) \oplus \varphi(E) . \quad (14.26)$$

The relation $K(\varphi) = K(\tilde{\varphi})$ implies that the mappings φ and $\tilde{\varphi}$ have the same eigenvectors and that the corresponding eigenvalues are complex conjugate. In fact, assume that e is an eigenvector of φ and that λ is the corresponding eigenvalue,

$$\varphi e = \lambda e .$$

Then e is contained in the kernel of $\varphi - \lambda \iota$. Since the mapping $\varphi - \lambda \iota$ is again normal, e must also be contained in the kernel of $\tilde{\varphi} - \bar{\lambda} \iota$, i. e.

$$\tilde{\varphi} e = \bar{\lambda} e .$$

In sec. 11.7 we have seen that a selfadjoint endomorphism of a real inner product space possesses always n eigenvectors which are mutually orthogonal. Now it will be shown that in a complex space the same assertion holds even for normal mapping. Consider the characteristic polynomial of φ. According to the fundamental theorem of algebra this polynomial must have a zero λ_1. Then λ_1 is an eigenvalue of φ. Let e_1 be a corresponding eigenvector and E_1 the orthogonal complement of e_1. The mapping φ induces an endomorphism in E_1. In fact, let y be an arbitrary vector of E_1. Then

$$(\varphi y, e_1) = (y, \tilde{\varphi} e_1) = (y, \bar{\lambda} e_1) = \lambda (y, e_1) = 0$$

and hence φy is contained in E_1. The induced endomorphism is obviously again normal and hence there exists an eigenvector e_2 in E_1. Continuing this way we finally obtain n eigenvectors e_ν $(\nu = 1 \ldots n)$ which are mutually orthogonal,

$$(e_\nu, e_\mu) = 0 \qquad (\nu \neq \mu) .$$

If these vectors are normed to length one, they form an orthonormal basis of E. Relative to this basis the matrix of φ has diagonal form with the eigenvalues in the main-diagonal,

$$\varphi e_\nu = \lambda_\nu e_\nu \qquad (\nu = 1 \ldots n) . \quad (14.27)$$

14.11. Selfadjoint and skew mappings. Let φ be a selfadjoint endomorphism of E, i. e. an endomorphism such that $\tilde{\varphi} = \varphi$. Then the relation (14.20) yields

$$(\varphi x, y) = (x, \varphi y) \qquad x, y \in E .$$

Replacing y by x we obtain

$$(\varphi x, x) = (x, \varphi x) = \overline{(\varphi x, x)}$$

showing that $(\varphi x, x)$ is real for every vector $x \in E$. This implies that all eigenvalues of a selfadjoint endomorphism are real. In fact, let e be an eigenvector and λ be the corresponding eigenvalue. Then $\varphi e = \lambda e$, whence

$$(\varphi e, e) = \lambda (e, e) .$$

Since $(\varphi e, e)$ and $(e, e) \neq 0$ are real, λ must be real.

Every selfadjoint mapping is obviously normal and hence there exists a system of n orthonormal eigenvectors. Relative to this system the matrix of φ has the form (14.27) where all numbers λ_ν are real.

The matrix of a selfadjoint endomorphism relative to an orthonormal basis is Hermitian.

An endomorphism φ of E is called *skew* if $\tilde{\varphi} = -\varphi$. In a unitary space there is no essential difference between selfadjoint and skew endomorphisms. In fact, the relation

$$\widetilde{i\varphi} = -i\tilde{\varphi}$$

shows that the multiplication by i associates with every selfadjoint endomorphism a skew endomorphism and conversely.

14.12. Unitary mappings. A *unitary mapping* is an endomorphism of E which preserves the inner product,

$$(\varphi x, \varphi y) = (x, y) \qquad x, y \in E . \tag{14.28}$$

The relation (14.28) implies for $y = x$ that

$$|\varphi x| = |x| \qquad x \in E$$

showing that every unitary mapping is regular and hence it is an automorphism of E. If the equation (14.28) is written in the form

$$(\varphi x, y) = (x, \varphi^{-1} y)$$

it shows that the inverse mapping of φ is equal to the adjoint mapping,

$$\tilde{\varphi} = \varphi^{-1}. \tag{14.29}$$

Passing over to the determinants we obtain

$$\det\varphi \cdot \overline{\det\varphi} = 1$$

whence

$$|\det\varphi| = 1 .$$

Every eigenvalue of a unitary mapping has the norm 1. In fact, the equation $\varphi e = \lambda e$ yields

$$|\varphi e| = |\lambda|\,|e|$$

whence $|\lambda| = 1$.

The equation (14.29) shows that a unitary map is normal. Hence, there exists an orthonormal basis e_ν $(\nu = 1 \ldots n)$ such that

$$\varphi e_\nu = \lambda_\nu e_\nu \qquad (\nu = 1 \ldots n)$$

where the λ_ν are complex numbers of absolute value one.

Problems: 1. Given an endomorphism $\varphi: E \to E$, show that the function

$$\Phi(x, y) = (\varphi x, y)$$

is sesquilinear. Conversely, prove that every such function determines an endomorphism and that the Hermitian conjugate function determines the adjoint endomorphism.

2. Show that the of all selfadjoint mappings forms a real linear space of dimension n^2.

3. Let φ be an endomorphism of a complex linear space E.

a) Prove that a positive-definite inner product can be introduced in E such that φ becomes a normal mapping if and only if the endomorphism φ has n linearly independent eigenvectors.

b) Prove that a positive definite inner product can be introduced such that φ is

i) selfadjoint
ii) skew
iii) unitary

if and only if in addition the following conditions are fulfilled in corresponding order:

i) all eigenvalues of φ are real
ii) all eigenvalues of φ are imaginary or zero
iii) all eigenvalues have the absolute value 1.

4. Denote by $S(E)$ the space of all selfadjoint mappings and by $A(E)$ the space of all skew mappings of the unitary space E.

Prove that a multiplication is defined in $S(E)$ and $A(E)$ by

$$[\varphi, \psi] = i(\varphi \circ \psi - \psi \circ \varphi) \qquad \varphi \in S(E), \psi \in S(E)$$

and

$$[\varphi, \psi] = \varphi \circ \psi - \psi \circ \varphi \qquad \varphi \in A(E), \psi \in A(E)$$

respectively and that these spaces become Lie-algebras under the above multiplications.

§ 4. Unitary mappings of the complex plane

14.13. Definition. In this paragraph we will study the unitary mappings of a 2-dimensional unitary space in further detail. Let τ be a unitary mapping of the complex plane C. Employing an orthonormal basis e_1, e_2 we can represent the mapping τ in the form

$$\begin{aligned} \tau e_1 &= \alpha e_1 + \beta e_2 \\ \tau e_2 &= \varepsilon(-\bar{\beta} e_1 + \bar{\alpha} e_2) \end{aligned} \tag{14.30}$$

where α, β and ε are complex numbers subject to the conditions

$$|\alpha|^2 + |\beta|^2 = 1$$

and

$$|\varepsilon| = 1 .$$

The equations (14.30) show that

$$\det \tau = \varepsilon .$$

We are particularly interested in the unitary mappings with the determinant $+1$. For every such mapping the equations (14.30) reduce to

$$\begin{aligned} \tau e_1 &= \alpha e_1 + \beta e_2 \\ \tau e_2 &= -\bar{\beta} e_1 + \bar{\alpha} e_2 \end{aligned} \qquad |\alpha|^2 + |\beta|^2 = 1 .$$

This implies that

$$\begin{aligned} \tau^{-1} e_1 &= \bar{\alpha} e_1 - \beta e_2 \\ \tau^{-1} e_2 &= \bar{\beta} e_1 + \alpha e_2 . \end{aligned}$$

Adding the above relations in the corresponding order we find that

$$(\tau + \tau^{-1})\, e_\nu = (\alpha + \bar{\alpha})\, e_\nu = \operatorname{tr} \tau \cdot e_\nu \qquad (\nu = 1, 2)$$

whence

$$\tau + \tau^{-1} = \iota \cdot \operatorname{tr} \tau . \tag{14.31}$$

The formula (14.31) implies that

$$(z, \tau z) + (z, \tau^{-1} z) = |z|^2 \operatorname{tr} \tau$$

for every vector $z \in C$. Observing that

$$(z, \tau^{-1} z) = (\tau z, z) = \overline{(z, \tau z)}$$

we thus obtain the relation

$$2 \operatorname{Re}(z, \tau z) = |z|^2 \operatorname{tr} \tau \qquad z \in C \tag{14.32}$$

showing that the real part of the inner product $(z, \tau z)$ depends only on the norm of z. (14.32) is the complex analogon of the relation (11.46) for a proper rotation of the real plane.

We finally note that the set of all unitary mappings with the determinant $+1$ forms a subgroup of the group of all unitary mappings.

14.14. The quaternion-algebra. Consider the set Q of all endomorphisms of the form

$$\varphi = \lambda\tau \tag{14.33}$$

where τ is a unitary mapping with the determinant 1 and λ is an arbitrary real non negative number. Given an orthonormal basis e_1, e_2 of C these mappings φ are in a one-to-one correspondence with all matrices of the form

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{14.34}$$

where α and β are arbitrary complex numbers. The sum of two matrices of the form (14.34) has again this form and the same holds for an arbitrary real multiple of the matrix (14.34). Consequently, the set Q is a (real) linear space.

This space obviously has the dimension 4. Moreover, with every two mappings $\varphi_1 \in Q$ and $\varphi_2 \in Q$ the product $\varphi_2 \circ \varphi_1$ is also contained in Q. Hence, Q is an algebra with the identity-map as unit-element. The algebra Q is even a *division-algebra,* i. e. every element $\varphi \neq 0$ possesses an inverse with respect to the multiplication. In fact, let $\varphi = \lambda\tau \neq 0$ be an element of Q. Then $\lambda \neq 0$ and hence $\varphi^{-1} = \frac{1}{\lambda}\tau^{-1}$ is the inverse of φ. The division-algebra thus obtained is called the *algebra of quaternions.*

It follows from (14.34) that

$$\varphi + \tilde{\varphi} = \iota \cdot \operatorname{tr}\varphi . \tag{14.35}$$

Replacing φ by $\tilde{\varphi}$ we find that

$$\tilde{\varphi} + \varphi = \iota \cdot \overline{\operatorname{tr}\varphi} .$$

These two equations yield

$$\operatorname{tr}\varphi = \overline{\operatorname{tr}\varphi}$$

showing that the trace of every endomorphism $\varphi \in Q$ is real.

Now consider the positive definite inner product

$$(\varphi, \psi) = \frac{1}{2}\operatorname{tr}(\varphi \circ \tilde{\psi}) \qquad \varphi, \psi \in Q \tag{14.36}$$

in Q (cf. sec. 14.9). Since the product $\varphi \circ \tilde{\psi}$ is contained in Q this bilinear function is real-valued. In other words, Q becomes a Euclidean space.

In view of (14.35) the inner product (14.36) can also be written as

$$(\varphi, \psi) = \frac{1}{2}(\operatorname{tr}\varphi \cdot \operatorname{tr}\psi - \operatorname{tr}\varphi \circ \psi) \qquad \varphi, \psi \in Q . \tag{14.37}$$

Inserting $\psi = \varphi$ in (14.36) we find

$$(\varphi, \varphi) = \frac{1}{2}\operatorname{tr}(\varphi \circ \tilde{\varphi}) .$$

Observing that

$$\varphi \circ \tilde{\varphi} = \det \varphi \cdot \iota$$

for every endomorphism $\varphi \in Q$ we thus obtain the formula

$$(\varphi, \varphi) = \det \varphi \qquad \varphi \in Q . \tag{14.38}$$

Substituting $\psi = \iota$ in (14.36) we see that

$$(\varphi, \iota) = \frac{1}{2} \operatorname{tr} \varphi . \tag{14.39}$$

Now it will be shown that the multiplication in Q and the inner product (14.36) are connected by the relations

$$(\varphi \circ \chi, \psi \circ \chi) = (\varphi, \psi) \, |\chi|^2 \tag{14.40}$$

and $\qquad \varphi, \psi, \chi \in Q$

$$(\chi \circ \varphi, \chi \circ \psi) = |\chi|^2 \, (\varphi, \psi) . \tag{14.41}$$

Without loss of generality we may assume that $|\chi| = 1$. Then $\tilde{\chi} = \chi^{-1}$ and we obtain

$$(\varphi \circ \chi, \psi \circ \chi) = \operatorname{tr}(\varphi \circ \chi \circ \tilde{\chi} \circ \tilde{\psi}) = \operatorname{tr}(\varphi \circ \chi \circ \chi^{-1} \circ \tilde{\psi}) = \operatorname{tr}(\varphi \circ \tilde{\psi}) = (\varphi, \psi)$$

and

$$(\chi \circ \varphi, \chi \circ \psi) = \operatorname{tr}(\chi \circ \varphi \circ \tilde{\psi} \circ \tilde{\chi}) = \operatorname{tr}(\chi \circ \varphi \circ \tilde{\psi} \circ \chi^{-1}) = \operatorname{tr}(\varphi \circ \tilde{\psi}) = (\varphi, \psi) .$$

14.15. The multiplication in *C*. Select a fixed unit-vector a in C. Then to every vector $z \in C$ there exists a unique mapping $\varphi_z \in Q$ such that $\varphi_z a = z$. This mapping is determined by the equations

$$\begin{aligned} \varphi_z a &= \alpha a + \beta b \\ \varphi_z b &= - \bar{\beta} a + \bar{\alpha} b \end{aligned} \tag{14.42}$$

where b is a unit-vector orthogonal to a and

$$z = \alpha a + \beta b .$$

The correspondence $z \to \varphi_z$ obviously satisfies the relation

$$\varphi_{\lambda z_1 + \mu z_2} = \lambda \varphi_{z_1} + \mu \varphi_{z_2}$$

for any two real numbers λ and μ. Hence, it defines a linear mapping of C onto the linear space Q, if C is considered as a 4-dimensional real linear space. This suggests to define a multiplication among the vectors $z \in C$ in the following way:

$$z_1 z_2 = \varphi_{z_2} z_1 . \tag{14.43}$$

Then

$$\varphi_{z_1 z_2} = \varphi_{z_2} \circ \varphi_{z_1} \qquad z_1, z_2 \in C . \tag{14.44}$$

In fact, the two mappings on both sides of (14.44) are both contained in Q and send a into the same vector. The relation (14.44) shows that the correspondence $z \to \varphi_z$ preservs the products. Consequently, the space C becomes a (real) division-algebra under the multiplication (14.43) and this algebra is isomorphic to the algebra of quaternions.

The equation (14.31) implies that

$$z + z^{-1} = 2\,(\varphi_z, \iota)\, a \tag{14.45}$$

for every unit-vector z.

In fact, if z is a unit-vector then φ_z is a unitary mapping with the determinant 1 and the formula (14.31) and (14.39) yield

$$z + z^{-1} = \varphi_z a + (\varphi_z a)^{-1} = \varphi_z a + \varphi_z^{-1} a = a \operatorname{tr} \varphi_z = 2a\,(\varphi_z, \iota)\,.$$

The complex valued inner product in C is the real valued inner product in Q by the equation

$$\operatorname{Re}\,(z_1, z_2) = (\varphi_{z_1}, \varphi_{z_2})\,. \tag{14.46}$$

To prove this we may again assume that z_1 and z_2 are unit-vectors. Then φ_{z_1} and φ_{z_2} are unitary mappings and we can write

$$(z_1, z_2) = (\varphi_{z_1} a, \varphi_{z_2} a) = (\varphi_{z_2}^{-1}\, \varphi_{z_1} a, a) = (\varphi_{z_1 z_2^{-1}} a, a)\,. \tag{14.47}$$

Since $\varphi_{z_1 z_2^{-1}}$ is also unitary the formula (14.32) yields

$$\begin{aligned} \operatorname{Re}\,(\varphi_{z_1 z_2^{-1}} a, a) &= \frac{1}{2} \operatorname{tr}\, \varphi_{z_1 z_2^{-1}} = \frac{1}{2} \operatorname{tr}\,(\varphi_{z_2}^{-1} \circ \varphi_{z_1}) = \frac{1}{2} \operatorname{tr}\,(\tilde{\varphi}_{z_2} \circ \varphi_{z_1}) \\ &= \frac{1}{2} \operatorname{tr}\,(\varphi_{z_1} \circ \tilde{\varphi}_{z_2}) = (\varphi_{z_1}, \varphi_{z_2})\,. \end{aligned} \tag{14.48}$$

The relations (14.47) and (14.48) imply (14.46).

Problems: 1. Assume that an orthonormal basis is chosen in C. Prove that the endomorphisms which correspond to the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$$

form an orthonormal basis of Q.

2. Show that an endomorphism $\varphi \in Q$ is skew if and only if $\operatorname{tr} \varphi = 0$.

3. Prove that an endomorphism $\varphi \in Q$ satisfies the equation

$$\varphi^2 = -\iota$$

if and only if

$$\det \varphi = 1 \quad \text{and} \quad \operatorname{tr} \varphi = 0\,.$$

4. Verify the formula

$$(z_1 z, z_2 z) = (z_1, z_2)\,|z|^2 \qquad z_1, z_2, z \in C\,.$$

§ 5. Application to the orthogonal group

14.16. Definition. Now consider a real 4-dimensional Euclidean space E and let ω be an isomorphism of C onto E such that

$$(\omega z_1, \omega z_2) = \mathrm{Re}\,(z_1, z_2)\,.^{*)} \tag{14.49}$$

Introduce a multiplication among the vectors of E by

$$x_1 x_2 = \omega\,(z_1 z_2) \quad \text{where} \quad x_i = \omega z_i \quad (i = 1, 2)\,. \tag{14.50}$$

Then E becomes a division-algebra and ω defines an algebraic isomorphism of C onto E. The unit-element of E obviously is the vector $e = \omega a$.

The relations (14.40) and (14.41) yield the formulas

$$(x_1 x, x_2 x) = (x_1, x_2)\,|x|^2 \tag{14.51}$$

and $\qquad x_1, x_2, x \in E\,.$

$$(x\, x_1, x\, x_2) = |x|^2\,(x_1, x_2) \tag{14.52}$$

In fact, let $x = \omega z$, $x_1 = \omega z_1$ and $x_2 = \omega z_2$. Then it follows from (14.50), (14.49) and (14.46) that

$$\begin{aligned}(x_1 x, x_2 x) = (\varphi_{z_1 z}, \varphi_{z_2 z}) &= (\varphi_z \circ \varphi_{z_1}, \varphi_z \circ \varphi_{z_2})\\ &= |\varphi_z|^2\,(\varphi_{z_1}, \varphi_{z_2}) = |x|^2\,(x_1, x_2)\,.\end{aligned}$$

The formula (14.52) is proved in the same way. Inserting $x_2 = x_1$ in (14.51) we obtain the relation

$$|x_1 x| = |x_1|\,|x| \tag{14.53}$$

showing that the norm is preserved under the multiplication. It follows from (14.45) that every unit-vector e satisfies the equation.

$$x + x^{-1} = 2\,(x, e)\,e, \quad |x| = 1\,. \tag{14.54}$$

The relations (14.53) and (14.54) show that a vector $x \in E$ satisfies the equation

$$x^2 = -\,e$$

if and only if

$$|x| = 1 \quad \text{and} \quad (x, e) = 0\,.$$

14.17. The relation to the exterior product. Let E_1 be the orthogonal complement of e. Then E_1 is a 3-dimensional subspace of E. It will be shown that a normed determinant-function is defined in E_1 by the equation

$$\Delta\,(y, y_1, y_2) = (y, y_1 y_2) \quad y \in E, y_1 \in E, y_2 \in E\,. \tag{14.55}$$

---

*) C is considered as a real 4-dimensional linear space. Such a mapping exists because every two Euclidean spaces of the some dimension are isometric.

To prove the skew symmetry of Δ we may assume that all three arguments are unit-vectors. Then it follows from (14.54) that $y^{-1} = -y$ and $y_j^{-1} = -y_j$ $(j = 1, 2)$. Using the formulas (14.51) and (14.52) we thus obtain

$$\begin{aligned}\Delta(y, y_1, y_2) = (y, y_1 y_2) = -(y^{-1}, y_1^{-1} y_2^{-1}) &= -(y^{-1}, (y_2 y_1)^{-1}) \\ &= -(y, y_2 y_1) = -\Delta(y, y_2, y_1).\end{aligned}$$

To prove the skew symmetry with respect to the first and the third argument we write

$$\begin{aligned}\Delta(y, y_1, y_2) + \Delta(y_2, y_1, y) &= (y, y_1 y_2) + (y_2, y_1 y) \\ &= (y y_2^{-1} + y_2 y^{-1}, y_1) = (q + q^{-1}, y_1)\end{aligned} \tag{14.56}$$

where

$$q = y y_2^{-1}.$$

Now q is a unit-vector and hence the formula (14.54) yields

$$q + q^{-1} = 2(q, e)\, e$$

showing that

$$(q + q^{-1}, y_1) = 2(q, e)(y_1, e) = 0. \tag{14.57}$$

The relations (14.56) and (14.57) imply that

$$\Delta(y, y_1, y_2) + \Delta(y_2, y_1, y) = 0.$$

It remains to be shown that Δ is a *normed* determinant-function.

Let e_1 and e_2 be two orthogonal unit-vectors of E_1. Then $e_i = e_i^{-1}$ $(i = 1, 2)$ and hence the formula (14.51) yields

$$(e_1 e_2, e) = -(e_1 e_2^{-1}, e) = -(e_1, e_2) = 0$$

showing that the product $e_1 e_2$ is again contained in E_1. Moreover, the relations (14.51) and (14.52) imply that.

$$(e_1 e_2, e_1) = (e_2, e) = 0 \quad \text{and} \quad (e_1 e_2, e_2) = (e_1, e) = 0.$$

Finally,

$$|e_1 e_2| = |e_1|\,|e_2| = 1.$$

Hence, the vectors $e_1, e_2, e_1 e_2$ form an orthonormal basis of E_1. Inserting these vectors into Δ we find that

$$\Delta(e_1 e_2, e_1, e_2) = (e_1 e_2, e_1 e_2) = 1.$$

Let us now introduce an orientation in the Euclidean space E_1 by the normed determinant-function Δ. It will be shown that then the product of two vectors $y_1 \in E_1$ and $y_2 \in E_1$ can be written as

$$y_1 y_2 = -(y_1, y_2)\, e + [y_1, y_2]. \tag{14.58}$$

Denote by π the orthogonal projection of E onto E_1. Then

$$y_1 y_2 = (y_1 y_2, e)\, e + \pi\,(y_1 y_2)\,.$$

We have to show that

$$(y_1 y_2, e) = -\,(y_1, y_2) \tag{14.59}$$

and

$$\pi\,(y_1 y_2) = [y_1, y_2]\,. \tag{14.60}$$

Let $y \in E_1$ be an arbitrary vector. Then

$$(y, \pi\,(y_1 y_2)) = (\pi y, y_1 y_2) = (y, y_1 y_2) = \Delta\,(y, y_1, y_2) = (y, [y_1, y_2])\,.$$

This implies the relation (14.60). To prove (14.59) we can assume that y_2 is a unit-vector. Then $y_2 = -\,y_2^{-1}$ and it follows from (14.51) that

$$(y_1 y_2, e) = -\,(y_1 y_2^{-1}, e) = -\,(y_1, y_2)\,.$$

Selecting a positive orthonormal basis e_ν ($\nu = 1, 2, 3$) of E_1 we obtain from (14.58) the well-known multiplication-table for the quaternions:

$$\begin{array}{lll} e^2 = e & & e_1 e_2 = -\,e_2 e_1 = e_3 \\ e\,e_\nu = e_\nu & (\nu = 1, 2, 3) & e_2 e_3 = -\,e_3 e_2 = e_1 \\ e_\nu^2 = -\,e & & e_3 e_1 = -\,e_1 e_3 = e_2 \end{array}$$

14.18. Representation of rotations. As an application of the quaternion-multiplication it will now be shown that every proper rotation of a Euclidean space of dimension 3 or 4 can be represented by the quaternion-product.

Let p be a fixed unit-vector of E and define the mapping τ by

$$\tau x = p x p^{-1} \qquad x \in E\,. \tag{14.61}$$

The relation

$$|\tau x| = |p\, x\, p^{-1}| = |x|$$

shows that τ is a rotation. It follows from (14.61) that $\tau e = e$ and hence τ induces a rotation in the 3-dimensional subspace E_1. The induced rotation is proper as follows from the equation

$$\begin{aligned} &\Delta\,(\tau y_1, \tau y_2, \tau y_3) = (\tau y_1, \tau y_2\, \tau y_3) \\ &= (p\, y_1 p^{-1}, p\, y_2 y_3 p^{-1}) = (y_1, y_2 y_3) = \Delta\,(y_1, y_2, y_3)\,. \end{aligned}$$

Let us now determine the axis and the angle of this rotation. The equations $\tau e = e$ and $\tau p = p$ imply that the orthogonal projection q of p onto the subspace E_1 is a fix-vector. This projection is given by

$$q = p - \lambda e \qquad \lambda = (p, e)\,. \tag{14.62}$$

From now on it will be assumed that $p \neq \pm e$ (otherwise τ is the identity-map). Then $q \neq 0$ and the norm of q is equal to

$$|q|^2 = 1 - \lambda^2 .$$

To find the rotation-angle consider the plane F in E_1 which is orthogonal to q. In this plane a proper rotation is induced by τ. Next it will be shown that, induced rotation can be written as

$$\tau z = (2\lambda^2 - 1)\, z + 2\lambda\, [q, z] \qquad z \in F . \tag{14.63}$$

The equations

$$p = \lambda e + q \quad \text{and} \quad p^{-1} = \lambda e - q$$

yield

$$p\, z\, p^{-1} = (\lambda e + q)\, z\, (\lambda e - q) = \lambda^2 z + \lambda\, (q z - z q) - q\, z\, q .$$

Using the formula (14.58) we obtain

$$q z - z q = 2\, [q, z]$$

and

$$q z + z q = -2\, (q, z) e = 0 .$$

The last equation shows that

$$q\, z\, q = -q^2 z = |q|^2 e z = (1 - \lambda^2)\, z .$$

We thus obtain

$$p\, z\, p^{-1} = (2\,\lambda^2 - 1)\, z + 2\lambda\, [q, z]$$

which proves the relation (14.63).

The orientation of E_1 and the vector q induce an orientation in the plane F (cf. sec. 5.3). Using this orientation we obtain for the rotation-angle θ the equations

$$\cos\ \theta = (z, \tau z)$$

and $\qquad (0 \leqq \theta < 2\pi) , \quad |z| = 1 .$

$$\sin\ \theta = \frac{1}{|q|} \Delta\, (q, z, \tau z) .$$

Substituting for τz the expression (14.63) we find that

$$\cos\ \theta = 2\,\lambda^2 - 1 \tag{14.64}$$

$$\sin\ \theta = \frac{2\lambda}{|q|} \Delta\, (q, z, [q, z]) = \frac{2\lambda}{|q|}\, |[q, z]|^2 = 2\lambda\, |q| = 2\lambda \sqrt{1 - \lambda^2} .$$

The equations (14.64) yield a simple relation between the rotation-angle θ and the angle ω between the vectors p and e. In fact, the angle ω is determined by the equation

$$\cos \omega = (p, e) \qquad (0 < \omega < \pi) .$$

Hence, the relations (14.64) can be written as

$$\cos\theta = 2\cos^2\omega - 1 = \cos 2\omega$$

and

$$\sin\theta = 2\cos\omega\sin\omega = \sin 2\omega\,^{*)}.$$

showing that $\theta = 2\omega$. Altogether we see that *the fix-axis of the rotation τ is determined by the vector q and the rotation-angle is twice the angle between p and e.*

It follows from this result that two unit-vectors p_1 and p_2 determine the same rotation if and only if $p_2 = \pm p_1$.

Now it will be shown that, conversely, every proper rotation σ of E_1 can be represented in the form (14.61). The case $\sigma = \iota$ may again be excluded. Let a be a unit-vector in the fix-axis of σ and F the plane orthogonal to a. Denote by ϑ $(0 < \vartheta < 2\pi)$ the rotation-angle with respect to the orientation induced by a. Now consider the unit-vector

$$p = e\cos\frac{\vartheta}{2} + a\sin\frac{\vartheta}{2}. \tag{14.65}$$

Then the rotation

$$\tau y = p\,y\,p^{-1} \quad y \in E_1$$

coincides with σ. To prove this let q be the vector determined by (14.62). The equations (14.62) and (14.65) yield

$$q = a\sin\frac{\vartheta}{2} \tag{14.66}$$

showing that τ and σ have the same axis. The rotation-angle θ of τ is given by

$$\theta = 2\frac{\vartheta}{2} = \vartheta.$$

This angle refers to the orientation of F which is induced by q. But the equation (14.66) shows that q is a positive multiple of a and hence these two vectors induce the same orientation in F. This implies that $\tau = \sigma$.

14.19. Rotations of the 4-dimensional space. Now it is simple to show that every proper rotation of the 4-dimensional space E can be represented in the form

$$\tau x = p\,x\,q$$

where p and q are unit-vectors. Consider the rotation

$$\tau_1 x = (\tau e)^{-1}\tau x. \tag{14.67}$$

---

*) The restriction $0 < \omega < \pi$ implies that $\sin\omega > 0$ and hence that $\sqrt{1 - \lambda_2} = +\sin\omega$.

Then

$$\tau_1 e = e$$

and hence τ_1 induces a rotation in the 3-dimensional subspace E_1. It now follows from the result of sec. 14.18 that τ_1 can be written as

$$\tau_1 x = p\, x\, p^{-1} \quad |p| = 1\,. \tag{14.68}$$

The equations (14.67) and (14.68) yield

$$\tau x = p\, x\, p^{-1} \tau e = p\, x\, q\,, \quad q = p^{-1} \tau e\,.$$

The unit-vectors p and q are determined by τ up to the sign. In fact, assume that

$$p_1\, x\, q_1 = p_2\, x\, q_2$$

for all vectors $x \in E$. Then

$$p_2^{-1} p_1 x\, q_1\, q_2^{-1} = x\,.$$

This implies that

$$p_2^{-1} p_1 = \varepsilon e \quad \text{and} \quad q_1 q_2^{-1} = \varepsilon e$$

where $\varepsilon = \pm 1$. Hence $p_2 = \varepsilon p_1$ and $q_2 = \varepsilon q_1$.

Problems: 1. Let $a \neq 0$ be a vector of E which is not a negative multiple of e. Prove that the equation $x^2 = a$ has exactly two solutions.

2. If $y_j\,(j = 1, 2, 3)$ are three vectors of E_1, prove that

$$y_1\, y_2\, y_3 = -\Delta\,(y_1, y_2, y_3)\, e - (y_1, y_2)\, y_3 + (y_1, y_3)\, y_2 - (y_2, y_3)\, y_1\,.$$

3. Let p be a unit-vector of E. Show that the rotation-vector (cf. sec. 11.23) of the rotation $\tau x = p\, x\, p^{-1}$ is given by

$$u = 2\lambda\,(p - \lambda e)\,,\; \lambda = (p, e)\,.$$

4. Introduce an orientation in E such that the vector e induces in E_1 the orientation defined by the determinant-function (14.55). Let $p \neq \pm e$ be a unit-vector. Denote by F the plane spanned by e and p and by $F^{\perp}$ the orthogonal plane. Introduce in E the orientation such that the vectors e, p form a positive basis and in $F^{\perp}$ the orientation induced by E and F. The rotations

$$\varphi x = p x \quad \psi x = x p$$

of E obviously leave the planes F and $F^{\perp}$ invariant and they coincide in E. Denote by ϑ the (common) rotation-angle in F and by $\vartheta_1^{\perp}$ and $\vartheta_2^{\perp}$ the rotation-angles in $F^{\perp}$. Prove that $\vartheta_1^{\perp} = \vartheta$ and $\vartheta_2^{\perp} = -\vartheta$.

5. Consider a skew endomorphism ψ of E.

a) Show that ψ can be written in the form

$$\psi x = p x + x q \quad p, q \in E_1$$

and that the vectors p and q are uniquely determined by ψ.

b) Show that ψ transforms E_1 into E_1 if and only if $q = -p$.
c) Establish the formula

$$\det\psi = (|p|^2 - |q|^2)^2.$$

6. Let a, b be an orthonormal basis of the complex plane C such that a is the unit-element of the multiplication defined in sec. 14.15. Show that the vectors

$$e = \omega a\,,\ e_1 = \omega(ia) \quad e_2 = \omega b\,,\ e_3 = \omega(ib)$$

form an orthonormal basis of E and that

$$\Delta\,(e_1, e_2, e_3) = -1\,.$$

7. Let φ be a skew endomorphism of C with the trace zero. Show that the skew endomorphism $\psi = \omega \circ \varphi \circ \omega^{-1}$ of E can be written as

$$\psi x = p x$$

where p is a vector of E_1.

If

$$\begin{pmatrix} \alpha & \beta + i\gamma \\ -\beta + i\gamma & -i\alpha \end{pmatrix}$$

is the matrix of φ relative to the orthonormal basis a, b (cf. prob. 6), show that

$$p = \alpha e_1 + \beta e_2 + \gamma e_3\,.$$

§ 6. Application to Lorentz-transformations

14.20. Selfadjoint endomorphisms of the complex plane. Consider the set of all selfadjoint mappings σ of the complex plane C. S is a real 4-dimensional linear space. In this space introduce an inner product by

$$\langle\sigma, \tau\rangle = \frac{1}{2}(\operatorname{tr}(\sigma\circ\tau) - \operatorname{tr}\sigma\cdot\operatorname{tr}\tau)\,. \tag{14.69}$$

This inner product is indefinite and has the index 3. To prove this we note first that

$$\langle\sigma, \sigma\rangle = \frac{1}{2}(\operatorname{tr}\sigma^2 - (\operatorname{tr}\sigma)^2) = -\det\sigma \tag{14.70}$$

and

$$\langle\sigma, \iota\rangle = -\frac{1}{2}\operatorname{tr}\sigma\,. \tag{14.71}$$

Now select an orthonormal basis z_1, z_2 of C and consider the endomorphisms σ_j ($j = 1, 2, 3$) which correspond to the *Pauli-matrices*

$$\sigma_1\colon \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_2\colon \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_3\colon \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Then it follows from (14.69) that

$$\langle \sigma_i, \sigma_j \rangle = \delta_{ij}$$

and

$$\langle \sigma_i, \iota \rangle = 0\,, \quad \langle \iota, \iota \rangle = -1\,.$$

These equations show that the endomorphisms $\iota, \sigma_1, \sigma_2, \sigma_3$ form an orthonormal basis of S with respect to the inner product (14.69) and that this inner product has the index 3.

The orthogonal complement of the identity-map consists of all seljadjoint endomorphisms with the trace zero.

14.21. The isomorphism between Q and S. Besides S let us now consider the 4-dimensional real linear space Q introduced in sec. 14.14. Then a linear mapping $\Omega\colon Q \to S$ is defined by

$$\Omega\,\varphi = \frac{1-i}{2}\,\iota\,\mathrm{tr}\,\varphi + i\,\varphi \qquad \varphi \in Q\,. \tag{14.72}$$

In fact, the equation (14.72) implies that

$$\widetilde{\Omega\varphi} = \frac{1+i}{2}\,\iota\,\mathrm{tr}\,\varphi - i\,\tilde{\varphi}$$

whence

$$\widetilde{\Omega\varphi} - \Omega\,\varphi = i\,\iota\,\mathrm{tr}\,\varphi - i\,(\varphi + \tilde{\varphi})\,. \tag{14.73}$$

Observing that

$$\tilde{\varphi} + \varphi = \iota\,\mathrm{tr}\,\varphi$$

we obtain from (14.73) the relation

$$\widetilde{\Omega\varphi} = \Omega\,\varphi$$

showing that $\Omega\,\varphi$ is selfadjoint and hence it is contained in S.

It follows immediately from (14.72) that

$$\Omega\,(\psi \circ \varphi \circ \psi^{-1}) = \psi \circ \Omega\,\varphi \circ \psi^{-1} \qquad \varphi, \psi \in Q \tag{14.74}$$

and

$$\mathrm{tr}\,\Omega\,\varphi = \mathrm{tr}\,\varphi \qquad \varphi \in Q\,. \tag{14.75}$$

Solving (14.72) with respect to φ we find the equation

$$\varphi = \frac{1+i}{2}\,\iota\,\mathrm{tr}\,\Omega\,\varphi - i\,\Omega\,\varphi$$

showing that Ω is an isomorphism of Q onto S. With the help of the isomorphism Ω the inner product (14.69) can be carried over into the space Q:

$$\langle \varphi, \psi \rangle = \langle \Omega\,\varphi, \Omega\,\psi \rangle\,. \tag{14.76}$$

This gives in view of (14.69)

$$\langle\varphi,\psi\rangle=\frac{1}{2}\{\operatorname{tr}(\Omega\varphi\circ\Omega\psi)-\operatorname{tr}\Omega\varphi\cdot\operatorname{tr}\Omega\psi\}$$
$$=\frac{1}{2}\{\operatorname{tr}\Omega\varphi\circ\Omega\psi-\operatorname{tr}\varphi\cdot\operatorname{tr}\psi\}.$$

Now

$$\Omega\varphi\circ\Omega\psi=\frac{(1-i)^2}{4}\iota\operatorname{tr}\varphi\operatorname{tr}\psi+\frac{1-i}{2}i(\psi\operatorname{tr}\varphi+\varphi\operatorname{tr}\psi)-\varphi\circ\psi$$

and consequently,

$$\operatorname{tr}\Omega\varphi\circ\Omega\psi=\operatorname{tr}\varphi\cdot\operatorname{tr}\psi-\operatorname{tr}(\varphi\circ\psi). \tag{14.77}$$

Combining the equations (14.76) and (14.77) we find the formula

$$\langle\varphi,\psi\rangle=-\frac{1}{2}\operatorname{tr}(\varphi\circ\psi)\qquad\varphi,\psi\in Q. \tag{14.78}$$

14.22. The transformations T_α. Now consider an arbitrary endomorphism α of the complex plane C such that $\det\alpha=1$. Then a linear transformation $T_\alpha\colon S\to S$ is defined by

$$T_\alpha\sigma=\alpha\circ\sigma\circ\tilde{\alpha}\qquad\sigma\in S. \tag{14.79}$$

In fact, the equation

$$\widetilde{T_\alpha\sigma}=\alpha\circ\tilde{\sigma}\circ\tilde{\alpha}=\alpha\circ\sigma\circ\tilde{\alpha}=T_\alpha\sigma$$

shows that the endomorphism $T_\alpha\sigma$ is again selfadjoint. The transformation T_α preserves the inner product (14.69) and hence it is a Lorentz-transformation:

$$\langle T_\alpha\sigma,T_\alpha\sigma\rangle=-\det T_\alpha\sigma=-\det(\alpha\circ\sigma\circ\tilde{\alpha})$$
$$=-\det\sigma\,|\det\alpha|^2=-\det\sigma=\langle\sigma,\sigma\rangle.$$

Every Lorentz-transformation obtained in this way is proper. To prove this let $\alpha(t)$ $(0\leqq t\leqq 1)$ be a continuous family of endomorphisms of C such that

$$\alpha(0)=\iota\quad\alpha(1)=\alpha\quad\text{and}\quad\det\alpha(t)=1\quad(0\leqq t\leqq 1).$$

It follows from the result of sec. 5.12 that such a family exists. The continuous function

$$\det T_{\alpha(t)}\qquad(0\leqq t\leqq 1)$$

is equal to ± 1 for every t. In particular

$$\det T_{\alpha(0)}=\det T_\iota=1.$$

This implies that

$$\det T_{\alpha(t)}=1\qquad(0\leqq t\leqq 1)$$

whence

$$\det T_\alpha=1.$$

The transformations T_α are also orthochroneous. To prove this, observe that

$$T_\alpha \iota = \alpha \circ \tilde{\alpha}$$

whence

$$\langle \iota, T_\alpha \iota \rangle = \langle \iota, \alpha \circ \tilde{\alpha} \rangle = -\frac{1}{2} \operatorname{tr}(\alpha \circ \tilde{\alpha}) < 0 .$$

This relation shows that the time-like vectors ι and $T_\alpha \iota$ are contained in the same cone (cf. sec. 12.23).

14.23. In this way every endomorphism α with determinant 1 defines a proper Lorentz-transformation T_α. Obviously,

$$T_{\alpha \circ \beta} = T_\alpha \circ T_\beta . \tag{14.80}$$

Now it will be shown that two transformations T_α and T_β coincide only if $\beta = \pm \alpha$. In view of (14.80) it is sufficient to prove that T_α is the identity-operator only if $\alpha = \pm \iota$. If T_α is the identity, then

$$\alpha \circ \sigma \circ \tilde{\alpha} = \sigma \quad \text{for every} \quad \sigma \in S . \tag{14.81}$$

Inserting $\sigma = \iota$ we find that $\alpha \circ \tilde{\alpha} = \iota$ whence $\alpha = \tilde{\alpha}^{-1}$. Now, the equation (14.81) implies that

$$\alpha \circ \sigma = \sigma \circ \alpha \quad \text{for every} \quad \sigma \in S . \tag{14.82}$$

To show that $\alpha = \pm \iota$ select an arbitrary unit-vector $e \in C$ and define the selfadjoint endomorphism σ by

$$\sigma z = (z, e)\, e \qquad\qquad z \in C .$$

Then

$$(\sigma \circ \alpha)\, e = (\alpha e, e)\, e \quad \text{and} \quad (\alpha \circ \sigma)\, e = \alpha e .$$

Employing (14.82) we find that

$$\alpha e = (\alpha e, e)\, e .$$

In other words, every vector αz is a multiple of z. Now it follows from the linearity that $\alpha = \lambda \iota$ where λ is a complex constant. Observing that $\det \alpha = 1$ we finally see that $\lambda = \pm 1$.

14.24. In this section it will be shown conversely that every proper orthochroneous Lorentz-transformation T can be represented in the form (14.79). Consider first the case that ι is a fix-vector of T,

$$T \iota = \iota .$$

Employing the isomorphism $\Omega : Q \to S$ (cf. sec. 14.21) we introduce the transformation

$$T' = \Omega^{-1} \circ T \circ \Omega \tag{14.83}$$

of Q. Obviously,

$$\langle T'\varphi, T'\psi\rangle = \langle\varphi, \psi\rangle \qquad \varphi, \psi \in Q \tag{14.84}$$

and

$$T'\iota = \iota\,. \tag{14.85}$$

Besides the inner product (14.78) we have in Q the positive inner product defined by (14.37). Comparing these two inner products we see that

$$(\varphi, \psi) = \langle\varphi, \psi\rangle + \frac{1}{2}\operatorname{tr}\varphi \operatorname{tr}\psi = \langle\varphi, \psi\rangle - 2\,\langle\varphi, \iota\rangle\,\langle\psi, \iota\rangle\,. \tag{14.86}$$

Now the formulas (14.84), (14.86) and (14.85) yield

$$(T'\varphi, T'\psi) = (\varphi, \psi) \qquad \varphi, \psi \in Q$$

showing that T' is also an isometry with respect to the positive definite inner product (14.37). Hence the result of sec. 14.18 applies to T': There exists a unit-vector $\beta \in Q$ such that

$$T'\varphi = \beta \circ \varphi \circ \beta^{-1}. \tag{14.87}$$

Using the formulas (14.83), (14.87) and (14.74) we thus obtain

$$T\sigma = (\Omega \circ T' \circ \Omega^{-1})\,\sigma = \Omega\,(\beta \circ \Omega^{-1}\sigma \circ \beta^{-1}) = \beta \circ \sigma \circ \beta^{-1}.$$

Since $\beta^{-1} = \tilde{\beta}$ this equation can be written in the form

$$T\sigma = \beta \circ \sigma \circ \tilde{\beta} = T_\beta\sigma\,.$$

The equation (14.38) finally shows that the endomorphism β has indeed the determinant 1,

$$\det\beta = (\beta, \beta) = 1\,.$$

In the general case where $T\iota \neq \iota$ consider the plane F generated by the vectors ι and $T\iota$. Let ω be a vector of F such that

$$\langle\iota, \omega\rangle = 0 \quad\text{and}\quad \langle\omega, \omega\rangle = 1\,. \tag{14.88}$$

Then

$$\langle\Omega^{-1}\omega, \iota\rangle = \langle\omega, \Omega\iota\rangle = \langle\omega, \iota\rangle = 0$$

and consequently,

$$\Omega^{-1}\omega \circ \Omega^{-1}\omega = -\,\iota\,. \tag{14.89}$$

At the same time it follows from (14.72) and the first equation (14.88) that

$$\Omega^{-1}\omega = \frac{1}{i}\,\omega\,. \tag{14.90}$$

The relations (14.89) and (14.90) yield

$$\omega \circ \omega = \iota\,.$$

By hypothesis, T preserves fore-cone and past-cone. Hence $T\iota$ can be written as

$$T\iota = \iota\cosh\theta + \omega\sinh\theta\,. \tag{14.91}$$

Let α be the selfadjoint endomorphism defined by

$$\alpha = \iota\cosh\frac{\theta}{2} + \omega\sinh\frac{\theta}{2}\,. \tag{14.92}$$

Then

$$\begin{aligned} T_\alpha \iota = \alpha\circ\tilde{\alpha} &= \iota\cosh^2\frac{\theta}{2} + 2\omega\cosh\frac{\theta}{2}\sinh\frac{\theta}{2} + \omega\circ\omega\sinh^2\frac{\theta}{2} \\ &= \iota\cosh\theta + \omega\sinh\theta\,. \end{aligned} \tag{14.93}$$

Comparing (14.91) and (14.93) we see that

$$T\iota = T_\alpha\iota\,.$$

This equation shows that the transformation $T_\alpha^{-1}\circ T$ leaves the vector ι invariant. As it has been shown already there exists an endomorphism $\beta\in Q$ of determinant 1 such that

$$T_\alpha^{-1}\circ T = T_\beta\,.$$

Hence,

$$T = T_\alpha\circ T_\beta = T_{\alpha\circ\beta}\,.$$

It remains to be proved that α has the determinant 1. But this follows from the equations (14.70), (14.92) and (14.88):

$$\det\alpha = -\langle\alpha,\alpha\rangle = -\langle\iota,\iota\rangle\cosh^2\frac{\theta}{2} - \langle\omega,\omega\rangle\sinh^2\frac{\theta}{2} = 1\,.$$

Problem: 1. Let α be the endomorphism of the complex plane defined by the matrix $\begin{pmatrix} 1 & 2i \\ -i & 3\end{pmatrix}$. Find the real 4×4 matrix which corresponds to the Lorentz-transformation T_α with respect to the basis ι, σ_1, σ_2, σ_3 (cf. sec. 14.20).

Chapter XV

Invariant subspaces

The eigenvalue problem concerns itself with the following question: Given an endomorphism σ of a linear space E find all the 1-dimensional invariant linear subspace, i. e. all vectors x for which $\sigma x = \lambda x$. In Chapter IV, § 5 we found that an endomorphism need not have eigenvectors. Therefore it is natural to generalize the problem by admitting invariant subspaces of higher dimension, i. e. by asking for linear subspaces E_1 of E with the property that $\sigma E_1\subset E_1$. Obviously the whole space E as well as the kernel of σ are such subspaces. In order to exclude trivial cases in the investigation of this question it will be necessary to require that the invariant subspaces are irreducible, i. e. that they can not be decomposed any further into invariant subspaces.

The purpose of this chapter is to construct a direct decomposition of E into irreducible invariant subspaces for an arbitrary endomorphism. This construction is based upon the decomposition of certain polynomials into prime-factors. In the first paragraph a few basic properties of the algebra of polynomials will be established.

§ 1. The algebra of polynomials

15.1. Polynomials. Consider the set P^∞ of all infinite sequences

$$f = (\alpha_0, \alpha_1 \ldots)$$

of real or complex numbers. Define the sum of two sequences

$$f = (\alpha_0, \alpha_1 \ldots)$$

and

$$g = (\beta_0, \beta_1 \ldots)$$

as the sequence

$$f + g = (\alpha_0 + \beta_0, \alpha_1 + \beta_1 \ldots)$$

and the sequence λf as

$$\lambda f = (\lambda \alpha_0, \lambda \alpha_1 \ldots) .$$

Then P^∞ becomes a linear space of infinite dimension. The zero-vector of this space is the sequence

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

Let us now introduce a multiplication in P^∞ in the following way: The product of the sequences f and g is defined as the sequence

$$fg = (\gamma_0, \gamma_1 \ldots)$$

where

$$\gamma_\lambda = \sum_{\nu+\mu=\lambda} \alpha_\nu \beta_\mu \qquad (\lambda = 0, 1 \ldots) . \qquad (15.1)$$

This multiplication is associative, distributive and commutative and thus P^∞ becomes a commutative algebra. The unit-element of this algebra is the sequence

$$1 = (1, 0, \ldots) .$$

It follows from (15.1) that $fg = 0$ if and only if $f = 0$ or $g = 0$.

A polynomial is a sequence f in which only finitely many numbers α_ν are different from zero. It follows from the definition of the operations in P^∞ that the sum and the product of two polynomials is again a polynomial. In other words, the set of all polynomials is a subalgebra of P^∞. This subalgebra will be denoted by $P[u]$.

The *degree* of a polynomial $f \neq 0$ is the largest integer n such that $\alpha_n \neq 0$. We do not assigne a degree to the zero-polynomial. The equation (15.1) implies that

$$\deg(fg) = \deg f + \deg g \qquad (f \neq 0, g \neq 0) .$$

A polynomial of degree n is called *monic* if $\alpha_n = 1$.

The polynomial

$$u = (0, 1, 0 \ldots)$$

plays a distinguished role in the algebra $P[u]$. Raising u to the p-th power we find that

$$u^p = (\underbrace{0 \ldots 0}_{p}, 1, 0 \ldots) \quad (p = 1, 2 \ldots) .$$

This formula suggests to define u^0 by

$$u^0 = (1, 0 \ldots) .$$

A polynomial

$$f = (\alpha_0, \alpha_1 \ldots \alpha_n, 0 \ldots)$$

of degree n can now be written as

$$f = \sum_{\nu=0}^{n} \alpha_\nu u^\nu . \tag{15.2}$$

In other words, the polynomials u^ν $(\nu = 0, 1 \ldots)$ form set of generators of the linear space $P[u]$. The numbers α_ν in the representation (15.2) are uniquely determined by the polynomial f. They are called the *coefficients* of f and the polynomial u is called the *indeterminate* in $P[u]$.

15.2. Ideals. A subset J of the algebra $P[u]$ is called an *ideal* if it has the following properties:

I_1: With any two polynomials f and g the polynomial $f + g$ is contained in J.

I_2: If $f \in J$ and g is an arbitrary polynomial (not necessarily contained in J) then $fg \in J$.

An ideal obviously is a subalgebra of $P[u]$ (but not conversely!). Every polynomial h generates an ideal J_h consisting of all polynomials of the form hg where g is an arbitrary polynomial. Conversely, it will be shown that every ideal J is generated in this way.

For this purpose we construct first the following *division algorithm* in the algebra $P[u]$: Let $f \neq 0$ and $g \neq 0$ be two polynomials such that

$\deg f \geqq \deg g$. Then there exist polynomials q and r such that

$$f = g\,q + r \tag{15.3}$$

and

$$\deg r < \deg g \quad \text{or} \quad r = 0\,.$$

We may assume that the polynomials f and g are monic,

$$f = u^n + \sum_{\nu=0}^{n-1} \alpha_\nu\, u^\nu, \quad g = u^m + \sum_{\mu=0}^{m-1} \beta_\mu u^\mu \quad n \geqq m\,.$$

Let

$$q_1 = u^{n-m}$$

and

$$r_1 = \sum_{\nu=0}^{n-1} \alpha_\nu u^\nu + \sum_{\mu=0}^{m-1} \beta_\mu u^\mu\,.$$

Then f can be written as

$$f = g\,q_1 + r_1 \tag{15.4}$$

where

$$\deg r_1 \leqq n - 1 \quad \text{or} \quad r_1 = 0\,.$$

If $\deg r_1 < m$ or $r_1 = 0$ the equation (15.4) gives the desired decomposition of f. Otherwise we can write r_1 in the form

$$r_1 = g q_2 + r_2$$

where

$$\deg r_2 \leqq \deg r_1 - 1 \quad \text{or} \quad r_2 = 0\,.$$

Continuing this process we ultimately obtain polynomials r_k, q_k such that

$$r_{k-1} = g q_k + r_k$$

where

$$\deg r_k < m \quad \text{or} \quad r_k = 0\,.$$

Adding (15.4) to the equations

$$r_{\nu-1} = g q_\nu + r_\nu \qquad (\nu = 2 \ldots k)$$

we find that

$$f = g \sum_{\nu=1}^{k} q_\nu + r_k\,.$$

This is a decomposition of the form (15.3).

Now let J be an arbitrary ideal. We may exclude the case that J consists of the zero-polynomial only. Let h be a monic polynomial of least degree in J. Then J is generated by h. In fact, let $f \neq 0$ be an arbitrary polynomial of J. Then f can be written as

$$f = hq + r \tag{15.5}$$

where

$$\deg r < \deg h \quad \text{or} \quad r = 0\,.$$

The conditions I_1 and I_2 imply that the polynomial

$$r = f - hq$$

is again contained in J. Since h is a polynomial of least degree in J the relation

$$\deg r < \deg h$$

is impossible and thus it follows that $r = 0$. Now the equation (15.5) shows that f is a multiple of h.

The monic polynomial h is uniquely determined by the ideal J. To show this, let h' be an arbitrary monic polynomial which generates J. Then there exist polynomials q_1 and q_2 such that

$$h' = hq_1 \quad \text{and} \quad h = h'q_2 .$$

These equations yield

$$h'h = h'h\, q_1 q_2$$

showing that q_1 and q_2 are polynomials of degree zero. We thus obtain the relation

$$h' = \lambda h$$

where λ is a scalar. Since h' and h are monic polynomials it follows that $\lambda = 1$, whence $h' = h$.

15.3. Polynomial-functions. Besides the polynomial in an indeterminate u we shall consider *polynomial-functions*, i. e. functions of the form

$$f(\lambda) = \sum_\nu \alpha_\nu \lambda^\nu \tag{15.6}$$

where λ is a real or a complex variable. With respect to the usual addition and multiplication the set of all polynomial-functions forms an algebra $P(\lambda)$. The substitution

$$\sum_\nu \alpha_\nu u^\nu \to \sum_\nu \alpha_\nu \lambda^\nu \tag{15.7}$$

defines a mapping of the algebra $P[u]$ onto the algebra $P(\lambda)$. It follows from the definition of the operations in $P[u]$ that this mapping preserves sum and product. Moreover, (15.7) is a one-to-one mapping. In fact, assume that

$$f = \sum_{\nu=0}^{n} \alpha_\nu u^\nu$$

is a polynomial which is mapped into the zero-function. Then

$$\sum_{\nu=0}^{n} \alpha_\nu \lambda^\nu = 0 \tag{15.8}$$

for all values of λ. The substitution $\lambda = 0$ gives $\alpha_0 = 0$. Now it follows from (15.8) that

$$\sum_{\nu=1}^{n} \alpha_\nu \lambda^{\nu-1} = 0 \quad \text{for every} \quad \lambda \neq 0 .$$

whence

$$\alpha_1 = -\sum_{\nu=2}^{n} \alpha_\nu \lambda^{\nu-1} \quad (\lambda \neq 0).$$

Passing over to the absolute values we obtain

$$|\alpha_1| \leqq |\lambda| \sum_{\nu=2}^{n} |\alpha_\nu| \qquad |\lambda| \leqq 1.$$

This relation holds for every arbitrarily small $|\lambda|$ and hence it implies that $\alpha_1 = 0$. Continuing this way we find that $\alpha_\nu = 0$ $(\nu = 0 \ldots n)$. Our result shows that the substitution (15.7) defines an isomorphism of the algebra $P[u]$ onto the algebra $P(\lambda)$.

Problems: 1. Define the mapping

$$\partial : P[u] \to P[u]$$

by

$$\partial \sum_{\nu=0}^{n} \alpha_\nu u^\nu = \sum_{\nu=1}^{n} \nu \alpha_\nu u^{\nu-1}.$$

Prove the formulas

$$\partial(f + g) = \partial f + \partial g$$

and

$$\partial(fg) = \partial f \cdot g + f \cdot \partial g.$$

2. Let ∂ be a linear mapping of the space $P[u]$ into itself and subject to the conditions

$$\partial(fg) = \partial f \cdot g + f \cdot \partial g$$

and

$$\partial u = 1.$$

Prove that ∂ is the mapping defined in problem 1.

§ 2. Polynomials of endomorphisms

15.4. Definition. Let E be a linear space and denote by $A(E)$ the algebra of all endomorphisms of E (cf. sec. 2.18). Given a fixed endomorphism σ of E, the mappings σ^ν $(\nu = 0, 1 \ldots; \sigma^0 = \iota)$ generate a subalgebra $A(\sigma)$ of $A(E)$. Since any two mappings σ^ν and σ^μ commute, $A(\sigma)$ is a *commutative* subalgebra. We now define a homomorphism $f \to f(\sigma)$ of $P[u]$ into $A(\sigma)$ by the substitution

$$\sum_\nu \alpha_\nu u^\nu \to \sum_\nu \alpha_\nu \sigma^\nu. \tag{15.9}$$

This mapping is indeed a homomorphism, i. e.

$$(f + g)(\sigma) = f(\sigma) + g(\sigma)$$

and

$$(fg)(\sigma) = f(\sigma) \circ g(\sigma).$$

Given a polynomial f denote by $K(f)$ the kernel of the endomorphism $f(\sigma)$. Then the subspace $K(f)$ of E is invariant under σ. In fact, assume that x is a vector of $K(f)$. Then

$$\sum_\nu \alpha_\nu \sigma^\nu x = 0 .$$

Application of σ to this relation yields

$$\sum_\nu \alpha_\nu \sigma^{\nu+1} x = 0$$

whence

$$f(\sigma)\, \sigma x = 0 .$$

15.5. Relations between divisibility and kernel. For any two polynomials f and g we have the relations

$$K(f) \cap K(g) \subset K(f+g) \tag{15.10}$$

and

$$K(f) + K(g) \subset K(fg) . \tag{15.11}$$

To prove (15.10), let x be a vector of $K(f) \cap K(g)$. Then $f(\sigma)\, x = 0$ and $g(\sigma)\, x = 0$, whence $(f+g)\,(\sigma)\, x = 0$. The relation (15.11) follows in a similar way: Every vector of $K(f) + K(g)$ can be written as

$$x = x_1 + x_2 \qquad x_1 \in K(f),\, x_2 \in K(g) .$$

This implies that

$$fg\,(\sigma)\, x = fg(\sigma)\, x_1 + fg(\sigma)\, x_2 = gf(\sigma)\, x_1 + fg(\sigma)\, x_2 = 0$$

whence $x \in K(fg)$.

It follows from (15.11) that $K(g) \subset K(f)$ if g is a divisor of f. Now consider two arbitrary polynomials f and g and their greatest common divisor d. It will be shown that

$$K(d) = K(f) \cap K(g) . \tag{15.12}$$

The formula (15.11) implies that $K(d) \subset K(f)$ and $K(d) \subset K(g)$, whence $K(d) \subset K(f) \cap K(g)$. On the other hand, there are two polynomials f_1 and g_1 such that

$$d = f f_1 + g g_1 .$$

Now it follows from (15.10) and (15.11) that

$$K(d) \supset K(f f_1) \cap K(g g_1) \supset K(f) \cap K(g) .$$

If f and g are relatively prime i. e. if the greatest common divisor is the polynomial $d = 1$ then the kernel $K(d)$ reduces to the zero-vector. This shows that $K(f) \cap K(g) = 0$ if f and g are relatively prime.

Next we show that

$$K(v) = K(f) + K(g)$$

if v is the least common multiple of f and g. From (15.11) we obtain

$$K(f) \subset K(v) \quad \text{and} \quad K(g) \subset K(v)$$

whence

$$K(f) + K(g) \subset K(v) .$$

To prove the above inclusion in the other direction let f_1 and g_1 be two polynomials such that

$$v = f f_1 \quad \text{and} \quad v = g g_1 .$$

The polynomials f_1 and g_1 are relatively prime, because otherwise v would not be the *least* common multiple of f and g. Consequently, there are two polynomials f_2 and g_2 with the property that

$$1 = f_1 f_2 + g_1 g_2 .$$

Inserting the endomorphism σ into this relation we obtain

$$\iota = f_1(\sigma) \circ f_2(\sigma) + g_1(\sigma) \circ g_2(\sigma) .$$

This equation shows that every vector $x \in E$ can be written as

$$x = y + z$$

where

$$y = f_1(\sigma) \circ f_2(\sigma)\, x \quad \text{and} \quad z = g_1(\sigma) \circ g_2(\sigma)\, x .$$

Now assume that $x \in K(v)$. Then

$$f(\sigma)\, y = f(\sigma) \circ f_1(\sigma) \circ f_2(\sigma)\, x = v(\sigma) \circ f_2(\sigma)\, x = f_2(\sigma) \circ v(\sigma)\, x = 0$$

and

$$g(\sigma)\, z = g(\sigma) \circ g_1(\sigma) \circ g_2(\sigma)\, x = v(\sigma) \circ g_2(\sigma)\, x = g_2(\sigma) \circ v(\sigma)\, x = 0 .$$

These equations show that $y \in K(f)$ and $z \in K(g)$ and consequently that

$$K(h) \subset K(f) + K(g) .$$

If the polynomials f and g are relatively prime, the least common multiple v is equal to the product fg. At the same time we know that then the kernels $K(f)$ and $K(g)$ have only the zero-vector in common. We thus obtain the direct decomposition

$$K(fg) = K(f) \oplus K(g) \quad (f, g \text{ relatively prime}) . \tag{15.13}$$

Repeated application of (15.13) gives the formula

$$K(f_1 \ldots f_r) = K(f_1) \oplus \cdots \oplus K(f_r)$$

where any two polynomials f_i and f_j $(i \neq j)$ are relatively prime.

15.6. The minimal polynomial. Let σ be a fixed endomorphism of E. Then every polynomial

$$f = \sum_\nu \alpha_\nu u^\nu$$

determines an endomorphism $f(\sigma)$ of E by

$$f(\sigma) = \sum_\nu \alpha_\nu \sigma^\nu .$$

The correspondence

$$f \to f(\sigma)$$

defines a homomorphism of the algebra $P[u]$ into the algebra $A(E)$. The kernel of this homomorphism is an ideal $J(\sigma)$. In fact, if $f(\sigma) = 0$ and $g(\sigma) = 0$, then

$$(f + g)(\sigma) = f(\sigma) + g(\sigma) = 0$$

and if $f(\sigma) = 0$ and g is an arbitrary polynomial, then

$$(f g)(\sigma) = f(\sigma) g(\sigma) = 0 .$$

The ideal $J(\sigma)$ does not consist of the zero-polynomial only. This follows from the fact that the linear space $P[u]$ has infinite dimension and $A(E)$ has the dimension n^2.

The ideal $J(\sigma)$ is generated by a monic polynomial μ (cf. sec. **15.2**). This polynomial is uniquely determined by $J(\sigma)$ and hence by σ. It is called the *minimal polynomial* of σ. The minimal polynomial of an endomorphism has at least the degree 1 unless E reduces to the zero-vector. To prove this assume that $\mu = 1$. Then the identity-mapping coincides with the zero-mapping which is only possible if E has the dimension zero.

The minimal polynomials of the identity-map and the zero-map are obviously given by

$$\mu = u - 1 \quad \text{and} \quad \mu = u$$

respectively.

Dual mappings have the same minimal polynomial. In fact, the relation

$$\langle \overset{*}{x}, \sigma x \rangle = \langle \sigma^* \overset{*}{x}, x \rangle$$

implies that

$$\langle \overset{*}{x}, f(\sigma) x \rangle = \langle f(\sigma^*) \overset{*}{x}, x \rangle$$

for every polynomial f. It follows from this relation that $f(\sigma^*) = 0$ whenever $f(\sigma) = 0$. In other words, the ideals $J(\sigma^*)$ and $J(\sigma)$ coincide and thus the minimal polynomials of σ^* and σ are the same.

15.7. Divisors of the minimal polynomial. Let f be an arbitrary polynomial and g be a divisor of f. It has been shown in sec. **15.5** that $K(g) \subset K(f)$. But $K(g)$ need not be a proper subspace of $K(f)$ even if the degree of g is less then the degree of f. As an example, let g be the minimal polynomial. Then $K(f) = K(g) = E$ for all multiples of g.

Now assume that f is a divisor of the minimal-polynomial. Then it will be shown that $K(g)$ is a proper subspace of $K(f)$ if g is a proper divisor of f. By hypothesis, there exists a polynomial h such that

$$\mu = f h .$$

Define the polynomial g_1 by

$$g_1 = g\,h\,.$$

Then the degree of g_1 is less than the degree of μ, because g is a proper divisor of f. This implies that $g_1(\sigma) \neq 0$. Hence, there exists a vector x_1 such that $g_1(\sigma)\,x_1 \neq 0$. Consider the vector $y = h(\sigma)\,x_1$. Then

$$f(\sigma)\,y = f(\sigma)\,h(\sigma)\,x_1 = \mu(\sigma)\,x_1 = 0$$

and

$$g(\sigma)\,y = g(\sigma)\,h(\sigma)\,x_1 = g_1(\sigma)\,x_1 \neq 0\,.$$

Hence, y is contained in $K(f)$ but is not contained in $K(g)$.

15.8. Decomposition of the minimal polynomial. Assume a decomposition

$$\mu = f_1 \ldots f_r$$

of the minimal polynomial μ into monic polynomials f_ϱ $(\varrho = 1 \ldots r)$ such that any two polynomials f_i and f_j $(i \neq j)$ are relatively prime. According to sec. 15.5 this induces a direct decomposition of the kernel $K(\mu) = E$ into the kernels $K(f_\varrho)$,

$$E = K(f_1) \oplus \cdots \oplus K(f_r)\,.$$

Every subspace $K(f_\varrho)$ is invariant under σ. Denote by σ_ϱ the induced endomorphisms. It will be shown that f_ϱ is the minimal polynomial of σ_ϱ. Let μ_ϱ be the minimal polynomial of σ_ϱ. Then the equation

$$f_\varrho(\sigma)\,x = 0 \quad x \in K(f_\varrho)$$

shows that f_ϱ is a multiple of μ_ϱ. So it is sufficient to show that $K(\mu_\varrho)$ is an improper subspace of $K(f_\varrho)$ (cf. sec. 15.7). But this is clear because

$$\mu_\varrho(\sigma)\,x = \mu_\varrho(\sigma_\varrho)\,x = 0$$

for every vector $x \in K(f_\varrho)$.

15.9. The relation to the eigenvalues. Every zero of the minimal polynomial μ is an eigenvalue of σ and conversely. To prove this, assume that $\mu(\lambda) = 0$ for a certain value λ. Then the polynomial

$$f = u - \lambda$$

is a divisor of μ. At the same time, the constant polynomial $g = 1$ is a proper divisor of f. This implies that $K(g)$ is a *proper* subspace of $K(f)$ (cf. sec. 15.7). In other words, $K(f)$ contains vectors $x \neq 0$. But every vector $x \neq 0$ of $K(f)$ is an eigenvector of σ with the eigenvalue λ. Conversely, let λ be an eigenvalue of σ. Then the intersection $K(f) \cap K(\mu) = K(f)$ does not consist of the zero-vector only. Hence, f and μ cannot be relatively prime. Since f has the degree one, this implies that f is a divisor of μ. Hence λ must be a zero of μ.

Problems: 1. Find the minimal polynomial of a proper and an improper rotation of the Euclidean plane and of the Euclidean 3-space.

2. Two endomorphisms σ_1 and σ_2 of a linear space E are called *similar* if there exists an automorphism α of E such that $\sigma_2 = \alpha^{-1} \circ \sigma_1 \circ \alpha$. Show that similar endomorphisms have the same minimal polynomial.

3. Prove that the constant term in the minimal polynomial is different from zero if and only if the endomorphism is regular.

4. Give an example which shows that the products $\sigma_2 \circ \sigma_1$ and $\sigma_1 \circ \sigma_2$ need not have the same minimal polynomial.

5. Assume a direct decomposition $E = E_1 \oplus E_2$ and two endomorphisms $\sigma_1 : E_1 \to E_1$ and $\sigma_2 : E_2 \to E_2$. Define the endomorphism $\sigma : E \to E$ by

$$\sigma x = \sigma_1 x_1 + \sigma_2 x_2 \quad \text{where} \quad x = x_1 + x_2, \quad x_1 \in E_1, x_2 \in E_2 .$$

Show that the minimal polynomial of σ is the least common multiple of the minimal polynomials of σ_1 and σ_2.

6. Show that the minimal polynomial μ of an endomorphism σ can be constructed in the following way: Select an arbitrary vector $x_1 \in E$ and determine the smallest integer k_1, such that the vectors $\sigma^\nu x_1$ $(\nu = 0 \ldots k_1)$ are linearly dependent,

$$\sum_{\nu=0}^{k_1} \lambda_\nu \sigma^\nu x_1 = 0 .$$

Define the polynomial f_1 by

$$f_1 = \sum_{\nu=0}^{k_1} \lambda_\nu u^\nu .$$

If the vectors $\sigma^\nu x_1$ $(\nu = 0 \ldots k_1)$ do not generate the space E select a vector x_2 which is not a linear combination of these vectors and apply the same construction to x_2. Let f_2 be the corresponding polynomial. Continue this procedure until the whole space E is exhausted. Then μ is the least common multiple of the polynomials f_ϱ.

7. Let σ be a rotation of an inner product space. Prove that the coefficients α_ν of the minimal polynomial of σ satisfy the relations

$$\alpha_\nu = \varepsilon \alpha_{k-\nu} \qquad (\nu = 0 \ldots k)$$

where $\varepsilon = \pm 1$ depending on whether the rotation is proper or improper.

8. Show that the minimal polynomial of a selfadjoint endomorphism of a unitary space has real coefficients.

9. Assume that a conjugation $z \to \bar{z}$ is defined in the complex linear space E (cf. 14.6). Let σ be an endomorphism of E such that $\overline{\sigma z} = \sigma \bar{z}$. Prove that the minimal polynomial of σ has real coefficients.

§ 3. Invariant subspaces

15.10. Relation between invariant subspaces of E and E^*. Let σ be an endomorphism of E and σ^* be the dual endomorphism. If E_1 is an invariant subspace of E, then the orthogonal complement $E_1^{\perp}$ is an invariant subspace of E^* under the dual mapping. In fact, let $\overset{*}{y}$ be a vector of $E_1^{\perp}$. Then

$$\langle \sigma^* \overset{*}{y}, y \rangle = \langle \overset{*}{y}, \sigma y \rangle = 0$$

for all vectors $y \in E_1$, whence $\sigma^* \overset{*}{y} \in E_1^{\perp}$.

It follows from the above remark that every direct decomposition

$$E = E_1 \oplus E_2$$

of E into invariant subspaces leads to the direct decomposition of E^* into the invariant subspaces $E_1^{\perp}$ and $E_2^{\perp}$,

$$E^* = E_1^{\perp} \oplus E_2^{\perp} .$$

Moreover, the spaces E_1, $E_2^{\perp}$ and E_2, $E_1^{\perp}$ respectively form dual pairs (cf. sec. 2.8).

$$E_1^{\perp} = E_2^* , \quad E_2^{\perp} = E_1^* .$$

Conversely, assume that two invariant subspaces $E_1 \subset E$ and $E_1^* \subset E^*$ are given which are dual to each other. Then the orthogonal complements $E_1^{\perp}$ and $(E_1^*)^{\perp}$ are again invariant under σ^* and σ respectively and we have the direct decompositions

$$E = E_1 \oplus (E_1^*)^{\perp} \quad \text{and} \quad E^* = E_1^* \oplus E_1^{\perp} . \tag{15.14}$$

To prove the relations (15.14) observe that

$$E_1 \cap (E_1^*)^{\perp} = 0 \tag{15.15}$$

because E_1 and E_1^* are dual. Moreover,

$$\dim E_1 + \dim (E_1^*)^{\perp} = \dim E_1^* + \dim (E_1^*)^{\perp} = \dim E . \tag{15.16}$$

Now the first relation (15.14) follows from (15.15) and (15.16). The second formula (15.14) is proved in the same way.

15.11. Nilpotent endomorphisms. An endomorphism σ of E is called *nilpotent* if the minimal polynomial of σ has the form u^k. The number k is called the *order* of σ. The order k never exceeds the dimension n of E. To prove this, let $a \in E$ be a vector such that $\sigma^{k-1} a \neq 0$. Then the vectors $a, \sigma a, \ldots \sigma^{k-1} a$ are linearly independent. In fact, assume that

$$\lambda_0 a + \lambda_1 \sigma a + \ldots + \lambda_{k-1} \sigma^{k-1} a = 0 . \tag{15.17}$$

Applying σ^{k-1} to (15.17) we obtain the equation

$$\lambda_0 \sigma^{k-1} a = 0$$

showing that $\lambda_0 = 0$. Hence, (15.17) reduces to

$$\lambda_1 \sigma a + \ldots \lambda_{k-1} \sigma^{k-1} a = 0 .$$

Applying σ^{k-2} to this equation we obtain

$$\lambda_1 \sigma^{k-1} a = 0$$

whence $\lambda_1 = 0$. Continuing this way we find that $\lambda_\nu = 0$ $(\nu = 0 \ldots k)$. Hence, the vectors $\sigma^\nu a\,(\nu = 0 \ldots k-1)$ are linearly independent. Consequently

$$k \leqq n .$$

In the following section we shall construct a decomposition of E into irreducible invariant subspaces for a nilpotent endomorphism.

15.12. The decomposition into irreducible subspaces. Let σ be a nilpotent endomorphism of order k. Then a direct decomposition of E into irreducible invariant subspaces can be constructed in the following way*): Select two vectors $\overset{*}{a} \in E^*$ and $a \in E$ such that

$$\langle \overset{*}{a}, \sigma^{k-1} a \rangle \neq 0 .$$

This is possible since $\sigma^{k-1} \neq 0$. Then the vectors

$$a, \sigma a \ldots \sigma^{k-1} a \quad \text{and} \quad \overset{*}{a}, \sigma^* \overset{*}{a} \ldots \sigma^{*k-1} \overset{*}{a}$$

are linearly independent and hence they generate two p-dimensional subspaces U and U^* of E and E^*. The subspaces U and U^* are obviously invariant under σ and σ^* respectively. Furthermore U and U^* are dual. In fact, assume that a vector

$$u = \sum_{\nu=0}^{k-1} \xi^\nu \sigma^\nu a \tag{15.18}$$

of U is orthogonal to U^*. Then

$$\langle \sigma^{*k-1} \overset{*}{a}, u \rangle = 0 . \tag{15.19}$$

The equations (15.18) and (15.19) imply that

$$\left\langle \sigma^{*k-1} \overset{*}{a}, \sum_{\nu=1}^{k-1} \xi^\nu \sigma^\nu a \right\rangle = \left\langle \overset{*}{a}, \sum_{\nu=0}^{k-1} \xi^\nu \sigma^{k+\nu-1} a \right\rangle = \xi^0 \langle \overset{*}{a}, \sigma^{k-1} a \rangle = 0$$

whence $\xi^0 = 0$. Next, we observe that

$$\langle \sigma^{*k-2} \overset{*}{a}, u \rangle = 0 .$$

Inserting (15.18) into this relation we obtain $\xi^1 = 0$. Continuing this way we find that $\xi^\nu = 0$ $(\nu = 0 \ldots k-1)$ whence $u = 0$. This proves that $U \cap (U^*)^\perp = 0$. In the same way it is shown that $U^* \cap U^\perp = 0$. These two relations imply that U and U^* are dual spaces.

---

* The above construction is due to V. Pták: Eine Bemerkung zur Jordanschen Normalform von Matrizen. Acta Sci. Math. Szeged, XVII (1956).

Now the result of sec. 15.10 asserts that E is the direct sum of the invariant subspaces U and $(U^*)^\perp$,

$$E = U \oplus (U^*)^\perp .$$

The induced endomorphism of $(U^*)^\perp$ is again nilpotent and hence the above construction can be applied to $(U^*)^\perp$. Now

$$\dim (U^*)^\perp < \dim E$$

and hence we obtain after a finite number of steps a direct decomposition

$$E = U_1 \oplus U_2 \oplus \cdots \oplus U_m \qquad U_1 = U \tag{15.20}$$

into invariant subspaces. It will be shown in the next section that the spaces U_μ are irreducible.

15.13. A criterion for irreducibility. *Let E be an n-dimensional linear space and σ be a nilpotent endomorphism of order k. Then E is irreducible if and only if $k = n$.*

We note first that $k \leqq n$ (cf. sec. 15.11). If $k < n$, the decomposition (15.20) consists of at least two subspaces because every subspace U_μ has a dimension $\leqq k$. Hence, E is reducible. To prove that E is irreducible if $k = n$, assume a direct decomposition

$$E = E_1 \oplus E_2 \tag{15.21}$$

into two invariant subspaces. The minimal polynomials μ_1 and μ_2 of the induced mappings have the form

$$\mu_1 = u^{k_1} \quad k_1 \leqq k \quad \text{and} \quad \mu_2 = u^{k_2} \quad k_2 \leqq k .$$

Now select a vector $x \in E$ such that $\sigma^{k-1} x \neq 0$. By hypothesis, x can be written as

$$x = x_1 + x_2 \qquad\qquad x_1 \in E_1, x_2 \in E_2 .$$

Applying σ^{k-1} we obtain

$$\sigma^{k-1} x = \sigma^{k-1} x_1 + \sigma^{k-1} x_2 . \tag{15.22}$$

Since $\sigma^{k-1} x \neq 0$, at least one of the vectors on the right hand-side of (15.22) must be different from zero. Assume that $\sigma^{k-1} x_1 \neq 0$; then $k \leqq k_1$ and, consequently,

$$\dim E_1 \geqq k_1 \geqq k = n = \dim E .$$

This relation shows that $E_1 = E$ and hence, the decomposition (15.21) is improper.

The above criterion shows that the spaces U_μ are irreducible and we obtain the following theorem: *To every nilpotent endomorphism σ of order k there exists a direct decomposition of E into irreducible invariant* *subspaces. In this decomposition occurs at least one subspace of dimension k. In every irreducible subspace there exists a basis relative to which the matrix of σ has the form*

$$\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \tag{15.23}$$

15.14. The number of the irreducible subspaces. It is clear from the construction in sec. 15.12 that the irreducible subspaces are generally not uniquely determined by the endomorphism. However, it can be shown that the *number* of the irreducible subspaces of a given dimension is determined by σ.

Denote by N_j the number of j-dimensional subspaces ($1 \leqq j \leqq k$) and by J_j the direct sum of all these spaces. Then we have the decomposition

$$E = J_1 \oplus J_2 \oplus \cdots \oplus J_k . \tag{15.24}$$

Passing over to the dimensions in (15.24) we find the relation

$$n = N_1 + 2N_2 + \cdots + kN_k .$$

Since every subspace J_j is invariant we obtain from (15.24)

$$\sigma(E) = \sigma(J_1) \oplus \sigma(J_2) \oplus \cdots \oplus \sigma(J_k) . \tag{15.25}$$

Now it follows from the form of the matrix (15.23) that the dimension of every irreducible subspace decreases by 1 under σ. This implies that

$$\dim \sigma(J_j) = (j-1)\, N_j \qquad (j = 1 \ldots k) . \tag{15.26}$$

The equations (15.25) and (15.26) yield

$$r(\sigma) = N_2 + 2\, N_3 + \cdots + (k-1)\, N_k .$$

Where $r(\sigma)$ denotes the rank of σ.

Applying σ to (15.25) and comparing the dimensions we obtain the equation

$$r(\sigma^2) = N_3 + 2N_4 + \cdots + (k-2)\, N_k$$

and in general,

$$r(\sigma^\lambda) = N_{\lambda+1} + 2\, N_{\lambda+2} + \ldots + (k-\lambda)\, N_k \qquad (\lambda = 0 \ldots k) . \tag{15.27}$$

Replacing λ by $\lambda + 1$ and $\lambda - 1$ respectively we find the relations

$$r\,(\sigma^{\lambda+1}) = N_{\lambda+2} + 2\, N_{\lambda+3} + \ldots + (k-\lambda-1)\, N_k \tag{15.28}$$

and

$$r\,(\sigma^{\lambda-1}) = N_\lambda + 2\, N_{\lambda+1} + \ldots + (k-\lambda+1)\, N_k . \tag{15.29}$$

Addition of (15.28) and (15.29) and use of (15.27) gives

$$r\,(\sigma^{\lambda+1}) + r\,(\sigma^{\lambda-1}) = N_\lambda + 2\,[N_{\lambda+1} + 2\, N_{\lambda+2} \ldots + (k-\lambda)\, N_k] = N_\lambda + 2r(\sigma^\lambda) .$$

We thus obtain the formula

$$N_\lambda = r(\sigma^{\lambda+1}) + r(\sigma^{\lambda-1}) - 2\,r(\sigma^\lambda) \qquad (\lambda = 1 \dots k)$$

which permits computing the numbers N_λ from the ranks of the endomorphisms σ^λ.

Problems: 1. Let σ be a nilpotent endomorphism of order k. Prove that E is irreducible if and only if

$$r(\sigma^\nu) = k - \nu \qquad (\nu = 1 \dots k)\,.$$

2. Prove that two nilpotent endomorphisms σ_1 and σ_2 of order k are similar (cf. § 2 prob. 2) if and only if

$$r(\sigma_1^\nu) = r(\sigma_2^\nu) \qquad (\nu = 1 \dots k)\,.$$

3. Show that two nilpotent endomorphisms of a 3-dimensional linear space are similar if and only they have the same minimal polynomial.

4. Prove that two nilpotent endomorphisms of a 6-dimensional linear space are similar if and only if they have the same minimal polynomial and the same rank.

5. Let E be a (real or complex) inner product space and σ be a nilpotent endomorphism. Prove that σ is similar to the adjoint endomorphism.

§ 4. The decomposition of a complex linear space

Combining the results of the two last paragraphs we now proceed to construct a decomposition into irreducible invariant subspaces for an arbitrary endomorphism. This construction consists of two steps:

1. decomposition of the minimal polynomial into its prime-factors,
2. irreducible decomposition of the invariant subspaces which correspond to these prime-factors.

The first step is particularly simple in the complex case because every complex polynomial can be decomposed into prime-factors of first order. The necessary modifications for a real linear space will be discussed in § 5.

15.15. Definition. Let σ be an endomorphism of the complex linear space E. Then the minimal polynomial μ can be written in the form

$$\mu = (u - \lambda_1)^{k_1} \dots (u - \lambda_r)^{k_r}, \tag{15.30}$$

where the λ_ϱ are the eigenvalues of σ. Denote by K_ϱ the kernel of the endomorphism

$$(\sigma - \lambda_\varrho\, \iota)^{k_\varrho} \qquad (\varrho = 1 \dots r)\,.$$

Then we have the direct decomposition

$$E = K_1 \oplus \dots \oplus K_r\,. \tag{15.31}$$

Every subspace K_ϱ is invariant under σ and hence σ induces an endomorphism $\sigma_\varrho: K_\varrho \to K_\varrho$. The minimal polynomial μ_ϱ of σ_ϱ is given by

$$\mu_\varrho = (u - \lambda_\varrho)^{k_\varrho}.$$

Define the endomorphism τ_ϱ by

$$\tau_\varrho = \sigma_\varrho - \lambda_\varrho \iota .$$

Then τ_ϱ is nilpotent and has the order k_ϱ. This implies that

$$k_\varrho \leqq \dim K_\varrho .$$

Adding these equations we find the relation

$$\deg \mu = \sum_\varrho k_\varrho \leqq \dim E$$

showing that the degree of the minimal polynomial never exceeds the dimension of E

To every subspace K_ϱ we can now apply the construction described in sec. 15.12. This yields a direct decomposition

$$K_\varrho = U_\varrho^{(1)} \oplus \cdots \oplus U_\varrho^{(m_\varrho)} \tag{15.32}$$

where every subspace $U_\varrho^{(\mu)}$ is invariant and irreducible with respect to τ_ϱ. But τ_ϱ and σ_ϱ have obviously the same invariant subspaces and hence the spaces $U_\varrho^{(\mu)}$ are also invariant and irreducible with respect to σ. Combining all the decompositions (15.32) we finally obtain a decomposition of the entire space E into irreducible subspaces.

15.16. The Jordan normal-form. In every irreducible subspace there exists a basis relative to which the matrix of σ has the form

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \cdot & & \\ & & & \cdot & 1 \\ & & & & \lambda \end{pmatrix}. \tag{15.33}$$

(cf. sec. 15.13). Putting all these bases together we obtain a basis of E in which the matrix of σ consists of submatrices of the form (15.33) following each other successively along the main-diagonal. This is the *Jordan normal-form* of a complex matrix.

15.17. Characteristic polynomial and minimal polynomial. Consider the direct decomposition

$$E = K_1 \oplus \cdots \oplus K_r$$

and define the polynomial χ by

$$\chi = (-1)^n (u - \lambda_1)^{n_1} \ldots (u - \lambda_r)^{n_r}, \tag{15.34}$$

where n_ϱ denotes the dimension of K_ϱ ($\varrho = 1 \ldots r$). Then χ has the degree $\sum\limits_{\varrho} n_\varrho = n$. We shall prove that χ is the characteristic polynomial of σ (cf. sec. 4.20) if this polynomial is considered as a polynomial in the indeterminate u. In other words, it will be shown that

$$\chi = \sum_{\nu} \alpha_\nu u^{n-\nu}$$

where the coefficients α_ν are given by (4.48). In order to do so it is sufficient to show that the two polynomial-functions

$$\chi(\lambda) = (-1)^n (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_r)^{n_r}$$

and

$$f(\lambda) = \det(\sigma - \lambda\iota) \tag{15.35}$$

coincide (cf. sec. 15.3). It follows from the above decomposition that

$$f(\lambda) = \det(\sigma_1 - \lambda\iota) \dots \det(\sigma_r - \lambda\iota)\,. \tag{15.36}$$

Using the Jordan normal-form of the matrix of σ_ϱ we find that

$$\det(\sigma_\varrho - \lambda\iota) = (\lambda_\varrho - \lambda)^{n_\varrho}. \tag{15.37}$$

The equations (15.36) and (15.37) yield

$$f(\lambda) = (\lambda_1 - \lambda)^{n_1} \dots (\lambda_r - \lambda)^{n_r}$$

showing that the polynomials $f(\lambda)$ and $\chi(\lambda)$ coincide.

Let us now compare the characteristic polynomial (15.37) with the minimal polynomial

$$\mu = (u - \lambda_1)^{k_1} \dots (u - \lambda_r)^{k_r}\,.$$

The relations $n_p \geqq k_p$ (cf. sec. 15.15) show that χ is a multiple of μ

$$\chi = \mu \cdot h\,.$$

Inserting $u = \sigma$ into this identity we find that

$$\chi(\sigma) = \mu(\sigma) \cdot h(\sigma) = 0\,.$$

In other words, *every endomorphism satisfies its own characteristic equation* (Cayley-Hamilton-theorem) *).

15.18. Irreducible spaces. As a consequence of our general results we can easily prove the following *irreducibility-criterion: An n-dimensional complex linear space E is irreducible with respect to a given endomorphism σ if and only if the minimal polynomial of σ has the form*

$$\mu = (u - \lambda)^n. \tag{15.38}$$

Assume first that E is irreducible. Then there can be only one factor in the decomposition (15.30). Consequently, μ has the form

$$\mu = (u - \lambda)^k.$$

Now it follows from the criterion of sec. 15.13 that $k = n$.

---

*) For a different proof of this theorem cf. Chap. IX, § 2, prob. 4.

Conversely, assume that the minimal polynomial of σ has the form (15.38). Then the same criterion asserts that E is irreducible with respect to the endomorphism $\sigma - \lambda\iota$. Hence, E is also irreducible with respect to σ.

15.19. Completely reducible spaces. The other extreme case is that E is *completely reducible*, i.e. that E can be decomposed into 1-dimensional invariant subspaces. *An n-dimensional complex linear space is completely reducible if and only if the minimal polynomial of σ has n different roots,*

$$\mu = (u - \lambda_1) \dots (u - \lambda_n) \qquad \lambda_i \neq \lambda_j \quad (i \neq j). \tag{15.39}$$

To prove this criterion assume that E is completely reducible. Let λ_ϱ $(\varrho = 1 \dots r)$ be all distinct eigenvalues of σ, and let E_ϱ $(\varrho = 1 \dots r)$ be the corresponding eigenspaces. Then the subspaces E_ϱ generate the entire space E,

$$E = E_1 \oplus \cdots \oplus E_r.$$

Hence, every vector x can be written as

$$x = \sum_\varrho x_\varrho \qquad x_\varrho \in E_\varrho \quad (\varrho = 1 \dots r). \tag{15.40}$$

Define the polynomial f by

$$f = (u - \lambda_1) \dots (u - \lambda_r).$$

Then it follows from (15.39) and (15.40) that

$$f(\sigma)\, x = 0.$$

This equation shows that f is a multiple of the minimal polynomial μ. On the other hand, the minimal polynomial contains the prime-factors $u - \lambda_\varrho$ $(\varrho = 1 \dots r)$ and hence μ is a multiple of f. This implies that $\mu = f$.

Problems: 1. Verify the Cayley-Hamilton-theorem for a 2×2-matrix by direct computation.

2. Show that two endomorphisms σ and τ of a complex linear space are similar if and only if the following conditions are satisfied:

a) σ and τ have the same minimal polynomial

$$\mu = (u - \lambda_1)^{k_1} \dots (u - \lambda_r)^{k_r}.$$

b) Let

$$E = K_1 \oplus \cdots \oplus K_r$$

be the decomposition of E with respect to the common minimal polynomial and denote by σ_ϱ and τ_ϱ the induced endomorphisms in K_ϱ $(\varrho = 1 \dots r)$. Then every two endomorphisms

$$(\sigma_\varrho - \lambda_\varrho \iota)^\nu \quad \text{and} \quad (\tau_\varrho - \lambda_\varrho \iota)^\nu \qquad (\nu = 1 \dots k_\varrho - 1)$$

have the same rank.

3. Prove that a unitary space is completely reducible with respect to a normal endomorphism.

4. Let σ be an endomorphism of a unitary space and $\tilde{\sigma}$ be the adjoint endomorphism. Prove that σ is normal if and only if there exists a polynomial f such that

$$\tilde{\sigma} = f(\sigma) .$$

5. Assume that σ is an endomorphism of a complex linear space such that $\sigma^p = \iota$ for some integer p. Show that E can be made into a unitary space in such a way that σ becomes a unitary mapping.

6. Given an endomorphism σ of E consider the decompositions

$$\chi = (-1)^n (u - \lambda_1)^{n_1} \dots (u - \lambda_r)^{n_r}$$

and

$$\mu = (u - \lambda_1)^{k_1} \dots (u - \lambda_r)^{k_r}$$

of the characteristic polynomial and the minimal polynomial. Show that the endomorphisms

$$(\sigma - \lambda_\varrho \iota)^{n_\varrho} \quad \text{and} \quad (\sigma - \lambda_\varrho)^{k_\varrho} \qquad (\varrho = 1 \dots r)$$

have the same kernel.

7. Let σ be an endomorphism of a complex linear space and λ_ν $(\nu = 1 \dots n)$ be the eigenvalues of σ (not necessarily distinct). Given an arbitrary polynomial f prove that the endomorphism $f(\sigma)$ has the eigenvalues $f(\lambda_\nu)$ $(\nu = 1 \dots n)$.

8. Show that two endomorphisms of a 3-dimensional complex linear space are similar if and only if they have the same minimal polynomial.

§ 5. The decomposition of a real linear space

In the case of a complex linear space the fundamental theorem of algebra asserts that all the prime-factors are of first degree. However, a polynomial over the real numbers contains in general prime-factors of degree 2. This fact requires a few modifications in the construction of the irreducible subspaces for an endomorphism of a real linear space.

15.20. Definition. Let E be a real linear space and σ be an endomorphism of E. Then the minimal polynomial of σ can be decomposed as

$$\mu = f_1 \dots f_r$$

where every polynomial f_ϱ has the form

$$(u - \lambda)^k \tag{15.41}$$

or

$$(u^2 + \alpha u + \beta)^k, \quad \alpha^2 - 4\beta < 0 . \tag{15.42}$$

We thus obtain the direct decomposition

$$E = K_1 \oplus K_2 \oplus \cdots \oplus K_r \tag{15.43}$$

where K_ϱ denotes the kernel of $f_\varrho(\sigma)$. Every subspace K_ϱ is invariant under σ and the minimal polynomial of the induced endomorphism is $f_\varrho(\sigma)$. A subspace K_ϱ which corresponds to a prime-factor of the form (15.41) can be decomposed into irreducible invariant subspaces by the construction given in sec. 15.12. Therefore we can restrict ourselves to endomorphisms whose minimal polynomials have the form (15.42).

15.21. Consider an endomorphism σ of E with the minimal polynomial

$$\mu = (u^2 + \alpha u + \beta)^k, \qquad \alpha^2 - 4\beta < 0 .$$

Then the endomorphism

$$\tau = \sigma^2 + \alpha\sigma + \beta\iota$$

is nilpotent of order k. Now we proceed as in sec. 15.12.: Select a vector $a \in E$ such that $\tau^{k-1}a \neq 0$ and consider the subspace U generated by the vectors $a, \tau a, \ldots, \tau^{k-1}a$. Then U is invariant under τ. This implies that the sum

$$V = U + \sigma(U)$$

is invariant under σ. In fact, let

$$v = u_1 + \sigma u_2 \qquad u_1, u_2 \in U$$

be a vector of V. Then

$$\sigma V = \sigma u_1 + \sigma^2 u_2 = \tau u_2 - \beta u_2 + \sigma(u_1 - \alpha u_2) \in U + \sigma(U) = V .$$

Next, we show that the subspaces U and $\sigma(U)$ have only the zero-vector in common. Assume a relation

$$\sum_{\nu=0}^{k-1} \xi_\nu \tau^\nu a = \sigma \sum_{\nu=0}^{k-1} \eta_\nu \tau^\nu a . \tag{15.44}$$

Applying τ^{k-1} we obtain

$$\xi_0 \tau^{k-1} a = \eta_0 \sigma \tau^{k-1} a .$$

This equation implies that $\xi_0 = 0$ and $\eta_0 = 0$ because otherwise, $\tau^{k-1}a$ would be an eigenvector of σ. But σ does not have eigenvectors because the minimal polynomial has no real roots. Now the relation (15.44) can be written as

$$\sum_{\nu=1}^{k-1} \xi_\nu \tau^\nu a = \sigma \sum_{\nu=1}^{k-1} \eta_\nu \tau^\nu a .$$

Applying τ^{k-2} to this equation we find that $\xi_1 = 0$ and $\eta_1 = 0$. Continuing this way we see that $\xi_\nu = 0$ and $\eta_\nu = 0$ $(\nu = 0 \ldots k-1)$.

To construct a complementary invariant subspace to V consider the dual space E^*. Since the vector $\tau^{k-1}a$ is not contained in σU the orthogonal complement $(\sigma U)^{\perp}$ is not contained in the orthogonal complement of $\tau^{k-1}a$. In other words, there exists a vector $\overset{*}{a} \in \sigma(U)^{\perp}$ such that

$$\langle \overset{*}{a}, \tau^{k-1}a \rangle \neq 0 .$$

Let U^* be the space generated by the vectors

$$\overset{*}{a}, \tau^* \overset{*}{a}, \ldots, \tau^{*k-1} \overset{*}{a} .$$

Then U^* is invariant under τ^* and

$$U^* \cap \sigma^*(U^*) = 0 .$$

Moreover, the spaces U^* and $\sigma(U)$ are orthogonal. To prove this, observe that

$$\langle \tau^{*\nu} \overset{*}{a}, \sigma\tau^{\mu} a \rangle = \langle \overset{*}{a}, \sigma\tau^{\nu+\mu} a \rangle ;$$

now

$$\overset{*}{a} \in \sigma(U)^{\perp} \quad \text{and} \quad \sigma\tau^{\mu+\nu} a \in \sigma(U)$$

and hence

$$\langle \tau^{*\nu} \overset{*}{a}, \sigma\tau^{\mu} a \rangle = \langle \overset{*}{a}, \sigma\tau^{\mu+\nu} a \rangle = 0 .$$

Now it follows that the spaces

$$V = U \oplus \sigma(U) \quad \text{and} \quad V^* = U^* \oplus \sigma^*(U^*)$$

are dual: Let $u_1 + \sigma u_2$ $(u_1, u_2 \in U)$ be a vector of V which is orthogonal to all vectors of V^*. Then

$$\langle \overset{*}{u}_1 + \sigma \overset{*}{u}_2, u_1 + \sigma u_2 \rangle = 0 \tag{15.45}$$

for all vectors $\overset{*}{u}_1 \in U^*$ and $\overset{*}{u}_2 \in U^*$. Observing that

$$\langle \overset{*}{u}_1, \sigma u_2 \rangle = 0 \quad \text{and} \quad \langle \sigma^* \overset{*}{u}_2, u_1 \rangle = \langle \overset{*}{u}_2, \sigma u_1 \rangle = 0$$

we obtain from (15.45)

$$\langle \overset{*}{u}_1, u_1 \rangle + \langle \sigma^* \overset{*}{u}_2, \sigma u_2 \rangle = 0$$

for all vectors $\overset{*}{u}_1 \in U^*$ and $\overset{*}{u}_2 \in U^*$. This implies that $u_1 = 0$ and $u_2 = 0$.

We thus have two dual subspaces $V \subset E$ and $V^* \subset E^*$ which are invariant under σ and σ^* respectively.

This yields the direct decomposition

$$E = V \oplus (V^*)^{\perp}$$

of E (cf. sec. 15.10).

The minimal polynomial of the induced endomorphism of $(V^*)^{\perp}$ has again the form

$$\mu_1 = (u^2 + \alpha u + \beta)^{k_1} \rightarrow k_1 \leqq k$$

and hence the above construction can be repeated with $(V^*)^{\perp}$. In this way we finally obtain a direct decomposition

$$E = V_1 \oplus V_2 \oplus \cdots \oplus V_m \to V_1 = V \tag{15.46}$$

into invariant subspaces. Every subspace V_μ is a direct sum $U_\mu \oplus \sigma(U_\mu)$ and hence it has even dimension. This implies especially that the dimension of V_μ is even. Since there is at least one subspace of the dimension $2k$ in the decomposition (15.46) it follows that $2k \leqq n$, where $n = \dim E$.

15.22. The irreducibility of the spaces V_μ. It remains to be shown that the subspaces V_μ are irreducible. This is a consequence of the following criterion: *Let σ be an endomorphism of an n-dimensional real linear space such that the minimal polynomial of σ has the form*

$$\mu = (u^2 + \alpha u + \beta)^k \qquad \alpha^2 - 4\beta < 0 .$$

Then E is irreducible if and only if $n = 2k$.

We note first that $n \geqq 2k$. If $n > 2k$, the decomposition (15.46) consists of at least two subspaces and hence E is reducible. Now assume that $n = 2k$. To show that E is then irreducible, assume a decomposition

$$E = E_1 \oplus E_2 \tag{15.47}$$

into two invariant subspaces. Denote by μ_1 and μ_2 the minimal polynomials of the induced endomorphisms. Then

$$\mu_1 = (u^2 + \alpha u + \beta)^{k_1}, \quad k_1 \leqq k \quad \text{and} \quad \mu_2 = (u^2 + \alpha u + \beta)^{k_2}, \quad k_2 \leqq k .$$

Let $a \in E$ be a vector such that $\tau^{k-1} a \neq 0$. This vector can be written as

$$a = a_1 + a_2 \qquad a_1 \in E_1,\ a_2 \in E_2 .$$

Applying τ^{k-1} we obtain

$$\tau^{k-1} a = \tau^{k-1} a_1 + \tau^{k-1} a_2 .$$

Since $\tau^{k-1} a \neq 0$, at least one of the vectors $\tau^{k-1} a_1$ and $\tau^{k-1} a_2$ must be different from zero. Assume that $\tau^{k-1} a_1 \neq 0$. Then $k - 1 \leqq k$. At the same time we know that $k_1 \leqq k$. These two relations imply that $k_1 = k$, whence

$$\dim E_1 \geqq 2k_1 = 2k = n = \dim E .$$

This relation shows that E_1 coincides with E, and consequently the decomposition (15.47) is improper.

15.23. The normal form of a real matrix. Let V be a real linear space of dimension $n = 2k$ and σ be an endomorphism of V having the minimal polynomial

$$\mu = (u^2 + \alpha u + \beta)^k \quad \alpha^2 - 4\beta < 0 .$$

Then there exists a basis of V relative to which the matrix of σ has the form

$$\begin{pmatrix} 0 & 1 & & & & & & \\ -\beta & -\alpha & 1 & & & & & \\ & & 0 & 1 & & & & \\ & & -\beta & -\alpha & \ddots & & & \\ & & & & \ddots & 1 & & \\ & & & & & 0 & 1 \\ & & & & & -\beta & -\alpha \end{pmatrix}. \tag{15.48}$$

In fact, V can be written as

$$V = U \oplus \sigma(U)$$

where U is a subspace generated by vectors of the form

$$a, \tau a, \ldots, \tau^{k-1} a, \quad \tau = \sigma^2 + \alpha\sigma + \beta\iota .$$

Then the vectors

$$a_\nu = \tau^\nu a, \qquad b_\nu = \sigma\tau^\nu a \qquad (\nu = 0\,1 \ldots k-1)$$

form a basis of V and the equations

$$\begin{aligned} &\sigma a_\nu = b_\nu \\ &\sigma b_\nu = a_{\nu+1} - \beta a_\nu - \alpha b_\nu \quad \nu = 1 \ldots k; \quad a_{k+1} = 0 \end{aligned} \tag{15.49}$$

show that the corresponding matrix has the form (15.48).

In the general case consider the direct decomposition of E into irreducible invariant subspaces. In every subspace there exists a basis relative to which the matrix of the induced endomorphism has one of the forms (15.33) and (15.48).

Combining the bases of all irreducible subspaces we thus obtain a basis of E in which the matrix of σ consists of submatrices of the forms (15.48) and (15.33) following each other successively along the main-diagonal.

Problems: 1. Assume that σ is an endomorphism of a 2-dimensional real linear space E such that E is irreducible. Prove that there exists a basis of E such that the matrix of σ has one of the forms

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \quad \beta \neq 0 .$$

2. Let σ be an endomorphism of a real 3-dimensional linear space such that E is irreducible. Show that there exists a basis of E such that the matrix of σ has the form

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

3. Prove the Cayley-Hamilton-theorem for an endomorphism of a real linear space.

4. Consider an endomorphism σ of a real linear space whose minimal polynomial has the form

$$\mu = (u^2 + \alpha u + \beta)^k \qquad \alpha^2 - 4\beta < 0 .$$

Prove that the number N_j of the $2j$-dimensional irreducible subspaces is given by

$$2N_j = r(\tau^{j+1}) + r(\tau^{j-1}) - 2r(\tau^j) \qquad (j = 1 \ldots k)$$

where

$$\tau = \sigma^2 + \alpha\sigma + \beta\iota .$$

5. Show that a real linear space is irreducible with respect to an endomorphism σ if and only if the minimal polynomial of σ has the form

$$\mu = (u - \lambda)^n \quad \text{or} \quad \mu = (u^2 + \alpha u + \beta)^k \quad \alpha^2 - 4\beta < 0 \quad n = 2k .$$

6. Prove that an n-dimensional real linear space is completely reducible with respect to σ if and only if the minimal polynomial has the form

$$\mu = (u - \lambda_1) \ldots (u - \lambda_n) \qquad \lambda_i \neq \lambda_j \quad (i \neq j) .$$

7. Show that every endomorphism of a real inner product space is similar to the adjoint endomorphism.

§ 6. Applications to inner product spaces

In this concluding paragraph we shall apply our general decomposition theorems to inner product spaces. Irreducible decompositions of an inner product space with respect to selfadjoint mappings, skew mappings and isometries have already been constructed in chap. XI.

Generalizing these results we shall now construct an irreducible decomposition for a *normal* endomorphism. Since a complex linear space is completely reducible with respect to a normal endomorphism (cf. sec. 14.10) we can restrict ourselves to real inner product spaces.

15.24. Normal mappings. First of all, we state the following properties of a normal mapping:

1. If σ is normal and f is an arbitrary polynomial with real coefficients, then $f(\sigma)$ is also normal.

2. A nilpotent normal mapping is the zero-mapping.

3. Let σ be a normal mapping such that

$$\sigma^2 + \alpha\sigma + \beta\iota = 0 , \qquad \alpha^2 - 4\beta < 0 .$$

Define the mappings ϱ and τ by

$$\varrho = \frac{1}{2}(\sigma + \tilde{\sigma}) \quad \text{and} \quad \tau = \frac{1}{2}(\sigma - \tilde{\sigma}) .$$

Then

$$\varrho = -\frac{\alpha}{2}\iota \qquad \tau = \left(\frac{\alpha^2}{4} - \beta\right)\iota \qquad \tilde{\sigma}\circ\sigma = \beta\iota . \tag{15.50}$$

4. Consider two irreducible monic polynomials p_1 and p_2 ($p_1 \neq p_2$) and denote by K_i the kernel of the mapping $p_i(\sigma)$ $(i = 1, 2)$. Then K_1 is orthogonal to K_2.

The first property follows immediately from the definition of a normal mapping. Property 2 is a consequence of the fact that all the mappings σ^ν $(\nu = 1, 2 \ldots)$ have the same rank.

To prove the formulas (15.50) we observe first that the mapping τ is regular. In fact, the kernel $K(\tau)$ consists of all vectors x with the property that

$$\tilde{\sigma}x = \sigma x .$$

Applying σ to this relation we find that

$$\sigma\,\tilde{\sigma}x = \sigma^2 x$$

whence

$$\tilde{\sigma}(\sigma x) = \sigma(\sigma x) .$$

This equation shows that $K(\tau)$ is invariant under σ. The induced endomorphism is obviously selfadjoint and consequently, there must be an eigenvector in $K(\tau)$ unless this space reduces to the zero-vector. But the inequality $\alpha^2 - 4\beta < 0$ implies that σ cannot have an eigenvector and consequently $K(\tau) = 0$.

Subtracting the equations

$$\sigma^2 + \alpha\sigma + \beta\iota = 0 \tag{15.51}$$

and

$$\tilde{\sigma}^2 + \alpha\tilde{\sigma} + \beta\iota = 0 \tag{15.52}$$

we find that

$$\tau \circ (2\varrho + \alpha\iota) = 0 .$$

This equation and the regularity of τ imply the first formula (15.50). Similarly we derive form (15.51) and (15.52) that

$$\tau \circ (\tilde{\sigma}\circ\sigma - \beta\iota) = 0$$

which gives the third relation (15.50). The second formula (15.50) follows from the first and the third one and the identity

$$(\tilde{\sigma} + \sigma)^2 - (\tilde{\sigma} - \sigma)^2 = 4\sigma\circ\tilde{\sigma} .$$

To prove the property 4, consider two arbitrary vectors $x_1 \in K_1$ and $x_2 \in K_2$. Now we distinguish three cases:

a) p_1 and p_2 both have the degree 1,

$$p_1 = u - \lambda_1, \; p_2 = u - \lambda_2 \qquad \lambda_1 \neq \lambda_2 .$$

Then

$$\sigma x_1 = \lambda_1 x_1 \quad \text{and} \quad \sigma x_2 = \lambda x_2 .$$

These equations yield

$$(\lambda_1 - \lambda_2)(x_1, x_2) = 0$$

whence $(x_1, x_2) = 0$.

b) p_1 has the degree 1 and p_2 has the degree 2,

$$p_1 = u - \lambda, \; p_2 = u^2 + \alpha u + \beta .$$

Then

$$\sigma x_1 = \lambda x_1 \quad \text{and} \quad \sigma^2 x_2 + \alpha \sigma x_2 + \beta x_2 = 0 .$$

It follows from these equations that

$$(\lambda^2 + \alpha\lambda + \beta)(x_1, x_2) = 0 .$$

The irreducibility of p_2 implies that $\lambda^2 + \alpha\lambda + \beta \neq 0$ and hence we obtain $(x_1, x_2) = 0$.

c) p_1 and p_2 both have the degree 2,

$$p_1 = u^2 + \alpha_1 u + \beta_1 \qquad p_2 = u^2 + \alpha_2 u + \beta_2 .$$

Then the relations (15.50) yield

$$\begin{aligned} \varrho x_1 &= -\frac{\alpha_1}{2} x_1 & \varrho x_2 &= -\frac{\alpha_2}{2} x_2 \\ \tau x_1 &= \left(\frac{\alpha_1^2}{4} - \beta_1\right) x_1 & \tau x_2 &= \left(\frac{\alpha_2^2}{4} - \beta_2\right) x_2 . \end{aligned} \tag{15.53}$$

Observing that the mappings ϱ and τ are selfadjoint we obtain from (15.53) the equations

$$(\alpha_1 - \alpha_2)(x_1, x_2) = 0 \quad \text{and} \quad (\beta_1 - \beta_2)(x_1, x_2) = 0 . \tag{15.54}$$

The hypothesis $p_1 \neq p_2$ implies that $\alpha_1 \neq \alpha_2$ or $\beta_1 \neq \beta_2$. Hence, it follows from (15.54) that $(x_1, x_2) = 0$.

15.25. The decomposition of E. We now proceed to the construction of a decomposition into irreducible subspaces. We note first that every prime-factor of the minimal polynomial occurs with the exponent 1. In fact, let f^k be a prime-factor of multiplicity k. Denote by σ_1 the endomorphism which is induced in the kernel of $f(\sigma)^k$. Then $f(\sigma_1)$ is normal and nilpotent of order k. Now the property 2 of sec. 15.24 implies that $k = 1$.

It follows from the above remark that the decomposition of the minimal polynomial of σ has the form

$$\mu = f_1 \cdots f_r \tag{15.55}$$

where the f_ϱ are polynomials of degree 1 or 2. From (15.55) we obtain the direct decomposition

$$E = K_1 \oplus \cdots \oplus K_r$$

of E where K_ϱ denotes the kernel of $f_\varrho(\sigma)$. The property 4 in sec. 15.24 implies that any two subspaces K_i and K_j $(i \neq j)$ are orthogonal.

If f_i has the degree 1, the subspace K_i is completely reducible. Now consider a polynomial f_i of degree 2. Then the induced endomorphism has a minimal polynomial of the form

$$\mu = u^2 + \alpha u + \beta \qquad \alpha^2 - 4\beta < 0 .$$

Not it follows from the third relation (15.50) that

$$(\sigma x, \sigma y) = (\tilde{\sigma}\sigma x, y) = \beta(x, y) \quad x, y \in K_i \quad (i = 1, 2) .$$

In other words, the induced mapping is homothetic. Hence, K_i can be decomposed into invariant planes which are mutually orthogonal.

We thus obtain an orthogonal decomposition of E into irreducible invariant subspaces of dimension 1 and 2. Now select an orthonormal basis in each one of the irreducible subspaces. Then these basis-vectors determine an orthonormal basis of E in which the matrix of σ has the form

$$\begin{pmatrix} \varrho_1 & \sigma_1 & & & & & & \\ -\sigma_1 & \varrho_1 & & & & & & \\ & & \ddots & & & & & \\ & & & \varrho_p & \sigma_p & & & \\ & & & -\sigma_p & \varrho_p & & & \\ & & & & & \lambda_1 & & \\ & & & & & & \ddots & \\ & & & & & & & \lambda_{n-2p} \end{pmatrix} \qquad \begin{aligned} 2\varrho_\nu &= -\alpha_\nu \\ \varrho_\nu^2 + \sigma_\nu^2 &= \beta_\nu \end{aligned}$$

15.26. Lorentz-transformations. As a second example we shall construct an irreducible decomposition of the Minkowski-space with respect to a Lorentz-transformation σ (cf. sec. 12.28). For the sake of simplicity we assume that the Lorentz-transformation is proper orthochroneous. The condition $\tilde{\sigma} = \sigma^{-1}$ implies that the characteristic polynomial of σ has the form

$$\chi = u^4 + u^3 + u^2 + u + 1 . \tag{15.56}$$

It follows from (15.56) that the inverse of every eigenvalue is again an eigenvalue. Since there exists at least one eigenvalue (cf. sec. 12.28) the polynomial (15.56) has at least one pair of real roots. Now we distinguish 3 cases:

I. The characteristic polynomial χ contains a prime-factor

$$u^2 + \alpha u + \beta \qquad \alpha^2 - 4\beta < 0$$

of second degree. Then consider the mapping

$$\tau = \sigma^2 + \alpha\sigma + \beta\iota .$$

The kernel of τ is an invariant subspace F of even dimension. This subspace has necessarily dimension 2 and hence it is a plane. The intersection of the plane F and the light-cone consists of two straight lines, one straight line, or the point 0 only. The two first cases are impossible because the plane F does not contain eigenvectors. Thus, the inner product must be positive definite in F and the induced endomorphism σ_1 is a proper Euclidean rotation*). Now consider the orthogonal complement $F^{\perp}$. The restriction of the inner product to $F^{\perp}$ has index 1. Hence, $F^{\perp}$ is a pseudo-Euclidean plane. Denote by σ_2 the induced endomorphism of $F^{\perp}$. The equation

$$\det\sigma = \det\sigma_1 \cdot \det\sigma_2$$

implies that $\det\sigma_2 = +1$, showing that σ_2 is a *proper* pseudo-Euclidean rotation. Choosing orthogonal bases in F and in $F^{\perp}$ we obtain an orthonormal basis of E in which the matrix of σ has the form

$$\begin{pmatrix} \cos\omega & \sin\omega & & \\ -\sin\omega & \cos\omega & & \\ & & \cosh\theta & \sinh\theta \\ & & \sinh\theta & \cosh\theta \end{pmatrix} \quad \omega \neq 0,\ \omega \neq \pi .$$

II. The characteristic polynomial has four real roots which are not all equal to 1. Then σ has two eigenvalues λ and $\frac{1}{\lambda}$ $(\lambda \neq 1)$. Let e and e' be the corresponding eigenvectors

$$\sigma e = \lambda e \qquad \sigma e' = \frac{1}{\lambda} e' .$$

The condition $\lambda \neq 1$ implies that e and e' are light-vectors. These vectors are linearly independent, whence $(e, e') \neq 0$ (cf. sec. 12.22). Let F be the plane generated by e and e' and

$$z = \xi e + \eta e'$$

be a vector of F. Then

$$(z, z) = 2(e, e')\,\xi\eta .$$

*) An improper rotation of F would have eigenvectors. This is in contradiction with the fact that the minimal polynomial of σ_1 is an irreducible polynomial of degree 2.

This equation shows that the induced inner product has index 1. The orthogonal complement $F^\perp$ is therefore a Euclidean plane and the induced mapping is a Euclidean rotation. The angle of this rotation must be 0 or π, because otherwise the characteristic polynomial of σ would contain an irreducible factor of second degree. Select orthonormal bases in F and $F^\perp$. These two bases determine an orthonormal basis of E in which the matrix of σ has the form

$$\begin{pmatrix} \cosh\theta & \sinh\theta & & \\ \sinh\theta & \cosh\theta & & \\ & & \varepsilon & 0 \\ & & 0 & \varepsilon \end{pmatrix} \qquad \theta \neq 0,\ \varepsilon = \pm 1 .$$

III. The characteristic polynomial has the root 1 with the multiplicity four. Then the minimal polynomial of σ has the form

$$\mu = (u-1)^k \qquad (1 \leqq k \leqq 4) .$$

If $k = 1$, σ reduces to the identity map. Next, it will be shown that the case $k = 2$ is impossible. If $k = 2$, it follows that

$$\sigma + \tilde{\sigma} = 2\iota$$

whence

$$(x, \sigma x) = (x, x) \qquad x \in E .$$

Inserting a light-vector l for x we see that $(l, \sigma l) = 0$. But two light-vectors can be orthogonal only if they are linearly dependent. We thus obtain $\sigma l = \lambda l$. Since σ does not have eigenvalues $\lambda \neq 1$, it follows that $\sigma l = \pm l$ for all light-vectors l. But this implies that σ is the identity. Hence the minimal polynomial is $u - 1$ in contradiction to our assumption $k = 2$.

Now consider the case $k \geqq 3$. As it has been shown in sec. 12.28 there exists an eigenvector e on the light-cone. The orthogonal complement E_1 of e is a 3-dimensional subspace of E which contains the light-vector e. The induced inner product has the rank and the index 2. Let F be a 2-dimensional subspace of E_1 in which the inner product is positive definite. Selecting an orthonormal basis e_1, e_2 in F we can write

$$\begin{aligned} \sigma e_1 &= e_1 \cos\omega + e_2 \sin\omega + \alpha_1 e \\ \sigma e_2 &= -e_1 \sin\omega + e_2 \cos\omega + \alpha_2 e \\ \sigma e &= \lambda e . \end{aligned} \qquad (15.57)$$

The coefficients α_1 and α_2 are not both zero. In fact, if $\alpha_1 = 0$ and $\alpha_2 = 0$, the plane F is invariant under σ and we have the direct decomposition $E = F \oplus F^\perp$ of E into two 2-dimensional invariant subspaces. This would imply that $k \leqq 2$.

Now consider the characteristic polynomial of the induced endomorphism $\sigma_1 : E_1 \to E_1$. Computing the characteristic polynomial from the matrix (15.57) we find that

$$\chi_1 = (u^2 - 2u\cos\omega + 1)(\lambda - u)\,. \tag{15.58}$$

At the same time we know that

$$\chi_1 = (1 - u)^3\,. \tag{15.59}$$

Comparing the polynomials (15.58) and (15.59) we find that $\omega = 0$ and $\lambda = 1$. Hence, the equations (15.57) reduce to

$$\begin{aligned} \sigma e_1 &= e_1 + \alpha_1 e \\ \sigma e_2 &= e_2 + \alpha_2 e \\ \sigma e &= e\,. \end{aligned}$$

Now consider the vector

$$y = \alpha_1 e_2 - \alpha_2 e_1\,.$$

Then

$$(y, y) = \alpha_1^2 + \alpha_2^2 > 0$$

and

$$\sigma y = \alpha_1 \sigma e_2 - \alpha_2 \sigma e_1 = \alpha_1 (e_2 + \alpha_2 e) - \alpha_2 (e_1 + \alpha_1 e) = y\,.$$

In other words, y is a space-like fix-vector of σ. Denote by Y the 1-dimensional subspace generated by y. Then we have the orthogonal decomposition

$$E = Y \oplus Y^\perp$$

into two invariant subspaces. The orthogonal complement $Y^\perp$ is a 3-dimensional pseudo-Euclidean space with index 2.

The subspace $Y^\perp$ is irreducible with respect to σ. This follows from our hypothesis that the degree of the minimal polynomial μ is $\geqq 3$. At the same time we see that μ can not have the degree 4 because then the space E would be irreducible.

Combining our results we see that the decomposition of a Minkowski-space E with respect to a proper orthochroneous Lorentz-transformation σ has one of the following forms:

I. E is completely reducible. Then σ is the identity.

II. E is the direct sum of an invariant Euclidean plane and an invariant pseudo-Euclidean plane. These planes are irreducible except for the case where the induced endomorphism is $\pm\iota$ (Euclidean plane) or ι (pseudo-Euclidean plane).

III. E is the direct sum of a space-like 1-dimensional subspace and an irreducible subspace of dimension 3 and index 2.

Problems: 1. Let E be a real linear space and σ an endomorphism of E. Prove that a positive definite inner product can be introduced in E such that σ becomes a normal mapping if and only if the following conditions are satisfied:

a) The space E can be decomposed into invariant subspaces of dimension 1 and 2.

b) If τ is the induced endomorphism in an irreducible subspace of dimension 2, then

$$\frac{1}{4}(\operatorname{tr}\tau)^2 - \det\tau < 0 .$$

2. Consider a 3-dimensional pseudo-Euclidean space E with the index 2. Let $l_i (i = 1, 2, 3)$ be three light-vectors such that.

$$(l_i, l_j) = 1 \qquad i \neq j .$$

Define the endomorphism σ by the equations

$$\begin{aligned} \sigma l_1 &= l_1 \\ \sigma l_2 &= \alpha(\alpha - 1)\, l_1 + \alpha l_2 + (1 - \alpha)\, l_3 \qquad \alpha \neq 1 \\ \sigma l_3 &= (\alpha - 2)\,(\alpha - 1)\, l_1 + (\alpha - 1)\, l_2 + (2 - \alpha)\, l_3 . \end{aligned}$$

Prove that σ is a rotation and that E is irreducible with respect to σ.

3. Let the Lorentz-transformation σ be defined by the matrix

$$\begin{pmatrix} \frac{1}{2} & \frac{2}{3} & -1 & -\frac{5}{6} \\ \frac{2}{3} & \frac{1}{9} & \frac{4}{3} & \frac{10}{9} \\ 1 & -\frac{4}{3} & 1 & \frac{5}{3} \\ \frac{5}{6} & -\frac{10}{9} & \frac{5}{3} & \frac{43}{18} \end{pmatrix}$$

Construct a decomposition into irreducible subspaces.

Bibliography

BAER, R.: Linear Algebra and Projective Geometry, New York: Academic Press Inc. 1952.

BIRKHOFF, G., and S. MAC LANE: A Survey of Modern Algebra. New York: The MacMillan Company 1944.

BOURBAKI, N.: Elements de mathematique. Premiere Partie, Livre II.

GANTMACHER, F. R.: Matrizenrechnung. Berlin: Deutscher Verlag der Wissenschaften 1958.

GRÖBNER, W.: Matrizenrechnung. München: R. Oldenbourg 1956.

HALMOS, P.: Finite Dimensional Vector-spaces. Princeton: D. Van Nostrand Co. 1958.

HOFFMAN, K., and R. KUNZE: Linear Algebra. Prentice Hall Inc. 1961.

JACOBSON, N.: Lectures in Abstract Algebra, II. Princeton: D. Van Nostrand Co. 1953.

KELLER, O. H.: Analytische Geometrie und lineare Algebra. Berlin: Deutscher Verlag der Wissenschaften 1957.

LICHNEROWICZ, A.: Lineare Algebra und Lineare Analysis. Berlin: Deutscher Verlag der Wissenschaften 1956.

NEVANLINNA, R., u. F.: Absolute Analysis. Grundlehren Math. Wiss. 102, Berlin-Göttingen-Heidelberg: Springer 1958.

PICKERT, G.: Analytische Geometrie. Leipzig: Akademische Verlagsgesellschaft 1953.

REICHARDT, H.: Vorlesungen über Vector- und Tensorrechnung. Berlin: Deutscher Verlag der Wissenschaften 1957.

SCHREIER, O., and E. SPERNER: Modern Algebra and Matrix Theory. New York: Chelsea Publishing Co. 1955.

SMIRNOW, W. I.: Lehrgang der höheren Mathematik Teil III, 1. Berlin: Deutscher Verlag der Wissenschaften 1958.

STOLL, R.: Linear Algebra and Matrix Theory. London-New York-Toronto: McGraw-Hill Book Publ. Comp. Inc. 1952.

VAN DER WAERDEN, B. L.: Algebra. Grundlehren Math. Wiss. Berlin-Göttingen-Heidelberg: Springer 1955.

Subject Index