PROJECTIVE GEOMETRY

PROJECTIVE GEOMETRY

of n dimensions

VOLUME TWO OF INTRODUCTION TO MODERN ALGEBRA AND MATRIX THEORY

BY

OTTO SCHREIER

EMANUEL SPERNER

CHELSEA PUBLISHING COMPANY

NEW YORK

First English-Language Edition

The present work and Introduction to Modern Algebra and Matrix Theory together constitute a complete and unabridged translation of the German-language book Einfuehrung in die analytische Geometrie und Algebra (volumes one and two) by Otto Schreier and Emanuel Sperner. The present work is translated by the late Professor Calvin A. Rogers.

Library of Congress Catalog Card No. 60-8967

Printed in the United States of America

EDITOR'S PREFACE

With the present publication of *Projective Geometry,* the project of translating the famous German-language textbook *Einfuhrung in die analytische Geometrie und Algebra,* by Otto Schreier and Emanuel Sperner, originally published in two volumes, is now complete. As is well known, the purpose of that textbook was to offer a course in Algebra and Analytic Geometry which, when supplemented with a course on the Calculus, would give the student all he needs for a profitable continuation of his studies in modern mathematics. The Preface to the German Edition (see below) gives a more detailed description of the two volumes.

The only change that has been made has been to divide the two volumes somewhat differently in order that they might be usable independently. The first volume and the early part of the second volume were combined into a single book under the title *Introduction to Modern Algebra and Matrix Theory.* The balance, consisting of the major portion of the second volume is published herewith as *Projective Geometry of n Dimensions.* The titles of the two books indicate their respective contents.

The chief prerequisite for reading the present book, aside from a few elementary facts about affine space and systems of linear equations, is a knowledge of the elements of matrix theory such as is contained, for example, in the first four sections of Chapter V (Linear Transformations and Matrices) of *Introduction to Modern Algebra and Matrix Theory.*

Professor Calvin A. Rogers, the translator of the present volume, died before the preparation of the manuscript for the press was begun. The numerous questions that always call for consultation between editor and translator were referred to Professor Abe Shenitzer, whom the Editor wishes to thank for his very considerable help. The Editor also wishes to thank Professor F. Steinhardt. The final form of the manuscript is, of course, the responsibility of the Editor alone.

FROM THE PREFACE TO THE GERMAN EDITION

Otto Schreier had planned, a few years ago, to have his lectures on Analytic Geometry and Algebra published in book form. Death overtook him in Hamburg on June 2, 1929, before he had really begun to carry out his plan. The task of doing this fell on me, his pupil. I had at my disposal some sets of lecture notes taken at Schreier's courses, as well as a detailed (if not quite complete) syllabus of his course drawn up at one time by Otto Schreier himself. Since then, I have also given the course myself, in Hamburg, gaining experience in the process.

In writing this textbook,[1] which is to be published in two volumes, I have followed Schreier's own presentation as closely as possible, so that it might retain the characteristics impressed on the subject matter in Otto Schreier's treatment. In particular, as regards choice and arrangement of material, I have followed Schreier's outline faithfully, except for a few changes of minor importance.

This textbook is motivated by the idea of offering the student, in two basic courses on Calculus and Analytic Geometry, all that he needs for a profitable continuation of his studies in accordance with modern requirements. It is evident that this implies a stronger emphasis than has been customary on algebra, in line with the recent developments in that subject.

The prerequisites for reading this book are few indeed. For the early parts, a knowledge of the real number system—such as is acquired in the first few lectures of almost any calculus course—is sufficient. The later chapters make use of some few theorems on continuity of real functions and on sequences of real numbers. These also will be familiar to the student from the calculus. In some sections which give intuitive interpretations of the subject matter, use is made of some well-known theorems of elementary geometry, whose derivation on an axiomatic basis would of course be beyond the scope of this text.

[1] See the preceding Editor's Preface.

What the book contains may be seen in outline by a glance at the table of contents. The student is urged not to neglect the exercises at the end of each section; among them will be found many an important addition to the material presented in the text.

• • •

The authors' earlier book on matrices has been incorporated into [Chapter V of *Introduction to Modern Algebra and Matrix Theory*],[1] with a few re-arrangements and omissions in order to achieve a more organic whole. The arrangement of material in this chapter is such that the first four sections of the chapter contain essentially all that is needed for [*Projective Geometry*].

• • •

To Mr. W. Blaschke (Hamburg) I owe a debt of gratitude for his continuous interest and help. I also wish to thank Messrs. O. Haupt (Erlangen) and K. Henke (Hamburg) for many valuable hints and suggestions. In preparing the manuscript, I have had the untiring assistance of my wife. For reading the proofs I am indebted to Mr. H. Bückner (Königsberg) in addition to those named above.

Königsberg, October 1935

EMANUEL SPERNER

[1] See the preceding Editor's Preface.

CONTENTS

CHAPTER I

n-DIMENSIONAL PROJECTIVE SPACE

For certain geometrical questions, whose study is central to this book, it is advantageous to extend affine (or euclidean) space by adding to it certain new points, the so-called points at infinity. This procedure is suggested by quite elementary geometrical facts. For example, in order to avoid the oftentimes awkward distinction between intersecting and parallel lines in a plane, we are tempted to ascribe to parallel lines a point of intersection 'at infinity.' Another case in point is afforded by the fact when one line of the affine plane is projected onto another by means of central projection[1] this does not in general establish a one-to-one correspondence between the points of these two lines, whereas it may be made into such a correspondence by an appropriate adjunction of points at infinity. The same is true for the central projection of two planes in space upon each other.

Our immediate task, then, will be to establish and to give a precise analytic description of the introduction of these points at infinity.

Extension of the Affine Plane to the Projective Plane

Because of its intuitive appeal, we shall start with the two-dimensional case.

We shall first of all introduce new coordinates in the affine plane (the so-called homogeneous coordinates). In doing this, we begin with

[1] The central projection upon each other of two lines g and h with respect to a center of projection S is defined by the following rule: P, on g, is taken as the image of Q, on h, and conversely, Q is taken as the image of P, if P, Q, and S lie on a line.

It is therefore clear that P_0 on g has no image point on h if P_0S is parallel to h. Similarly, Q_0 has no image point on g if Q_0S is parallel to g. If we let P_0 correspond to *one* point at infinity on h and Q_0 to *one* point at infinity on g, then exactly one point of h is associated with each point of g, and conversely linear coordinates and hence take as our starting point a *fixed* linear coordinate system in the plane. We get in this way a definite one-to-one correspondence between the points of the plane and the ordered pairs of real numbers. If a point P has the coordinates x_1, x_2, we write $P = (x_1, x_2)$.

Next, we consider all the *ordered triples* of real numbers (ξ_0, ξ_1, ξ_2) for which $\xi_0 \neq 0$. These number triples and the points of the plane are now put into correspondence by means of the following rule:

$P = (x_1, x_2)$ and an ordered triple (ξ_0, ξ_1, ξ_2) with $\xi_0 \neq 0$ are to correspond to each other if and only if:

$$x_1 = \frac{\xi_1}{\xi_0}, \quad x_2 = \frac{\xi_2}{\xi_0}.$$

It follows immediately from this that to each triple (ξ_0, ξ_1, ξ_2) there corresponds only one point, namely, the point with linear coordinates $\frac{\xi_1}{\xi_0}, \frac{\xi_2}{\xi_0}$. On the other hand, to each point $P = (x_1, x_2)$ there correspond infinitely many number-triples. For, the point P obviously corresponds to the triples (ξ_0, ξ_1, ξ_2) and $(\lambda\xi_0, \lambda\xi_1, \lambda\xi_2)$ for arbitrary real $\lambda \neq 0$, since

$$\frac{\xi_i}{\xi_0} = \frac{\lambda\,\xi_i}{\lambda\,\xi_0} \quad (i = 1, 2).$$

Furthermore, the following holds: If two number-triples (ξ_0, ξ_1, ξ_2) and (ξ'_0, ξ'_1, ξ'_2) with $\xi_0, \xi'_0 \neq 0$, correspond to the same point, then there exists a $\lambda \neq 0$ such that $\xi'_i = \lambda\,\xi_i,\ i = 0, 1, 2$. For from $\frac{\xi'_i}{\xi'_0} = \frac{\xi_i}{\xi_0}$ $(i = 1, 2)$ it follows immediately that $\xi'_i = \frac{\xi'_0}{\xi_0}\,\xi_i$. Thus, $\frac{\xi'_0}{\xi_0}$ is the desired λ.

Hence, it is also evident that all the triples corresponding to the same point may be obtained from a given one of them (ξ_0, ξ_1, ξ_2) by multiplying it by an arbitrary real $\lambda \neq 0$.

In particular, all the triples associated with the point $P = (x_1, x_2)$ are of the form $(\lambda, \lambda x_1, \lambda x_2)$, since $(1, x_1, x_2)$ is one particular triple of this kind.

Since the numbers ξ_i of one of our triples (ξ_0, ξ_1, ξ_2) uniquely determine the corresponding point, we may regard them as the coordinates of that point. The coordinates introduced with the help of this correspondence are called *homogeneous coordinates*, or *ratio coordinates* (since they are determined only up to a common constant of proportionality).

Now let $P = (x_1, x_2)$ be a fixed point in the plane, distinct, however, from the origin (Fig. 1). If we now set $Q = (\lambda x_1, \lambda x_2)$ and let λ vary from $+1$ to $+\infty$, then the point Q moves along the line g determined by the points O and P (Fig. 1), from P outward to infinity (in the direction of the arrow).

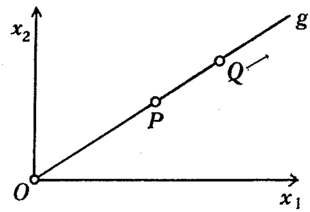

Fig. 1.

We can take $\xi_0 = \frac{1}{\lambda}$, $\xi_1 = x_1$, $\xi_2 = x_2$ as homogeneous coordinates for Q. Then as $\lambda \to \infty$, we have $\xi_0 \to 0, \xi_1 \to x_1, \xi_2 \to x_2$. We are accordingly led to look upon $0, x_1, x_2$ as the homogeneous coordinates of a point *'at infinity'* (or *improper* point). It is clear that only the ratio of the three coordinates is of significance here, for instead of considering the homogeneous coordinates of Q to be $\frac{1}{\lambda}$, x_1, x_2, we could equally well have thought of them as being $\frac{\varrho}{\lambda}, \varrho x_1, \varrho x_2$, with any fixed ϱ (independent of λ). Upon passing to the limit as $\lambda \to \infty$, we then obtain $0, \varrho x_1, \varrho x_2$ as coordinates of the point at infinity of g.

In all of this, the point $P = (x_1, x_2)$ must be different from the origin. Hence we can ascribe to the triple $(0, 0, 0)$ neither a point in the finite part of the plane nor a point at infinity. For this reason, we once and for all rule out the triple $(0, 0, 0)$; it shall not designate any point whatever.

Finally, then, we have the following definition:

Every triple $(0, \xi_1, \xi_2)$ in which not both ξ_1 and ξ_2 vanish is called a point at infinity or, better, an improper point of the plane. Two improper points $(0, \xi_1, \xi_2)$ and $(0, \xi_1', \xi_2')$ are said to be equal (or coincident) whenever there exists a $\lambda \neq 0$ such that $\xi_1 = \lambda \xi_1'$, $\xi_2 = \lambda \xi_2'$.

The plane obtained by the adjunction of these improper points is called the **projective plane**.

In contradistinction to the improper points, all the points of the projective plane that can be represented by a coordinate triple (ξ_0, ξ_1, ξ_2) with $\xi_0 \neq 0$ are called *proper* points. The totality of proper points is called, as before, the *affine plane*.

We now wish to extend these definitions still further. To each line through the origin of the fixed linear coordinate system we have already assigned a point at infinity. Now, is it desirable to do the same for an arbitrary line, and how can this be accomplished? In order to decide, let us first consider the following question: Can the equation of a line be written in homogeneous coordinates?

In the affine plane a line g can always be represented by an equation of the form

$$a_0 + a_1 x_1 + a_2 x_2 = 0. \tag{1}$$

That is to say, the totality of points whose *linear* coordinates x_1, x_2 satisfy equation (1), fill out a line in the affine plane. Now if P is a proper point of the line g, with linear coordinates x_1, x_2 and homogeneous coordinates ξ_0, ξ_1, ξ_2, then it follows, by the substitution of

$$x_1 = \frac{\xi_1}{\xi_0}, \qquad x_2 = \frac{\xi_2}{\xi_0}$$

into (1), that ξ_0, ξ_1, ξ_2 satisfy the equation

$$a_0 \xi_0 + a_1 \xi_1 + a_2 \xi_2 = 0. \tag{2}$$

And the converse is also true. If ξ_0, ξ_1, ξ_2 satisfy equation (2) and $\xi_0 \neq 0$, then $x_1 = \dfrac{\xi_1}{\xi_0}$, $x_2 = \dfrac{\xi_2}{\xi_0}$ satisfy equation (1); that is to say, the triple (ξ_0, ξ_1, ξ_2) represents a point of g.

Thus, we see that equation (2) is satisfied by *all* those triples and *only* those triples (ξ_0, ξ_1, ξ_2), with $\xi_0 \neq 0$, which represent (proper) points of g.[2]

The following definition now suggests itself: *All* those improper points and *only* those improper points whose coordinates satisfy (2) shall belong to g.

How many improper points is that? We claim: Exactly one. For if (ξ_0, ξ_1, ξ_2) is one such point, then the ξ_i must satisfy the following equations:

$$\begin{aligned} a_0 \xi_0 + a_1 \xi_1 + a_2 \xi_2 &= 0, \\ \xi_0 \qquad\qquad\quad &= 0. \end{aligned} \tag{3}$$

This is a system of homogeneous equations in the three unknowns, ξ_0, ξ_1, ξ_2. The rank of the matrix of (3) is 2. For, a_1 and a_2 cannot both

[2] Equation (2) of our line g is homogeneous in the ξ_i, and a similar situation obtains when the equation of any curve is written in terms of these new coordinates; hence the name '*homogeneous coordinates.*'

vanish; else (1) would not represent the equation of a line. According to § 6 of *Modern Algebra*,[3] the totality of the vector solutions $\{\xi_0, \xi_1, \xi_2\}$ of (3) form a one-dimensional linear vector space. That is to say, all the vector solutions are multiples of a *fixed* one among them. This implies, however, that all the triples (ξ_0, ξ_1, ξ_2) that are solutions of (3), with the exception of $(0, 0, 0)$, represent the *same* point in the projective plane (and moreover, by virtue of the second equation of (3), an *improper* point).

We now ask, conversely: Does every homogeneous equation of the form (2) represent a line? Up to now we have seen this to be the case only for such equations of the form (2) as are derivable from an equation of the form (1). In (1), however, a_1 and a_2 must not vanish simultaneously. Let us now consider the case $a_1 = a_2 = 0$. Then (2) reduces to

$$a_0 \xi_0 = 0. \tag{4}$$

If $a_0 = 0$ also, then of course equation (4) no longer represents a line (since (4) is then satisfied by every point of the plane). Thus, let $a_0 \neq 0$. Then (4) is equivalent to:

$$\xi_0 = 0.$$

That is to say: The points that satisfy (4) are precisely all the points at infinity.

Now, for the sake of simplicity, we make the following definition.

The totality of all improper points is called the improper line (or *the line at infinity*).

Thus, we have: *Every homogeneous equation* (2) *in which not all three coefficients vanish, represents a line.*

Now, what can be said about the intersection of two lines conceived of in this extended sense? Let g and h be two lines, g being given by equation (2) and h by

$$b_0 \xi_0 + b_1 \xi_1 + b_2 \xi_2 = 0. \tag{5}$$

The points common to g and h satisfy both equations (2) and (5) and thus are the solutions of the system

[3] *Introduction to Modern Algebra and Matrix Theory,* by O. Schreier and E. Sperner. See Editor's Preface to the present work.

$$(6)\qquad \begin{aligned} a_0\,\xi_0 + a_1\,\xi_1 + a_2\,\xi_2 &= 0,\\ b_0\,\xi_0 + b_1\,\xi_1 + b_2\,\xi_2 &= 0. \end{aligned}$$

The matrix of this system of homogeneous linear equations can have rank 1 or 2.

In the first case, the totality of solutions of (6) is identical with that of each of (2) or (5) separately, that is, the two lines are identical.

In the second case, the vector solutions $\{\xi_0, \xi_1, \xi_2\}$ constitute a one-dimensional linear vector space; that is, there exists exactly one point whose homogeneous coordinates satisfy both the equations (6).

We have thus shown that *any two distinct lines of the projective plane intersect in exactly one point.*

Consequently, parallel lines must also intersect in a point. However, since such lines can have no proper point in common, this point of intersection must be an improper point. From the fact that each line has but one improper point, it follows, in addition, that:

Parallel lines all go through one and the same point at infinity.

On the other hand, non-parallel lines have a finite point of intersection. Their improper points must therefore necessarily be distinct (since two lines have *only one* point of intersection).

In what follows, the definitions that we have adopted for the plane will be generalized to n dimensions ($n > 0$ an arbitrary integer).

n-Dimensional Projective Space

We proceed in complete analogy to the two-dimensional case. We first define homogeneous coordinates in affine R_n by establishing a relation between the points $P = (x_1, x_2, \ldots, x_n)$ of affine R_n and the *ordered* $(n+1)$-tuples $(\xi_0, \xi_1, \ldots, \xi_n)$ of real numbers in which $\xi_0 \neq 0$. This we do in accordance with the following rule.

$P = (x_1, x_2, \ldots, x_n)$ and $(\xi_0, \xi_1, \ldots, \xi_n)$ shall be said to correspond if and only if $x_i = \dfrac{\xi_i}{\xi_0}$ for all $i = 1, 2, \ldots, n$.

According to this rule, we see that, precisely as in the two-dimensional case, just one point of R_n corresponds to each $(n+1)$-tuple $(\xi_0, \xi_1, \ldots, \xi_n)$ with $\xi_0 \neq 0$. Furthermore, two $(n+1)$-tuples $(\xi_0, \xi_1, \ldots, \xi_n)$ and $(\xi'_0, \xi'_1, \ldots, \xi'_n)$ correspond to the same point if and only if there exists a $\lambda \neq 0$ such that $\xi'_0 = \lambda\,\xi_0,\ \xi'_1 = \lambda\,\xi_1,\ \cdots,\ \xi'_n = \lambda\,\xi_n$.

If $P = (x_1, x_2, \ldots, x_n)$ and $(\xi_0, \xi_1, \ldots, \xi_n)$ correspond in accordance with this rule, then the ξ_i are called the *homogeneous coordinates* (or *ratio coordinates*) of P.

The homogeneous coordinates of a point are determined only up to a constant of proportionality; they determine the point, however, uniquely. Let us now adopt the following notation: If a point P has the homogeneous coordinates $\xi_0, \xi_1, \ldots, \xi_n$, we shall write $P = [\xi_0, \xi_1, \ldots, \xi_n]$.[4]

Our final step is the adjunction of the improper points. We adjoin to affine R_n the previously excluded $(n+1)$-tuples $[\xi_0, \xi_1, \ldots, \xi_n]$ in which $\xi_0 = 0$, but in which *not all* the ξ_i vanish simultaneously, and these $(n+1)$-tuples will also be called points; in contradistinction to the '*proper*' points that we have been discussing hitherto, we shall call these new points '*improper*' points (or points '*at infinity*'). Our definition of equality for the improper points (in analogy to that for the proper points) will be as follows: $P = [0, \xi_1, \xi_2, \cdots, \xi_n]$ and $Q = [0, \xi_1', \xi_2', \cdots, \xi_n']$ will be *equal* if and only if there exists a $\lambda \neq 0$ such that $\xi_i = \lambda \xi_i'$ for $i = 1, 2, \ldots, n$.

The extension of affine R_n obtained by adjoining the improper points in this way will be referred to as n-**dimensional projective space** and will be denoted by $\boldsymbol{P}_n$.

We can summarize by saying: Projective P_n consists of the totality of non-trivial[5] ordered $(n+1)$-tuples of real numbers $[\xi_0, \xi_1, \cdots, \xi_n]$, where two such $(n+1)$-tuples $[\xi_0, \xi_1, \cdots, \xi_n]$ and $[\xi_0', \xi_1', \cdots, \xi_n']$ are said to be *equal* (or *coincident*) if and only if there exists a $\lambda \neq 0$ such that $\xi_i = \lambda \xi_i'$ for $i = 0, 1, \ldots, n$.[6]

If $P = [\xi_0, \xi_1, \ldots, \xi_n]$ is a proper point and $x_i = \dfrac{\xi_i}{\xi_0}$ $(i = 1, 2, \ldots, n)$, that is, if $P = (x_1, x_2, \ldots, x_n)$, then we call the x_i the *affine*, or *non-homogeneous*, coordinates of P, in contradistinction to the *homogeneous* coordinates ξ_i.

Now, what is to be understood by a linear subspace in P_n? In affine R_n, a linear subspace of dimension r can always be represented by a system of linear equations

---

[4] We have chosen brackets to avoid confusion with the points of $(n+1)$-dimensional affine space, which we always write in parentheses.

[5] We mean by this the $(n+1)$-tuples in which not all the ξ_i vanish simultaneously. As in the two-dimensional case, we shall once and for all exclude the 'trivial' $(n+1)$-tuple $[0, 0, \ldots, 0]$; it shall not designate any point whatever.

[6] The essential difference between projective P_n and $(n+1)$-dimensional affine R_{n+1} lies in the way in which equality of two points is defined.

$$
(7) \qquad \begin{aligned} a_{10} + a_{11}\, x_1 + \cdots + a_{1n}\, x_n &= 0, \\ a_{20} + a_{21}\, x_1 + \cdots + a_{2n}\, x_n &= 0, \\ \cdots\cdots\cdots\cdots\cdots & \\ a_{m0} + a_{m1}\, x_1 + \cdots + a_{mn}\, x_n &= 0. \end{aligned}
$$

Here the rank of the coefficient matrix and that of the augmented matrix in (7) are both equal to $n - r$.

Substituting $x_i = \dfrac{\xi_i}{\xi_0}$ into (7), we see that the homogeneous coordinates $\xi_0, \xi_1, \ldots, \xi_n$ of a point of this subspace satisfy the homogeneous system of equations

$$
(8) \qquad \sum_{k=0}^{n} a_{ik}\, \xi_k = 0, \qquad i = 1, 2, \cdots, m
$$

and conversely. The rank of the matrix of (8) is equal to $n - r$.

This suggests the following definition:

The totality of the points $[\xi_0, \xi_1, \ldots, \xi_n]$ of P that satisfy a system of equations of the form (8) having a matrix of rank $n - r$ is called an *r-dimensional linear space.* We must assume that the rank of this matrix is $\leq n$. For otherwise, the only solution of (8) would be the trivial solution $0, 0, \ldots, 0$, which does not represent a point of P_n.

We now introduce a few more terms (in analogy to R_n).

A linear space of dimension 1 is called a *line*; a linear space of dimension 2 is called a *plane*; and a linear space of dimension $n - 1$ is called a *hyperplane.* In the case of a hyperplane, the rank of the matrix of (8) is equal to 1, so that in this case one equation always suffices.

A linear space of dimension 0 consists of but a *single* point. For in this case the rank of (8) is equal to n, that is, all the vector solutions $\{\xi_0, \xi_1, \ldots, \xi_n\}$ of (8) are multiples of some fixed one among them ($\neq 0$). Thus, in this case all the $(n+1)$-tuples $[\xi_0, \xi_1, \ldots, \xi_n]$ that are solutions of (8) represent the same point in P_n.

The only linear space of dimension n is the entire projective space P_n. For the rank of the system of equations (8) is then equal to 0, that is, all the coefficients are equal to zero, and this means that *every* point of P_n satisfies the system (8).

It is by no means true that every linear space contains proper points. For if, beginning with any system of equations such as (8), we construct the corresponding system (7), it may happen that (7) is not solvable. In this case (and in this case only), the linear space given by (8) consists *entirely* of improper points. The space itself will then be called 'improper.'

An example of an improper space is the improper hyperplane $\xi_0 = 0$, which consists of all the improper points of P_n and no others.

If L is a linear space of dimension r in P_n that does contain proper points, then the totality of all the proper points of L constitutes a linear space of dimension r in R_n, as is immediately seen by passing from (8) to the system of equations (7). The totality of improper points of L, on the other hand, forms an improper $(r-1)$-dimensional linear space in P_n. For if L is given by (8), then the improper points of L satisfy the system of equations obtained from (8) by adjoining the equation $\xi_0 = 0$, that is to say, the system

(9)
$$\sum_{k=0}^{n} a_{ik}\,\xi_k = 0, \qquad i = 1, 2, \cdots, m,$$
$$\xi_0 = 0.$$

We now must show that the matrix of (9), that is,

(10)
$$\begin{pmatrix} a_{10} & a_{11} & a_{12} & \cdots & a_{1n} \\ a_{20} & a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m0} & a_{m1} & a_{m2} & \cdots & a_{mn} \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

has rank $n-(r-1)$. The matrix of (8), that is, (10) without the last row, is of rank $n-r$ (since L is of dimension r). Now if (10) also had rank $n-r$, it would follow that both systems of equations, (8) and (9), have the same solutions. But this cannot be, since (8) has solutions with $\xi_0 \neq 0$ (for L was assumed to contain some proper points), whereas $\xi_0 = 0$ must hold for every solution of (9). Thus, the rank of (10) is $n-r-1$.

We shall next investigate what happens in P_n to the concept of *parallelism* as we know it in affine R_n.

To this end, let us first make the following observations. Let L be a linear space of dimension r in P_n. Let L be defined by the system of equations (8) and let L moreover contain proper points. Let L' be the totality of proper points of L. This means that L' is a linear space of dimension r in R_n. The affine (that is, non-homogeneous) coordinates

$x_i = \frac{\xi_i}{\xi_0}$ $(i = 1, 2, \ldots, n)$ of a point of L' satisfy the system of equations (7).

The totality of vectors $\overrightarrow{PQ}$ of affine R_n having the property that the initial point P as well as the terminal point Q lies in L' form a linear vector space L of dimension r (cf. *Modern Algebra*, § 5). Every vector* $\mathfrak{x} = \{x_1, x_2, \ldots, x_n\}$ of L satisfies[7] the system of homogeneous equations

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

If we now compare the vectors $\{x_1, x_2, \ldots, x_n\}$ of R_n and the improper points $[0, x_1, x_2, \ldots, x_n]$ of P_n, we see that if $\{x_1, x_2, \ldots, x_n\}$ is a vector solution of (11), then $[0, x_1, x_2, \ldots, x_n]$ is a point solution of (8), and conversely. That is to say: If $\{x_1, x_2, \ldots, x_n\}$ is a vector of L, then $[0, x_1, x_2, \ldots, x_n]$ is an *improper* point of L, and conversely.

We now wish to apply this natural relation between the vectors of L and the improper points of L to *two* linear spaces L_1 and L_2. Let L_i', as before, be the totality of all proper points of L_i $(i = 1, 2)$, neither L_1' nor L_2' being empty. Furthermore, let the totality of vectors of R_n whose initial and terminal points lie in L_i' be denoted by L_i. Finally, let us denote by L_i'' $(i = 1, 2)$ the totality of improper points of L_i.

Now, let L_1' be parallel to L_2'. Then, according to the definition of parallelism, *at least* one of the relations $\text{L}_1 \subset \text{L}_2$ or $\text{L}_2 \subset \text{L}_1$ holds.[8] Suppose, say, that $\text{L}_1 \subset \text{L}_2$. Then it follows immediately from the above relation that $L_1'' \subset L_2''$. In the same way, it follows conversely from $L_1'' \subset L_2''$ that $\text{L}_1 \subset \text{L}_2$. Thus, we see the following:

Parallelism of L_1' and L_2' implies that at least one of the two relations $L_1'' \subset L_2''$ or $L_2'' \subset L_1''$ holds, i.e., that all the improper points of one of the two L_i belong to the other.

In the special case that L_1 and L_2 are of the same dimension, L_1' and L_2' are parallel if and only if $L_1'' = L_2''$ (for in this case, $\text{L}_1 = \text{L}_2$).

---

* The symbol $\mathfrak{x}$ is the German x. The symbol L may be read as '*small-cap l*' (i.e., *small* capital L).

[7] Cf. *Modern Algebra*, § 6.

[8] As is well known, the notation '$\text{L}_1 \subset \text{L}_2$' is used to express that L_1 is contained in L_2. The notation '$\text{L}_2 \supset \text{L}_1$' means the same.

If L_1 and L_2 are lines, then L_i' $(i=1,2)$ consists of a single point. From this, it follows that parallel lines intersect in a *single improper* point.

In what follows, we make a slight extension of our fundamental assumptions.

Up to this point we have only permitted real numbers as coordinates of the points of P_n. However, it is highly advantageous on occasion to allow complex numbers as well. To keep our terminology straight, we make the following definitions:

The totality of the homogeneous ordered $(n+1)$-tuples of real numbers is called the *real space* P_n, or *real* P_n.

The totality of the homogeneous ordered $(n+1)$-tuples of complex numbers is called the *complex projective n-dimensional space*, or *complex* P_n.

A single $(n+1)$-tuple of complex numbers is called a *point* and is written $[\xi_0, \xi_1, \ldots, \xi_n]$. The complex numbers ξ_i are then called the *coordinates* of this point.

The definition of equality in complex P_n is completely analogous to that in real P_n: Two points $[\xi_0, \xi_1, \ldots, \xi_n]$ and $[\xi_0', \xi_1', \ldots, \xi_n']$ are said to be *equal* if and only if there exists a *complex* number $\lambda \neq 0$ such that $\xi_i' = \lambda \xi_i$ for $i = 0, 1, 2, \ldots, n$.

Let us furthermore agree on the following convention: When we speak of projective n-dimensional space, or P_n, simply, and without qualification, then whatever we say is understood to hold true for real P_n as well as for complex P_n.

The definitions given of a linear subspace for real P_n also hold for complex P_n. That is to say: An r-dimensional linear space in complex P_n $(0 \leqq r \leqq n)$ will mean the totality of points $[\xi_0, \xi_1, \ldots, \xi_n]$ of complex P_n that satisfy a system of homogeneous linear equations of the form (8), where the rank of the matrix of (8) is equal to $n-r$. Of course, the coefficients in (8) are now allowed to be complex numbers.

In conclusion, we should like to offer an example of a theorem that holds true equally for real and for complex P_n:

The intersection of two linear spaces is either empty or is itself a linear space.[9]

---

[9] One can also speak about *projective P_n over an arbitrary field* F, if we take the coordinates from some arbitrary field F. Not all of the results of this book hold true for so general a concept of projective space (in particular, the later developments do not). However, the content of Chapters II-IV, for example, does hold true even in this general framework, as does that of Chapter V, with the exception of Theorems 7 and 8.

Exercises

1. Let P_n be the complex and P_n^* the real n-dimensional projective space. Also, let L be a linear subspace of P_n and let r be the dimension of L. Let L^* be the totality of the points of L that belong to P_n^*. Show the following:

a) L^* is a *linear* subspace of P_n^*;
b) The dimension s of L^* satisfies the inequality: $2r - n \leqq s \leqq r$;
c) s can actually take on any of the values between $2r - n$ and r, provided it is $\geqq 0$.

2. Let the notation be the same as in Exercise 1. Furthermore, let L_1 be a second linear subspace of P_n with dimension r_1. Let L_1^* be formed analogously to L^*, i.e., L_1^* is the intersection of L_1 with P_n^*. Let the dimension of L_1^* be s_1. Show that if $r = s$, $r_1 = s_1$, and $L^* = L_1^*$, then $L = L_1$.

3. Let the notation again be the same as in Exercise 1. L is said to be 'real' if it is possible to find for L a system of homogeneous linear equations with purely *real* coefficients. Show that:

a) L is real if and only if $r = s$.
b) L is real if and only if for every point $[\xi_0, \xi_1, \cdots, \xi_n]$ of L, the 'complex conjugate point' $[\bar{\xi}_0, \bar{\xi}_1, \cdots, \bar{\xi}_n]$ also belongs to L.

4. Let $Q_1 = [\eta_0, \eta_1, \cdots, \eta_n]$ and $Q_2 = [\xi_0, \xi_1, \cdots, \xi_n]$ be two points in P_n. Let $n+1$ continuous functions $f_0(t), f_1(t), \ldots, f_n(t)$ be given on $0 \leqq t \leqq 1$ such that $f_i(0) = \eta_i$, $f_i(1) = \xi_i$ for $i = 0, 1, 2, \ldots, n$, and such that for no t do all the $f_i(t)$ vanish simultaneously. We then define a *continuous path* joining Q_1 and Q_2 as the totality of points $[f_0(t), f_1(t), \ldots, f_n(t)]$ for all $0 \leqq t \leqq 1$.

Now, let a hyperplane in P_n be given by the equation:

$$a_0\xi_0 + a_1\xi_1 + \cdots + a_n\xi_n = 0.$$

Show that: If Q_1 and Q_2 are any two points *not* on the hyperplane, then there is a continuous path joining Q_1 and Q_2 that does not intersect the hyperplane.

Hint: Tentatively set, say, $f_i(t) = (1-t)\eta_i + t\xi_i$ and normalize the η_i, ξ_i suitably.

CHAPTER II

GENERAL PROJECTIVE COORDINATES

In order to utilize the vector calculus in our investigation of the projective space P_n, we shall now study a certain fairly obvious correspondence between, on the one hand, the points of real and of complex P_n, and, on the other hand, the vectors of the real and the complex $(n+1)$-dimensional vector space[1] V_{n+1}, respectively. This correspondence is defined as follows:

A point Q of P_n and[2] a vector* $\mathfrak{x} = \{\xi_0, \xi_1, \ldots, \xi_n\}$ of V_{n+1} shall be said to correspond if and only if $Q = [\xi_0, \xi_1, \ldots, \xi_n]$, i.e., if $\xi_0, \xi_1, \ldots, \xi_n$ are homogeneous coordinates of Q.[3]

If Q and $\mathfrak{x}$ correspond to each other in this sense, then $\mathfrak{x}$ is called a *coordinate vector* of Q.

According to this rule, there corresponds to every non-vanishing vector $\mathfrak{x}$ of V_{n+1} exactly *one* point of P_n. Only the null vector has *no* point of P_n corresponding to it. Further, according to the definition of equality for points of P_n it follows that two vectors* $\mathfrak{x}$ and $\mathfrak{y}$ correspond to the same point of P_n if and only if there exists a $\lambda \neq 0$ such that $\mathfrak{x} = \lambda\mathfrak{y}$; in

---

[1] We recall the definition of the n-dimensional vector space V_n over a field F given in § 21 of *Modern Algebra* [*Introduction to Modern Algebra and Matrix Theory,* by O. Schreier and E. Sperner; see the Editor's Preface to the present work]. In particular, the V_n over the field of real numbers is called the real n-dimensional vector space, and the V_n over the field of complex numbers is called the complex n-dimensional vector space.

[2] ξ_0 is to be thought of as the first component of $\mathfrak{x}$, ξ_1 as the second component and, in general, ξ_i as the $(i+1)$-th component.

[3] We speak here simply of P_n, without qualification. Hence according to the convention that we adopted toward the end of the last chapter, P_n can refer equally to real projective n-dimensional space and to complex projective n-dimensional space. Of course, in either case we have to take the corresponding (real or complex) V_{n+1}.

* The symbol $\mathfrak{x}$ is German lower case x; the symbol $\mathfrak{y}$ is German lower case y.

other words, if $\mathfrak{x}$ is a coordinate vector of Q, then all vectors $\lambda\mathfrak{x}$ with $\lambda \neq 0$, and these vectors *only*, are also coordinate vectors of Q.[4]

Thus, to each point Q of P_n there correspond, exactly, all of the non-null vectors of a *one*-dimensional linear vector space in V_{n+1}. We call this vector space L_Q.

It is clear that $\mathrm{L}_Q \neq \mathrm{L}_{Q'}$ if $Q \neq Q'$. Thus, the relation $Q \rightleftarrows \mathrm{L}_Q$ is a one-to-one correspondence between the points of P_n and the one-dimensional linear vector spaces of V_{n+1}.

Let us now investigate the following problem: What does the totality of coordinate vectors of all the points of a linear space look like?

In order to answer this, let us consider a system of homogeneous equations of the form

$$\begin{aligned} a_{10}\,\xi_0 + a_{11}\,\xi_1 + \cdots + a_{1n}\,\xi_n &= 0, \\ a_{20}\,\xi_0 + a_{21}\,\xi_1 + \cdots + a_{2n}\,\xi_n &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m0}\,\xi_0 + a_{m1}\,\xi_1 + \cdots + a_{mn}\,\xi_n &= 0. \end{aligned}$$

Let the rank of the matrix (a_{ik}) be $n - r$. Then this system of equations represents an r-dimensional linear space L in P_n; but in V_{n+1} it represents an $(r+1)$-dimensional linear vector space L. Thus, we have:

THEOREM 1. *The totality of the coordinate vectors of all the points of an r-dimensional linear space constitutes an $(r+1)$-dimensional linear vector space in V_{n+1}, with the omission of the null vector. And conversely, all the points whose coordinate vectors belong to a given $(r+1)$-dimensional linear vector space of V_{n+1} constitute an r-dimensional linear space of P_n.*

The linear dependence or independence of the coordinate vectors has important implications for the corresponding points. Let $Q_1, Q_2, \ldots, Q_k$ be k points in P_n and let $\mathfrak{x}_i$ be a coordinate vector of Q_i $(i = 1, 2, \ldots, k)$. The $\mathfrak{x}_i$ are not uniquely determined. We may, if we wish, replace each of the $\mathfrak{x}_i$ by $\lambda_i\mathfrak{x}_i$, with $\lambda_i \neq 0$. Can this in any way alter the linear dependence or independence of the coordinate vectors of Q_i? *No.* For an equation in the $\lambda_i\mathfrak{x}_i$

$$a_1(\lambda_1\,\mathfrak{x}_1) + a_2(\lambda_2\,\mathfrak{x}_2) + \cdots + a_k(\lambda_k\,\mathfrak{x}_k) = 0$$

[4] Here, λ has to run through all the real values $\neq 0$, in the case of real P_n and V_{n+1}, and through all the complex values $\neq 0$, in the case of complex P_n and V_{n+1}.

(the a_i here may be either real or complex numbers) can be looked upon as an equation in the $\mathfrak{x}_i$, and conversely.

Thus, the vector systems $\mathfrak{x}_1, \mathfrak{x}_2, \ldots, \mathfrak{x}_k$ and $\lambda_1 \mathfrak{x}_1, \lambda_2 \mathfrak{x}_2, \ldots, \lambda_k \mathfrak{x}_k$ ($\lambda_i \neq 0$) are always either both linearly dependent or both linearly independent.

We see, then, that the linear dependence or independence of the coordinate vectors of the points $Q_1, Q_2, \ldots, Q_k$ does not depend on the choice of those vectors but is a property *of the points themselves.* We can accordingly introduce the following concept.

A finite number of points of P_n is said to be linearly dependent or linearly independent according as the coordinate vectors of these points are linearly dependent or independent.

Now let there be given a further point Q, with coordinate vector $\mathfrak{x}$. If $\mathfrak{x}$ can be represented by a linear combination of the $\mathfrak{x}_i$, say,

$$\mathfrak{x} = \sum_{i=1}^{k} a_i \mathfrak{x}_i, \tag{1}$$

then the point Q is said to be a *linear combination* of the points $Q_1, Q_2, \ldots, Q_k$.[5]

In what follows, let $\mathfrak{x}_i$ be *fixed* coordinate vectors of the Q_i. Let the maximal number of linearly independent points among $Q_1, Q_2, \ldots, Q_k$ be q. Then q has the same meaning for the $\mathfrak{x}_i$. Thus, we know that the linear vector space L spanned by the $\mathfrak{x}_i$ in V_{n+1} is of dimension q (cf. *Modern Algebra,* § 4, Theorem 5). Moreover, since every linear vector space that contains all the $\mathfrak{x}_i$ must also contain L, it follows immediately that:

There is *no* linear vector space of dimension $q-1$ that contains all the $\mathfrak{x}_i$. There is, on the other hand, exactly *one* linear vector space of dimension q in which all the $\mathfrak{x}_i$ are contained.

This statement, together with Theorem 1, yields at once the following theorem about the points $Q_1, Q_2, \ldots, Q_k$:

---

[5] If we replace each of the x_i by $\lambda_i \mathfrak{x}_i$, with $\lambda_i \neq 0$, and perhaps even replace $\mathfrak{x}$ by $\lambda\mathfrak{x}$ ($\lambda \neq 0$), we can rewrite (1) in the form

$$\lambda \mathfrak{x} = \sum_{i=1}^{k} \left(\lambda \frac{a_i}{\lambda_i}\right) (\lambda_i \mathfrak{x}_i).$$

Thus the concept of 'linear combination of points' is again shown to be independent of the choice of the coordinate vectors, and this justifies the definition.

THEOREM 2. *If q is the maximal number of linearly independent points among the k points $Q_1, Q_2, \ldots, Q_k$, then there is no linear space of dimension $q-2$ that contains all the Q_i. There is, however, exactly one linear space of dimension $q-1$ to which all the Q_i belong.*

Let L designate the linear space of dimension $q-1$ that contains all the Q_i. It consists of all the points whose coordinate vectors lie in L.

If we now call the linear space of *smallest* dimension that contains any given k points $Q_1, Q_2, \ldots, Q_k$ of P_n the *spanning space* of the Q_i, we have:

THEOREM 3. *The spanning space of the Q_i consists precisely of the totality of all linear combinations of the Q_i.*

This is true because the linear vector space spanned by the $\mathfrak{x}_i$ ($\mathfrak{x}_i$ continues to denote a coordinate vector of Q_i) consists precisely of the totality of all linear combinations $\mu_1 \mathfrak{x}_1 + \mu_2 \mathfrak{x}_2 + \cdots + \mu_k \mathfrak{x}_k$.

The following is a direct consequence of Theorem 2:

THEOREM 4. *The spanning space L of k points, $Q_1, Q_2, \ldots, Q_k$ of P_n is of dimension $k-1$ if the Q_i are linearly independent; otherwise, L is of dimension at most $k-2$.*

Applying Theorem 4 to a special case, we have:

Two linearly dependent points coincide; two linearly independent points are distinct. Three linearly dependent points lie on a straight line; three linearly independent points do not. Four linearly dependent points lie in a plane; four linearly independent points do not. $n+1$ linearly dependent points lie in a hyperplane; $n+1$ linearly independent points, however, do not.

As a further special case of Theorem 3, we have:

If Q_1 and Q_2 are two linearly independent points and $\mathfrak{x}_1$ and $\mathfrak{x}_2$ are their respective coordinate vectors, then the coordinate vectors of all the points on the line determined by Q_1 and Q_2 are of the form $\mu_1 \mathfrak{x}_1 + \mu_2 \mathfrak{x}_2$.

In like manner, the coordinate vectors of all the points of the plane determined by three linearly independent points Q_1, Q_2, Q_3, with coordinate vectors $\mathfrak{x}_1, \mathfrak{x}_2, \mathfrak{x}_3$, may be obtained in the form $\mu_1 \mathfrak{x}_1 + \mu_2 \mathfrak{x}_2 + \mu_3 \mathfrak{x}_3$, and all the points of the hyperplane determined by n linearly independent points $Q_1, Q_2, \ldots, Q_n$, with coordinate vectors $\mathfrak{x}_i$, may be obtained by taking all the linear combinations $\mu_1 \mathfrak{x}_1 + \mu_2 \mathfrak{x}_2 + \cdots + \mu_n \mathfrak{x}_n$.

The question of determining the linear space of least dimension that contains a number of given points is but a special case of the more general question as to the linear space of least dimension that contains a finite number of linear spaces. For the sake of later applications, we here take up this more general question for the case of two given linear spaces. For this purpose we merely need to adapt to linear spaces a theorem on linear vector spaces, namely, Theorem 7 of § 4 of *Modern Algebra.*[6]

First, we introduce the following convenient notation: If a linear space L of P_n and a linear vector space L of V_{n+1} are related in the way set forth in Theorem 1 above, then we shall write: $L \rightleftarrows \mathrm{L}$.

Now, let L_1 and L_2 be two given linear spaces in P_n and let L_1 and L_2 be the linear spaces of V_{n+1} for which the relations $L_1 \rightleftarrows \mathrm{L}_1$ and $L_2 \rightleftarrows \mathrm{L}_2$ hold. Let D be the intersection of L_1 and L_2, and S their sum. Then let $D \rightleftarrows \mathrm{D}$ and $S \rightleftarrows \mathrm{S}$. D can be empty; this happens when D is dimension 0 and thus consists of the null vector alone. Let the dimensions of L_1 and L_2 be r_1 and r_2, respectively, and let the dimensions of S and D be s and d, respectively. If D is empty, it will be convenient to set $d = -1$.[7] Then it follows that L_1 and L_2 are of dimension $r_1 + 1$ and $r_2 + 1$, respectively, and S and D of dimension $s + 1$ and $d + 1$, respectively.

It is clear that D is the intersection of L_1 and L_2. But what is S? We claim that:

Every linear space that contains L_1 and L_2 both, must also contain S.

As proof, let L be a linear space such that $L \supset L_1$ and $L \supset L_2$. Determine L by the correspondence $\mathrm{L} \rightleftarrows L$. Then $\mathrm{L} \supset \mathrm{L}_1$, L_2, and from this it follows that $\mathrm{L} \supset \mathrm{S}$, since L contains every vector $\mathfrak{a}_1 + \mathfrak{a}_2$ for which $\mathfrak{a}_1 \subset \mathrm{L}_1$ and $\mathfrak{a}_2 \subset \mathrm{L}_2$. Hence, we finally have: $L \supset S$.

We thus see the following: There is *no* linear space of dimension $< s$ that contains both L_1 and L_2. But there is *exactly* one such space of dimension s, and that space is S.

S is accordingly characterized, independently of S, as the linear space of least dimension that contains both L_1 and L_2. Again, we call S the *spanning space* of L_1 and L_2, as a generalization of the definition of a spanning space of points given above on p. 26.

According to *Modern Algebra,* § 4, Theorem 7, the following relation exists among the dimensions of L_1, L_2, D, and S: $(r_1 + 1) + (r_2 + 1) = (d + 1) + (s + 1)$, i.e.,

$$r_1 + r_2 = d + s.$$

---

[6] See Chapter I, footnote 3.

[7] For, as a result, the discussion that follows will held true also in the case in which D is of dimension 0.

From this, we have:

THEOREM 5. *If D is the intersection of two linear spaces L_1 and L_2 and S is the spanning space of L_1 and L_2, and if, moreover, r_1, r_2, and s are the respective dimensions of L_1, L_2, and S, then $r_1 + r_2 - s$ is the dimension of D, if D is not empty; on the other hand, if D is empty, then $r_1 + r_2 - s = -1$.*

In particular, if $r_1 + r_2 \geqq n$, then $r_1 + r_2 - s \geqq 0$, since $s \leqq n$, i.e., *the intersection of L_1 and L_2 is certainly not empty in this case and indeed is at least $(r_1 + r_2 - n)$-dimensional.* For example, *the intersection of a hyperplane and a line is never empty, and the intersection of a hyperplane and a plane is of dimension at least one.*

Another immediate consequence of Theorem 5, which we shall often make use of, is the following:

Two lines of P_n which lie in a plane (i.e., in a 2-dimensional linear space) always have a non-empty intersection and accordingly either have a point in common or are identical.

For, the spanning space of two such lines is at most 2-dimensional, so that $s \leqq 2$, and from this it follows that $d = r_1 + r_2 - s \geqq 0$.

The correspondence between the points of P_n and the vectors of V_{n+1} that we have been making use of until now is susceptible of another important generalization, which will lead to the concept of a general projective coordinate system.

We begin with a linear space L of dimension r in P_n. Let $Q_0, Q_1, \ldots, Q_r$ be $r + 1$ linearly independent points of L (such points exist, by Theorem 1), and let $\mathfrak{x}_i$ be a *fixed* coordinate vector of Q_i (for $i = 0, 1, \ldots, r$).

We then define a *correspondence between the points of L and the $(r + 1)$-tuples (of real or complex numbers)*[8] according to the following rule:

A point Q of L and an $(r + 1)$-tuple $\mu_0, \mu_1, \ldots, \mu_r$ shall be said to correspond if and only if $\sum_{i=0}^{r} \mu_i \mathfrak{x}_i$ is a coordinate vector of Q.

It is clear from Theorems 2 and 3 that this correspondence yields *every* point of L and also every $(r + 1)$-tuple with the exception of the one in which $\mu_0 = \mu_1 = \ldots = \mu_r = 0$.

---

[8] This is of course meant in the following sense: $(r + 1)$-tuples of *real* numbers when we are dealing with real P_n and of *complex* numbers when we are dealing with complex P_n.

Now, if two $(r+1)$-tuples $\mu_0, \mu_1, \ldots, \mu_r$ and $\mu_0', \mu_1', \ldots, \mu_r'$ correspond to the same point, then there must exist a $\lambda \neq 0$ for which

$$\sum_{i=0}^{r} \mu_i' \mathfrak{x}_i = \lambda \sum_{i=0}^{r} \mu_i \mathfrak{x}_i \quad \text{or} \quad \sum_{i=0}^{r} (\mu_i' - \lambda \mu_i)\, \mathfrak{x}_i = 0.$$

By virtue of the linear independence of the $\mathfrak{x}_i$, it follows that $\mu_i' = \lambda \mu_i$ for $i = 0, 1, \ldots, r$. The converse is trivial. Hence:

Two $(r+1)$-tuples correspond to the same point of L if and only if they differ solely by a constant of proportionality $\lambda \neq 0$.

This correspondence is thus of the same type as that between $(n+1)$-tuples and the points of P_n. Since the $\mathfrak{x}_i$ play an essential role in the definition, the correspondence will of course depend on the Q_i. If we change the Q_i, the correspondence will certainly change. But even more is true: The correspondence depends not only on the choice of the Q_i, but on the choice of the $\mathfrak{x}_i$ as well. If, holding the Q_i fixed, we change the $\mathfrak{x}_i$, then the correspondence may also change. We now examine this interrelation more closely.

For this purpose, let $\mathfrak{x}_i$ and $\mathfrak{y}_i$ $(i = 0, 1, \ldots, r)$ be two systems of coordinate vectors of the points Q_i. We then set up the correspondence defined above in two ways—first, using the $\mathfrak{x}_i$, second using the $\mathfrak{y}_i$. Now, under what condition would we get the same correspondence? This is the case only if, for *every* $(r+1)$-tuple $\mu_0, \mu_1, \ldots, \mu_r$, the vectors $\sum_{i=0}^{r} \mu_i \mathfrak{x}_i$ and $\sum_{i=0}^{r} \mu_i \mathfrak{y}_i$ are always coordinate vectors of the *same* point. In particular, this must be true for $\mu_0 = \mu_1 = \ldots = \mu_r = 1$. This means that a $\lambda \neq 0$ must exist such that

$$\lambda \sum_{i=0}^{r} \mathfrak{x}_i = \sum_{i=0}^{r} \mathfrak{y}_i. \tag{2}$$

Thus, equation (2) is a necessary condition.

But it is also sufficient. For assume that (2) is satisfied. Since *both* the systems $\mathfrak{x}_i$ and $\mathfrak{y}_i$ are coordinate vectors of Q_i, then for every i there exists a $\lambda_i \neq 0$ such that $\mathfrak{y}_i = \lambda_i \mathfrak{x}_i$. If we substitute in (2), we get: $\sum_{i=0}^{r} (\lambda - \lambda_i)\, \mathfrak{x}_i = 0$. By the linear independence of the $\mathfrak{x}_i$, it follows that $\lambda_i = \lambda$ for all i. That means that

$$\mathfrak{y}_i = \lambda \mathfrak{x}_i \tag{3}$$

for $i = 0, 1, \ldots, r$. Then if the linear combinations $\sum_{i=0}^{r} \mu_i \mathfrak{x}_i$ and $\sum_{i=0}^{r} \mu_i \mathfrak{y}_i$ are formed with the same μ_i, it follows from (3) that

$$\sum_{i=0}^{r} \mu_i \mathfrak{y}_i = \lambda \sum_{i=0}^{r} \mu_i \mathfrak{x}_i,$$

i.e., $\sum_{i=0}^{r} \mu_i \mathfrak{y}_i$ and $\sum_{i=0}^{r} \mu_i \mathfrak{x}_i$ are always coordinate vectors of the same point. In other words: The correspondences we set up with the aid of the $\mathfrak{x}_i$ and the $\mathfrak{y}_i$ are identical under assumption (2).

Condition (2) has a simple meaning, namely: The correspondence established with the aid of the $\mathfrak{x}_i$ makes the $(r+1)$-tuple $\mu_0 = \mu_1 = \ldots = \mu_r = 1$ correspond to the same point as does the correspondence established with the aid of the $\mathfrak{y}_i$. And conversely, the coordinate vectors $\mathfrak{x}_i$ and $\mathfrak{y}_i$ employed in constructing the two correspondences that make the same point correspond to the $(r+1)$-tuple $\mu_0 = \mu_1 = \ldots = \mu_r = 1$, satisfy equation (2).

Thus, we see that among all the possible correspondences which can be set up by different choices of the $\mathfrak{x}_i$ (the Q_i being fixed), no two are distinct which make the same point correspond to the $(r+1)$-tuple $\mu_0 = \mu_1 = \ldots = \mu_r = 1$. This means that:

The correspondence is uniquely determined by the points Q_i and the point E that corresponds to the $(r+1)$-tuple $\mu_0 = \mu_1 = \ldots = \mu_r = 1$.

It is desirable to introduce some nomenclature at this point.

If a correspondence is set up of the kind described, we say that *a* ***projective coordinate system*** *has been introduced into L.* Since such a coordinate system is determined by the points Q_i and E, we designate it by the symbol $(Q_0, Q_1, \ldots, Q_r \mid E)$. The numbers $\mu_0, \mu_1, \ldots, \mu_r$ which correspond to a point, are called the *coordinates of this point in the coordinate system* $(Q_0, Q_1, \ldots, Q_r \mid E)$. They are unique only up to a common constant of proportionality. The points $Q_0, Q_1, \ldots, Q_r$ are called the *fundamental points* of the coordinate system. The coordinates of the points Q_i in the $(Q_0, Q_1, \ldots, Q_r \mid E)$ system are all 0 except for the $(i+1)$-st coordinate, which is equal to 1. The point E, all of whose coordinates are equal to 1, is called the *unit point.*

The $r+1$ points Q_i, taken in their totality, are also called the *fundamental simplex*; in particular, in the case $r = 2$ (three points), the *fundamental triangle,* and in the case $r = 3$ (four points), the *fundamental tetrahedron.*

We already know that we can choose any $r+1$ linearly independent points of L as the fundamental points of a projective coordinate system.

But we do not yet know the extent to which E may be chosen arbitrarily. This question is equivalent to the following: Given Q_i and E, can the $\mathfrak{x}_i$ (as coordinate vectors of Q_i) always be so chosen that

$$\mathfrak{e} = \mathfrak{x}_0 + \mathfrak{x}_1 + \cdots + \mathfrak{x}_r$$

is a coordinate vector of E?

From the equation

$$\mathfrak{e} = \mathfrak{x}_0 + \mathfrak{x}_1 + \cdots + \mathfrak{x}_r \tag{4}$$

it follows that *every* $r+1$ of the vectors $\mathfrak{x}_0, \mathfrak{x}_1, \ldots, \mathfrak{x}_r, \mathfrak{e}$ must be linearly independent. For $\mathfrak{x}_0, \mathfrak{x}_1, \ldots, \mathfrak{x}_r$, this is true by assumption. However, if $\mathfrak{x}_0, \mathfrak{x}_1, \ldots, \mathfrak{x}_{r-1}, \mathfrak{e}$, say, were linearly dependent, then (by *Modern Algebra,* § 3, Theorem 5) a relation of the form

$$\mathfrak{e} = \lambda_0 \mathfrak{x}_0 + \lambda_1 \mathfrak{x}_1 + \cdots + \lambda_{r-1} \mathfrak{x}_{r-1} \tag{5}$$

would exist. By elimination of $\mathfrak{e}$ between (4) and (5), this would imply

$$(1-\lambda_0)\mathfrak{x}_0 + (1-\lambda_1)\mathfrak{x}_1 + \cdots + (1-\lambda_{r-1})\mathfrak{x}_{r-1} + \mathfrak{x}_r = 0.$$

This contradicts the linear independence of the $\mathfrak{x}_0, \mathfrak{x}_1, \ldots, \mathfrak{x}_r$. Thus, we see that:

A necessary condition for the existence of a projective coordinate system $(Q_0, Q_1, \ldots, Q_r \mid E)$ *in* L *is that not only the* Q_i, *but also every* $r+1$ *of the points* $Q_0, Q_1, \ldots, Q_r, E$ *be linearly independent.*

It turns out that this condition is also sufficient. To prove this, assume the condition satisfied. Furthermore, let $\mathfrak{e}$ be any coordinate vector of E and let $\mathfrak{y}_0, \mathfrak{y}_1, \ldots, \mathfrak{y}_r$ be any coordinate vectors whatsoever of $Q_0, Q_1, \ldots, Q_r$. By Theorem 3, we can write

$$\mathfrak{e} = \varrho_0 \mathfrak{y}_0 + \varrho_1 \mathfrak{y}_1 + \cdots + \varrho_r \mathfrak{y}_r. \tag{6}$$

None of the ϱ_i in this equation can equal zero. For, $\varrho_i = 0$ (for a fixed i) would imply that the vectors $\mathfrak{e}, \mathfrak{y}_0, \mathfrak{y}_1, \cdots, \mathfrak{y}_{i-1}, \mathfrak{y}_{i+1}, \cdots, \mathfrak{y}_r$ are linearly dependent, contrary to assumption. But since $\varrho_i \neq 0$, it follows that $\mathfrak{x}_i = \varrho_i \mathfrak{y}_i$ is also a coordinate vector of Q_i $(i = 0, 1, \ldots, r)$. With these $\mathfrak{x}_i$, (6) becomes

$$\mathfrak{e} = \mathfrak{x}_0 + \mathfrak{x}_1 + \cdots + \mathfrak{x}_r.$$

This proves, then, that there does exist a coordinate system with the Q_i as fundamental points and E as unit point. We have thus shown the following:

To the symbol $(Q_0, Q_1, \ldots, Q_r \mid E)$ there corresponds a projective coordinate system if and only if every $r+1$ of the points $Q_0, Q_1, \ldots, Q_r, E$ are linearly independent.

If $r = 2$ (i.e., if L is a plane), this condition has the following geometrical meaning: The points Q_0, Q_1, Q_2 must form a triangle, and E may not be taken on any of the three sides or their extensions. A corresponding statement holds for any r.

Let us look a bit more closely at the special case $r = n$ (i.e., $L = P_n$). In this case, $(Q_0, Q_1, \ldots, Q_n \mid E)$ has the meaning of a new coordinate system for *the whole of P_n*.

There is this to be observed: To the points of P_n which are given by the $(n+1)$-tuple $[\xi_0, \xi_1, \ldots, \xi_n]$ there correspond, by definition, the coordinates $\xi_0, \xi_1, \ldots, \xi_n$. In a given projective coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$, however, a point $[\xi_0, \xi_1, \ldots, \xi_n]$ will in general be assigned new coordinates, *different* from the ξ_i.

Nevertheless, the '*natural*' coordinates ξ_i of a point $[\xi_0, \xi_1, \ldots, \xi_n]$ can actually be interpreted as coordinates in a certain particular projective coordinate system. We obtain this coordinate system by choosing the points $Q_0, Q_1, \ldots, Q_n$ in such a way that for every $Q_i = [\xi_0^{(i)}, \xi_1^{(i)}, \cdots, \xi_n^{(i)}]$ we have

$$\xi_0^{(i)} = \xi_1^{(i)} = \cdots = \xi_{i-1}^{(i)} = \xi_{i+1}^{(i)} = \cdots = \xi_n^{(i)} = 0, \quad \xi_i^{(i)} = 1.$$

Let E be the point $[1, 1, \ldots, 1]$. The $n+1$ *unit* vectors $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_{n+1}$ of V_{n+1} can then be chosen as the coordinate vectors of the Q_i. In fact, $\sum_{i=1}^{n+1} \mathfrak{e}_i$ is then a coordinate vector of E. In this special coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$, the projective coordinates of a point $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ are just the ξ_i.

Of the fundamental points of the projective coordinate system just given, Q_0 is the only *proper* point. All the other Q_i are improper (because their first coordinate is 0). We now wish to make clear the position of the Q_i in relation to the affine R_n from which P_n was obtained by extension. All the affine (non-homogeneous) coordinates of the point $Q_0 = [1, 0, \ldots, 0]$ are 0. Thus, Q_0 is the origin of R_n. On the other hand, for $i = 1, 2, \ldots, n$, Q_i is the improper point of the x_i-axis of R_n. For, the x_i-axis is given in affine coordinates by the equations

$$(7)\ x_1 = 0, \quad x_2 = 0, \cdots, x_{i-1} = 0, \quad x_{i+1} = 0, \cdots, x_n = 0,$$

and hence, in homogeneous coordinates:[9]

$$(8)\ \xi_1 = 0, \quad \xi_2 = 0, \cdots, \xi_{i-1} = 0, \quad \xi_{i+1} = 0, \cdots, \xi_n = 0.$$

The coordinates of the Q_i do, in fact, satisfy equations (8).

To avoid ambiguity, let us expressly agree that the notation $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ shall always mean the following: Q *is* the homogeneous $(n+1)$-tuple $[\xi_0, \xi_1, \ldots, \xi_n]$. In other words, we shall use the notation $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ only if the ξ_i are the coordinates of Q in the special coordinate system just mentioned, in which the origin of R_n and the improper points of the coordinate axes are the fundamental points and $[1, 1, \ldots, 1]$ is the unit point.[1] We shall make a corresponding restriction on the use of the term 'coordinate vector.' Hereafter, '$\mathfrak{x} = \{\xi_0, \xi_1, \ldots, \xi_n\}$ is a coordinate vector of Q' shall always mean $Q = [\xi_0, \xi_1, \ldots, \xi_n]$.

When a projective coordinate system for the whole of P_n is under consideration, matrices may easily be used for purposes of description (or definition). This may be done as follows.

Let there be given a projective coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ in P_n. Let $\mathfrak{x}_i$ be coordinate vectors of the Q_i, so that $\sum_{i=0}^{n} \mathfrak{x}_i$ is a coordinate vector of E.

Now, if $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ is any point in P_n (according to our convention, the ξ_i have reference to our special coordinate system), and if $\eta_0, \eta_1, \ldots, \eta_n$ are its coordinates in the system $(Q_0, Q_1, \ldots, Q_n \mid E)$, then, by definition, $\sum_{i=0}^{n} \eta_i \mathfrak{x}_i$ is a coordinate vector of Q. Hence, there exists a $\lambda \neq 0$ such that

$$(9) \qquad \sum_{i=0}^{n} \eta_i \mathfrak{x}_i = \lambda \{\xi_0, \xi_1, \cdots, \xi_n\}.$$

[9] Equations (7) and (8) are related to each other in the same way as equations (7) and (8) of the preceding chapter to each other.

[1] Compare the corresponding agreement for affine R_n on p. 117 of *Modern Algebra.*

This equation can easily be written in matrix form. To do this, let $\mathfrak{x}_i = \{x_{0i}, x_{1i}, \cdots, x_{ni}\}$ and let

$$(x_{ik}) = \begin{pmatrix} x_{00} & x_{01} & \cdots & x_{0n} \\ x_{10} & x_{11} & \cdots & x_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ x_{n0} & x_{n1} & \cdots & x_{nn} \end{pmatrix}.$$

In addition, let the symbols (η) and (ξ) represent (as in *Modern Algebra*, § 22, p. 293) the matrices

$$(10) \qquad (\eta) = \begin{pmatrix} \eta_0 & 0 & \cdots & 0 \\ \eta_1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \eta_n & 0 & \cdots & 0 \end{pmatrix}, \qquad (\xi) = \begin{pmatrix} \xi_0 & 0 & \cdots & 0 \\ \xi_1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \xi_n & 0 & \cdots & 0 \end{pmatrix}.$$

Then the matrix equation

$$(11) \qquad (x_{ik}) \cdot (\eta) = \lambda \cdot (\xi)$$

is completely equivalent[2] to (9).

The matrix (x_{ik}) is non-singular, since the Q_i, and hence also the $\mathfrak{x}_i$, are linearly independent. Consequently the inverse matrix $(x_{ik})^{-1}$ exists. Thus, from (11) it follows that

$$(12) \qquad (\eta) = \lambda \cdot (x_{ik})^{-1} (\xi).$$

If we now set $X = (x_{ik})^{-1}$, we can state the following theorem:

THEOREM 6. *For every projective coordinate system* $(Q_0, Q_1, \ldots, Q_n \mid E)$ *of* P_n *a non-singular* $(n+1)$*-rowed square matrix* X *can be found such that the coordinates* η_i *of every point* $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ *are given in that coordinate system by* $(\eta) = \lambda X (\xi)$.

In this equation λ may be assigned any arbitrary value $\neq 0$. The $\eta_0, \eta_1, \ldots, \eta_n$ thus obtained will, regardless, always be homogeneous coordinates of Q in $(Q_0, Q_1, \ldots, Q_n \mid E)$. It goes without saying that X depends only on the coordinate system and not on Q.

[2] Equivalent in the following sense: The components of the left-hand and right-hand sides of (9) are equal to the elements of the first column of the left-hand and right-hand sides, respectively, of (11).

Conversely, for every given non-singular matrix X there exists a coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ for which the two are related in the way described in Theorem 6.

For one need only make an appropriate choice of $(Q_0, Q_1, \ldots, Q_n \mid E)$ in order to obtain, in (9), any desired system of linearly independent vectors $\mathfrak{x}_0, \mathfrak{x}_1, \ldots, \mathfrak{x}_n$ and, hence, in (12), any desired matrix $X = (x_{ik})^{-1}$.

At this point we propose to make another important *convention*, which will hold for the remainder of this volume. It will very often happen that we shall wish to construct the first matrix of (10) using the coordinates $\eta_0, \eta_1, \ldots, \eta_n$ of a point Q. We shall therefore reserve the notation (η) for this matrix and shall refer to it as the *coordinate matrix* of Q in the coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$. We shall of course also use other letters, such as ζ_i, ξ_i, instead of the η_i and shall then denote the corresponding matrices by (ζ), (ξ).

By use of Theorem 6 we can easily get a clear picture of a given **transformation of coordinates.** Let $(Q_0, Q_1, \ldots, Q_n \mid E)$ and

$$(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*),$$

say, be two given coordinate systems, and let the point $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ have the coordinates $\eta_0, \eta_1, \ldots, \eta_n$ in $(Q_0, Q_1, \ldots, Q_n \mid E)$ and the coordinates $\eta_0^*, \eta_1^*, \ldots, \eta_n^*$ in $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$.

We now pose the following question: Given the η_i^*, how do we compute the η_i, and vice versa?

By Theorem 6, there exist two matrices X and X^* such that

$$(\eta) = \lambda X(\xi), \tag{13}$$

$$(\eta^*) = \lambda^* X^*(\xi). \tag{14}$$

By elimination of (ξ) between the two equations, we obtain

$$(\eta) = \varrho X \cdot X^{*-1}(\eta^*), \tag{15}$$

$$(\eta^*) = \sigma X^* \cdot X^{-1}(\eta), \tag{16}$$

where $\varrho = \lambda/\lambda^* = 1/\sigma$.

Equations (15) and (16) give us the desired relation between the η_i and the η_i^*. Hence we call equations (15) and (16) the *equations of transformation* between the two coordinate systems.

If we now set $XX^{*-1} = T = (t_{ik})$ and $X^*X^{-1} = T^{-1} = (s_{ik})$, we see that the matrix equations (15) and (16) are equivalent to the two systems of equations

$$\eta_i = \varrho \cdot \sum_{k=0}^{n} t_{ik}\eta_k^*, \qquad (i = 0, 1, 2, \cdots, n), \tag{17}$$

$$\eta_i^* = \sigma \cdot \sum_{k=0}^{n} s_{ik}\eta_k, \qquad (i = 0, 1, 2, \cdots, n). \tag{18}$$

Thus, we have:

THEOREM 7. *If the coordinates of a point Q in one projective coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ are denoted by $\eta_0, \eta_1, \ldots, \eta_n$, and denoted by $\eta_0^*, \eta_1^*, \ldots, \eta_n^*$ in a second projective coordinate system $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$, then the relation between the η_i^* and the η_i is given by equations* (17) *and* (18) (*in which the t_{ik} and the s_{ik} are constants independent of the η_i and the η_i^*, with $\mid t_{ik} \mid \neq 0$ and $\mid s_{ik} \mid \neq 0$*).

The presence of the factors ϱ, σ in (15) and (16) (something similar will happen quite frequently) is due to the fact that the η_i and η_i^* are uniquely determined only up to a constant of proportionality. If the η_i^* are a particular choice of coordinates for Q in $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$, then (15) always yields homogeneous coordinates η_i for the point Q in $(Q_0, Q_1, \ldots, Q_n \mid E)$, no matter how $\varrho \neq 0$ is chosen.

A system of equations of the form (17) is called a *non-singular homogeneous linear substitution.* It is important to note that *every* non-singular homogeneous linear substitution can be looked upon as a transformation of coordinates, i.e., can occur as a system of equations of transformation for passing from one coordinate system to another. To prove this, we need to show that, given a coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ and a system of equations of the form (17), we can find a new coordinate system $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$ for which (17) is precisely the equations of transformation for passing from the one system to the other.

Let the given coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ be related to the matrix X in the way set forth in Theorem 6; and let the desired coordinate system $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$ be determined by the matrix X^*, still to be found. Then X^* and X and the given matrix $T = (t_{ik})$ of (17) must be related by $T = XX^{*-1}$. This means we have only to set $X^* = T^{-1}X$. Then X^*, as a non-singular matrix,[3] is indeed the defining matrix of a coordinate system $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$. We summarize this result in Theorem 8:

[3] X^* is non-singular because T and X, by assumption, are non-singular; cf. *Modern Algebra*, § 22, Theorem 6 (p. 297).

THEOREM 8. *For a given coordinate system* $(Q_0, Q_1, \ldots, Q_n \mid E)$ *of* P_n *and a given system of equations* (17) *with non-singular matrix* $T = (t_{ik})$, *there always exists a second projective coordinate system* $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$ *such that the coordinates* η_i *of an arbitrary point in the first system will be related to its coordinates* η_i^* *in the second system precisely by the equations* (17).

Using Theorem 6, we should now like to ascertain how a linear space of dimension r can be represented in an arbitrary projective coordinate system.

Let $(Q_0, Q_1, \ldots, Q_n \mid E)$ be a given projective coordinate system and let X be the corresponding matrix which, as in (13), gives the coordinates $\eta_0, \eta_1, \ldots, \eta_n$ of an arbitrary point $[\xi_0, \xi_1, \ldots, \xi_n]$ of P_n in the coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$.

An r-dimensional linear space L was defined as the totality of all the points $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ whose coordinates ξ_i satisfy a system of homogeneous equations of rank $n - r$:[4]

$$(19) \qquad \sum_{k=0}^{n} a_{ik}\,\xi_k = 0, \qquad (i = 1, 2, \cdots, s).$$

We may assume that s, the number of equations in (19), is $\leqq n$.[5] We then extend the matrix of (19) by the adjunction of $n + 1 - s$ rows consisting entirely of zeros, thus obtaining the following $(n + 1)$-rowed square matrix:

$$A = \begin{pmatrix} a_{10} & a_{11} & \cdots & a_{1n} \\ a_{20} & a_{21} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{s0} & a_{s1} & \cdots & a_{sn} \\ 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

By use of the second matrix of (10), we can now write the system of equations (19) in matrix form as:

---

[4] By the *rank of a system of equations* is meant, of course, the rank of its matrix.

[5] Since (19) is of rank $n - r$, any $n - r$ linearly independent equations of (19), for example, represent the same linear space. Cf., for example, *Modern Algebra*, p. 104.

(20) $$A \cdot (\xi) = O.$$

By substituting $(\xi) = X^{-1}(\eta)/\lambda$ from (13), (20) becomes (upon dropping the factor $1/\lambda$):

(21) $$A \cdot X^{-1}(\eta) = O.$$

Since, conversely, we can also obtain (20) once again from (21) by the use of (13), we see the following:

If the point $Q = [\xi_0, \xi_1, \ldots, \xi_n]$ has the coordinates $\eta_0, \eta_1, \ldots, \eta_n$ in $(Q_0, Q_1, \ldots, Q_n \mid E)$, then the ξ_i satisfy the matrix equation (20) *if and only if* the η_i satisfy equation (21).

That is to say, (21) *is the system of equations for the linear space L in the coordinate system* $(Q_0, Q_1, \ldots, Q_n \mid E)$.

Since A and X^{-1} are constant matrices (independent of the η_i), (21) is again a system of linear equations. Now, what is the rank of $A \cdot X^{-1}$? It is immediately seen that the *row* vectors of the matrix AX^{-1} are linear combinations of the row vectors of X^{-1}, with the elements of A acting as coefficients. Since X^{-1} is non-singular, and its row vectors are therefore linearly independent, it follows at once, from § 10, Theorem 1, of *Modern Algebra,* that the rank of AX^{-1} is equal to the rank of A.[6] Thus, we have Theorem 9.

THEOREM 9. *A linear space of dimension r can always be represented in a projective coordinate system of P_n by a system of linear equations of rank $n - r$.*

If, conversely, we start with (21) as a *given* system of equations, then we can always find a corresponding system of the form (20). For this is merely to say that, given the matrix $A \cdot X^{-1}$, we can compute $A = (A \cdot X^{-1})X$. Thus, we have:

THEOREM 10. *Every system of linear equations of rank $n - r$ in the η_i, where the η_i are coordinates in the projective coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$, represents a linear space of dimension r.*

As a final application of Theorem 6, we now prove the following theorem:

---

[6] This result shows that Theorem 1 of § 10 of *Modern Algebra* is equivalent to the following theorem on matrices: *If one of the two n-rowed square matrices A and B, say A, is non-singular, then the products $A \cdot B$ and $B \cdot A$ have the same rank as B.*

THEOREM 11. *If* $S_1, S_2, \ldots, S_k$ *are any* k *points of* P_n *and* $\eta_{i0}, \eta_{i1}, \eta_{i2}, \cdots, \eta_{in}$ *are the coordinates of* S_i *in a projective coordinate system* $(Q_0, Q_1, \ldots, Q_n \mid E)$, *then the rank of the matrix*

$$\begin{pmatrix} \eta_{10} & \eta_{11} & \cdots & \eta_{1n} \\ \eta_{20} & \eta_{21} & \cdots & \eta_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \eta_{k0} & \eta_{k1} & \cdots & \eta_{kn} \end{pmatrix} \tag{22}$$

is equal to the maximal number of linearly independent points among $S_1, S_2, \ldots, S_k$.

Proof: Since neither the rank of the matrix in question nor the maximal number of linearly independent points can ever exceed $n+1$, we immediately see the following: If Theorem 11 were false for $k > n+1$, then among the S_i we could choose h points $S_{r_1}, S_{r_2}, \cdots, S_{r_h}$, with $h \leqq n+1$, for which Theorem 11 is false.[7] Thus, it suffices to prove Theorem 11 for $k \leqq n+1$.

Now, let $S_i = [\xi_{i0}, \xi_{i1}, \cdots, \xi_{in}]$ for $i = 1, 2, \ldots, k$, and let X be the matrix associated, in accordance with Theorem 6, with $(Q_0, Q_1, \ldots, Q_n \mid E)$; moreover, let the ξ_{ik} be so chosen that

$$\begin{pmatrix} \eta_{i0} & 0 & \cdots & 0 \\ \eta_{i1} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \eta_{in} & 0 & \cdots & 0 \end{pmatrix} = X \cdot \begin{pmatrix} \xi_{i0} & 0 & \cdots & 0 \\ \xi_{i1} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \xi_{in} & 0 & \cdots & 0 \end{pmatrix} \tag{23}$$

(i.e., with $\lambda = 1$). We then form the matrix product

$$X \cdot \begin{pmatrix} \xi_{10} & \xi_{20} & \cdots & \xi_{k0} & \overbrace{0 \quad \cdots \quad 0}^{n+1-k \text{ columns}} \\ \xi_{11} & \xi_{21} & \cdots & \xi_{k1} & 0 \quad \cdots \quad 0 \\ \vdots & \vdots & & \vdots & \vdots \qquad \vdots \\ \xi_{1n} & \xi_{2n} & \cdots & \xi_{kn} & 0 \quad \cdots \quad 0 \end{pmatrix}. \tag{24}$$

Now, the rank of the second (right-hand) factor of (24) is equal to the maximal number of linearly independent points among the points S_i.

[7] Let k be $> n+1$, let q be the rank of (22) and q' the maximal number of linearly independent points among the S_i, and let $q \neq q'$. Clearly, we have $q \leqq n+1$ and $q' \leqq n+1$. If $q' > q$, then Theorem 11 is false for any q' linearly independent points S_i. If, however, $q > q'$, then we choose q *such* points S_i *for which* the corresponding submatrix of (22) has rank q.

Since, however, the rank of the second factor of (24) is equal to the rank of (24) itself[8] and since, by (23), the product (24) is equal to the matrix

$$(25) \qquad \begin{pmatrix} \eta_{10} & \eta_{20} & \cdots & \eta_{k0} & 0 & \cdots & 0 \\ \eta_{11} & \eta_{21} & \cdots & \eta_{k1} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \eta_{1n} & \eta_{2n} & \cdots & \eta_{kn} & 0 & \cdots & 0 \end{pmatrix},$$

we have that the rank of (25), and therefore also the rank of (22), is equal to the maximal number of linearly independent points among the S_i, which was to be proved.

From Theorem 11 it follows immediately that the row vectors of (22) are linearly dependent or independent according as the points $S_1, S_2, \ldots, S_k$ are linearly dependent or independent. This can be expressed as follows:

Linear dependence, and linear independence, of points of P_n is invariant under a transformation of coordinates.

Exercises

1. Show that the matrix X which belongs in accordance with Theorem 6 to each projective coordinate system, is uniquely determined by that coordinate system up to an arbitrary numerical factor $\neq 0$.

2. Two projective coordinate systems in *real* P_n, say $(Q_0, Q_1, \ldots, Q_n \mid E)$ and $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$, are said to be *continuously deformable* into each other if $(n+1)$-dimensional (real) vectors $\mathfrak{f}_0(t), \mathfrak{f}_1(t), \cdots, \mathfrak{f}_n(t)$ can be found which are continuous functions of a parameter t in the interval $0 \leqq t \leqq 1$ and which, in addition, exhibit the following properties:

(1) the $n+1$ vectors $\mathfrak{f}_0(t), \mathfrak{f}_1(t), \cdots, \mathfrak{f}_n(t)$ are linearly independent for every t in the interval $0 \leqq t \leqq 1$;

(2) for $t=0$, $\mathfrak{f}_i(Q)$ is a coordinate vector of Q_i; (for all $i=0,1,\ldots,n$), and $\sum_{i=0}^{n} \mathfrak{f}_i(0)$ is a coordinate vector of E;

(3) for $t=1$, $\mathfrak{f}_i(1)$ is a coordinate vector of Q_i^* (for all $i=0,1,\ldots,n$), and $\sum_{i=0}^{n} \mathfrak{f}_i(1)$ is a coordinate vector of E^*.

Show the following:

a) If $(\nu_0, \nu_1, \cdots, \nu_n)$ is a permutation of the digits $(0,1,2,\ldots,n)$ such that $\operatorname{sgn}(\nu_0, \nu_1, \ldots, \nu_n)$ is $+1$ (cf. *Modern Algebra* § 9, p. 88), then the co-

[8] See footnote 6 above.

ordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ is always continuously deformable into $(Q_{\nu_0}, Q_{\nu_1}, \cdots, Q_{\nu_n} \mid E)$. (In this latter coordinate system the same Q_i enter, only permuted.)

b) If $\operatorname{sgn}(\nu_0, \nu_1, \cdots, \nu_n) = -1$, then $(Q_0, Q_1, \ldots, Q_n \mid E)$ is continuously deformable into $(Q_{\nu_0}, Q_{\nu_1}, \cdots, Q_{\nu_n} \mid E)$ if and only if n, the dimension of the P_n under consideration, is an *even* number.

On account of b), the real spaces $P_1, P_3, P_5, \ldots$ are called *orientable* and the spaces $P_2, P_4, P_6, \ldots$ *non-orientable.*

Hint. If the continuous deformation called for in a) and b) is possible, then both $\mathfrak{f}_i(1)$ and $\mathfrak{f}_{\nu_i}(0)$ must be a coordinate vector of Q_{ν_i}, whence $\mathfrak{f}_i(1) = \lambda_i \mathfrak{f}_{\nu_i}(0)$. Since, moreover, $\sum_{i=0}^{n} \mathfrak{f}_i(0)$ and $\sum_{i=0}^{n} \mathfrak{f}_i(1)$ are both coordinate vectors of E, we must have $\sum_{i=0}^{n} \mathfrak{f}_i(1) = \lambda \sum_{i=0}^{n} \mathfrak{f}_i(0)$. It then follows from the linear independence of the Q_i that $\lambda = \lambda_0 = \lambda_1 = \cdots = \lambda_n$. Accordingly, for the determinants

$$D(\mathfrak{f}_0(1), \mathfrak{f}_1(1), \cdots, \mathfrak{f}_n(1))$$

and

$$D(\mathfrak{f}_0(0), \mathfrak{f}_1(0), \cdots, \mathfrak{f}_n(0)):$$

we have the relation:

$$\begin{aligned} D(\mathfrak{f}_0(1), \mathfrak{f}_1(1), \cdots, \mathfrak{f}_n(1)) &= D(\lambda_0 \mathfrak{f}_{\nu_0}(0), \lambda_1 \mathfrak{f}_{\nu_1}(0), \cdots, \lambda_n \mathfrak{f}_{\nu_n}(0)) \\ &= \lambda^{n+1} \operatorname{sgn}(\nu_0, \nu_1, \cdots, \nu_n) \cdot D(\mathfrak{f}_0(0), \mathfrak{f}_1(0), \cdots, \mathfrak{f}_n(0)). \end{aligned}$$

From this and *Modern Algebra,* § 10, Theorem 4, it is seen that the necessary and sufficient condition for the deformation to be possible is that there exist a real λ such that $\lambda^{n+1} \operatorname{sgn}(\nu_0, \nu_1, \cdots, \nu_n)$ be > 0.

CHAPTER III

HYPERPLANE COORDINATES. THE DUALITY PRINCIPLE

A hyperplane in P_n is given by one equation:

$$u_0 \xi_0 + u_1 \xi_1 + \cdots + u_n \xi_n = 0. \tag{1}$$

In this equation we may *not* have all of the u_i equal to zero. The hyperplane is of course uniquely determined by the $n+1$ coefficients u_i. Now, to what extent are the u_i, conversely, determined by the hyperplane? Let us compare (1) with, say,

$$v_0 \xi_0 + v_1 \xi_1 + \cdots + v_n \xi_n = 0 \tag{2}$$

under the assumption that (1) and (2) represent the same hyperplane. Then the intersection of (1) and (2) is just the hyperplane itself, and thus is itself $(n-1)$-dimensional. That is to say, the matrix

$$\begin{pmatrix} u_0 & u_1 & \cdots & u_n \\ v_0 & v_1 & \cdots & v_n \end{pmatrix}$$

must have rank 1. Hence,[1] there exists a $\lambda \neq 0$ such that $u_i = \lambda v_i$ for all $i = 0, 1, \ldots, n$.

Consequently, the u_i are determined by the hyperplane up to a common constant of proportionality. We have in this way thus obtained a relation of the same kind between the totality of hyperplanes and homogeneous $(n+1)$-tuples as between the points of P_n and the homogeneous $(n+1)$-tuples. For this reason, the coefficients $u_0, u_1, \ldots, u_n$ are also called the *(homogeneous) coordinates* (or *hyperplane coordinates*) *of the hyperplane represented by* (1). For abbreviation, we shall designate a hyperplane with coordinates $u_0, u_1, \ldots, u_n$ by the symbol $\langle u_0, u_1, \ldots, u_n \rangle$.

---

[1] Since $\{u_0, u_1, \ldots, u_n\}$ and $\{v_0, v_1, \ldots, v_n\}$ are two non-vanishing linearly dependent vectors, each is a multiple of the other.

The fact that we can associate coordinates with the hyperplanes of P_n in the same way as with the points of P_n has important consequences. For, all the theorems that state anything about points of P_n, about collections of such points, and about relations among them are after all nothing but statements about homogeneous $(n+1)$-tuples, collections of such $(n+1)$-tuples, and *algebraic* relations among them. But, as we have just seen, a homogeneous $(n+1)$-tuple may be interpreted not only as a point but, equally well, as a hyperplane. Under this latter interpretation, our theorems will yield corresponding results about hyperplanes and collections of hyperplanes. Two statements which exactly correspond in this way, the one being expressed in point coordinates and dealing with points, the other in hyperplane coordinates and dealing with hyperplanes, are said to be **duals** (or *reciprocals*) of each other. Dual statements are nothing but different interpretations of one and the same algebraic result. The fact of this possible double interpretation is called **the principle of duality.**

In order to understand the concept and significance of this—as yet, purely formal—principle, we need to form some sort of intuitive idea of the collection of hyperplanes in question.[2] To this end, let us first investigate the following question: What is the dual of a linear space?

We begin with the simplest case, that of the hyperplane itself. A hyperplane is defined to be a collection of points; specifically, it is the totality of all the points $[\xi_0, \xi_1, \ldots, \xi_n]$ whose coordinates ξ_i satisfy a homogeneous linear equation. The dual concept would accordingly be that of the totality of all the hyperplanes $\langle u_0, u_1, \ldots, u_n\rangle$ whose coordinates u_i satisfy a homogeneous linear equation.

Let

$$(3) \qquad \gamma_0 u_0 + \gamma_1 u_1 + \cdots + \gamma_n u_n = 0$$

be such an equation, with given γ_i not all of which are zero. Observe that (by the definition of hyperplane coordinates) a point

$$[\xi_0, \xi_1, \ldots, \xi_n]$$

will lie on the hyperplane $\langle u_0, u_1, \ldots, u_n\rangle$ if and only if

[2] The usefulness of the duality principle will, of course, depend in large measure on whether or not the manifold of hyperplanes in question and the relations among them can be given an intuitive interpretation. This is the case, as we shall soon see, with the coordinates as we have chosen them.

$$u_0\xi_0 + u_1\xi_1 + \cdots + u_n\xi_n = 0.$$

This equation is therefore referred to as the *incidence condition.* Accordingly, (3) signifies the incidence of the fixed point $[\gamma_0, \gamma_1, \ldots, \gamma_n]$ with *every* hyperplane of the collection under consideration. Hence, we have

THEOREM 1. *The collection of hyperplanes defined by* (3) *consists of all the hyperplanes that pass through the point* $[\gamma_0, \gamma_1, \ldots, \gamma_n]$ *and no others.*

Such a collection is called a **hyperbundle**. The common point is called the *kernel* or *carrier* of the hyperbundle.

The dual concept of a general linear space is now the totality of hyperplanes $\langle u_0, u_1, \ldots, u_n\rangle$ whose coordinates u_i satisfy a system of homogeneous linear equations

$$(4)\qquad \begin{aligned} \gamma_{10}u_0 + \gamma_{11}u_1 + \cdots + \gamma_{1n}u_n &= 0, \\ \gamma_{20}u_0 + \gamma_{21}u_1 + \cdots + \gamma_{2n}u_n &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \gamma_{s0}u_0 + \gamma_{s1}u_1 + \cdots + \gamma_{sn}u_n &= 0. \end{aligned}$$

Let the rank of (4) be $n-r$. Then we call the totality of hyperplanes that satisfy (4) a *linear bundle of dimension* r. If we now denote by S_i (for $i = 1, 2, \ldots, s$) the point $[\gamma_{i0}, \gamma_{i1}, \cdots, \gamma_{in}]$ (which thus has the coefficients of the i-th equation in (4) as coördinates) and denote by L the linear space of least dimension (i.e., of dimension $n-r-1$) containing all the S_i, then we state the following theorem:

THEOREM 2. *The linear bundle represented by* (4) *consists precisely of all the hyperplanes that contain L. Conversely, every point common to all the hyperplanes of this bundle belongs to L.*[3]

L is called the *kernel* or *carrier* of the bundle.

Proof: The equations (4) simply mean that every hyperplane of the bundle contains the points $S_1, S_2, \ldots, S_s$ (incidence condition). A hyperplane that contains all the S_i, however, also contains L, which is, after all, the set of all linear combinations of the S_i. Thus, the first part of Theorem 2 is proved. Moreover, if $[\gamma_{s+1,0}, \gamma_{s+1,1}, \cdots, \gamma_{s+1,n}]$ is any point that lies on all the hyperplanes of our bundle, then equations (4), supplemented by the equation

[3] Thus, L is *precisely* the intersection of all the hyperplanes of this bundle.

$$\gamma_{s+1,0}\, u_0 + \gamma_{s+1,1}\, u_1 + \cdots + \gamma_{s+1,n}\, u_n = 0,$$

must still represent the same bundle; i.e., the matrix (γ_{ik}) $(i = 1, 2, \ldots, s+1;\ k = 0, 1, \ldots, n)$ has the same rank $n - r$ as (4). Consequently, the vector $\{\gamma_{s+1,0}, \gamma_{s+1,1}, \cdots, \gamma_{s+1,n}\}$ is necessarily a linear combination of the s vectors $\{\gamma_{i0}, \gamma_{i1}, \cdots, \gamma_{in}\}$ $(i = 1, 2, \ldots, s)$.[4] Thus, the point $[\gamma_{s+1,0}, \gamma_{s+1,1}, \cdots, \gamma_{s+1,n}]$ belongs to L.

A zero-dimensional linear bundle is a single hyperplane. This hyperplane itself, thought of as a linear space, is the carrier of this zero-dimensional bundle. A one-dimensional linear bundle is also referred to as a *pencil*, and its carrier is of dimension $n - 2$. And above we have already called an $(n - 1)$-dimensional bundle, a hyperbundle. Its carrier is a point. Finally, an n-dimensional bundle consists of the totality of all the hyperplanes of P_n. In this case, the carrier is empty.

It is clear that *every* linear space L can occur as the kernel of a bundle, for the coefficients γ_{ik} in (4), which determine L, can be chosen arbitrarily. Thus, we immediately have the converse of Theorem 2, namely,

Theorem 2a. *The totality of all hyperplanes containing a given linear space L represents a bundle, i.e., it can be represented in turn by a system of homogeneous linear equations.*

From the above meaning of linear bundles we shall now obtain a further result, concerning the incidence relations of bundles, which is important for the application of the principle of duality. Let B_1 and B_2 be two bundles such that $B_1 \subset B_2$; that is, let every hyperplane of the bundle B_1 belong to the bundle B_2. Let the kernels of B_1 and B_2 be K_1 and K_2, respectively. The relation $B_1 \subset B_2$ then implies that every hyperplane that contains the linear space K_1 also contains K_2, i.e., every point of K_2 belongs to *all* the hyperplanes of B_1 and hence, according to Theorem 2, to K_1 as well. In symbols: $K_2 \subset K_1$. Conversely, it follows in similar fashion from $K_2 \subset K_1$ that $B_1 \subset B_2$. We thus see that the relations $B_1 \subset B_2$ and $K_2 \subset K_1$ are *equivalent*.

Let us now apply the principle of duality to this. Let L_1 and L_2 be two linear spaces such that $L_1 \subset L_2$. The relation $L_1 \subset L_2$ implies a certain algebraic relation among the systems of equations defining L_1 and

[4] For it is (by, for example, *Modern Algebra*, § 3, Theorem 5) a linear combination of $n - r$ linearly independent vectors from among

$$\{\gamma_{10}, \gamma_{11}, \ldots, \gamma_{1n}\}, \{\gamma_{20}, \gamma_{21}, \ldots, \gamma_{2n}\}, \ldots, \{\gamma_{s0}, \gamma_{s1}, \ldots, \gamma_{sn}\}.$$

L_2. According to the principle of duality, the systems of equations can also be interpreted as linear bundles B_1 and B_2. The algebraic relation in question between the systems of equations then means for B_1 and B_2 as well that, necessarily, $B_1 \subset B_2$.

Let the kernels of B_1 and B_2 be K_1 and K_2, respectively. It follows that $K_2 \subset K_1$. Since the relations $B_1 \subset B_2$ and $K_2 \subset K_1$ are equivalent, either of them may be looked upon as the dual of $L_1 \subset L_2$.

Now consider the dimensions involved. Let the dimensions of L_1 and L_2 be r_1 and r_2, respectively. Then B_1 and B_2 also have the respective dimensions r_1 and r_2, but K_1 and K_2 have the dimensions $(n - r_1 - 1)$ and $(n - r_2 - 1)$, respectively.

Thus, if two linear spaces L_1 and L_2 with dimensions r_1 and r_2 occur in a theorem, and $L_1 \subset L_2$, then for the dual of the theorem we can either replace L_i by a bundle B_i of dimension r_i $(i = 1, 2)$ and the relation $L_1 \subset L_2$ by $B_1 \subset B_2$ or we can equally well, if suitable, replace L_i by a linear space K_i of dimension $n - r_i - 1$ $(i = 1, 2)$ and the relation $L_1 \subset L_2$ by $K_1 \supset K_2$.

This demonstrates the truth of the following important special case of the duality principle:

A theorem concerning incidence relations among linear spaces of projective P_n remains valid if we replace every dimension r that occurs in the theorem by $n - r - 1$, every relation $\subset$ (= 'is contained in') by $\supset$ (= 'contains'), and every relation $\supset$ by $\subset$.

By specializing this result in turn we obtain still other important consequences. Suppose that the linear spaces dealt with are all of dimension 0 or $n - 1$; thus, we are concerned solely with points and hyperplanes. $n - 1$ and 0 are then dual dimensions, i.e., the hyperplanes and the points are dual geometrical constructs. For this case, then, the above result reads as follows:

A theorem dealing with incidence relations among points and hyperplanes of projective P_n remains valid if the word 'point' is replaced throughout by 'hyperplane,' the word 'hyperplane' by 'point,' the relation $\subset$ by $\supset$, and the relation $\supset$ by $\subset$.

In projective P_2 there are only zero-dimensional and one-dimensional linear subspaces, namely points and lines (= hyperplanes). According to what we have just said, we must here replace 'point' by 'line' and 'line' by 'point' and interchange the relations as above. Analogously, in projective P_3 we would need to replace 'point' by 'plane,' 'line' by 'line,' and 'plane' by 'point.'

Let us now make a first application of our knowledge by dualizing some of the theorems of Chapter II on linear dependence and independence of points. Of course, the definitions of linear dependence and independence of points carry over verbatim to hyperplanes. Likewise, the correspondence between points and vectors carries over into a relation between the hyperplanes of P_n and the vectors of V_{n+1}. Also, it is clear what must be understood by a coordinate vector of a hyperplane. Then Theorems 1 and 3 of Chap. II allow of immediate dualization. We shall only formulate the dual of Theorems 2 and 4. According to the rules for dualizing given above, these theorems become:

The DUAL of Theorem 2 of Chap. II: *Let $h_1, h_2, \ldots, h_k$ be k hyperplanes in P_n, and let q be the maximal number of linearly independent hyperplanes among them. Then there is* ***no*** *linear space of dimension $n-(q-2)-1=n-q+1$ that is contained in all the h_i, but there is* ***exactly one*** *linear space of dimension $n-(q-1)-1=n-q$ that is contained in all the h_i.*

The DUAL of Theorem 4 of Chap. II: *If the hyperplanes $h_1, h_2, \ldots, h_k$ are linearly dependent, then there exists at least one linear space of dimension $n-k+1$ that is contained in all the h_i; in the other case, the (uniquely determined) linear space of largest dimension contained in all the h_i is of dimension $n-k$.*

We shall now illustrate the principle of duality by an *example* in the projective plane (i.e., in P_2). We choose as our example the *Theorem of Desargues,* which can be stated as follows:

Let the correspondingly numbered vertices S_1, S_2, S_3 and *T_1, T_2, T_3 of two triangles*[5] *be joined and the sides s_1, s_2, s_3 and t_1, t_2, t_3 be extended (where s_i and t_i are, respectively, the sides opposite S_i and T_i for $i=1, 2, 3$; Fig. 2). Then if the lines joining the pairs of points S_i and T_i $(i=1, 2, 3)$ meet in a point Z, then the three points of intersection U_i of the pairs of lines s_i and t_i $(i=1, 2, 3)$ lie on a line.*

For the purposes of the proof we restrict ourselves to the case in which each of the pairs of points S_i, T_i $(i=1, 2, 3)$ consists of two *distinct* points and in which $s_i \neq t_i$ for every i. In the other cases, both the hypothesis and the conclusion of the theorem are trivially satisfied.[6] Let the coordi-

[5] Triangle = three linearly independent points.

[6] Because of the possible indeterminacy of the lines joining S_i and T_i as well as those joining the U_i, the theorem is to be interpreted in these exceptional cases as follows: If there is a point Z such that S_i, T_i, and Z all lie on a line for each value of i, then there also exists a line g such that s_i, t_i, and g all belong to a pencil for each value of i. The general case, to be sure, is also correctly given by this formulation. But we have chosen the above wording because of its greater intuitive appeal.

nate vector of Z be denoted by $\mathfrak{z}$, that of S_i by $\mathfrak{x}_i$ $(i = 1, 2, 3)$ and that of T_i by $\mathfrak{y}_i$ $(i = 1, 2, 3)$.

Then $\mathfrak{z}$ can be represented by a linear combination of the $\mathfrak{x}_i$ and $\mathfrak{y}_i$ for each $i = 1, 2, 3$, say:

$$\mathfrak{z} = \lambda_1 \mathfrak{x}_1 + \mu_1 \mathfrak{y}_1 = \lambda_2 \mathfrak{x}_2 + \mu_2 \mathfrak{y}_2 = \lambda_3 \mathfrak{x}_3 + \mu_3 \mathfrak{y}_3 . \tag{5}$$

Hence,

$$\begin{aligned} \lambda_1 \mathfrak{x}_1 - \lambda_2 \mathfrak{x}_2 &= \mu_2 \mathfrak{y}_2 - \mu_1 \mathfrak{y}_1 . \\ \lambda_2 \mathfrak{x}_2 - \lambda_3 \mathfrak{x}_3 &= \mu_3 \mathfrak{y}_3 - \mu_2 \mathfrak{y}_2 , \\ \lambda_3 \mathfrak{x}_3 - \lambda_1 \mathfrak{x}_1 &= \mu_1 \mathfrak{y}_1 - \mu_3 \mathfrak{y}_3 . \end{aligned} \tag{6}$$

This means the following: The point with the coordinate vector $\lambda_1\mathfrak{x}_1 - \lambda_2\mathfrak{x}_2$ is both a linear combination of S_1 and S_2 and also (by the first equation of (6)) a linear combination of T_1 and T_2. It must therefore lie

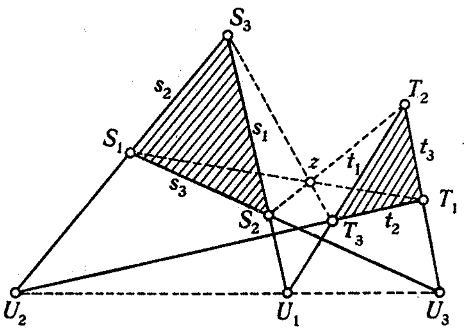

Fig. 2

on both s_3 and t_3 and so must coincide with the point U_3 (Fig. 2). In exactly the same way, it follows that $\lambda_2 \mathfrak{x}_2 - \lambda_3 \mathfrak{x}_3$ is a coordinate vector of U_1 and $\lambda_3 \mathfrak{x}_3 - \lambda_1 \mathfrak{x}_1$ a coordinate vector of U_2.

The obvious identity

$$(\lambda_1 \mathfrak{x}_1 - \lambda_2 \mathfrak{x}_2) + (\lambda_2 \mathfrak{x}_2 - \lambda_3 \mathfrak{x}_3) + (\lambda_3 \mathfrak{x}_3 - \lambda_1 \mathfrak{x}_1) = 0$$

then shows that U_1, U_2, U_3 are linearly dependent and therefore lie on a line, as was to be proved.

If we now dualize the theorem of Desargues, we obtain:

If the points of intersection of the pairs of lines s_i, t_i are collinear, then the lines connecting the pairs of points S_i, T_i are concurrent at a point in Z.

In this case, the dual of the theorem is its exact converse!

The definition of a general projective coordinate system as given in Chapter II can, of course, be applied to bundles of hyperplanes. This is not, however, of great importance. But another question that *is* significant is the following:

Let $(Q_0, Q_1, \ldots, Q_n \mid E)$ and $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$ be two projective coordinate systems in P_n. Let h be a fixed hyperplane. According to Theorem 9 of Chap. II, this may be represented in each of the coordinate systems by a linear equation, say by

$$u_0 \xi_0 + u_1 \xi_1 + \cdots + u_n \xi_n = 0$$

in $(Q_0, Q_1, \ldots, Q_n \mid E)$ and by

$$u_0^* \xi_0^* + u_1^* \xi_1^* + \cdots + u_n^* \xi_n^* = 0$$

in $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$, where ξ_i and ξ_i^* designate *point* coordinates in $(Q_0, Q_1, \ldots, Q_n \mid E)$ and $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$, respectively. Then the u_i are uniquely determined up to a constant of proportionality, as are the u_i^*. The u_i and the u_i^* are referred to as the *hyperplane coordinates in the coordinate system* $(Q_0, Q_1, \ldots, Q_n \mid E)$ and $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$, respectively.

Let us now consider the question, What relation exists between the u_i and the u_i^*?

By Theorem 7, Chap. II, we know that there exists a non-singular matrix T such that

$$(\xi^*) = \varrho \cdot T \cdot (\xi), \tag{7}$$

where (ξ^*) and (ξ) are matrices analogous to those in (10) of Chap. II. If in the same way we construct matrices (u) and (u^*) with the u_i and u_i^*, respectively, we can also write the equations of hyperplane h in the two coordinate systems in matrix form: $(u)'(\xi) = O$ in $(Q_0, Q_1, \cdots, Q_n \mid E)$ and $(u^*)'(\xi^*) = O$ in $(Q_0^*, Q_1^*, \cdots, Q_n^* \mid E^*)$.[7] If we substitute (7) in the second of these equations, we obtain:

$$\varrho \cdot (u^*)' \cdot T \cdot (\xi) = O. \tag{8}$$

As follows from its derivation, this equation will be satisfied if and only if the ξ_i are the coordinates of a point of h in the coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$. The matrix $\varrho \cdot (u^*)' \cdot T$ has the following form: The second to the $(n + 1)$-st rows inclusive consist entirely of zeros. The

[7] A' denotes the transpose of A. (This is in conformity with *Modern Algebra*, p. 301.)

first row therefore contains just the hyperplane coordinates of h in $(Q_0, Q_1, \ldots, Q_n \mid E)$. Thus, except for a constant of proportionality $(\neq 0)$, $\varrho(u^*)' \cdot T$ is equal to $(u)'$. Combining the constant of proportionality with ϱ, we may write

$$(u)' = \lambda (u^*)' \cdot T,$$

or, by *Modern Algebra*, § 22,

$$(u) = \lambda T' (u^*), \tag{9}$$

$$(u^*) = \mu (T')^{-1}(u) = \mu (T^{-1})' \cdot (u). \tag{10}$$

The equations in (10) parallel (7). They also represent a non-singular linear substitution, but with a different matrix. We call the substitution (10) *contragredient* to (7), and we accordingly say:

Hyperplane coordinates and point coordinates transform contragrediently to each other upon passage to a new projective coordinate system.

We should like to conclude this chapter by going briefly into the question of the natural limits of the applicability of the principle of duality. There is of course no intrinsic reason why we cannot set up a dual in hyperplane coordinates to match every definition and relation in point coordinates. However, it turns out that for a number of concepts the dual has no particular geometrical meaning. Take, for example, the concept of improper points. In many connections (for example, parallelism), these points play a special role among the points of P_n. But how about their dual set of hyperplanes? Since a point $[\xi_0, \xi_1, \ldots, \xi_n]$ is improper if $\xi_0 = 0$, the corresponding dual is the set of hyperplanes $\langle u_0, u_1, \ldots, u_n \rangle$ with $u_0 = 0$. These hyperplanes have no particular geometrical distinctiveness, however. They do not even include the improper hyperplane, whose coordinates, in fact, are $\langle 1, 0, \ldots, 0 \rangle$. Hence the dualization of statements that involve improper points or concepts derived from them (e.g., parallelism) is without geometrical meaning or importance.[8]

We shall return to these non-dualizable concepts in a later chapter, and we shall then have the means of determining more exactly the scope of the principle of duality (cf. Chap. XI, pp. 173–174).

[8] Any attempt to establish the improper points as dual of the (single) improper hyperplane (by a different choice of coordinates, say) is, of course, doomed to failure, since the mere difference in *number* of the improper elements (at least, for $n \geqq 2$) destroys the symmetry of the hypotheses about points and hyperplanes.

Exercises

1. Let two triples of points, say 1, 2, 3 and 1′, 2′, 3′, be given in the projective plane and let them, moreover, be in *perspective position*; that is to say, there exists some point Z such that i, i', and Z are collinear for each i. Denote the line joining the two points i and k by g_{ik}, the line joining the points i' and k' by $g_{i'k'}$, and that joining the points i and k' by $g_{ik'}$.

Consider the following six pairs of lines:

$$g_{12},\ g_{1'2'};\qquad g_{23},\ g_{2'3'};\qquad g_{13},\ g_{1'3'};$$
$$g_{12'},\ g_{21'};\qquad g_{23'},\ g_{32'};\qquad g_{13'},\ g_{31'}.$$

Each pair will in general determine a point of intersection.

Show that there are four (new) lines whose six points of intersection are identical with the six points of intersection of our six pairs of lines. What does the dual theorem state?

As a special case, if the triples 1, 2, 3, and 1′, 2′, 3′ each lie on a straight line, some of the above six points of intersection coincide. Show that the four distinct remaining points of intersection all lie on a single line. What does the dual theorem state?

2. Let two lines g_1 and g_2 with point of intersection S be given in the projective plane. Let 1, 2, 3 be three points on g_1 and 1′, 2′, 3′ be three points on g_2. Let $g_{ik'}$ be the line joining i with k'. Extend the lines $g_{12'}$ and $g_{21'}$ until they intersect. Call the point of intersection S_3. Likewise, let S_2 be the point of intersection of $g_{13'}$ and $g_{31'}$ and S_1 that of $g_{23'}$ and $g_{32'}$.

Show that the following statement is a direct consequence of the special case of Exercise 1: If *two* of the points S_1, S_2, S_3 are collinear with S, then so is the third.

State and prove the dual.

CHAPTER IV

THE CROSS RATIO

Let g be a line in (real or complex) P_n and let $(Q_1, Q_2 \mid E)$ be a projective coordinate system on g (cf. Chap. II). Also let Q be a point on g distinct from Q_1, with the coordinates λ_1, λ_2 in $(Q_1, Q_2 \mid E)$. Since $Q \neq Q_1$, it follows that $\lambda_2 \neq 0$, and consequently the fraction $\lambda_1 : \lambda_2$ is completely determined by the four points Q_1, Q_2, E, and Q. It is called the *cross ratio* (also, anharmonic ratio and double ratio) of the four points Q_1, Q_2, E, Q and is denoted by the symbol $\mathcal{R}(Q_1\, Q_2\, E\, Q)$. Thus,

$$\mathcal{R}\,(Q_1\, Q_2\, E\, Q) = \frac{\lambda_1}{\lambda_2}. \tag{1}$$

In this definition, the coordinate vectors $\mathfrak{x}$ and $\mathfrak{y}$ of Q_1 and Q_2 respectively are normalized in such a way that the coordinate vector of E is equal to $\mathfrak{x} + \mathfrak{y}$ (by virtue of the fact that $(Q_1, Q_2 \mid E)$ was to be a projective coordinate system). The coordinate vector of Q is, of course, $\lambda_1\mathfrak{x} + \lambda_2\mathfrak{y}$. But let us refrain from taking this special normalization into account. Thus let $\mu_1\mathfrak{x} + \mu_2\mathfrak{y}$, say, be a coordinate vector of E and $\lambda_1\mathfrak{x} + \lambda_2\mathfrak{y}$, once again, a coordinate vector of Q. Now what is the cross ratio $\mathcal{R}(Q_1\, Q_2\, E\, Q)$, computed in terms of the four numbers $\mu_1, \mu_2, \lambda_1, \lambda_2$?

To bring this within the scope of the preceding definition, let $\mu_1\,\mathfrak{x} = \mathfrak{x}'$ and $\mu_2\,\mathfrak{y} = \mathfrak{y}'$.[1] Then the coordinate vector of Q can be written as follows as a linear combination of $\mathfrak{x}'$ and $\mathfrak{y}'$:

$$\frac{\lambda_1}{\mu_1}(\mu_1\,\mathfrak{x}) + \frac{\lambda_2}{\mu_2}(\mu_2\,\mathfrak{y}) = \frac{\lambda_1}{\mu_1}\mathfrak{x}' + \frac{\lambda_2}{\mu_2}\mathfrak{y}'.$$

By our first definition, we thus have

$$\mathcal{R}\,(Q_1\, Q_2\, E\, Q) = \frac{\lambda_1}{\mu_1} : \frac{\lambda_2}{\mu_2}. \tag{2}$$

[1] For E now does have $\mu_1\,\mathfrak{x} + \mu_2\,\mathfrak{y} = \mathfrak{x}' + \mathfrak{y}'$ as coordinate vector.

Note that the order in which the four points occur in the symbol $\mathcal{R}(Q_1\, Q_2\, E\, Q)$ is essential, for the definition of the cross ratio is by no means symmetric in the four points.

It follows directly from the definition that if $Q \neq Q'$, then $\mathcal{R}(Q_1\, Q_2\, E\, Q) \neq \mathcal{R}(Q_1\, Q_2\, E\, Q')$.

In (1), Q can stand for any point on the line g, with the exception of Q_1. In order to eliminate this exceptional role of Q_1, at least in the notation, we occasionally write, symbolically,[2] $\mathcal{R}(Q_1\, Q_2\, E\, Q_1) = \infty$.

In order to investigate how the cross ratio depends upon the order of the four points, we begin with a more general question. Let S_1, S_2, S_3, S_4 be four distinct points on g and let all of the S_i, moreover, be distinct from the points Q_1, Q_2, E.[3] Also, let $\varkappa_i = \mathcal{R}(Q_1\, Q_2\, E\, S_i)$ be given for $i = 1, 2, 3, 4$. Then what is $\mathcal{R}(S_1\, S_2\, S_3\, S_4)$ in terms of the $\varkappa_i$?

To answer this question, we must determine the coordinates of S_4 in the coordinate system $(S_1\, S_2 \mid S_3)$. To this end, let $\mathfrak{x}$ and $\mathfrak{y}$ again denote the coordinate vectors of Q_1 and Q_2 respectively and $\mathfrak{x} + \mathfrak{y}$ a coordinate vector of E. By definition of the cross ratio, $\mathfrak{z}_i = \varkappa_i\, \mathfrak{x} + \mathfrak{y}$ is a coordinate vector of S_i. From this we immediately obtain the following equations:

$$(\varkappa_1 - \varkappa_2)\, \mathfrak{z}_3 = (\varkappa_3 - \varkappa_2)\, \mathfrak{z}_1 + (\varkappa_1 - \varkappa_3)\, \mathfrak{z}_2\,,$$
$$(\varkappa_1 - \varkappa_2)\, \mathfrak{z}_4 = (\varkappa_4 - \varkappa_2)\, \mathfrak{z}_1 + (\varkappa_1 - \varkappa_4)\, \mathfrak{z}_2\,.$$

Since all of the $\varkappa_i$ are different[4] from each other, we have at once from (2):

$$\mathcal{R}(S_1\, S_2\, S_3\, S_4) = \frac{\varkappa_4 - \varkappa_2}{\varkappa_3 - \varkappa_2} : \frac{\varkappa_1 - \varkappa_4}{\varkappa_1 - \varkappa_3}\,,$$

or

$$\mathcal{R}(S_1\, S_2\, S_3\, S_4) = \frac{\varkappa_3 - \varkappa_1}{\varkappa_3 - \varkappa_2} : \frac{\varkappa_4 - \varkappa_1}{\varkappa_4 - \varkappa_2}\,. \tag{3}$$

By the use of formula (3), we can easily get a clear picture of all the possible values that the cross ratio can have under various orderings of the points S_1, S_2, S_3, S_4. All these values may be obtained by permutation of the subscripts in the right-hand side of (3). Among these permutations there are certainly some that leave the right-hand side of (3)

---

[2] Here ∞ is to be understood as a pure symbol and by no means as a number of infinitely large value. The notation $\mathcal{R}(Q_1\, Q_2\, E\, Q) = \infty$ merely means that $Q = Q_1$.

[3] For any given S_i, it is always possible to choose the coordinate system so that this condition is satisfied.

[4] Since the S_i, by assumption, are all distinct from each other.

unaltered: the identity permutation, for example, which leaves all the subscripts the same, has this property. Now, all the permutations that leave the expression on the right-hand side of (3) unaltered constitute a subgroup of the complete permutation group[5] on the four given points S_1, S_2, S_3, S_4, because the product of any two permutations with this property, itself has this property. Call this subgroup H.

We then have the following result: *Two permutations of the subscripts in the right-hand side of* (3) *yield the same result if and only if, as elements of the complete permutation group, they belong to the same left coset of* H. For if τ_1 and τ_2 are two permutations having the same effect on the right-hand side of (3), then $\tau_1^{-1}\tau_2$ leaves the right-hand side of (3) unaltered. Thus, $\tau_1^{-1}\cdot\tau_2 = \sigma$, with σ in H. From this it follows that $\tau_2 = \tau_1\cdot\sigma$, i.e., τ_1 and τ_2 belong to the same left coset of H. Conversely, two elements of a given left coset of H always yield the same result, since the application of a product $\tau\cdot\sigma$, with σ in H and τ an arbitrary permutation, yields the same result as the application of τ alone.[6]

We can easily determine H. First of all, observe that the four permutations that take (1, 2, 3, 4) into

$$(1,2,3,4),\quad (2,1,4,3),\quad (3,4,1,2),\quad (4,3,2,1)$$

leave the right-hand side of (3) unaltered. Thus, the order of H is $\geqq 4$. On the other hand, the complete permutation group on four points has order 24, so that the index of H in the complete permutation group must be $\leqq 6$. But by what has been said above, this index must also be the number of *different* expressions obtainable from the right-hand side of (3) by a permutation of the subscripts. Thus, if we can produce six permutations which, when applied to (3), give completely different results, then the index of H must be exactly equal to 6 and, consequently, the order of H exactly equal to 4. But six such permutations can indeed be found. We can take, for instance, the permutations that carry (1, 2, 3, 4) into

$$(1,2,3,4),\quad (1,2,4,3),\quad (1,3,2,4),\quad (1,3,4,2),$$
$$(1,4,2,3),\quad (1,4,3,2).$$

Finally, if we determine the elements of all six cosets and denote $\mathcal{R}(S_1 S_2 S_3 S_4)$ by σ, we obtain

[5] For the definition of complete permutation group, see *Modern Algebra*, p. 256. (The special case of the complete permutation group that we are dealing with here is also called the *symmetric group* on four elements.)

[6] Since σ does not alter the right-hand side of (3).

$$
(4)\quad
\begin{aligned}
\sigma &= \mathcal{R}(S_1 S_2 S_3 S_4) = \mathcal{R}(S_2 S_1 S_4 S_3) = \mathcal{R}(S_3 S_4 S_1 S_2) \\
&= \mathcal{R}(S_4 S_3 S_2 S_1), \\
\frac{1}{\sigma} &= \mathcal{R}(S_1 S_2 S_4 S_3) = \mathcal{R}(S_2 S_1 S_3 S_4) = \mathcal{R}(S_4 S_3 S_1 S_2) \\
&= \mathcal{R}(S_3 S_4 S_2 S_1), \\
1-\sigma &= \mathcal{R}(S_1 S_3 S_2 S_4) = \mathcal{R}(S_3 S_1 S_4 S_2) = \mathcal{R}(S_2 S_4 S_1 S_3) \\
&= \mathcal{R}(S_4 S_2 S_3 S_1), \\
\frac{\sigma-1}{\sigma} &= \mathcal{R}(S_1 S_4 S_2 S_3) = \mathcal{R}(S_4 S_1 S_3 S_2) = \mathcal{R}(S_2 S_3 S_1 S_4) \\
&= \mathcal{R}(S_3 S_2 S_4 S_1), \\
\frac{1}{1-\sigma} &= \mathcal{R}(S_1 S_3 S_4 S_2) = \mathcal{R}(S_3 S_1 S_2 S_4) = \mathcal{R}(S_4 S_2 S_1 S_3) \\
&= \mathcal{R}(S_2 S_4 S_3 S_1), \\
\frac{\sigma}{\sigma-1} &= \mathcal{R}(S_1 S_4 S_3 S_2) = \mathcal{R}(S_4 S_1 S_2 S_3) = \mathcal{R}(S_3 S_2 S_1 S_4) \\
&= \mathcal{R}(S_2 S_3 S_4 S_1).
\end{aligned}
$$

Now when do any of these values $\sigma, \frac{1}{\sigma}$, etc., coincide with any of the others? To find these cases, it is sufficient to investigate the possibility that σ be equal to any one of the remaining five values, $\frac{1}{\sigma}$, $1-\sigma$, $\frac{\sigma-1}{\sigma}$, $\frac{1}{1-\sigma}$, and $\frac{\sigma}{\sigma-1}$.[7]

By setting σ equal to $\frac{1}{\sigma}$, we get $\sigma^2 = 1$, or $\sigma = \pm 1$. $\sigma = +1$ means[8] that $S_3 = S_4$. This is a degenerate case, which we shall exclude for the present, since we assumed to begin with that the S_i are all distinct. $\sigma = -1$ yields

$$\sigma = \frac{1}{\sigma} = -1, \quad 1-\sigma = \frac{\sigma-1}{\sigma} = 2, \quad \frac{1}{1-\sigma} = \frac{\sigma}{\sigma-1} = \frac{1}{2}.$$

Thus we obtain only three different values for the cross ratio, namely $-1, 2, 1/2$. In this case, the four points S_1, S_2, S_3, S_4 are called a **harmonic set**.

If the points S_1, S_2, S_3, S_4 of a harmonic set are arranged in such an order that $\mathcal{R}(S_1 S_2 S_3 S_4) = -1$, then *the pairs of points S_1, S_2 and S_3, S_4 are said to separate each other harmonically.*

[7] Since any one of the six possible values can be denoted by σ, the remaining five values are then, by (4), given by $\frac{1}{\sigma}$, $1-\sigma$, $\frac{\sigma-1}{\sigma}$, $\frac{1}{1-\sigma}$ and $\frac{\sigma}{\sigma-1}$.

[8] Cf. the defining equation (1) of the cross ratio.

If we put $\sigma = 1 - \sigma$, we again get a harmonic set. And likewise, putting σ equal to $\frac{\sigma}{\sigma - 1}$ yields, in the non-trivial case, only another harmonic set.

On the other hand, by setting σ equal to either of the two values $\frac{1}{1-\sigma}$, $\frac{\sigma - 1}{\sigma}$, we get the equation $\sigma^2 - \sigma + 1 = 0$ for σ. This equation has *no* real solution. This case is therefore ruled out in real P_n; but in complex P_n, it is possible. σ can then be either of the roots of

$$\sigma^2 - \sigma + 1 = 0,$$

i.e.,

$$\sigma = \frac{1 \pm i\sqrt{3}}{2}.$$

We then get

$$\sigma = \frac{1}{1-\sigma} = \frac{\sigma - 1}{\sigma} = \frac{1 + i\sqrt{3}}{2}, \quad \frac{1}{\sigma} = 1 - \sigma = \frac{\sigma}{\sigma - 1} = \frac{1 - i\sqrt{3}}{2},$$

where $\sqrt{3}$ in each instance stands for, say, the positive root. Thus we get only two distinct values for the cross ratio. In this case, the four points S_1, S_2, S_3, S_4 are said to form an *equianharmonic set.*

The equianharmonic case is not of great interest for us, but we shall soon be concerned with the harmonic case in some detail.

It is customary to extend the validity of formula (3) to that case in which only three of the points S_i are distinct from each other, and this is done by stipulating that

$$\mathcal{R}(S_1 S_2 S_3 S_4) = \frac{(x_3 - x_1)(x_4 - x_2)}{(x_3 - x_2)(x_4 - x_1)}$$

provided $x_3 - x_2$ and $x_4 - x_1$ are $\neq 0$, but that

$$\mathcal{R}(S_1 S_2 S_3 S_4) = \infty$$

if either $x_3 - x_2$ or $x_4 - x_1$ is 0.[9] In particular, $\mathcal{R}(S_1 S_2 S_3 S_1) = \infty$, which is in accord with our earlier agreement.

By virtue of this extension of our definition, there is one more case in which the cross ratio assumes only three different values, namely, that in which two of the points S_i coincide. If, for example, $S_3 = S_4$, then

[9] Only one at most of the four differences

$$x_3 - x_1, \; x_3 - x_2, \; x_4 - x_1, \; x_4 - x_2$$

can vanish, since there are at least three distinct points.

$$\sigma = \frac{1}{\sigma} = 1, \quad 1-\sigma = \frac{\sigma-1}{\sigma} = 0, \quad \frac{1}{1-\sigma} = \frac{\sigma}{\sigma-1} = \infty.$$

In the following, we examine the dual to the above definitions and arguments. To this end, instead of starting out, as on page 52, with the line g, we must start out with a pencil b of hyperplanes, i.e., with a *one-dimensional* linear bundle (in P_2, a pencil of lines; in P_3, a pencil of planes). Also, let q_1, q_2, and e be three hyperplanes of b with the respective coordinate vectors $\mathfrak{u}_1$, $\mathfrak{u}_2$, and $\mathfrak{u}_1 + \mathfrak{u}_2$. Then let $q \neq q_1$ be a further hyperplane of b, with a coordinate vector $\mathfrak{v}$, and let $\mathfrak{v} = \lambda_1 \mathfrak{u}_1 + \lambda_2 \mathfrak{u}_2$. Then $\lambda_1 : \lambda_2$ will be the cross ratio of the four hyperplanes q_1, q_2, e, q, and in analogy to (1), we shall set

$$\mathcal{R}(q_1\, q_2\, e\, q) = \frac{\lambda_1}{\lambda_2}. \tag{5}$$

The rest of what we have already said as regards points now carries over word for word; we need merely replace everywhere the expression 'point on the line g' by 'hyperplane of the pencil b.'

Equations (4) in particular remain valid if we take the S_i to mean four distinct planes of the pencil b. If $\mathcal{R}(h_1\, h_2\, h_3\, h_4) = -1$, we say, as before, that *the four hyperplanes h_i constitute a harmonic set,* or that *the pairs of hyperplanes h_1, h_2 and h_3, h_4 separate each other harmonically.*

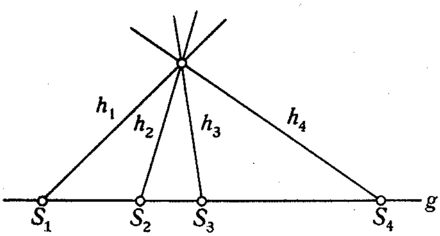

Fig. 3

The following is a fact of particular importance. Let us consider, in P_2, *four lines* h_1, h_2, h_3, h_4 of a pencil (Fig. 3), and let their points of intersection with another line g, *not* of the pencil, be respectively S_1, S_2, S_3, S_4. Then

$$\mathcal{R}(h_1\, h_2\, h_3\, h_4) = \mathcal{R}(S_1\, S_2\, S_3\, S_4).$$

The proof that we shall now give will, of course, be of the n-dimensional generalization of this theorem. In this case, h_1, h_2, h_3, h_4 shall mean four hyperplanes of the pencil b, and g, a *line that does not pass through the carrier*[1] *L of this pencil.*

[1] The carrier L is an $(n-2)$-dimensional linear space (Chap. III); for example, in P_2 it is a point, and in P_3, a line.

It follows at once that g cannot belong to any hyperplane h of b. For if $g \subset h$, then the spanning space of g and L would also be contained in h, i.e., the spanning space would be at most $(n-1)$-dimensional. By Theorem 5 of Chap. II, it would then follow (if we recall that g is one-dimensional and L $(n-2)$-dimensional) that the intersection of g and L is non-empty. But this contradicts our assumption.

On the other hand, the intersection of g and h can never be empty, and so must consist of exactly one point.

And conversely, every point of g lies on exactly one hyperplane of b. For since the intersection of L with any point Q of g is always empty,[2] the spanning space of L and Q, by Theorem 5 of Chap. II, is of dimension $n-1$, i.e., the spanning space is a hyperplane. As a spanning space, however, this hyperplane is uniquely determined and moreover also belongs to b, since it contains L.

Now if S_i is the point of intersection of h_i with g, we may state the following theorem.

THEOREM 1. *If, for $i=1, 2, 3$, the points S_i of g are incident with h_i, then the equation*

$$\mathcal{R}(S_1\, S_2\, S_3\, S_4) = \mathcal{R}(h_1\, h_2\, h_3\, h_4)$$

is a necessary and sufficient condition for the incidence of S_4 with h_4.

The theorem is trivial if two of the elements h_i or S_i coincide. For the coincidence of two elements h_i, h_k necessarily entails the coincidence of the correspondingly numbered S_i, S_k, and conversely. Furthermore, the coincidence of two elements implies a definite value for the cross ratio (namely, 0, 1, or ∞; cf. p. 55).

We may therefore assume the S_i, and the h_i as well, to be distinct. Then let the coordinate vectors of S_1, S_2, S_3 and of h_1, h_2, h_3 be respectively $\mathfrak{x}_1, \mathfrak{x}_2, \mathfrak{x}_3$ and $\mathfrak{u}_1, \mathfrak{u}_2, \mathfrak{u}_3$, and let them, moreover, be normalized in such a way that $\mathfrak{x}_3 = \mathfrak{x}_1 + \mathfrak{x}_2$ and $\mathfrak{u}_3 = \mathfrak{u}_1 + \mathfrak{u}_2$. If we set $\mathcal{R}(S_1\, S_2\, S_3\, S_4) = \lambda$ and $\mathcal{R}(h_1\, h_2\, h_3\, h_4) = \mu$, then by the definition of cross ratio, the vectors $\lambda\mathfrak{x}_1 + \mathfrak{x}_2$ and $\mu\,\mathfrak{u}_1 + \mathfrak{u}_2$ are the coordinate vectors of S_4 and h_4, respectively.

Since, by assumption, S_i is incident with h_i for $i=1, 2, 3$, we have[3]

$$(6) \qquad \mathfrak{u}_1\mathfrak{x}_1 = 0, \qquad \mathfrak{u}_2\mathfrak{x}_2 = 0, \qquad (\mathfrak{u}_1 + \mathfrak{u}_2)(\mathfrak{x}_1 + \mathfrak{x}_2) = 0.$$

From the last equation of (6), in conjunction with the first two, it follows that

[2] I.e., Q does not lie in L.

[3] Here $\mathfrak{u}_i\mathfrak{x}_i$ denotes the scalar product of the two vectors.

$$\mathfrak{u}_1 \mathfrak{x}_2 + \mathfrak{u}_2 \mathfrak{x}_1 = 0. \tag{7}$$

The incidence in question of S_4 and h_4 is determined by the behaviour of the expression $(\mu\, \mathfrak{u}_1 + \mathfrak{u}_2)(\lambda\, \mathfrak{x}_1 + \mathfrak{x}_2)$.

From (6) and (7), we obtain

$$(\mu\, \mathfrak{u}_1 + \mathfrak{u}_2)(\lambda\, \mathfrak{x}_1 + \mathfrak{x}_2) = (\mu - \lambda)\, \mathfrak{u}_1 \mathfrak{x}_2. \tag{8}$$

Now $\mathfrak{u}_1\mathfrak{x}_2$ is certainly $\neq 0$, since S_2 is not incident with h_1. Hence, the left-hand side of (8) will vanish *if and only if* $\mu = \lambda$. Thus, Theorem 1 is proved.

If we assign to each point of the line g that hyperplane of the pencil b which goes through the point, and conversely, then we have, by definition, a *one-to-one* correspondence between the points of g and the hyperplanes of b. Such a correspondence is called a *perspectivity,* and we shall say: *g and b are perspective* or *in perspective.*[4]

Theorem 1 can now be rephrased as follows:

The cross ratio is invariant under perspectivities between g and b.

This theorem may be generalized still further. Consider two lines g and g' and a pencil b that is perspective to both. Then a one-to-one correspondence between g and g' is defined by associating with each other the points of g and of g' that lie in one and the same hyperplane. Such a relation between g and g' is likewise called a perspectivity. Of course, the cross ratio is also invariant under perspectivities of this kind. Perspectivities between two pencils b and b' may be defined dually.

For the sake of later application, we give here the following *example* of a perspectivity:

Let e be a (two-dimensional) plane of P_n $(n \geqq 2)$ and let $g, \bar{g}$ be two lines in e. Let Q be a point in e that does not lie either on g or on $\bar{g}$. We now define a one-to-one correspondence between g and $\bar{g}$ as follows:

The image point of a point S of g shall be the point of intersection $\bar{S}$ of $\bar{g}$ with the line SQ.

We claim that *this mapping is a perspectivity.*

Proof: If $n = 2$, then $e = P_2$, and our assertion is trivial (for it is then actually equivalent to the definition of perspectivity).

[4] It must not be forgotten that for such a correspondence to be possible, g must not pass through the kernel of b.

If $n > 2$, then we select a linear space L of dimension $n - 2$, having the point Q but nothing else in common with e.[5] Then g and $\bar{g}$ have no points in common with L. All the hyperplanes that contain L form a pencil b. A hyperplane of b passing through a point S of g contains the entire straight line SQ and therefore the image point $\bar{S}$. We have thus shown that the correspondence defined above is a perspectivity.

Let us take up, as our next topic, a geometrical interpretation of the harmonic set.

Let S_1, S_2, S_3, S_4 be a harmonic set on a line g. Then S_4 is uniquely determined by the points S_1, S_2, S_3 and the condition $\mathcal{R}(S_1 S_2 S_3 S_4) = -1$.

S_4 can be found by a simple geometrical construction. This is accomplished as follows (Fig. 4): Let A be an auxiliary point not on g, and let

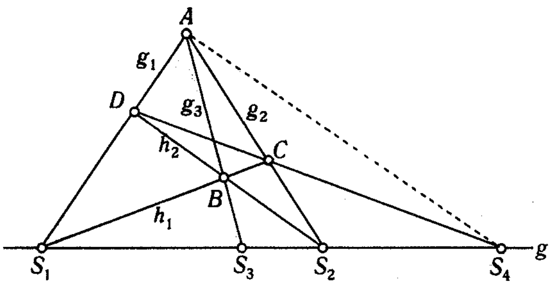

Fig. 4

g_1, g_2, g_3 be the lines joining A with S_1, S_2, S_3, respectively. Let B be any point on g_3 distinct from A and S_3, and denote by h_1 the line joining B with S_1 and by h_2 the line joining B with S_2.

The lines h_1 and g_2 are certainly not identical, for h_1 contains the point S_1, which does not lie on g_2.[6] On the other hand, h_1 and g_2, like all the elements of the construction, lie in the plane determined by S_1, S_2, A, i.e., in a *two*-dimensional linear space. This plane is the spanning space of h_1 and g_2. By Theorem 5 of Chap. II, h_1 and g_2 therefore have just *one* point of intersection, say C. C is certainly different from S_2 and A,

[5] The existence of such a linear space can be seen as follows: Let Q_1, Q_2 be points of e such that Q, Q_1, Q_2 are linearly independent. In addition, let $Q_3, Q_4, \ldots, Q_n$ be points of P_n such that $Q, Q_1, Q_2, Q_3, \ldots, Q_n$ in their totality are still linearly independent. Then the linear space L of dimension $n - 2$ containing the points $Q, Q_3, Q_4, \ldots, Q_n$ gives the desired result. For, the spanning space of L and e is the whole of P_n, whence the intersection of L and e, in accordance with Theorem 5 of Chap. II, is zero-dimensional and is thus equal to Q.

[6] For A, S_1, S_2 are linearly independent by assumption.

for otherwise B would have to coincide with S_3 or with A, contrary to our assumption.

In exactly the same way it can be shown that h_2 and g_1 have in common just one point D, distinct from S_1 and A.

The line joining C and D is therefore different from g and consequently intersects g in a single point.

We assert the following: *This point of intersection of CD and g is precisely the fourth harmonic point S_4.*

Proof: We take S_1, S_2, A as fundamental points and B as unit point of a projective coordinate system in the plane determined by S_1, S_2, A. Let the coordinate vectors of S_1, S_2, A be $\mathfrak{x}, \mathfrak{y}, \mathfrak{z}$, respectively. Then $\mathfrak{x} + \mathfrak{y} + \mathfrak{z}$ is a coordinate vector of B.

Now since the vector $\mathfrak{x} + \mathfrak{y}$ is a linear combination of $\mathfrak{x}$ and $\mathfrak{y}$ and also a linear combination of $\mathfrak{x} + \mathfrak{y} + \mathfrak{z}$ and $\mathfrak{z}$, the point with coordinate vector $\mathfrak{x} + \mathfrak{y}$ must lie both on g and on g_3 and hence must be the uniquely determined point of intersection S_3. Thus $\mathfrak{x} + \mathfrak{y}$ is a coordinate vector of S_3. In the very same way, it can be shown that $\mathfrak{x} + \mathfrak{z}$ is a coordinate vector of D, and $\mathfrak{y} + \mathfrak{z}$ a coordinate vector of C.

We have still to find a coordinate vector for the point of intersection of g with the line through C and D. Because of the uniqueness of this point of intersection, it is once again a matter of finding a vector which is both a linear combination of $\mathfrak{x} + \mathfrak{z}$ (point D) and $\mathfrak{y} + \mathfrak{z}$ (point C) and a linear combination of $\mathfrak{x}$ and $\mathfrak{y}$. Such a vector is $\mathfrak{x} - \mathfrak{y}$.

Thus we see that the four points S_1, S_2, S_3, S_4 have $\mathfrak{x}$, $\mathfrak{y}$, $\mathfrak{x} + \mathfrak{y}$, $\mathfrak{x} - \mathfrak{y}$ respectively as coordinate vectors. But from this it follows that $\mathcal{R}(S_1\, S_2\, S_3\, S_4) = -1$, which was to be proved.

The above construction of the point S_4, the so-called *fourth harmonic point*, employs just two basic operations (carried out a number of times and in a definite sequence), namely:

1. Passing a line through two points;
2. Forming the intersection of two lines.

Constructions of this kind, which involve these two operations only, are called *linear constructions.* Thus we can state:

THEOREM 2. *The fourth harmonic point can be obtained by means of a linear construction.*

The possibility of the above construction is also called the *Theorem of the Complete Quadrilateral.* This name comes about in the following way:

A *complete quadrilateral* is defined as the plane figure consisting of four lines (no three of which go through the same point) and their six points of intersection. The lines are called sides of the quadrilateral and the points of intersection are called its vertices. Two vertices which have no side in common are called opposite vertices.

Let us consider, for example, Fig. 4 (p. 60). g_1, g_2, h_1, h_2 are the sides of a complete quadrilateral and A, B, C, D, S_1, S_2 are its vertices. A, B and C, D and S_1, S_2 are then the three pairs of opposite vertices.

The lines joining opposite vertices are called diagonals. Thus, g, g_3, and the line joining C and D are the diagonals of the quadrilateral in Fig. 4. Of the four points S_1, S_2, S_3, S_4 on the diagonal g, the first pair are the opposite vertices that lie on g, the second pair are the points of intersection of g with the other two diagonals.

By use of these concepts, the fact that $\mathcal{R}(S_1\, S_2\, S_3\, S_4) = -1$ can now be expressed as follows:

THEOREM OF THE COMPLETE QUADRILATERAL: *On every*[7] *diagonal of a complete quadrilateral, the pair of opposite vertices on that diagonal and the pair of points of intersection with the other two diagonals separate each other harmonically.*

The theorem of the complete quadrilateral, as well as the foregoing construction, can be dualized in the projective plane (i.e., in P_2). As regards the construction, we leave that to the reader, and we confine ourselves to mentioning that the concept dual to that of the complete quadrilateral is called the '*complete quadrangle.*' Its definition is immediate:

Four points, every three of which are linearly independent, together with the six lines joining them, constitute a complete quadrangle.

The points are again called vertices and the lines are again called sides of the quadrangle. Two sides which have no vertex in common are called opposite sides, and the points of intersection of two opposite sides are called diagonal points.

The *Theorem of the Complete Quadrangle* then reads as follows: *At each diagonal point of a complete quadrangle, the opposite sides (which pass through it) and the lines to the two other diagonal points separate each other harmonically.*

If, for example, we regard the points S_1, S_2, C, D in Fig. 4 as vertices of a complete quadrangle, then A is a diagonal point and g_1, g_2 har-

[7] Not only on g but, by reason of symmetry, on *every* one of the three diagonals.

monically separate g_3 and the line from A to S_4. The same result could have been obtained from Theorem 1 (cf. Fig. 3).

But it is not only the harmonic cross ratio that has geometrical properties; the general cross ratio has geometrical properties also, a few of which we shall now derive. If in the relation

$$(9) \qquad \mathcal{R}(S_1\, S_2\, S_3\, T) = a$$

the S_1, S_2, S_3 are thought of as *three distinct fixed* points and T as a variable point on a line g, then to each $T \neq S_1$ there belongs a uniquely determined a, and to each a there belongs a uniquely determined T; that is, (9) sets up a one-to-one correspondence between the numbers a of the (real or complex) ground field and all the points $T \neq S_1$ of g.

Let a, b be two numbers of the ground field and let T_a, T_b be the points of g that correspond to them in accordance with (9). Let T_{a+b} be the point corresponding to the sum $a + b$ and let $T_{a \cdot b}$ be the point corresponding to the product $a \cdot b$. Surprisingly enough, it is possible to obtain T_{a+b} and $T_{a \cdot b}$ from T_a and T_b by means of geometrical constructions—and, what is more, by linear constructions alone.

In order to see this, we first resort to heuristic considerations in P_2 (i.e., in the projective plane). Consider an ordinary rectangular coordi-

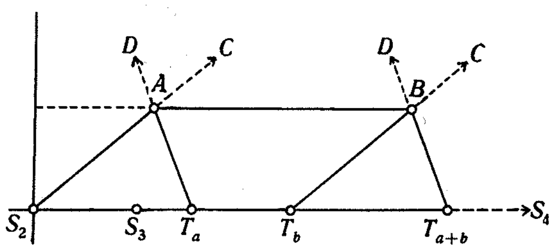

Fig. 5

nate system to be given in the plane so that (cf. Chap. I) every *proper* point $[\xi_0, \xi_1, \xi_2]$ of P_2 is represented by a point with the rectangular (nonhomogeneous) coordinates $\xi_1/\xi_0, \xi_2/\xi_0$.

Let us take the x_1-axis as our line g and let us choose for the points S_1, S_2, and S_3 the improper point of the x_1-axis, the origin, and the unit point of the x_1-axis, respectively, i.e., let

$$S_1 = [0, 1, 0], \qquad S_2 = [1, 0, 0] = (0, 0), \qquad S_3 = [1, 1, 0] = (1, 0)$$

(cf. Fig. 5). Now if any proper point on the x_1-axis be given, say T_a, with abscissa a, so that $T_a = [1, a, 0]$, then we have at once the relation

$$\{1, a, 0\} = a\{0, 1, 0\} + \{1, 0, 0\}$$

among the coordinate vectors of T_a, S_1, and S_2. This means that $\mathcal{R}(S_1 S_2 S_3 T_a) = a$. Thus, the cross ratio has here a simple geometrical meaning: $\mathcal{R}(S_1 S_2 S_3 T_a)$ is equal to the abscissa of T_a.

In this special case, we can now immediately give a construction that will yield T_{a+b} with abscissa $a + b$, i.e., $\mathcal{R}(S_1 S_2 S_3 T_{a+b}) = a + b$, when T_a, T_b are given. This construction can be read off directly from Fig. 5. Point A in Fig. 5 is an auxiliary point and can be any arbitrary *proper* point of the plane. S_2A is parallel to T_bB, and $\overline{S_2A} = \overline{T_bB}$.[8] Finally, AT_a is parallel to BT_{a+b}. The correctness of our construction follows from the fact that $\overline{S_2 T_a} = \overline{T_b T_{a+b}}$.

In this construction let us think of the improper line as an auxiliary line h passing through S_1. Let the intersection of S_2A and h (that is, the improper point on the line S_2A) be denoted by C, and the intersection of T_aA and h, by D. Then we know that T_bB also goes through C and $T_{a+b}B$ through D. Furthermore, AB, being parallel to the x_1-axis, passes through S_1. Accordingly, we can look upon the construction in Fig. 5 as the realization of the following operations:

First, we pass a line through S_2 and A, and determine its intersection C with h. Then we draw the two lines through T_b and C and through A and S_1 and determine their point of intersection B. Finally, we connect T_a with A and form the intersection of this line with h. Let D be the point of intersection. If we then pass a line through B and D, its point of intersection with g gives us the desired point T_{a+b}.

In this form, in which the word 'parallel' no longer appears, the construction is of general validity. Indeed, let us return to P_n. There too our starting point was a line g and three fixed points S_1, S_2, S_3 on g. Now let us draw through S_1 any auxiliary line h distinct from g, and let us choose some fixed auxiliary point A, not on g or on h, but in the plane determined by them.[9] If we now carry out in P_n (in the plane determined by g and h) the above operations, word for word as described in the preceding paragraph (cf. Fig. 6), we assert that the following is true for the point T_{a+b} thus constructed:

---

[8] As in *Modern Algebra*, $\overline{S_2A}$ denotes the length of the segment S_2A.

[9] The spanning space of two lines that intersect in a point is of dimension 2, by Theorem 5, Chap. II, and accordingly is a plane.

$$\mathcal{R}(S_1\, S_2\, S_3\, T_{a+b}) = \mathcal{R}(S_1\, S_2\, S_3\, T_a) + \mathcal{R}(S_1\, S_2\, S_3\, T_b).$$

Proof: Let the coordinate vectors of S_1, S_2, S_3 be $\mathfrak{x}, \mathfrak{y}, \mathfrak{x}+\mathfrak{y}$, respectively, and the coordinate vectors of A and C, $\mathfrak{z}$ and $\mathfrak{y}+\mathfrak{z}$, respectively.[1] By assumption, we have $\mathcal{R}(S_1\, S_2\, S_3\, T_a) = a$ and $\mathcal{R}(S_1\, S_2\, S_3\, T_b) = b$, where a, b are two numbers of the ground field. Hence (by definition of the cross ratio), $a\mathfrak{x}+\mathfrak{y}$ is a coordinate vector of T_a and $b\mathfrak{x}+\mathfrak{y}$, a coordinate vector of T_b.

The point B is uniquely[2] determined as the point of intersection of CT_b with AS_1. Hence, any vector that is a linear combination of $\mathfrak{x}$ and $\mathfrak{z}$

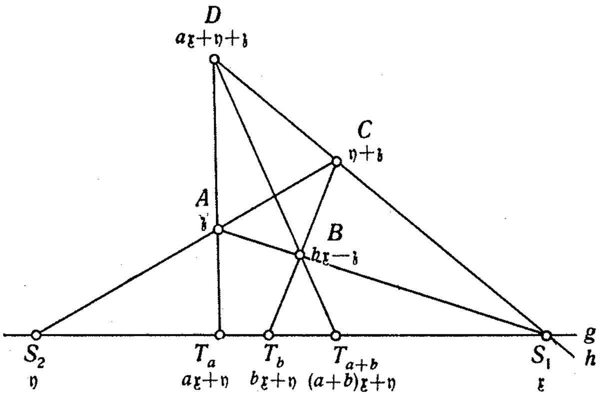

Fig. 6

and at one and the same time a linear combination of $b\mathfrak{x}+\mathfrak{y}$ and $\mathfrak{y}+\mathfrak{z}$ must be a coordinate vector of B. Such a linear combination is $b\mathfrak{x}-\mathfrak{z}$.

Similarly, $a\mathfrak{x}+\mathfrak{y}+\mathfrak{z}$, being a linear combination of $\mathfrak{x}$ and $\mathfrak{y}+\mathfrak{z}$ on the one hand and of $\mathfrak{z}$ and $a\mathfrak{x}+\mathfrak{y}$ on the other, is a coordinate vector of D, the uniquely determined point of intersection of AT_a with CS_1 (CS_1 is the auxiliary line h).

Finally, T_{a+b}, the uniquely determined point of intersection of DB and S_1S_2, has the coordinate vector $(a+b)\mathfrak{x}+\mathfrak{y}$, since this vector is both a linear combination of $\mathfrak{x}$ and $\mathfrak{y}$ and of $a\mathfrak{x}+\mathfrak{y}+\mathfrak{z}$ and $b\mathfrak{x}-\mathfrak{z}$.

---

[1] The possibility of such a choice is clear. For if we first set the coordinate vector of A equal to $\mathfrak{z}$, then the coordinate vector of C can certainly be represented in the form $\lambda\mathfrak{y}+\mu\bar{\mathfrak{z}}$. Both λ and μ are different from 0, inasmuch as S_2, A, and C are all distinct. Thus $\mathfrak{y}+\frac{\mu}{\lambda}\bar{\mathfrak{z}}$ is a coordinate vector of C. We then merely need set $\frac{\mu}{\lambda}\bar{\mathfrak{z}} = \mathfrak{z}$ to obtain the desired normalization.

[2] Only *one* point of intersection exists, since the lines through CT_b and AS_1 are distinct from each other.

It follows that

$$\mathcal{R}(S_1\, S_2\, S_3\, T_{a+b}) = a+b,$$

as was to be shown.

The question of constructing $T_{a \cdot b}$ may be disposed of in an entirely similar way. We first proceed heuristically, as we did above.

Again, let g be the x_1-axis of the projective plane; and let $S_1 = [0, 1, 0]$, the improper point of g; $S_2 = [1, 0, 0] = (0, 0)$; and

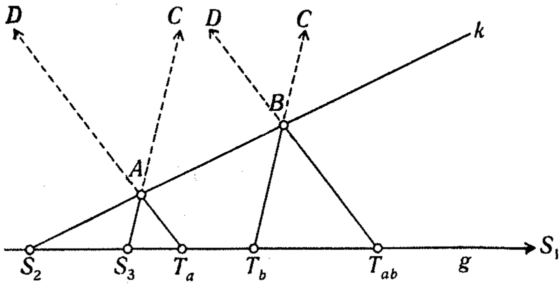

Fig. 7

$S_3 = [1, 1, 0] = (1, 0)$. Furthermore, let $T_a = [1, a, 0] = (a, 0)$ and $T_b = [1, b, 0] = (b, 0)$. From this it follows that $\mathcal{R}(S_1\, S_2\, S_3\, T_a) = a$ and $\mathcal{R}(S_1\, S_2\, S_3\, T_b) = b$.

Fig. 7 then shows how the point $T_{a \cdot b}$ with abscissa $a \cdot b$ is to be found. In Fig. 7, k is an auxiliary line through S_2, distinct from g but otherwise arbitrary. A is any auxiliary point on k (but distinct from S_2, and proper). Also, AS_3 is parallel to BT_b and AT_a parallel to $BT_{a \cdot b}$. The validity of the construction follows from the equation

$$\begin{aligned} \overline{S_2\, A} : \overline{S_2\, B} &= \overline{S_2\, S_3} : \overline{S_2\, T_b} = \overline{S_2\, T_a} : \overline{S_2\, T_{ab}} \\ &= 1 : b = a : \overline{S_2\, T_{ab}}, \end{aligned}$$

that is,

$$\overline{S_2\, T_{ab}} = a\, b.$$

Now let us again use the line at infinity h as an auxiliary line through S_1, denoting by C the intersection of h with S_3A and by D the intersection of h with T_aA. Then the intersection of T_bB with h is also equal to C and that of $T_{ab}\, B$ with h, to D.

We can therefore interpret the construction of Figure 7 as the carrying out of the following operations.

First we pass a line through S_3 and the auxiliary point A and determine its intersection C with h. Then we pass a line through T_a and A and find its point of intersection D with h. Finally, we connect C with T_b and determine the point of intersection B of CT_b and k. If we then pass a line through B and D, its point of intersection with g gives us the desired point T_{ab}.

In this way we have again obtained a form of the construction which has general validity. Returning to our line g in P_n with the three fixed points S_1, S_2, S_3 and given T_a, T_b $(\neq S_1)$, with $\mathcal{R}(S_1\, S_2\, S_3\, T_a) = a$ and $\mathcal{R}(S_1\, S_2\, S_3\, T_b) = b$, we must proceed as follows. We first draw an auxiliary line $h \neq g$ through S_1; second, another auxiliary line $k \neq g$ through S_2; third, we choose on k an auxiliary point A distinct from S_2 and not lying on h. Then we carry out, word for word, the operations described in the preceding paragraph. For the point T_{ab} thus obtained, the following holds:

$$\mathcal{R}(S_1\, S_2\, S_3\, T_{ab}) = \mathcal{R}(S_1\, S_2\, S_3\, T_a) \cdot \mathcal{R}(S_1\, S_2\, S_3\, T_b).$$

The proof is carried out in exactly the same way as in the construction of T_{a+b} and can be read off from Figure 8 if one determines the indicated

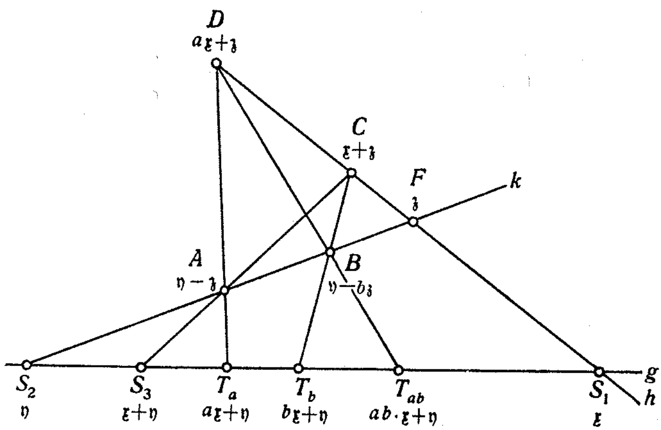

Fig. 8

coordinate vectors of the points in the following order: S_1, S_2, S_3, T_a, T_b, F (F is the intersection of h with k), C, A, D, B, T_{ab}.

Exercises

1. Let four distinct points Q_1, Q_2, Q_3, Q_4 be given on a line of real projective P_n and, moreover, let the Q_i all be proper points, so that they also represent points in affine R_n. Let the idea of euclidean length be introduced into affine R_n and denote the distance between Q_i and Q_k by $\overline{Q_i Q_k}$. Show that

$$|Dv(Q_1 Q_2 Q_3 Q_4)| = \frac{\overline{Q_3 Q_1} \cdot \overline{Q_4 Q_2}}{\overline{Q_3 Q_2} \cdot \overline{Q_4 Q_1}}.$$

2. In the construction of Figure 6 on page 65 interchange the roles of T_a and T_b (leaving the auxiliary elements A and h fixed). Since $a + b = b + a$, the outcome of the construction as thus altered must of course be the same final T_{a+b}. The fact that the two configurations of lines of the two constructions end in the same point gives a geometrical theorem. Convince yourself that this is precisely the theorem of Exercise 2 of Chapter III.

If, similarly, the construction in Figure 8 is repeated with the roles of T_a and T_b interchanged, the auxiliary elements A, h, k being held fixed, then the commutative law of multiplication gives the so-called Theorem of Pascal for a pair of lines: *If the vertices of a hexagon lie alternately on two lines, then its opposite sides intersect in three points of a line.* Carry out the two constructions and find Pascal's Hexagon.

3a) In Figure 8 on page 67, the fact that the construction is independent of the choice of the auxiliary elements h, k, A is equivalent to the following theorem:

If the vertices of two complete quadrangles A, B, C, D and A', B', C', D' are considered as corresponding to each other in that order, A with A', B with B', C with C', D with D', then the following is true for the six points of intersection of corresponding sides (i.e., for the intersection of AB with $A'B'$, of AC with $A'C'$, etc.): If five of them lie on a line g, then so does the sixth.

Remark: This theorem can also be proved easily as a consequence of the Theorem of Desargues.

b) In Figure 6 on page 65, the fact that the construction is independent of the choice of the auxiliary elements h, A is equivalent to a special case of the theorem just mentioned. (Two of the points of intersection coincide in this case, namely, at S_1.)

CHAPTER V

PROJECTIVITIES

In this chapter we shall study certain important mappings between linear spaces. Let L and L^* be two linear spaces of P_n and let there be defined a one-to-one mapping between them—that is, let there exist a correspondence between the points of L and those of L^* such that exactly one point of L^* corresponds to each point of L, and conversely.

Besides this, we impose the following condition on our mapping: Linearly dependent points shall always be carried into linearly dependent points and linearly independent points always into linearly independent points. This means that if $Q_1, Q_2, \ldots, Q_k$ is any finite number of points of L and if $Q_1^*, Q_2^*, \ldots, Q_k^*$ are their images in L^*, then the Q_i and the Q_i^* are always either linearly dependent or linearly independent together.

It follows immediately that a mapping defined in this way can exist between two linear spaces L and L^* only if L and L^* are of the same dimension.

As proof, let s be the dimension of L and s^* the dimension of L^*. Then we may assume, say, that $s \geqq s^*$, and we need to show that $s > s^*$ is impossible. If we now choose $s+1$ linearly independent points in L, say $Q_1, Q_2, \ldots, Q_{s+1}$, then their images $Q_1^*, Q_2^*, \ldots, Q_{s+1}^*$ in L^* must also be linearly independent, i.e., s^* must be *at least* equal to s, as was to be shown.

We can assume from now on that $s = s^*$ for two linear spaces L and L^*, between which there exists a mapping with the required properties. If $s > 1$, we shall see that our requirement singles out from the set of all one-to-one mappings a class of special mappings that can be described exactly. These mappings we call **projectivities** (also **collineations** or *projective relations*) between L and L^*.

In case $s = 1$, however, our condition tells us nothing that is not already implicit in the mapping being one to one. For, the maximal number of linearly independent points in L, or in L^*, is 2. But for two points linear independence or dependence merely means being distinct or being coincident.

In order to obtain in the case $s = 1$ mappings similar to those obtainable in the case $s > 1$, a further condition must be imposed. This will be done later, after we first have considered the case $s > 1$.

Projective Relations between Two Linear Spaces with Dimensions greater than 1

From now on, until page 79, let us assume that $s = s^* > 1$, where s and s^* are the dimensions of L and L^*. Moreover, let us consider as given[1] a fixed projectivity between L and L^*.

As a first property of such a projectivity we state the following:

THEOREM 1. *Let L_1 be a linear subspace of L. For each point in L_1 find the image point in L^*. The totality of these image points again fills out a linear subspace $L_1{}^*$ of L^* with the same dimension as L_1.*

Proof: Let r be the dimension of L_1. Then if the $r + 1$ points $Q_0, Q_1, \ldots, Q_r$ of L_1 are linearly independent, so also are their image points $Q_0^*, Q_1^*, \ldots, Q_r^*$ in $L_1{}^*$. Let Q be some other point in L_1 and Q^* its image point. Then according to Theorem 3 of Chapter II, Q belongs to L_1 if and only if Q is a linear combination of the Q_i, i.e. (by *Modern Algebra,* § 3, Theorem 4), if and only if the points $Q_0, Q_1, \ldots, Q_r, Q$ are linearly dependent. It therefore also holds that Q^* belongs to $L_1{}^*$ if and only if $Q_0^*, Q_1^*, \ldots, Q_r^*, Q^*$ are linearly dependent, i.e. (by *Modern Algebra,* § 3, Theorem 5), if Q^* is a linear combination of the $Q_i{}^*$. Therefore $L_1{}^*$ consists precisely of all linear combinations of the $Q_i{}^*$ and is thus a linear space, and indeed, the linear space of least dimension containing all the $Q_i{}^*$. As such, it has dimension r. Thus, Theorem 1 is proved.

We call $L_1{}^*$ the *image space* or simply the *image* of L_1.

Let L_2 be another subspace of L, and $L_2{}^*$ its image (in L^*). If the intersection of L_1 and L_2 is empty, then since the mapping is one-to-one, the intersection of $L_1{}^*$ and $L_2{}^*$ must also be empty. In the opposite case, it follows, likewise from the one-to-one nature of the mapping, that there is a point in the intersection of $L_1{}^*$ and $L_2{}^*$ corresponding to each point in the intersection of L_1 and L_2, and conversely. From Theorem 1 we then immediately have Theorem 2.

---

[1] In order to eliminate any possible doubt as to the existence of such a correspondence, it may be pointed out that for $L = L^*$ the identity correspondence (i.e., the correspondence that associates each point with itself) is a projectivity. We shall shortly become acquainted with more general examples.

THEOREM 2. *The intersection of L_1^* and L_2^* has the same dimension as the intersection of L_1 and L_2.*[2]

Now let us consider how, if at all, the cross ratio is altered by our projectivity. Let g be a line in L and g^* its image line. Also, let $(S_1, S_2 \mid S_3)$ be a coordinate system on g and $(S_1^*, S_2^* \mid S_3^*)$ a coordinate system on g^*, formed with the image points of the S_i. Then, T and T^* being two further points corresponding to each other on g and g^*, respectively, we should like to compare the two values $a = \mathcal{R}(S_1 S_2 S_3 T)$ and $a^* = \mathcal{R}(S_1^* S_2^* S_3^* T^*)$.

Since $a \rightleftarrows T$ and $a^* \rightleftarrows T^*$ are one-to-one correspondences between, on the one hand, the elements of the ground field F, and, on the other hand, the points of g other than s_1 and the points of g^* other than S_1^*, respectively, and since in addition $T \rightleftarrows T^*$ is a one-to-one correspondence between the points of g other than S_1 and the points of g^* other than S_1^*, it follows that $a \rightleftarrows a^*$ is a one-to-one mapping of F onto itself.

We shall write $a^* = \sigma(a)$, $a = \sigma^{-1}(a^*)$. The function σ thus defined has two very remarkable properties, namely: For any two elements a, b of the field,

$$\sigma(a+b) = \sigma(a) + \sigma(b), \tag{1}$$
$$\sigma(a \cdot b) = \sigma(a) \cdot \sigma(b). \tag{2}$$

Proof: Let T_a, T_b, T_{a+b} be three points of g (all distinct from S_1) for which

$$\mathcal{R}(S_1 S_2 S_3 T_a) = a, \quad \mathcal{R}(S_1 S_2 S_3 T_b) = b, \quad \mathcal{R}(S_1 S_2 S_3 T_{a+b}) = a+b.$$

We know that the point T_{a+b} can then be found by the linear construction given in Fig. 6, on p. 65. Let us now think of such a construction as being carried out within L, i.e., all the points and lines of Fig. 6 so chosen that they lie in L. This is possible because the construction of Fig. 6 can be carried out in an arbitrary plane containing g and thus, in particular, in a plane lying wholly in L (and containing g). Such a plane exists, for the dimension of L is by assumption $\geqq 2$. We then apply the given projectivity to all the points and lines of the construction. According to Theorem 1, the image of this figure as a whole will itself lie in a plane. Furthermore, the images of two lines that intersect in a point will again have exactly one point in common, namely the image point of the point of intersection (Theorem 2). From this we see that Fig. 6 maps under

[2] Note that if the intersection is of dimension 0, Theorem 2 follows solely from the fact that the correspondence is one-to-one.

a projectivity into an exactly analogous figure; indeed, we obtain the image figure by simply putting an asterisk on each designation in Fig. 6. But that means that the image point T^*_{a+b} of T_{a+b} can be found from $T_a{}^*$ and $T_b{}^*$ by exactly the same construction as T_{a+b} from T_a and T_b. From the meaning of this construction it then necessarily follows that

$$(3) \qquad \mathcal{R}(S_1^* \, S_2^* \, S_3^* \, T^*_{a+b}) = \mathcal{R}(S_1^* \, S_2^* \, S_3^* \, T_a^*) + \mathcal{R}(S_1^* \, S_2^* \, S_3^* \, T_b^*)$$

But by the definition of σ,

$$\mathcal{R}(S_1^* \, S_2^* \, S_3^* \, T^*_{a+b}) = \sigma(a+b),$$

$$\mathcal{R}(S_1^* \, S_2^* \, S_3^* \, T_a^*) = \sigma(a), \qquad \mathcal{R}(S_1^* \, S_2^* \, S_3^* \, T_b^*) = \sigma(b),$$

and a comparison of this with (3) immediately yields relation (1), the relation to be proved.

Relation (2) follows in an exactly similar way from the fact that the point $T_{a \cdot b}$ can likewise be found from T_a and T_b by a linear construction (namely, the construction in Fig. 8 on p. 67).

What we have just proved for σ we can, of course, also show for the inverse mapping σ^{-1}, since the assumptions for σ^{-1} are exactly the same as for σ. Thus, the following also hold:

$$(4) \qquad \sigma^{-1}(a+b) = \sigma^{-1}(a) + \sigma^{-1}(b),$$

$$(5) \qquad \sigma^{-1}(a \cdot b) = \sigma^{-1}(a) \cdot \sigma^{-1}(b).$$

It should be noted, however, that equations (4) and (5) follow just as well from (1) and (2) and the fact that σ is a one-to-one mapping. For if $\sigma^{-1}(a)$ and $\sigma^{-1}(b)$ are substituted for a and b respectively in (1) and (2), it follows that

$$\sigma(\sigma^{-1}(a) + \sigma^{-1}(b)) = a + b, \quad \sigma(\sigma^{-1}(a) \cdot \sigma^{-1}(b)) = a \cdot b.$$

But these are equivalent to (4) and (5).

Any one-to-one mapping $a^* = \sigma(a)$ of a field onto itself which satisfies relations (1) and (2) is called an **automorphism** (of the field).

From (1) and (2) we can derive some results that will presently be of use. To begin with, we have

$$(6) \qquad \sigma(0) = 0, \qquad \sigma(1) = 1.$$

For, we note that, for any elements a of the field, $a + 0 = 0$. It follows from this that $\sigma(a) + \sigma(0) = \sigma(a)$ and hence that $\sigma(0) = 0$. Since the correspondence σ is one to one, we must have $\sigma(a) \neq 0$ for $a \neq 0$. Now, since the equation $\sigma(a) = \sigma(a) \cdot \sigma(1)$ holds for every a, it follows that $\sigma(1) = 1$.

If the element a in (1) is replaced by $a - b$, we now obtain

$$\sigma(a - b) = \sigma(a) - \sigma(b) \tag{7}$$

and by substituting a/b $(b \neq 0)$ for a in (2), we obtain

$$\sigma\left(\frac{a}{b}\right) = \frac{\sigma(a)}{\sigma(b)}. \tag{8}$$

In defining σ we have to choose the three fixed points S_1, S_2, S_3. Therefore, it would appear as though σ depended on the choice of the S_i. This is not actually the case, however. To show this, we shall now give σ a meaning independent of the S_i by means of the following theorem.

THEOREM 3. *If T_1, T_2, T_3, T_4 are any four points of the line g for which $\mathcal{R}(T_1 T_2 T_3 T_4)$ is defined and $\neq \infty$, then the following is always true of the four image points T_i^*:*

$$\mathcal{R}(T_1^* T_2^* T_3^* T_4^*) = \sigma[\mathcal{R}(T_1 T_2 T_3 T_4)].$$

Proof: We may assume that all the T_i are distinct. For if two of the T_i coincide, then $\mathcal{R}(T_1 T_2 T_3 T_4)$ must be either 0 or 1, since the value ∞ has been excluded.[3] Then $\mathcal{R}(T_1^* T_2^* T_3^* T_4^*)$ must have the same value (i.e., either 0 or 1),[4] so that in these special cases the equality stated in the theorem is true (by (6)).

Now consider first the case in which all the T_i are different from S_1. Then $\varkappa_i = \mathcal{R}(S_1 S_2 S_3 T_i)$ is a number of the field for each $i = 1, 2, 3, 4$, and according to formula (3) of Chapter IV, we have

$$\mathcal{R}(T_1 T_2 T_3 T_4) = \frac{\varkappa_3 - \varkappa_1}{\varkappa_3 - \varkappa_2} : \frac{\varkappa_4 - \varkappa_1}{\varkappa_4 - \varkappa_2}. \tag{9}$$

Moreover, by the definition of σ,

$$\mathcal{R}(S_1^* S_2^* S_3^* T_i^*) = \sigma(\varkappa_i),$$

[3] The cross ratio was defined only for the case in which at least three points were distinct; indeed, $\mathcal{R}(T_1T_2T_3T_4)$ was $= 0$ if $T_1 = T_3$ or $T_2 = T_4$; it was $= 1$ if $T_1 = T_2$ or $T_3 = T_4$; and it was $= \infty$ if $T_2 = T_3$ or $T_1 = T_4$.

[4] For it follows from $T_i = T_k$ that $T_i^* = T_k^*$.

and hence

$$(10) \qquad \mathcal{R}(T_1^* \, T_2^* \, T_3^* \, T_4^*) = \frac{\sigma(\varkappa_3) - \sigma(\varkappa_1)}{\sigma(\varkappa_3) - \sigma(\varkappa_2)} : \frac{\sigma(\varkappa_4) - \sigma(\varkappa_1)}{\sigma(\varkappa_4) - \sigma(\varkappa_2)}.$$

Repeated application of (7) and (8) now yields immediately that

$$(11) \quad \mathcal{R}(T_1^* T_2^* T_3^* T_4^*) = \sigma\left[\frac{\varkappa_3 - \varkappa_1}{\varkappa_3 - \varkappa_2} : \frac{\varkappa_4 - \varkappa_1}{\varkappa_4 - \varkappa_2}\right] = \sigma[\mathcal{R}(T_1 T_2 T_3 T_4)],$$

which proves our theorem for the case in which $T_i \neq S_1$ $(i = 1, 2, 3, 4)$.

Now if any one of the T_i coincides with S_1 we may always take $S_1 = T_1$, for since

$$\begin{aligned} \mathcal{R}(T_1 \; T_2 \; T_3 \; T_4) &= \mathcal{R}(T_2 \; T_1 \; T_4 \; T_3) \\ = \mathcal{R}(T_3 \; T_4 \; T_1 \; T_2) &= \mathcal{R}(T_4 \; T_3 \; T_2 \; T_1) \end{aligned}$$

(cf. Chap. IV, formula (4)), we see that any of the four points can be made to occupy the first place. Then it can be proved, the proof being analogous to that of equation (3) of Chap. IV, that[5]

$$(12) \qquad \mathcal{R}(T_1 \; T_2 \; T_3 \; T_4) = \frac{\varkappa_4 - \varkappa_2}{\varkappa_3 - \varkappa_2}.$$

From (12) and the analogous formula for the T_i^* it now follows, just as above from (9) and (10), that our theorem is true for the case $S_1 = T_1$. Thus, Theorem 3 is completely proved.

Theorem 3 shows that to every line g in L there belongs a definite automorphism of the ground field which tells us how the given projectivity affects the cross ratio of four points on g. We now wish to show, further, that (for the fixed projectivity) the *same* automorphism belongs to *all* the lines of L.

---

[5] If $\mathfrak{x}$, $\mathfrak{y}$, $\mathfrak{x} + \mathfrak{y}$ are the coordinate vectors of S_1, S_2, S_3, then those of T_1, T_2, T_3, T_4 are $\mathfrak{x}$, $\varkappa_2 \mathfrak{x} + \mathfrak{y}$, $\varkappa_3 \mathfrak{x} + \mathfrak{y}$, $\varkappa_4 \mathfrak{x} + \mathfrak{y}$, respectively. By formula (2) of Chap. IV, (12) then follows at once from the identities

$$\varkappa_3 \mathfrak{x} + \mathfrak{y} = (\varkappa_3 - \varkappa_2)\mathfrak{x} + (\varkappa_2 \mathfrak{x} + \mathfrak{y})$$

and

$$\varkappa_4 \mathfrak{x} + \mathfrak{y} = (\varkappa_4 - \varkappa_2)\mathfrak{x} + (\varkappa_2 \mathfrak{x} + \mathfrak{y}).$$

It will suffice, for this purpose, to prove that two lines which intersect in a point always have the same automorphism. For, the general case can always be reduced to this special case by comparing the automorphism of two lines g, $\bar{g}$ that do *not* intersect with the automorphism of a third line $\bar{\bar{g}}$ which intersects both g and $\bar{g}$.

Let g and $\bar{g}$, then, be two intersecting lines. To show that the automorphisms associated with g and $\bar{g}$ under the projectivity are the same, we must, according to the meaning of these automorphisms, show the following:

If S_1, S_2, S_3, S_4 and $\bar{S}_1$, $\bar{S}_2$, $\bar{S}_3$, $\bar{S}_4$ are any four points on g and $\bar{g}$, respectively, such that $\mathcal{R}(S_1\, S_2\, S_3\, S_4) = \mathcal{R}(\bar{S}_1\, \bar{S}_2\, \bar{S}_3\, \bar{S}_4)$ and if S_i^*, $\bar{S}_i^*$ $(i = 1, 2, 3, 4)$ are their respective images on the image lines g^* and $\bar{g}^*$, then it also always holds that $\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, S_4^*) = \mathcal{R}(\bar{S}_1^*\, \bar{S}_2^*\, \bar{S}_3^*\, \bar{S}_4^*)$.

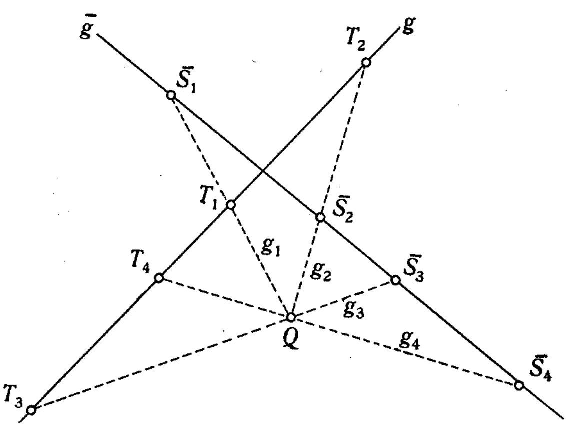

Fig. 9

The space spanned by g and $\bar{g}$ is a two-dimensional plane e in L. Let Q be a point of e that lies neither on g nor on $\bar{g}$. Denote the lines joining Q with the $\bar{S}_i$ by g_i (Fig. 9), and let T_i be the point of intersection of g_i with g. Then by Chap. IV, Theorem 1,

$$\mathcal{R}(\bar{S}_1\, \bar{S}_2\, \bar{S}_3\, \bar{S}_4) = \mathcal{R}(T_1\, T_2\, T_3\, T_4)$$

and thus also

$$\mathcal{R}(S_1\, S_2\, S_3\, S_4) = \mathcal{R}(T_1\, T_2\, T_3\, T_4).$$

From this, by Theorem 3 of the present chapter,

$$\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, S_4^*) = \mathcal{R}(T_1^*\, T_2^*\, T_3^*\, T_4^*). \tag{13}$$

If we now apply the given projectivity to all the points and lines of Fig. 9, we obtain, according to Theorems 1 and 2, an exactly analogous figure, from which it follows that

$$\mathcal{R}(T_1^*\, T_2^*\, T_3^*\, T_4^*) = \mathcal{R}(\bar{S}_1^*\, \bar{S}_2^*\, \bar{S}_3^*\, \bar{S}_4^*). \tag{14}$$

Comparison of (13) and (14) then yields the desired result.

Thus we have proved the following theorem.

THEOREM 4. *To each projectivity there belongs a uniquely determined automorphism σ with the property that for any four points T_i ($i = 1, 2, 3, 4$) of L which lie on a line and for which $\mathcal{R}(T_1\, T_2\, T_3\, T_4)$ is a number of the field,*

$$\mathcal{R}\,(T_1^*\; T_2^*\; T_3^*\; T_4^*) = \sigma\,[\,\mathcal{R}\,(T_1\; T_2\; T_3\; T_4)].$$

Now let us consider a projective coordinate system $(Q_0, Q_1, \ldots, Q_s \mid E)$ in L. Let Q_i^*, E^* be the image points in L^*. Since every $s + 1$ of the points $Q_0^*, Q_1^*, \ldots, Q_s^*, E^*$ must be linearly independent, it follows that $(Q_0^*, Q_1^*, \ldots, Q_s^* \mid E^*)$ is a well-defined projective coordinate system in L^*.

We now assert that the projectivity is uniquely determined by the automorphism σ, as given in Theorem 4, and by the choice of the $s + 2$ image points Q_i^*, E^* of the Q_i, E; or in other words,

THEOREM 5. *If $(Q_0, Q_1, \ldots, Q_s \mid E)$ and $(Q_0^*, Q_1^*, \ldots, Q_s^* \mid E^*)$ are two fixed projective coordinate systems in L and L^* respectively and σ is a given automorphism of* F, *then there exists one and only one projectivity between L and L^* for which σ has the meaning given in Theorem 4 and which, in addition, carries the Q_i into the Q_i^* and E into E^*.*

Proof: We first prove the *uniqueness* of the projectivity in question. Let us consider, as before, a projectivity that satisfies the requirements of Theorem 5. Then we want to show that under the assumptions of Theorem 5 alone the image point in L^* of every point in L is uniquely determined.

We shall break up this uniqueness proof into three parts. Let us first consider the line g_{ik} determined by a fixed two of the points Q_i, Q_k ($i \neq k$) and the $(s - 1)$-dimensional linear space L_{ik} determined by E and all the Q_ν with $\nu \neq i, k$. L_{ik} and g_{ik} have exactly one point in common;[6] call it E_{ik}. It follows at once that the image point E_{ik}^* is uniquely determined. For the image line g_{ik}^* must pass through Q_i^* and Q_k^* and is therefore uniquely determined. The same is true of the image space L_{ik}^*, since by Theorem 1 it is $(s - 1)$-dimensional and moreover must contain E^* and all the Q_ν^* with $\nu \neq i, k$. Thus E_{ik}^* is also uniquely determined as the point of intersection of g_{ik}^* and L_{ik}^*.

---

[6] This exists, as we have so often seen before, by Theorem 5, Chap. II, because the spanning space is L itself.

We next show the uniqueness of the image point of any point Q_{ik} on the line g_{ik}. By our assumption regarding the meaning of σ for our projectivity, we have

$$\mathcal{R}(Q_i^* \, Q_k^* \, E_{ik}^* \, Q_{ik}^*) = \sigma \ (\mathcal{R}(Q_i \, Q_k \, E_{ik} \, Q_{ik})).$$

This equation already determines the image point Q_{ik}^* uniquely, however.

In the third, and final part of our proof, we prove the uniqueness of the image for a general point Q. Consider all the $(s-1)$-dimensional linear spaces that pass through at least $s-1$ of the fundamental points Q_ν and also through Q.[7] These spaces have only the *one* point Q in common. To see this, let us show that for every point $Q' \neq Q$ there can be found one of these linear spaces that does not contain Q'. The following statement is equivalent to this: For every Q, Q' there can be found $s-1$ fundamental points such that the totality of these $s+1$ points is linearly independent. But this is an immediate consequence of the Steinitz Replacement Theorem[8] applied to the coordinate vectors of our points. The Steinitz Replacement Theorem then states that among the $s+1$ linearly independent fundamental points we can always find two that can be replaced by Q, Q' without disturbing the linear independence.[9]

If we can now show that the images of the spaces under consideration are all uniquely determined, then the same will be true of the image of Q, as the sole point of intersection of these spaces. To show the uniqueness of the image spaces it suffices to consider any one of them, say the linear space $L_{s-1,s}$ which passes through Q and through $Q_0, Q_1, \ldots, Q_{s-2}$. Form the intersection of the line $g_{s-1,s}$ determined by Q_{s-1} and Q_s with this linear space. The point of intersection is uniquely determined; call it $Q_{s-1,s}$. This point, together with $Q_0, Q_1, \ldots, Q_{s-2}$, in turn uniquely determines our linear space $L_{s-1,s}$.[1] But we already know that $Q_{s-1,s}$, as a point of the line $g_{s-1,s}$, has a unique image point $Q_{s-1,s}^*$. The image space $L_{s-1,s}^*$ is consequently uniquely determined by the points $Q_0^*, Q_1^*, \cdots, Q_{s-2}^*, Q_{s-1,s}^*$.

---

[7] In general, the number of such linear spaces is finite $\left[= \binom{s+1}{2}\right]$; in certain cases, however, there may be an infinite number: namely, if there are $s-1$ fundamental points that, together with Q, form a linearly dependent set.

[8] Cf. *Modern Algebra*, p. 20.

[9] The coordinate vectors of Q, Q' are certainly linearly independent, by virtue of the assumption $Q' \neq Q$.

[1] For the points $Q_{s-1,s}, Q_0, Q_1, \cdots, Q_{s-2}$ are linearly independent, since $Q_{s-1,s}$, as a point of the line $g_{s-1\,s}$, cannot belong to the $(s-2)$-dimensional linear space determined by $Q_0, Q_1, \cdots, Q_{s-2}$.

The uniqueness has thus been proved. We now turn to proving the *existence* of the projectivity in question. This we shall do by giving such a mapping explicitly; indeed, we assert that for arbitrary given Q_i, Q_i^*, E, E^*, and σ, a projectivity satisfying the requirements of Theorem 5 is determined by the following condition:

To a point Q in L with coordinates $\xi_0, \xi_1, \ldots, \xi_s$ in the coordinate system $(Q_0, Q_1, \ldots, Q_s \mid E)$ there shall correspond, as image point, the point Q^* in L^* with the coordinates $\sigma(\xi_0)$, $\sigma(\xi_1)$, $\ldots$, $\sigma(\xi_s)$ in the coordinate system $(Q_0^*, Q_1^*, \ldots, Q_s^* \mid E^*)$.

It remains to be shown that the mapping thus defined is truly a projectivity. First, that the mapping is *one-to-one*, is trivial. For, the coordinates ξ_i of the point Q in the one coordinate system and the coordinates ξ_i^* of the image point Q^* in the other coordinate system determine each other uniquely in accordance with the relations[2] $\xi_i^* = \sigma(\xi_i), \xi_i = \sigma^{-1}(\xi_i^*)$.

Further, consider any k points $S_1, S_2, \ldots, S_k$ in L. Let $S_1^*, S_2^*, \ldots, S_k^*$ be their image points in L^*. Also, let $\xi_{i0}, \xi_{i1}, \ldots, \xi_{is}$ be the coordinates of the point S_i in the coordinate system $(Q_0, Q_1, \ldots, Q_s \mid E)$ and let $\sigma(\xi_{i0})$, $\sigma(\xi_{i1})$, $\ldots$, $\sigma(\xi_{is})$ be the coordinates of S_i^* in $(Q_0^*, Q_1^*, \ldots, Q_s^* \mid E^*)$. If the S_i are now linearly dependent, there exists k constants $\lambda_1, \lambda_2, \ldots, \lambda_k$, not all zero, for which

$$\sum_{i=1}^{k} \lambda_i \xi_{i\nu} = 0$$

is true for all $\nu = 0, 1, \ldots, s$. By applying σ to both sides of this equation, we find that

$$\sum_{i=1}^{k} \sigma(\lambda_i) \cdot \sigma(\xi_{i\nu}) = 0$$

holds for all ν. But this means that the S_i^* must likewise be linearly dependent, for certainly not all the $\sigma(\lambda_i)$ are equal to zero.

Since we can show, conversely, by use of the inverse automorphism σ^{-1}, that the linear dependence of the S_i^* implies the linear dependence of the S_i, it follows that the S_i and the S_i^* are either linearly dependent or linearly independent together. Our mapping is thus seen to be a projectivity, and Theorem 5 is proved in full.

---

[2] If the ξ_i be multiplied by a constant λ, the $\sigma(\xi_i)$ are multiplied by $\sigma(\lambda)$, which gives the same point. A similar argument holds in the other direction.

For future reference, we summarize the results of the last part of the proof in

THEOREM 6. *The uniquely determined projectivity of Theorem 5 can be characterized as follows: If a point Q has the coordinates $\xi_0, \xi_1, \ldots, \xi_s$ in the coordinate system $(Q_0, Q_1, \ldots, Q_s \mid E)$, then its image point Q^* has the coordinates $\sigma(\xi_0)$, $\sigma(\xi_1)$, $\cdots$, $\sigma(\xi_s)$ in the coordinate system $(Q_0^*, Q_1^*, \ldots, Q_s^* \mid E^*)$.*

Projective Relations between two Lines

Let us now examine the case in which the linear space L and its image L^* both have dimension 1, that is, in which both are lines in P_n. As already pointed out earlier in this chapter (p. 69), to require that linear dependence, or linear independence, be invariant is of no use to us in the present case. However, we can require instead that the harmonic set[3] be invariant. Hence we make the following definition:

A projectivity between the one-dimensional spaces L and L^* is a one-to-one mapping of L on L^* that always takes harmonic sets into harmonic sets.

This means that if S_1, S_2, S_3, S_4 are four points of L and S_1^*, S_2^*, S_3^* S_4^* are their image points on L^*, then $\mathcal{R}(S_1\, S_2\, S_3\, S_4) = -1$ always implies that $\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, S_4^*) = -1$. It is not necessary to require the converse, that $\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, S_4^*) = -1$ shall imply $\mathcal{R}(S_1\, S_2\, S_3\, S_4) = -1$, for this follows as a consequence of the first requirement. For if $\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, S_4^*) = -1$, there is certainly a uniquely determined point Q for which $\mathcal{R}(S_1\, S_2\, S_3\, Q) = -1$ (S_i being the original points of the S_i^*). But from this it follows that $\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, Q^*) = -1$ as well, so that $Q^* = S_4^*$; i.e., $Q = S_4$ and, accordingly, $\mathcal{R}(S_1\, S_2\, S_3\, S_4) = -1$.

Let S_1, S_2, S_3 be three different fixed points on the line L and let S_1^*, S_2^*, S_3^* be their image points on L^*. If T is then any further point $\neq S_1$ on L and T^* is its image, we set up a correspondence between the values $a = \mathcal{R}(S_1\, S_2\, S_3\, T)$ and $a^* = \mathcal{R}(S_1^*\, S_2^*\, S_3^*\, T^*)$.

The one-to-one mapping of the ground field onto itself which is thus defined we again denote, as in the preceding section, by $a^* = \sigma(a)$. We wish to prove *that σ is an automorphism.*

---

[3] Of course, the harmonic set is invariant also under a projectivity between linear spaces with dimensions > 1. This follows directly from Theorem 4, since for every automorphism σ it always holds that $\sigma(-1) = -1$.

To this end, let T_a and T_b be two points $\neq S_1$ on L with

$$\mathcal{R}(S_1\, S_2\, S_3\, T_a) = a, \qquad \mathcal{R}(S_1\, S_2\, S_3\, T_b) = b.$$

Also, let T_a^*, T_b^* be their image points on L^*, whence we have

$$\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, T_a^*) = \sigma(a), \qquad \mathcal{R}(S_1^*\, S_2^*\, S_3^*\, T_b^*) = \sigma(b).$$

As coordinate vectors of the points S_1, S_2, S_3, S_1^*, S_2^*, S_3^* we may take $\mathfrak{x}$, $\mathfrak{y}$, $\mathfrak{x}+\mathfrak{y}$, $\mathfrak{x}^*$, $\mathfrak{y}^*$, $\mathfrak{x}^*+\mathfrak{y}^*$, respectively. It then follows that $a\mathfrak{x}+\mathfrak{y}$, $b\mathfrak{x}+\mathfrak{y}$, $\sigma(a)\mathfrak{x}^*+\mathfrak{y}^*$, $\sigma(b)\mathfrak{x}^*+\mathfrak{y}^*$ are coordinate vectors of T_a, T_b, T_a^*, T_b^*, respectively.

In addition to the points T_a, T_b of L let us now consider the points with coordinate vectors $(a\mathfrak{x}+\mathfrak{y})+(b\mathfrak{x}+\mathfrak{y})$ and $(a\mathfrak{x}+\mathfrak{y})-(b\mathfrak{x}+\mathfrak{y})$. Call the first of these points T. The second vector represents the point S_1—to be sure, only if $a \neq b$, which we shall accordingly assume for the moment. From the form of the coordinate vectors we can easily derive the following relations:

$$\mathcal{R}(S_1\, S_2\, S_3\, T) = \frac{a+b}{2}, \tag{15}$$

$$\mathcal{R}(T_a\, T_b\, S_1\, T) = -1. \tag{16}$$

The second equation is meaningful, because we have assumed that $a \neq b$, i.e., $T_a \neq T_b$.

Similarly, in L^*, let us consider, besides T_a^* and T_b^*, the points with the coordinate vectors $(\sigma(a)\mathfrak{x}^*+\mathfrak{y}^*)+(\sigma(b)\mathfrak{x}^*+\mathfrak{y}^*)$ and $(\sigma(a)\mathfrak{x}^*+\mathfrak{y}^*) - (\sigma(b)\mathfrak{x}^*+\mathfrak{y}^*)$. The second is S_1^*; call the first one U. Then it again follows that

$$\mathcal{R}(S_1^*\, S_2^*\, S_3^*\, U) = \frac{\sigma(a)+\sigma(b)}{2}, \tag{17}$$

$$\mathcal{R}(T_a^*\, T_b^*\, S_1^*\, U) = -1. \tag{18}$$

A comparison of equations (16) and (18) shows at once that $U = T^*$, i.e., that U is the image of T.[4] Taking this into consideration, we see that equations (15) and (17) immediately yield

[4] Owing to our assumption, it follows from (16) that the image point T^*, when substituted for U in (18), will satisfy that equation; on the other hand, there exists only *one* point U that satisfies the equation.

$$\sigma\left(\frac{a+b}{2}\right) = \frac{\sigma(a)+\sigma(b)}{2}. \tag{19}$$

This equation, however, states essentially the same thing as the first of the defining equations of an automorphism. First, observe that equation (19), proved for $a \neq b$, is also trivially true for $a = b$. If we set $b = 0$, it follows that $\sigma(a/2) = \sigma(a)/2$ for arbitrary a, and applying this to the left-hand side of (19) yields

$$\sigma(a+b) = \sigma(a)+\sigma(b). \tag{20}$$

In order to obtain the second of the defining equations of an automorphism, we proceed similarly. In addition to the points T_a, T_b we now consider the points with coordinate vectors $b(a\mathfrak{x}+\mathfrak{y})+a(b\mathfrak{x}+\mathfrak{y})$ and $b(a\mathfrak{x}+\mathfrak{y})-a(b\mathfrak{x}+\mathfrak{y})$. This time, the second point is S_2 (again under the assumption that $a \neq b$), and the first is a point V. It follows that

$$\mathcal{R}(S_1\, S_2\, S_3\, V) = \frac{2\,ab}{a+b}, \tag{21}$$

$$\mathcal{R}(T_a\, T_b S_2\, V) = -1. \tag{22}$$

In a similar way it may be shown that there exists a point Z for which the following hold:

$$\mathcal{R}(S_1^*\, S_2^*\, S_3^* Z) = \frac{2\,\sigma(a)\,\sigma(b)}{\sigma(a)+\sigma(b)}, \tag{23}$$

$$\mathcal{R}(T_a^*\, T_b^*\, S_2^* Z) = -1. \tag{24}$$

Comparison of (22) and (24) shows that $Z = V^*$, whence it follows from (21) and (23) that

$$\sigma\left(\frac{2\,ab}{a+b}\right) = \frac{2\,\sigma(a)\,\sigma(b)}{\sigma(a)+\sigma(b)}. \tag{25}$$

Of course, this equation only has meaning if $a+b \neq 0$. Its derivation, as remarked, is valid only for $a \neq b$. But on the other hand, (25) is, again, trivial if $a = b$.

Now, in (25) let us set $a = 1+c$ and $b = 1-c$, where c can be any element of the ground field. Then $a+b = 2$; and since $\sigma(1) = 1$,[5] it follows from (20) that

[5] This means merely that S_3^* is the image of S_3.

$$\sigma(a)+\sigma(b) = \sigma(a+b) = \sigma(2) = \sigma(1)+\sigma(1) = 2.$$

Hence, from (25) we obtain

$$\sigma(1-c^2) = \sigma(1+c)\cdot\sigma(1-c)$$

or, using (20),[6]

$$\sigma(1)-\sigma(c^2) = [\sigma(1)+\sigma(c)]\,[\sigma(1)-\sigma(c)].$$

From this we readily compute that

$$\sigma(c^2) = [\sigma(c)]^2. \tag{26}$$

Let us now apply σ to the identity

$$a\cdot b = \frac{1}{2}[(a+b)^2-a^2-b^2].$$

Use of the equation $\sigma(a/2)=\sigma(a)/2$, proved above, and of (20) and (26), then yields

$$\sigma(a\cdot b) = \frac{1}{2}([\sigma(a+b)]^2-[\sigma(a)]^2-[\sigma(b)]^2).$$

With the aid of $[\sigma(a+b)]^2 = [\sigma(a)+\sigma(b)]^2$, we have, finally,

$$\sigma(a\cdot b) = \sigma(a)\cdot\sigma(b).$$

This was our goal; σ is now seen to be an automorphism.

We can draw the same conclusions as before from what we have just proved. First, exactly as in Theorem 3, it follows that σ is independent of the choice of the three points S_1, S_2, S_3 used to define it and is determined by the projectivity alone. σ again has the meaning it had in Theorem 4. The same argument shows that Theorem 6 holds true. For what is affirmed in the course of the proof of Theorem 3 implies for our present one-dimensional case precisely this: If the law of formation given in Theorem 6 holds good for a one-to-one mapping between two lines (whether as the result of proof or by assumption), then the statement of Theorem 3 also holds. From the *assumption* of that law of formation it also follows, in particular, that every harmonic set is mapped into a harmonic set.

---

[6] We have seen earlier that from (20) it follows that $\sigma(b)+\sigma(a-b)=\sigma(a)$, i.e., that $\sigma(a-b) = \sigma(a)-\sigma(b)$.

Lastly, Theorem 5 is also true. The only thing remaining to be proved, namely, the uniqueness of a projectivity determined under the assumptions of this theorem, now follows at once in the same manner as the second part of the uniqueness proof of Theorem 5.[7]

Projectivities in REAL P_n

The most important projectivities for geometry are those in which the automorphism σ is the identity mapping of the ground field onto itself, i.e., those in which $\sigma(a) = a$ for all a. Such projectivities are called *linear* and are characterized by the fact that they leave all cross ratios invariant. The following very interesting and basic theorem holds: *In* ***real*** *P_n there exist no projectivities other than linear projectivities.* In other words, we shall prove:

THEOREM 7. *In the field of real numbers the one and only automorphism is the identity mapping.*

Proof: Let σ be an automorphism in the field of real numbers. It is easily seen at once that $\sigma(a) = a$ *for every integer a.* We already know this to be true for $a = 0$ and $a = 1$. If a is now an integer > 1, then by (1),

$$\sigma(a) = \sigma(\underbrace{1 + 1 + \cdots + 1}_{a \text{ times}}) = \underbrace{\sigma(1) + \sigma(1) + \cdots + \sigma(1)}_{a \text{ times}} = a.$$

That the equation $\sigma(a) = a$ holds for a negative integer a then follows immediately from

$$\sigma(-a) = \sigma(0 - a) = \sigma(0) - \sigma(a) = -\sigma(a).$$

Furthermore, $\sigma(a) = a$ holds likewise *for every rational number a.* This is an immediate consequence of formula (8). For, a rational number a is the quotient of two integers, say $a = b/c$. Thus,

$$\sigma(a) = \sigma\left(\frac{b}{c}\right) = \frac{\sigma(b)}{\sigma(c)} = \frac{b}{c} = a.$$

[7] Cf. p. 77.

As our next step we prove that *if* $a > 0$, *then* $\sigma(a) > 0$ *as well, and if* $a < 0$ *then* $\sigma(a) < 0$ *as well.* Since every positive real number is the square of another real number, but a negative number never is, it will suffice to show that the property of being or not being a square is *invariant* under automorphism. But this is trivial. For by (2) it follows from $a = b^2$ that $\sigma(a) = [\sigma(b)]^2$. And if a is not a square, then $\sigma(a)$ also cannot be a square, for from $\sigma(a) = c^2$ it would follow that $a = \sigma^{-1}(c^2) = [\sigma^{-1}(c)]^2$.

Now let us assume there exists a number for which $B \neq \sigma(B)$. We may then take $\sigma(B) < B$. (For if $\sigma(B) > B$, then $\sigma(-B) < -B$.) Then there is certainly some *rational* number a such that[8]

$$\sigma(B) < a < B. \tag{27}$$

Then, since $\sigma(a) = a$, it follows from (27) that

$$\sigma(B - a) = \sigma(B) - a < 0,$$

whereas on the other hand $B - a > 0$. But this is in contradiction to what we just proved.

Thus $\sigma(B) = B$ for all B, as was to be proved.

Let us make clear the meaning of this result for Theorems 5 and 6. The automorphism σ which enters into these theorems must, in real P_n, always be the identity automorphism. Hence, for the projectivity between two s-dimensional linear spaces L and L^* of *real* P_n, Theorem 5 takes on the following form:

THEOREM 8. *If* $Q_0, Q_1, \cdots, Q_s, Q_{s+1}$ *are any* $s+2$ *points of* L *and* $Q_0^*, Q_1^*, \cdots, Q_s^*, Q_{s+1}^*$ *any* $s+2$ *points of* L^* *such that, in both cases, every* $s+1$ *of the points are linearly independent, then it follows that there is one and only one projectivity between* L *and* L^* *which, for every* $i = 0, 1, \ldots, s+1$, *maps the point* Q_i *into* Q_i^*.

Theorem 8 (or its equivalent) is sometimes referred to as *The Fundamental Theorem of Projective Geometry.* It is valid in *complex* P_n only if the additional requirement be made that the projectivity be linear. For in the field of complex numbers there most certainly are automorphisms that are different from the identity mapping.[9] For example, the mapping that sends every complex number into its conjugate is such an automorphism.

---

[8] If E, say, is a rational number such that $E < B - \sigma(B)$, then there exists (at least) one *integer* n such that $\sigma(B) < n \cdot E < B$.

[9] Indeed, there are infinitely many distinct automorphisms in the field of the complex numbers.

Finally, as regards Theorem 6, the description of a projectivity in real P_n now takes on a particularly simple form. It is seen that in the reals Theorem 6 becomes: *The image point* Q^* *has the same coordinates in the coordinate system* $(Q_0^*, Q_1^*, \ldots, Q_s^* \mid E^*)$ *as* Q *does in the coordinate system* $(Q_0, Q_1, \ldots, Q_s \mid E)$.

All the observations of this section can of course be dualized. In dualizing, we obtain mappings of bundles of hyperplanes onto each other. The details of the reinterpretation are easy to carry out and are left to the reader. One case, however, may be emphasized for purposes of later application, namely, that in which two pencils (one-dimensional bundles) are mapped onto each other. It is clear what is meant by a projectivity between two such pencils. The definition is the exact dual of that on p. 79. If we confine ourselves, moreover, to *linear* projectivities, then the duals of Theorems 5 and 6 read as follows:

THEOREM 9. *If* g_1, g_2, g_3 *are three distinct hyperplanes of a pencil* b_1 *and* h_1, h_2, h_3 *likewise three hyperplanes of a pencil* b_2, *then there exists one and only one linear projectivity between* b_1 *and* b_2 *which takes* g_i *into* h_i, *for* $i = 1, 2, 3$. *Furthermore, if the coordinate vectors* $\mathfrak{u}_i$, $\mathfrak{v}_i$ *of* g_i, h_i *are chosen so that* $\mathfrak{u}_3 = \mathfrak{u}_1 + \mathfrak{u}_2$, $\mathfrak{v}_3 = \mathfrak{v}_1 + \mathfrak{v}_2$, *then the mapping* $\lambda\mathfrak{u}_1 + \mu\mathfrak{u}_2 \rightleftarrows \lambda\mathfrak{v}_1 + \mu\mathfrak{v}_2$ *yields precisely the desired projectivity.*

Exercises

1. In Chapter IV (p. 59), perspectivity between two lines was defined. It follows from Theorem 1 of that chapter that every such perspectivity is a *linear* projectivity. The concept of perspectivity can be generalized in the following way.

Let L and L^* be two linear spaces in P_n of the same dimension r. Also, let a linear space M be given which has no point in common with either L or L^*. Let the dimension of M be k. In addition, we think of a one-to-one mapping of L onto L^* as being given. If P is in L and P^* is its image point in L^*, let $V(P, P^*, M)$ designate the linear space of least dimension that contains P, P^*, and M (i.e., the spanning space of P, P^*, and M). For every P in L, inasmuch as no P lies in M, this space has dimension greater than that of M, and thus has dimension at least $k+1$. The given one-to-one correspondence between L and L^* is then called a *perspectivity with center* M if for *every* P of L the dimension of $V(P, P^*, M)$ is equal to $k+1$ (and thus is only greater by one than the dimension of M). If $r = 1$ and $k = n-2$, we obtain the original definition given on p. 59, whereas if $r = 1$ and $k = 0$ (i.e., M is a point) we recover the example given on p. 59.

Show the following:

a) $V(P, P^*, M)$ has only the point P in common with L and only the point P^* in common with L^*.

b) If s is the dimension of the spanning space of L and L^*, then $s - r - 1 \leqq k \leqq n - r - 1$.

c) Now take $s > r$; if a fixed perspectivity with any center is given between L and L^*, then for *each* k that is admissible under b), it is possible to find a linear space M of dimension k that can likewise serve as a center for the given perspectivity.

d) Every perspectivity with a center M is a linear projectivity.

2. As in Exercise 1, let L and L^* be two linear spaces of P_n of the same dimension r. Assume $L \neq L^*$. Let D designate the intersection of L and L^*, and set aside the case in which D is empty. In the other case, let d be the dimension of D.

Now let a *linear* projectivity between L and L^* be given that maps each point of D (if such exists) into itself. Show that this projectivity is a perspectivity (in the sense of Exercise 1).

Hint: First consider the question for the case in which L, L^* are two lines in P_3.

3. Again let L, L^* be two linear spaces of the same dimension r. Let $\tilde{L}_k$ designate the totality of the k-dimensional subspaces of L, and $\tilde{L}_k{}^*$ the corresponding totality for the space L^*. By Theorem 1, every projectivity between L and L^* induces a one-to-one correspondence between $\tilde{L}_k$ and $\tilde{L}_k{}^*$, for every k. According to Theorem 2, this correspondence has the property that the dimension of the intersection of a *finite number* of spaces of $\tilde{L}_k$ is invariant.

Conversely, if for any fixed $k \geqq 1$ there is given a correspondence of this kind between $\tilde{L}_k$ and $\tilde{L}_k{}^*$, then there exists one and only one projectivity which induces that correspondence.

4. The defining property of a projectivity may be weakened by requiring the invariance of linear independence or dependence of k points not for every k, but merely for a *fixed* k such that $3 \leqq k \leqq r + 1$.

5. Let a one-to-one mapping be given of an r-dimensional linear space L on another r-dimensional linear space L^*. Let it have the property that for some definite fixed integer k, which necessarily is $\geqq 3$ and $\leqq r + 1$, it always maps k linearly dependent points of L into k linearly dependent points of L^*. (This is a weaker requirement than that of Exercise 4.) Prove the following:

a) If $P_1, P_2, \ldots, P_h$ are any points of L (h an arbitrary integer) and $P_1{}^*, P_2{}^*, \ldots, P_h{}^*$ are their images in L^*, if, further, V is the spanning space of the P_i and, lastly, if V^* is the image set of V, then every linear subspace of L^* containing the points $P_1{}^*, P_2{}^*, \ldots, P_h{}^*$ also contains V^*. (The proof is easy by mathematical induction.)

b) The given mapping is a projectivity between L and L^* (use a)).

CHAPTER VI

LINEAR PROJECTIVITIES OF P_n ONTO ITSELF

In the preceding chapter, we considered general projectivities between two arbitrary s-dimensional linear spaces L, L^* of P_n, where $1 \leqq s \leqq n$. In this chapter, we shall discuss in greater detail the special case in which $s = n$, i.e., $L = L^* = P_n$. In other words, from now on we consider projectivities that map (real or complex) P_n onto itself.

Such a projectivity can be described in a simple way by a system of equations with respect to a single coordinate system of P_n. To see how this is done, let us consider as given a fixed projective coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ in P_n. Furthermore, let (ξ) be the coordinate matrix[1] of a point S in this coordinate system. Let S^* be the image point of S under the given projectivity and (ξ^*) the coordinate matrix of S^* (again with reference to the *same* coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$). We pose the question: How may (ξ^*) be computed, (ξ) being given?

Of course, we shall try to make use of Theorem 6 of Chap. V. Accordingly, we employ the notation used there: Q_i^* is the image point of Q_i $(i = 0, 1, \ldots, n)$ and E^*, the image point of E. Let σ be the automorphism that belongs, by Chap. V, Theorem 4, to our projectivity.

According to Chap. V, Theorem 6, $(\sigma(\xi))$ is the coordinate matrix of S^* in the coordinate system $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$. In order to find the coordinates of S^* in the coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$, let us recall the formula for transformation of coordinates from Chapter II (p. 36). According to that formula, there exists an $(n+1)$-by-$(n+1)$ matrix $T = (t_{ik})$, with $|T| \neq 0$, such that

$$\xi_i^* = \varrho \cdot \sum_{k=0}^{n} t_{ik}\, \sigma(\xi_k), \qquad i = 0, 1, \cdots, n, \tag{1}$$

or, in matrix form,

---

[1] I.e., S has the coordinates $(Q_0, Q_1, \cdots, Q_n \mid E)$ in $\xi_0, \xi_1, \cdots, \xi_n$. Cf. Chap. II, p. 35.

$$(\xi^*) = \varrho \cdot T \cdot (\sigma(\xi)). \tag{2}$$

In (1) or (2) we have already found the desired system of equations for our projectivity.

Conversely, every system of equations (1) represents a projectivity. To see this, we need merely show that by suitable choice of projectivity the matrix (t_{ik}) can be made equal to any arbitrary prescribed non-singular matrix. But this is an immediate consequence of Chap. II, Theorem 8 and Chap. V, Theorem 5. For we obtained (t_{ik}) as the matrix of a coordinate system. To make it equal to a prescribed matrix, we need merely, according to Chap. II, Theorem 8, alter the points Q_i^*, E^* appropriately. By Chap. V, Theorem 5, however, such an alteration of Q_i^*, E^* is always possible by an appropriate choice of projectivity.

Now let us specialize our discussion still further to *linear* projectivities of P_n onto itself. Thus, σ becomes the identity automorphism, and equations (1) and (2) take on the linear homogeneous form

$$\xi_i^* = \varrho \sum_{k=0}^{n} t_{ik}\, \xi_k, \qquad i = 0, 1, \cdots, n, \tag{3}$$

$$(\xi^*) = \varrho \cdot T \cdot (\xi). \tag{4}$$

The matrix T that appears in (4), is, in a fixed coordinate system, uniquely determined up to an arbitrary factor ϱ by the linear collineation.

For let, say,

$$(\xi^*) = \tilde{\varrho} \cdot \tilde{T} \cdot (\xi) \tag{5}$$

be a second system of equations of the form (4), with non-singular matrix $\tilde{T}$, that represents the same linear projectivity in the same coordinate system. The *inverse* projectivity will then be represented by

$$(\xi) = \frac{1}{\tilde{\varrho}} \cdot \tilde{T}^{-1} \cdot (\xi^*). \tag{6}$$

The product[2] of (4) and (6) must be the identity transformation. That is to say, if we substitute for the ξ_i in the right-hand side of (4) any $n+1$ numbers, compute the corresponding ξ_i^*, and in turn substitute these ξ_i^* into (6), we must always obtain, in the first column of the left-hand side

[2] The product of two one-to-one mappings is defined in *Modern Algebra*, § 19.

of (6), $n + 1$ numbers which differ from the ξ_i merely by a common factor. Thus it follows that the matrix $\tilde{T}^{-1} \cdot T \cdot (\xi)$ is always equal, up to a numerical factor $\neq 0$, to (ξ) itself. But this means that the equation

$$(\eta) = \tilde{T}^{-1} \cdot T \cdot (\xi) \tag{7}$$

represents, in $(n + 1)$-dimensional vector space, a linear transformation that maps every vector into a multiple of itself.[3] Since this would then be true, in particular, for every basis vector, it follows at once that the matrix $\tilde{T}^{-1} \cdot T$ must have diagonal form (cf. the rule for constructing the matrix of a linear transformation given in *Modern Algebra,* p. 286), say:

$$\tilde{T}^{-1} \cdot T = \begin{pmatrix} \alpha_0 & & & 0 \\ & \alpha_1 & & \\ & & \ddots & \\ 0 & & & \alpha_n \end{pmatrix}.$$

If we could now show that $\alpha_0 = \alpha_1 = \cdots = \alpha_n$, the last equation would immediately tell us that $\tilde{T}$ is equal to a constant factor times T. The equality of the α_i, however, may be shown as follows. We set $\xi_0 = \xi_1 = \cdots = \xi_n = 1$ in the right-hand side of (7), whence $\eta_0 = \alpha_0, \eta_1 = \alpha_1, \cdots, \eta_n = \alpha_n$. Since the η_i, however, can only differ from the ξ_i by a common constant of proportionality, all the α_i must be equal.

Now, how does the system of equations of a linear projectivity change under a transformation of coordinates? Let (4), say, and

$$(\eta^*) = \lambda \cdot U \cdot (\eta) \tag{8}$$

($U = (u_{ik})$, a non-singular matrix) be the systems of equations for a linear collineation in two different coordinate systems. Let $V = (v_{ik})$ be the matrix of the equations for the transformation of coordinates, so that

$$(\eta) = \mu \cdot V \cdot (\xi), \qquad (\eta^*) = \nu \cdot V \cdot (\xi^*)\,, \tag{9}$$

where (η) and (ξ) represent the coordinate matrices of a point S in the two different coordinate systems and (η^*), (ξ^*), similarly, the coordinate

[3] That is not to say, however, that the same multiple is involved throughout. For all we know thus far, two vectors $\mathfrak{x}$, $\mathfrak{y}$ might go over into $c_1 \cdot \mathfrak{x}$, $c_2 \cdot \mathfrak{y}$, with $c_1 \neq c_2$.

matrices of the image point S^*. Then it follows, upon substituting from (9) into (8) $\left(\text{and setting } \frac{\lambda \cdot \mu}{\nu} = \tau\right)$, that

$$(10) \qquad (\xi^*) = \tau \cdot V^{-1} \cdot U \cdot V \cdot (\xi)$$

must also be a system of equations for our projectivity in the same coordinate system as that in which (4) holds. Hence,

$$(11) \qquad V^{-1} \cdot U \cdot V = \alpha \cdot T,$$

$$(12) \qquad U = \alpha \cdot V \cdot T \cdot V^{-1}.$$

Thus, as with linear transformations, U is obtained from T essentially (i.e., aside from the arbitrary constant of proportionality α) by transformation with the matrix of the coordinate system. If we think of one of the coordinate systems as fixed and the other as variable, we can make V equal to any preassigned arbitrary non-singular matrix by suitable choice of the variable system (Chap. II, Theorem 8). We can also express this as follows: *If T is the matrix of a linear projectivity in a certain projective coordinate systems and V is any given non-singular square matrix, then a projective coordinate system can always be found in which $V \cdot T \cdot V^{-1}$ is the matrix of the given projectivity.*

In Chapter V of *Modern Algebra* we saw how a square matrix could be reduced to certain simple forms, the so-called normal forms, by transformation with a suitable non-singular matrix. We can now make use of this for the study of linear projectivities. It is quite clear that it will be advantageous to study a linear collineation in a coordinate system with reference to which the equations take on the simplest possible form. To show how easy it is to deduce geometrical properties of the linear projectivity in this way, we give a few examples.

As a first example, let us consider the *diagonal* form. What are the implications for a linear projectivity if its matrix in some coordinate system is susceptible of being put into diagonal form? The system of equations of the projectivity will have the form

$$(13) \qquad \begin{aligned} \xi_0^* &= \varrho \cdot \alpha_0 \cdot \xi_0, \\ \xi_1^* &= \varrho \cdot \alpha_1 \cdot \xi_1, \\ &\cdots\cdots \\ \xi_n^* &= \varrho \cdot \alpha_n \cdot \xi_n \end{aligned}$$

where ϱ is an arbitrary non-zero constant (the same, however, for all $n+1$ equations), while $\alpha_0, \alpha_1, \ldots, \alpha_n$ are *fixed* numbers $\neq 0$.[4] From the equations (13) we readily deduce that *the* $n+1$ *fundamental points*[5] *of our coordinate system are fixed points of the linear projectivity.*

Conversely, if a linear collineation is given having $n+1$ linearly independent fixed points, we can take these fixed points as the fundamental points of a coordinate system. Let

$$\xi_i^* = \varrho \cdot \sum_{k=0}^{n} t_{ik}\, \xi_k, \qquad i = 0, 1, \ldots, n, \tag{14}$$

say, be the system of equations of the linear projectivity in this coordinate system. From these equations let the coordinates of the image points of the h-th fundamental point $(0 \leqq h \leqq n)$ be computed:[6] $t_{0h}, t_{1h}, \ldots, t_{nh}$. From the invariance of the fundamental point considered it follows that $t_{ih} = 0$ for $i \neq h$ and $t_{hh} \neq 0$. I.e., the matrix of (14) has diagonal form.

We have thus proved the following:

A linear projectivity in P_n has $n+1$ linearly independent fixed points if and only if its matrix can be put into diagonal form.

As a second example, we shall make use of the *Jordan normal form* (*Modern Algebra,* § 26) in order to get a general picture of the linear projectivities of P_1, i.e., of straight lines. We know that a matrix over the field of complex numbers can always be transformed into Jordan normal form, but that a matrix over the field of real numbers cannot always be so transformed. Accordingly, let us first consider *complex* P_1. The matrix of a linear projectivity of P_1 is a 2-by-2 matrix. Therefore (*Modern Algebra,* § 26), the possible Jordan normal forms are

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}, \qquad \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}. \tag{15}$$

Thus, in complex P_1 a projective coordinate system can always be found in which the matrix of a given linear projectivity takes one of the two forms (15).

---

[4] The α_i must all be $\neq 0$, since the matrix of a linear projectivity is always *non-singular.*

[5] The unit point, of course, need not be a fixed point. If all the fundamental points *and* the unit point are fixed points, then the linear projectivity is the identity mapping (Chap. V, Theorem 5).

[6] We have dropped the factor ϱ common to all the coordinates.

It is clear from previous considerations that the matrix of a fixed linear projectivity is not capable of both forms (15). For otherwise there would be an equation of the kind (12) between the two matrices (15) (in which the matrices (15) play the roles of T and U in (12)). That is, a multiple of the first matrix of (15) would be obtainable from the second by transformation. But that is impossible, because the elementary divisors of the corresponding characteristic matrices are distinct.

The two cases can easily be distinguished geometrically. From our previous discussion, we know that the *first case* can occur if and only if there are *two distinct fixed points.* Consequently, in the second case there can be only one fixed point at most. But such a fixed point does occur, namely the point with coordinates 1, 0 in that coordinate system in which the matrix of the linear projectivity has the form of the second matrix of (15). The *second case* is thus characterized by the fact that there exists *one and only one fixed point.*

Let us now consider a linear projectivity in *real* P_1, which of course has a real system of equations with a real matrix in a projective coordinate system of real P_1. By *Modern Algebra,* § 26, a matrix over the reals can be transformed into Jordan normal form if and only if its characteristic polynomial can be completely decomposed into *real* linear factors. Thus, besides the again possible cases (15), we have the still further case in which the characteristic polynomial has two complex roots, which, by *Modern Algebra,* § 17, Theorem 10, are then necessarily conjugate.

In order to consider this last case more closely, let us think of the matrix of the linear projectivity as being fixed[7] in a fixed coordinate system, say $A = (a_{ik})$. Let the roots of the characteristic polynomial be $\alpha = a + ib$, $\bar{\alpha} = a - ib$ (a, b real, $b \neq 0$). The characteristic polynomial itself is then $\chi(u) = (u - \alpha)(u - \bar{\alpha}) = u^2 - 2au + a^2 + b^2$. The elementary divisors of the characteristic matrix of A also can be found immediately. For, on the one hand, each of these elementary divisors must be a divisor of $\chi(u)$ (cf., for example, *Modern Algebra,* § 25, formula (9)), and, on the other hand, $\chi(u)$ has no proper real divisors. The elementary divisors of the characteristic matrix of A are therefore 1, $\chi(u)$. The same is true for *every* real matrix with characteristic polynomial $\chi(u)$. By *Modern Algebra,* § 26, Theorem 2, A may therefore be transformed by a *real* transformation into any real matrix with characteristic polynomial $\chi(u)$. This must be true, in particular, for the matrix

[7] I.e., let even the arbitrary factor ϱ be thought of, for the sake of simplicity, as fixed.

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \tag{16}$$

which we can therefore use in this case as a (real) normal form for our linear projectivity.

From this normal form we can again easily determine the number of fixed points of our linear collineation. For such a fixed point (with coordinates ξ_0, ξ_1), we must have, by (16),

$$\begin{aligned} \lambda \xi_0 &= a\xi_0 + b\xi_1, \\ \lambda \xi_1 &= -b\xi_0 + a\xi_1, \end{aligned} \tag{17}$$

where λ must necessarily be real. Since we can also write (17) in the form

$$\begin{aligned} \xi_0 \cdot (a - \lambda) + \xi_1 \cdot b &= 0, \\ \xi_0 \cdot (-b) + \xi_1 \cdot (a - \lambda) &= 0, \end{aligned} \tag{18}$$

we see that $\begin{vmatrix} a-\lambda & b \\ -b & a-\lambda \end{vmatrix} = 0$, i.e., that the *real* number λ must be a root of the characteristic polynomial $\chi(u)$ of (16). But $\chi(u)$ has no real zeros. Consequently, our linear projectivity can have *no* real fixed point in this case.[8]

In summary, we can say:

The matrix of a linear projectivity of real P_1 can be made to assume one of the following three forms by suitable choice of coordinate system.

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}, \quad \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}, \quad \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}. \tag{19}$$

The first case occurs if and only if there are at least[9] two (real) fixed points, the second when there is exactly one fixed point, and the third when there is no (real) fixed point.

The three cases (in the order in which they occur in (19)) are sometimes referred to as *hyperbolic, parabolic,* and *elliptic,* respectively.

---

[8] Equations (17) have *complex* solutions; one need merely set $\lambda = \alpha$ or $\lambda = \bar{\alpha}$. And in fact, if ξ_0, ξ_1 is a solution for $\lambda = \alpha$, then $\bar{\xi}_0, \bar{\xi}_1$ is a solution for $\lambda = \bar{\alpha}$. For this reason, we frequently speak of *two complex conjugate fixed points.*

[9] If more than two fixed points occur, the linear projectivity is the identity mapping.

Finally, let us consider the *involutory linear projectivities* of projective P_1 (real or complex). A one-to-one mapping of any set onto itself is called *involutory,* or an *involution,* if its square is the identity mapping.

In particular, the square of a linear projectivity $(\xi^*) = \varrho \cdot A \cdot (\xi)$ is given by the equation $(\xi^*) = \lambda \cdot A^2 \cdot (\xi)$. The requirement that such a projectivity be involutory then becomes: A^2 must be equal, aside from a non-zero numerical factor, to the unit matrix.

To apply this to P_1, let us form the squares of the matrices (19). They are, respectively,

$$(20) \qquad \begin{pmatrix} \alpha_1^2 & 0 \\ 0 & \alpha_2^2 \end{pmatrix}, \quad \begin{pmatrix} \alpha^2 & 2\alpha \\ 0 & \alpha^2 \end{pmatrix}, \quad \begin{pmatrix} \alpha^2 - \beta^2 & 2\alpha\beta \\ -2\alpha\beta & \alpha^2 - \beta^2 \end{pmatrix}.$$

The second of these matrices can not equal $\varrho \cdot E$ $(\varrho \neq 0)$ for any value of α. Thus, a linear projectivity with exactly one fixed point can never be an involution. In *complex* P_1, therefore, only the case of the first matrix (19) remains. By (20), it represents an involution if and only if $\alpha_1^2 = \alpha_2^2 = \varrho$. Hence the first matrix (19) then has the form

$$\begin{pmatrix} \pm\sqrt{\varrho} & 0 \\ 0 & \pm\sqrt{\varrho} \end{pmatrix}.$$

If the sign before the radical is the same in both cases, the mapping is the identity. Since we can moreover always divide through by $+\sqrt{\varrho}$ or $-\sqrt{\varrho}$, we see the following:

The matrix of any involuntary linear projectivity of complex P_1, other than the identity, can always be put in the form

$$(21) \qquad \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In real P_1 we also need to consider the third matrix of (19) or (20). Since β is necessarily different from zero in this case,[1] the third matrix of (20) can equal $\varrho \cdot E$ $(\varrho \neq 0)$ only if $\alpha = 0$. From this, we see the following:

[1] Otherwise we just have the case of the *first* matrix of (19) over again.

In real P_1 the matrix of an involutory linear projectivity other than the identity can be put either into the form (21) *or the form*

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{22}$$

In the case of the matrix (21), the fixed points are the fundamental points of the coordinate system. Denote them by Q_0, Q_1. Let S be any point $\neq Q_0, Q_1$ and let x_1, x_2 be its coordinates with respect to the coordinate system in question. Its image point S^* then has, by (21), the coordinates $x_1, -x_2$. From this, by Chapter IV, we have

$$\mathcal{R}(Q_0, Q_1, S, S^*) = -1.$$

This tells us that *every pair of corresponding points S, S^* separate the fixed points harmonically.*

Conversely, a mapping that sends the point with coordinates x_0, x_1 in the coordinate system $(Q_0, Q_1 \mid E)$ (the unit point E being arbitrary) into the point with coordinates $x_0, -x_1$ is an involutory linear projectivity of the desired sort. Thus, it follows at once that *for two given points Q_0, Q_1 there always exists one and only one involutory linear projectivity, distinct from the identity, which has Q_0, Q_1 as fixed points.* In this form, the theorem is valid for real as well as for complex P_1.

The totality of the linear projectivities of P_n constitute a subgroup of the complete permutation group of P_n (*Modern Algebra*, § 19). For, the product of two linear collineations as well as the inverse of a linear collineation is always itself a linear collineation. The group of all linear projectivities of P_n is called the *linear* **projective group**, and the terms *real* or *complex* projective group are used depending on whether the projectivities of real or of complex P_n are meant.

An important subgroup of the projective group consists of those linear projectivities that transform the proper points of P_n into proper points and the improper points into improper points. In *real* P_n these projectivities are closely related to affine transformations; in fact, we have the following:

A linear projectivity of real P_n that takes every improper point into an improper point induces an affine transformation in affine R_n (i.e., in the totality of the proper points of real P_n).

To prove this, let us take as our projective coordinate system the one in which each point has its natural coordinates. That is to say, we choose the $n+1$ points $[1, 0, 0, \ldots, 0]$, $[0, 1, 0, \ldots, 0]$, $\ldots$, $[0, 0, \ldots, 0, 1]$ as fundamental points and $[1, 1, \ldots, 1]$ as unit point.[2] Then, as we know, $\xi_0 = 0$ is the equation of the improper hyperplane. Now let

$$(23) \qquad \xi_i^* = \varrho \cdot \sum_{k=0}^{n} t_{ik}\, \xi_k, \qquad i = 0, 1, \cdots, n,$$

be the system of equations in this coordinate system of a linear projectivity that maps the improper hyperplane onto itself. If in the right-hand side of (23) we put

$$\xi_0 = \xi_1 = \cdots = \xi_{h-1} = 0, \quad \xi_h = 1, \quad \xi_{h+1} = \xi_{h+2} = \cdots = \xi_n = 0$$

then, on the left-hand side we must have $\xi_0^* \neq 0$ for the case $h = 0$, but $\xi_0^* = 0$ for the cases $h > 0$. But this is possible only if $t_{00} \neq 0$, $t_{01} = t_{02} = \cdots = t_{0n} = 0$. Thus, the system of equations (23), when written out, has the form

$$(24) \qquad \begin{aligned} \xi_0^* &= \varrho\,(t_{00}\, \xi_0), \\ \xi_1^* &= \varrho\,(t_{10}\, \xi_0 + t_{11}\, \xi_1 + \cdots + t_{1n}\, \xi_n), \\ &\cdots\cdots\cdots\cdots\cdots \\ \xi_n^* &= \varrho\,(t_{n0}\, \xi_0 + t_{n1}\, \xi_1 + \cdots + t_{nn}\, \xi_n). \end{aligned}$$

Conversely, it is easy to see that a linear projectivity whose system of equations in our 'natural' coordinate system has the form (24), carries every proper point into a proper point and every improper point into an improper point. Thus, the form (24) of the system of equations is characteristic for a linear projectivity of this kind.

Let us now consider the linear collineation (24) only as regards how it affects the proper points. Since for such points both ξ_0 and ξ_0^* are $\neq 0$, we can divide all the equations (24) by the first of them, thus obtaining for the last n equations the form

$$(25) \quad \frac{\xi_i^*}{\xi_0^*} = \frac{t_{i0}}{t_{00}} + \frac{t_{i1}}{t_{00}}\left(\frac{\xi_1}{\xi_0}\right) + \frac{t_{i2}}{t_{00}}\left(\frac{\xi_2}{\xi_0}\right) + \cdots + \frac{t_{in}}{t_{00}}\left(\frac{\xi_n}{\xi_0}\right), \; i = 1, 2, \cdots, n.$$

If in this we now set $\dfrac{\xi_i^*}{\xi_0^*} = x_i^*$, $\dfrac{\xi_i}{\xi_0} = x_i$, we see that the mapping

[2] Cf. Chap. II, p. 32.

induced in R_n can be represented in a certain linear coordinate system[3] by the system of equations

$$(26) \qquad x_i^* = \frac{t_{i0}}{t_{00}} + \sum_{k=1}^{n} \frac{t_{ik}}{t_{00}} x_k, \qquad i = 1, 2, \cdots, n.$$

But this is the system of equations of an affine transformation and, moreover, because of the one-to-one nature of the mapping, of a non-singular affine transformation. (Cf. *Modern Algebra,* § 13.)

All non-singular affine transformations can be obtained in this way. For by a suitable choice of (24), (26) can be made the system of equations of any arbitrarily prescribed non-singular affine transformation of affine R_n.

Can it happen that two *different* linear projectivities of P_n both of which leave the improper plane (as a whole) invariant will induce the *same* transformation in affine R_n? No! For each time we can choose a projective coordinate system in P_n whose fundamental points and unit point are all proper points. Then if two linear collineations coincide in their effects on the proper points, the fundamental points and unit point must have the same images under both collineations. From this, however, it follows at once, by Chap. V, Theorem 5 (σ is now the identity automorphism), that the two linear projectivities must be identical throughout the whole of P_n.

In summary we can say:

Every non-singular affine transformation of affine R_n is induced by one and only one linear projectivity of real projective P_n.

Because of this relation, the linear projectivities that map the improper plane onto itself are themselves called *affine transformations of P_n*,[4] and the totality of these projectivities is referred to as the ***affine group*** *of P_n*. We also agree that this latter term is to apply not only to real P_n but to *complex* P_n as well.

We leave as an exercise for the reader the simple but not particularly important dualization of the discussion in the foregoing paragraphs. It should be noted, however, that the dualization of our last results, concerning affine transformations, is impossible—or at least, geometrically meaningless—because of the use of the improper elements (cf. Chap. III).

---

[3] This is the natural (affine) coordinate system, in which every point $(x_1, x_2, \ldots, x_n)$ has the x_i themselves as coordinates. Cf. Chap. I, p. 16.

[4] We here drop a modifier like 'non-singular,' since in *projective* P_n we never consider anything but one-to-one mappings.

Exercises

1. Show that there exist exactly two automorphisms of the field of complex numbers which map every real number into a real number, namely, the identity automorphism and the automorphism that takes every number into its complex conjugate.

The projectivities formed with the latter automorphism, in conformity with Theorem 6 of Chapter V, are called *anti-collineations.*

2. Every point $[\xi_0, \xi_1, \cdots, \xi_n]$ of real P_n also belongs to complex P_n. Such a point is called a *real* point of complex P_n. A projectivity of complex P_n onto itself that maps every real point into a real point induces a linear projectivity in real P_n. Every linear projectivity of real P_n can be induced by exactly two projectivities of complex P_n, namely a linear collineation and an anti-collineation.

3. Let T be an n-by-n square matrix and let $e_1(u), e_2(u), \ldots, e_n(u)$ be the elementary divisors of its characteristic matrix $T - uE$. Furthermore, let ν_i be the degree of $e_i(u)$ $(i = 1, 2, \ldots, n)$ and let ϱ be a constant (an element of the field).

From $e_i(u)$ we derive a new polynomial by replacing the indeterminate u by $\dfrac{u}{\varrho}$ and then multiplying the result by ϱ^{ν_i}. The resulting polynomial, which again has leading coefficient 1, we accordingly denote by $\varrho^{\nu_i} \cdot e_i\left(\dfrac{u}{\varrho}\right)$. Show that the n polynomials $\varrho^{\nu_i} \cdot e_i\left(\dfrac{u}{\varrho}\right)$, $i = 1, 2, \ldots, n$ are the elementary divisors of the characteristic matrix of $\varrho \cdot T$.

What relation exists between the Jordan normal forms of T and $\varrho \cdot T$ (in case their Jordan normal forms exist)?

4. Let A, B be two $(n+1)$-by-$(n+1)$ square matrices with real or complex numbers as elements. Let $e_0(u), e_1(u), \ldots, e_n(u)$ be the elementary divisors of $A - uE$ and $\bar{e}_0(u), \bar{e}_1(u), \ldots, \bar{e}_n(u)$ those of $B - uE$, and let ν_i be the degree of $e_i(u)$ $(i = 0, 1, \ldots, n)$. The following is an easy consequence of Exercise 3.

A and B can appear as matrices of one and the same linear projectivity of P_n if and only if there exists a single constant $\varrho \neq 0$ such that for all $i = 0, 1, \ldots, n$, simultaneously, we have $\bar{e}_i(u) = \varrho^{\nu_i} \cdot e_i\left(\dfrac{u}{\varrho}\right)$.

5. If $\bar{\xi}_i$ denotes the complex conjugate of ξ_i, then the point $[\bar{\xi}_0, \bar{\xi}_1, \cdots, \bar{\xi}_n]$ is called the complex conjugate of $[\xi_0, \xi_1, \cdots, \xi_n]$.

Consider in complex P_1 an involutory linear projectivity different from the identity and having no *real* fixed point. Show that the linear projectivity will map each real point into another real point if and only if the fixed points are complex conjugates.

6. Let h be a fixed hyperplane in P_n. Give an example of a linear projectivity which leaves *every* point of h fixed, but has *no* fixed point outside of h.

CHAPTER VII

CORRELATIONS

Besides the mapping of two sets of points on each other and the dual mapping between sets of hyperplanes, mappings of sets of points onto sets of hyperplanes also play a certain role in projective geometry. We restrict ourselves here to the case in which, on the one hand, the set of points is all the points of P_n and, on the other hand, the set of hyperplanes is all the hyperplanes of P_n.

We now wish to consider a given mapping (correspondence) which associates one and only one hyperplane with every point of P_n. To get something other than mappings of points, we must assume $n \geqq 2$, since the hyperplanes of P_1 are themselves merely points. Furthermore, we require that our mapping be one-to-one—that is, of such a nature that *every* hyperplane is the image of one and only one point. Finally, we require, in analogy to projectivities, that if $Q_1, Q_2, \ldots, Q_k$ are any finite number of points of P_n and $h_1, h_2, \ldots, h_k$ are their image hyperplanes, then $h_1, h_2, \ldots, h_k$ shall always be linearly dependent if and only if this is also true of $Q_1, Q_2, \ldots, Q_k$. Such a mapping is called a **correlation.**

If the point $[\xi_0, \xi_1, \ldots, \xi_n]$ is associated under our correlation with the hyperplane $\langle u_0, u_1, \ldots, u_n \rangle$, we shall write

$$(1) \qquad [\xi_0, \xi_1, \cdots, \xi_n] \rightleftarrows \langle u_0, u_1, \cdots, u_n \rangle .$$

To facilitate the investigation of this mapping, let us at the same time consider the point-point mapping

$$(2) \qquad [\xi_0, \xi_1, \cdots, \xi_n] \rightleftarrows [u_0, u_1, \cdots, u_n]$$

of P_n onto itself, a mapping which associates those points whose coordinates ξ_i, u_i satisfy the relation (1). By virtue of the conditions imposed on the correlation (1), it is clear that the mapping (2) is a projectivity of P_n onto itself.[1]

[1] Keep in mind that we have assumed $n \geqq 2$.

As we have seen in the preceding chapter, there must exist an $(n+1)$-by-$(n+1)$ matrix $T = (t_{ik})$, with $|T| \neq 0$, and an automorphism σ of the ground field such that the mapping (2) can be represented by

$$(u) = \varrho \cdot T \cdot (\sigma(\xi)) \tag{3}$$

(see (2), Chap. VI). But since exactly the same numbers u_i always appear in the mapping (1) as appear in (2), it follows that (3) also represents the mapping denoted by (1). Conversely, if we define the mapping (1) by a preassigned system of equations of the kind given in (3), the mapping will necessarily satisfy our conditions.

This derivation of equation (3) for a correlation remains valid when we interpret the ξ_i and u_i as point and hyperplane coordinates respectively in any projective coordinate system.

By comparing mappings (1) and (2), we can immediately draw some further conclusions. Consider an r-dimensional linear space L. All the image points of the points of L under the mapping (2) again constitute, by Chap. V, Theorem 1, an r-dimensional linear space, i.e., their coordinates satisfy a system of homogeneous linear equations of rank $n - r$. This fact is not altered if we interpret the u_i, not as point coordinates, but as hyperplane coordinates. We thus have Theorem 1.

THEOREM 1. *Under a correlation there corresponds to an r-dimensional linear space of P_n exactly one r-dimensional linear bundle as the totality of its image elements, and conversely.*

If we interpret equation (3) as a projectivity, in accordance with (2), the automorphism σ tells us, by Chap. V, Theorem 3, how the cross ratio alters under the mapping. The same must therefore be true if we interpret the u_i as hyperplane coordinates in accordance with (1). We thus have Theorem 2.

THEOREM 2. *If Q_1, Q_2, Q_3, Q_4 are four points of a line with $\mathcal{R}(Q_1 Q_2 Q_3 Q_4) \neq \infty$ and h_1, h_2, h_3, h_4 are their image hyperplanes as given by mapping* (3), *then* $\mathcal{R}(h_1 h_2 h_3 h_4) = \sigma[\mathcal{R}(Q_1 Q_2 Q_3 Q_4)]$.

Mappings of the type (3) in which σ is the identity automorphism are, again, of particular geometrical importance. They are the counterpart to linear collineations and are called *linear correlations*. By Theorem 2, a correlation is linear if and only if it leaves all cross ratios invariant.

The system of equations for a linear correlation takes on the linear homogeneous form

$$(u) = \varrho \cdot T \cdot (\xi). \tag{4}$$

This holds for any arbitrary projective coordinate system. In a fixed coordinate system the matrix T is uniquely determined up to an arbitrary constant ϱ ($\neq 0$).[2]

Let us now consider how the equations of a *fixed* linear correlation in two *different* coordinate systems are related. Let (4), say, be the equation of a correlation in one coordinate system and

$$(v) = \bar{\varrho} \cdot U \cdot (\eta) \tag{5}$$

its equation in a second coordinate system. Let V be the matrix of the equations of transformation, so that $(\eta) = \lambda \cdot V \cdot (\xi)$ and consequently, by Chap. III, equation (10), $(v) = \mu \cdot (V')^{-1} \cdot (u)$. If we substitute this in (5), we see that our correlation, in the coordinate system with respect to which (4) was given, may be represented by the equation[3]

$$(u) = \varrho \cdot V' U V (\xi). \tag{6}$$

Therefore a number $\alpha \neq 0$ must exist such that

$$T = \alpha \cdot V' U V. \tag{7}$$

But formula (7) is just the relation we were looking for.

We now seek to determine those points of P_n which, under the linear correlation (4), lie in their own image hyperplanes. The incidence requirement for a point with coordinate matrix (ξ) and a hyperplane with coordinate matrix (u) is, in matrix form,

$$(u)' \cdot (\xi) = O. \tag{8}$$

If we replace (u) here by its expression from (4), we obtain (since $\varrho \neq 0$) the equation

$$(\xi)' \cdot T' \cdot (\xi) = O \tag{9}$$

[2] For two equations of form (4) representing the same linear correlation must represent the same linear collineation if the u_i as well as the ξ_i be interpreted as point coordinates. From Chapter VI, however, we know that the matrix of a linear collineation is, except for ϱ, uniquely determined.

[3] We have again written ϱ in place of $\dfrac{\bar{\varrho} \cdot \lambda}{\mu}$.

as the necessary and sufficient condition that a point (ξ) must satisfy if it is to lie in its own image hyperplane. This condition may also be written in the form

$$\sum_{i,k=0}^{n} t_{ik}\,\xi_i\,\xi_k = 0. \tag{10}$$

In analogous fashion, by putting $(\xi) = 1/\varrho \cdot T^{-1}(u)$ in (8), we obtain the equation

$$(u)'\,T^{-1}(u) = 0 \tag{11}$$

as the condition that hyperplane (u) pass through its own image point.

It can happen that equation (9), or (10), holds for *every* point of P_n. The linear correlation is then called a *null system.* What conditions would have to be imposed on the matrix of correlation (4) for it to represent a null system? Since, in that case, the left-hand side of (10) must then vanish for every value of ξ_i, it must do so, in particular, if for a fixed i, we set

$$\xi_0 = \xi_1 = \cdots = \xi_{i-1} = 0,\ \xi_i = 1,\ \xi_{i+1} = \xi_{i+2} = \cdots = \xi_n = 0.$$

From this, it follows that for every i,

$$t_{ii} = 0, \qquad i = 0, 1, \cdots, n. \tag{12}$$

Furthermore, if for two different but fixed subscripts i, k in (10), we set $\xi_i = \xi_k = 1$, and $\xi_h = 0$ for $h \neq i$, k, it follows that

$$t_{ii} + t_{kk} + t_{ik} + t_{ki} = 0, \qquad i, k = 0, 1, \cdots, n;\ i \neq k. \tag{13}$$

From (12) and (13) we see we must have for *every* pair of subscripts i, k (and this is also true for $i = k$)

$$t_{ik} = -t_{ki}, \qquad i, k = 0, 1, \cdots, n. \tag{14}$$

The relations (14) represent a *necessary condition* that (4) be a null system. But, in addition, they are also *sufficient.* For if they are satisfied for *every* pair of subscripts i, k, it immediately follows that $t_{ii} = 0$ for $i = 0, 1, \ldots, n$. Moreover, for every pair i, k with $i \neq k$, the terms $t_{ik}\xi_i\xi_k$ and $t_{ki}\xi_i\xi_k$ in (10) cancel each other. Thus, (10) is identically satisfied in the ξ_i.

A matrix $T = (t_{ik})$ that satisfies condition (14) is called *skew-symmetric*. Thus we may state the following.

THEOREM 3. *The linear correlation* (4) *is a null system if and only if T is skew-symmetric.*

The condition (14) can also be written in matrix form as

$$T + T' = O. \tag{15}$$

The question as to the *existence of null systems* is still not decided, however, but is merely reduced to the following: Do *non-singular* skew-symmetric matrices exist? If the $(n+1)$-by-$(n+1)$ matrix $T = (t_{ik})$ is skew-symmetric, it follows from $T' = (-t_{ik})$ that

$$|T'| = (-1)^{n+1}|T|. \tag{16}$$

However, since $|T'| = |T|$ for every square matrix, from (16) it follows, for *even* n, that $|T| = -|T|$, and so $|T| = 0$. Thus we see that *in even-dimensional spaces P_n, there are no null spaces.*

The reverse is true if n is *odd*. For then $n + 1$ is even, and consequently $h = (n+1)/2$ is a whole number. The matrix

$$\begin{pmatrix} 0 & & & & & & +1 \\ & & & & & \cdot & \\ & & & & \cdot & & \\ & & & +1 & & & \\ & & -1 & & & & \\ & \cdot & & & & & \\ -1 & & & & & & 0 \end{pmatrix}, \qquad \begin{matrix} +1 \text{ occurs } h \text{ times} \\ -1 \text{ occurs } h \text{ times} \end{matrix} \tag{17}$$

furnishes an example of a non-singular skew-symmetric $(n+1)$-by-$(n+1)$ matrix; in this matrix only the elements of the *secondary* diagonal (from upper right to lower left) need be different from zero.

THEOREM 4. *Null systems exist in the spaces $P_3, P_5, P_7, \ldots$, but not in the spaces $P_2, P_4, P_6, \ldots$.*

Let there now be given a hyperplane h, whose coordinates with reference to a fixed projective coordinate system we shall denote by $v_0, v_1, \ldots, v_n$. If we look for the image hyperplanes of each point of h under a given linear correlation, these hyperplanes, by Theorem 1, constitute a hyperbundle with a single point Q as kernel. It is easy to compute the coordinates $\eta_0, \eta_1, \ldots, \eta_n$ of Q.

To this end, let us write the equation of h in matrix form:

$$(v)' \cdot (\xi) = O. \tag{18}$$

Let the equation of the linear correlation be given by (4). If we solve (4) for (ξ) and substitute it in (18), we see that the coordinates $u_0, u_1, \ldots, u_n$ of each one of the image hyperplanes must satisfy the equation

$$(v)' \cdot T^{-1} \cdot (u) = O. \tag{19}$$

This equation, which is a linear homogeneous equation in the u_i, does, in fact, represent a hyperbundle. By Chap. III, Theorem 1, the kernel of this hyperbundle has the coefficients of equation (19) as coordinates. Thus, we must have

$$(\eta) = \varrho \cdot (T^{-1})' \cdot (v) = \varrho \cdot (T')^{-1} \cdot (v), \tag{20}$$

or[4]

$$(v) = \varrho \cdot T' \cdot (\eta). \tag{21}$$

Since this system of equations again is of the form (4), we have Theorem 5.

THEOREM 5. *If, starting out with a given linear correlation* (4), *a correspondence is set up between each hyperplane and the carrier of its image hyperbundle, this correspondence is itself a linear correlation and is represented by equation* (21).

In general, the correlation (21) will be different from the correlation (4). If (4) and (21) represent the same linear correlation, the correlation is called *involutory*.[5] When will this occur? As we know, T and T' must

---

[4] We have once again replaced the factor $1/\varrho$, which appears on the left-hand side when we solve (20) for (v), by ϱ, since ϱ, and hence $1/\varrho$ also, are entirely arbitrary.

[5] The term may be justified as follows. A linear correlation can be looked upon as a one-to-one mapping of the set of all points *and* hyperplanes onto itself, in which the image of a point is the hyperplane given by (4) and the image of a hyperplane is the point given by (20). In the involutory case, the square of *this* mapping is the identity.

then be identical up to a constant of proportionality. Thus, there must exist a $\lambda \neq 0$ such that

$$T' = \lambda \cdot T. \tag{22}$$

Now, equation (22) remains valid if we replace the matrices by their transposes. Hence,

$$T = \lambda \cdot T' \tag{23}$$

also. If we substitute in (23) the expression for T' from (22), we obtain

$$T = \lambda^2 \cdot T. \tag{24}$$

Since not all the elements of the *non-singular* matrix T can be equal to 0, it follows from (24) that $\lambda^2 = 1$, and so $\lambda = \pm 1$. Thus, we have an involutory linear correlation if and only if either $T' = -T$ or $T' = T$.

We have already dealt with the skew-symmetric case $T' = -T$. These linear correlations are the null systems. In the case $T' = T$—in other words, when the matrix is symmetric—the linear correlation is called a *polarity*.[6] In this terminology, we can now state the following theorem.

Theorem 6. *The only involutory linear correlations are the null systems and the polarities.*

Exercises

1. Let σ and τ be two one-to-one mappings in real P_n of the totality of points onto the totality of hyperplanes. Let σ and τ be related as follows: If the hyperplane u denotes the image of a point P under σ, and the hyperplane v the image of a point Q under τ, then the incidence of P and v always implies the incidence of Q and u. Show that σ and τ are necessarily linear correlations and are related in the way described in Theorem 5.

2. Let $(u) = \varrho\, T(\xi)$ be a given linear correlation. Consider also the correlation $(u) = \varrho\, T'(\xi)$ related to it as in Theorem 5. If, beginning with an arbitrary point P, we pass to the image hyperplane u by the first correlation and thence from this hyperplane u to its image point Q under the second correlation, then the mapping $P \to Q$ is a linear collineation ω. Also consider the inverse collineation ω^{-1} (i.e., $Q \to P$). The matrices of ω and ω^{-1} in a fixed coordinate system are (after appropriate normalization) transformable into each other (in the sense of *Modern Algebra,* § 26, Theorem 2).

[6] The existence of polarities is trivial; the identity mapping is always a polarity.

CHAPTER VIII

HYPERSURFACES OF THE SECOND ORDER

In the last chapter, upon inquiring into the condition under which a point lies in its own image hyperplane under a given linear correlation of P_n, we obtained an equation of the form

$$(1) \qquad \sum_{i,k=0}^{n} b_{ik}\, \xi_i\, \xi_k = 0.$$

From here on, the subject of our consideration will be the set of points represented by such an equation. The totality of points whose coordinates $\xi_0, \xi_1, \ldots, \xi_n$ (in a fixed projective coordinate system) satisfy a (fixed) equation of the form (1) is called a **hypersurface of the second order** (in P_3 and P_2, a *surface of the second order* and *curve of the second order*, respectively). The equation itself is said to be *homogeneous of the second degree.*[1]

The $(n+1)$-by-$(n+1)$ matrix (b_{ik}), formed from the coefficients of (1), is called the matrix of the hypersurface. It is easy to see that we can always take this matrix to be symmetric. For if we use the given b_{ik} to define $(n+1)^2$ new numbers a_{ik} by means of the equations[2]

[1] This name arises from the fact that only terms of the second degree appear in the equation—either the square of a *single* variable or the product of *two* variables. More generally, a function $f(\xi_0, \xi_1, \ldots, \xi_n)$ of $n+1$ variables is said to be *homogeneous of degree* k if, for arbitrary λ and ξ_i, $f(\lambda\xi_0, \lambda\xi_1, \cdots, \lambda\xi_n)$ always equals $\lambda^k f(\xi_0, \xi_1, \cdots, \xi_n)$. It is a consequence of homogeneity that the equation

$$f(\xi_0, \xi_1, \cdots, \xi_n) = 0$$

is either satisfied, or not satisfied, by all the coordinate vectors of a point simultaneously.

[2] In particular, $a_{ii} = b_{ii}$ for $i = 0, 1, \ldots, n$.

(2) $$a_{ik} = \frac{b_{ik} + b_{ki}}{2}, \qquad i, k = 0, 1, \cdots, n,$$

then, for arbitrary ξ_i, we have

$$\sum_{i,k=0}^{n} b_{ik}\, \xi_i\, \xi_k = \sum_{i,k=0}^{n} a_{ik}\, \xi_i\, \xi_k,$$

and hence the equation

(3) $$\sum_{i,k=0}^{n} a_{ik}\, \xi_i\, \xi_k = 0$$

represents exactly the same hypersurface as (1). But for the a_{ik} defined in this way it is indeed true that $a_{ik} = a_{ki}$ for every pair of subscripts i, k.

Let us agree that in future the matrix of a hypersurface of the second order *will always be assumed to be symmetric* without our making special mention of this fact in every particular case.

It is often advantageous to express equation (3) in matrix form. If we set $A = (a_{ik})$ and let (ξ) denote a coordinate matrix, the equation

(4) $$(\xi)' \cdot A \cdot (\xi) = O$$

represents the same hypersurface as (3). The condition $a_{ik} = a_{ki}$ can then be expressed in matrix form as $A = A'$.

How is the equation of a hypersurface of the second order altered under a transformation of coordinates? If we express our result in the form (4), it is easy to see what happens. If we let ξ_i and ξ_i^* be the coordinates of a point in two different projective coordinate systems, we know that a relation of the form

$$(\xi) = \varrho\, T(\xi^*)$$

must exist, where $T = (t_{ik})$ is a certain non-singular matrix. By substituting this expression in (4), we see at once that:

If (4) is the equation of a fixed hypersurface of the second order in a given projective coordinate system, then for any other coordinate system a non-singular matrix T can be found such that the equation

$$(5) \qquad (\xi^*)' \, T' A \, T (\xi^*) = O$$

represents the same hypersurface in that other coordinate system.[3]

Intersection with a Line

The first question to which we address ourselves is this: How many points of intersection does a line have with a hypersurface of second order? Let us think of the line as determined by two of its points, whose coordinate matrices in the initial coordinate system are (η) and (ζ). Then the coordinate matrix of an arbitrary point of the line has (by Chap. II, Theorem 3) the form $\lambda\,(\eta) + \mu\,(\zeta)$ for suitable values of λ, μ. Our question then amounts to the determination of those λ and μ for which the matrix $\lambda\,(\eta) + \mu\,(\zeta)$ satisfies equation (4). If we replace (ξ) in (4) by the matrix $\lambda\,(\eta) + \mu\,(\zeta)$, we obtain the following condition on λ and μ:

$$\lambda^2\,(\eta)'\,A\,(\eta) + \lambda \cdot \mu\,[(\eta)'\,A\,(\zeta) + (\zeta)'\,A\,(\eta)] + \mu^2\,(\zeta)'\,A\,(\zeta) = O.$$

The matrix $(\eta)'A(\zeta)$ which appears here is symmetric, because it contains only zeros except for the first entry in the first row (which is equal to the second sum in equation (7) below). Since A is moreover assumed symmetric, it follows that $(\eta)'\,A\,(\zeta) = ((\eta)'\,A\,(\zeta))' = (\zeta)'\,A\,(\eta)$. Hence, the last equation can also be written in the form

$$(6) \qquad \lambda^2\,(\eta)'\,A\,(\eta) + 2\,\lambda\,\mu\,(\eta)'\,A\,(\zeta) + \mu^2\,(\zeta)'\,A\,(\zeta) = O.$$

To find the points of intersection it suffices to determine the numbers λ, μ only up to a common factor or, in other words, to determine the ratio

---

[3] We cannot say here that every equation of the hypersurface in this coordinate system differs from (5) only by a multiplicative constant. For, in real P_n it is not in general true that the equation of a hypersurface of the second order is determined by this up to a multiplicative constant. For example, the equations $\xi_0^2 + \xi_1^2 = 0$ and $\xi_0^2 + 2\,\xi_1^2 = 0$ represent the same set of points in *real* P_2 (namely, the single point $\xi_0 = \xi_1 = 0$, $\xi_2 = 1$). Since we have no immediate need for a discussion of this question, we shall postpone it to the chapter after next, where it can be more easily handled. Cf. Chap. X, Theorem 5.

of the two numbers. Let us first assume that neither of the two points (η), (ζ) lies on the hypersurface. Then neither λ nor μ can vanish at a point of intersection, and λ/μ is a well-defined real or complex number $(\neq 0)$. Furthermore, since $(\eta)'A(\eta) \neq O$ and also $(\zeta)'A(\zeta) \neq O$,[4] (6) represents in the present case a quadratic equation in λ/μ; for (6) is equivalent[5] to

$$(7)\quad \left(\frac{\lambda}{\mu}\right)^2 \cdot \left(\sum_{i,k=0}^{n} a_{ik}\,\eta_i\,\eta_k\right) + 2\cdot\frac{\lambda}{\mu}\cdot\left(\sum_{i,k=0}^{n} a_{ik}\,\eta_i\,\zeta_k\right) + \sum_{i,k=0}^{n} a_{ik}\,\zeta_i\,\zeta_k = 0.$$

Since to every solution for λ/μ there corresponds a point of intersection, we see that:

If we consider the hypersurface (3) *in complex* P_n, *then it has in general two points of intersection with our line* (*which goes through the two points* (η), (ζ), *both assumed not to lie on the hypersurface*). *It can also happen, however, that there is only one point of intersection, which occurs when equation* (7) *has a double root* (*i.e., a solution of multiplicity* 2); *in this latter case, it is customary to speak of two coincident points of intersection.*

If, on the other hand, we consider (3) *in real* P_n—in which case the coefficients of (3) as well as those of (7) must be assumed to be real—*only one additional case can occur, namely that there exists no point of intersection, which happens when* (7) *has no real solution for* λ/μ.

There still remains to be considered the possibility that at least one of the two points (η), (ζ) lies on the hypersurface. Suppose this is (η), say; then $(\eta)'A(\eta) = 0$. We distinguish four cases, depending on whether or not the other two coefficients in (6) vanish.

I. $(\eta)'A(\zeta) = (\zeta)'A(\zeta) = O$. Then *every* pair of values for λ, μ satisfy equation (6), i.e., the entire line lies in the hypersurface.

II. $(\eta)'A(\zeta) = O$; $(\zeta)'A(\zeta) \neq O$. Then the equation is satisfied only for $\mu = 0$. That is to say, we now have the single point of intersection (counted twice), $\lambda = 1$, $\mu = 0$.

III. $(\eta)'A(\zeta) \neq O$; $(\zeta)'A(\zeta) = O$. There are exactly two points of intersection, namely, $\lambda = 1$, $\mu = 0$ and $\lambda = 0$, $\mu = 1$.

IV. $(\eta)'A(\zeta) \neq O$; $(\zeta)'A(\zeta) \neq O$. Again there are two points of intersection, namely, $\lambda = 1$, $\mu = 0$ and $\lambda = (\zeta)'A(\zeta)$, $\mu = -2(\eta)'A(\zeta)$.

---

[4] Because the points (η), (ζ) do *not* belong to the hypersurface (4).

[5] The coefficients of equation (7) are the first elements of the first row of the corresponding matrices of (6). The remaining elements of the matrices of (6) consist only of zeros.

In all four cases the results are valid for the real case as well as for the complex case. Let us put these results in tabular form for future reference.

POINTS OF INTERSECTION OF THE LINE THROUGH (η), (ζ) WITH THE HYPERSURFACE $(\xi)'A(\xi) = 0$ UNDER THE ASSUMPTION $(\eta)'A(\eta) = 0$.

	Value of $(\eta)'A(\zeta)$	Value of $(\zeta)'A(\zeta)$	Number of points of intersection
I	0	0	The entire line lies in the hypersurface
II	0	$\neq 0$	One point of intersection (counted twice)
III and IV	$\neq 0$	arbitrary	Two distinct points of intersection

The Tangents to a Hypersurface of the Second Order at a Point

If a line either has just one point in common with a hypersurface of the second order or lies entirely in the hypersurface, the line is called a **tangent** to the hypersurface. In this case, the points of intersection are also referred to as *points of tangency.* We now wish to determine all the tangents that pass through a given point of a hypersurface of the second order.

Let the hypersurface again be given by (4) and let (η) denote the fixed point on the hypersurface. Our problem can then be phrased as follows: To find all the points (ζ), distinct from (η), for which the line joining (η) and (ζ) is a tangent to the hypersurface. Since reference to the above table shows that tangents exist only in cases I and II and that these cases are characterized by the equation

$$(8) \qquad (\eta)'A(\zeta) = 0,$$

we see that the desired points are simply the solutions (ζ) of equation (8).

With a fixed (η), condition (8) amounts to a homogeneous linear equation in $\zeta_0, \zeta_1, \cdots, \zeta_n$. If not all the coefficients[6] of this equation vanish, then (8) defines a hyperplane. Since this hyperplane contains all[7] the tangents through the point (η), it is called the **tangent hyperplane** *at the point* (η).

[6] That is, the elements of the matrix $(\eta)'\cdot A$.

[7] If we were working in P_1 and $(\eta)'A \neq 0$, equation (8) could only be satisfied by a single point, namely $(\zeta) = (\eta)$. In this case, there is *no* tangent at the point (η). In P_2, assuming $(\eta)'A \neq 0$, (8) represents a line, actually a tangent line, the only one through the point (η). If $n > 2$ there is always an infinite number of tangents at each point of a hypersurface of the second order.

Can it happen that all the coefficients of equation (8), i.e., all the elements of the matrix $(\eta)'A$, vanish? This would mean that

$$\sum_{i=0}^{n} a_{ik}\,\eta_i = 0 \tag{9}$$

for $k = 0, 1, \ldots, n$. Since the η_i, being coordinates of a point of P_n, cannot all be zero, (9) can only be true if the determinant $|A| = |a_{ik}| = 0$. But if, conversely, $|A| = 0$, there always exists a non-trivial solution $\eta_0, \eta_1, \ldots, \eta_n$ of (9). Moreover, since it always follows from $(\eta)'A = O$ that $(\eta)'A(\eta) = O$ as well, every point (η) that satisfies (9) is also a point of the hypersurface (4).

The vanishing of all the coefficients of equation (8) means that every line through (η) is a tangent to the hypersurface. A point (η) of this kind is called a *double point* of the hypersurface. Thus, $|A| = 0$ is the necessary and sufficient condition for the existence of double points.

The most important of these results are summed up in the following theorems.

THEOREM 1. *If (η) is a point of the hypersurface $(\xi)'A(\xi) = O$, but not a double point, then the equation*[8]

$$(\eta)' \cdot A \cdot (\xi) = O$$

represents the (uniquely determined) tangent hyperplane at the point (η).

THEOREM 2. *The hypersurface $(\xi)'A(\xi) = O$ has double points (i.e., points at which the tangent plane is not uniquely determined) if and only if $|A| = 0$. The double points are then given by the solutions of equation (9), and so fill out a linear space of dimension $n - r$, where r is the rank of A.*

If $|A| = 0$, i.e., if double points exist, the hypersurface is called *degenerate*; otherwise, *non-degenerate.*

Theorem 2 shows that the rank of the matrix A is independent of the coordinate system as well as of the choice of the equation of the hypersurface. Hence we can call the rank of A the *rank of the hypersurface* defined by $(\xi)'A(\xi) = O$.

---

[8] We again denote the variable coordinates by ξ_i instead of ζ_i.

Tangents to a Hypersurface of the Second Order from a Point Exterior to the Hypersurface

Let us now start with a point with coordinate matrix (ζ) which does *not* lie on the hypersurface (4). We are thus assuming that $(\zeta)'A(\zeta) \neq O$. We again wish to determine all the tangents that go through (ζ). None of these tangents can lie entirely in the hypersurface (since (ζ) does not). Consequently each of them has but *one* point of tangency with the hypersurface, and our problem will be solved if we can specify all these points of tangency.

Therefore let us consider, in addition to (ζ), one other point of the hypersurface (4). Let its coordinate matrix be (η). Under what circumstances will the line through (η) and (ζ) be a tangent line? A glance at the table on p. 110 shows at once, since $(\zeta)'A(\zeta) \neq O$, that this can only occur in Case II. The necessary and sufficient condition for this case is given by the equation

$$(\eta)'A(\zeta) = O. \tag{10}$$

Equation (10), which, by virtue of $(\eta)'A(\zeta) = (\zeta)'A(\eta)$, we can also write in the form

$$(\zeta)'A(\eta) = O \tag{11}$$

can not in the present case hold identically in the η_i, since, for example, it is not satisfied, by assumption, for $(\eta) = (\zeta)$. Thus, equation (11), or (10), always represents a hyperplane. The desired points of tangency are the points of intersection of this hyperplane with the hypersurface (4).

If, for the sake of uniformity, we again write ξ_i for the variable coordinates η_i in (11), we can state the following:

THEOREM 3. *If (ζ) is a fixed point not on the hypersurface $(\xi)'A(\xi) = O$, the points of tangency of all the tangents through (ζ) are given by the points of intersection of the hyperplane*

$$(\zeta)'A(\xi) = O \tag{12}$$

with the hypersurface $(\xi)'A(\xi) = O$.

It should be clearly noted that such points of intersection need *not* always exist in the case of the *reals*. If no such points exist, then there can exist no tangents through (ζ). In the *complex* case, on the other hand, such points of intersection and corresponding tangents always exist.[9]

[9] For in the complex domain every line which lies in the hyperplane, for example, has points in common with the hypersurface.

The Polar

The hyperplane (12) has still other geometrical properties. To discover them, let us assume in what follows that (ζ) is *not* a point of the hypersurface (4) and consider those points of the hyperplane $(\zeta)'A(\xi) = O$ that, like (ζ), do *not* lie on the hypersurface (4). Let (η) be such a point of the hyperplane. The line[1] through the point (η) and the fixed point (ζ) will have either no points of intersection with the hypersurface (4) or *exactly two*. In the first case, which can only occur in real P_n, there is nothing more to say. In the second case, however, we can take the coordinate matrix of the point of intersection to be of the form $\lambda(\eta) + \mu(\zeta)$. Then, for a point of intersection, the λ and μ must satisfy equation (6). The middle term in this equation, however, now vanishes.[2] Since by assumption, $(\eta)'A(\eta) \neq O$ and $(\zeta)'A(\zeta) \neq O$, and hence neither λ nor μ can be zero, it follows that for a point of intersection

$$\frac{\lambda}{\mu} = \pm\sqrt{-\frac{(\zeta)'A(\zeta)}{(\eta)'A(\eta)}}.$$

From this we see that if $\lambda(\eta) + \mu(\zeta)$ is one of the points of intersection, then $\lambda(\eta) - \mu(\zeta)$ is the other. The cross ratio of the four points (η), (ζ), $\lambda(\eta) + \mu(\zeta)$, $\lambda(\eta) - \mu(\zeta)$ is equal to -1, i.e., they form a harmonic set. Thus we have Theorem 4.

Theorem 4. *If (ζ) is a point not on the hypersurface $(\xi)'A(\xi) = O$ and if g is a line through (ζ) which is not tangent to the hypersurface (but does intersect it), then the point (ζ) and the point of intersection of g with the hyperplane $(\zeta)'A(\xi) = O$, on the one hand, and the points of intersection of g with the hypersurface $(\xi)'A(\xi) = O$, on the other hand, separate each other harmonically.*

Because of its remarkable properties, the hyperplane $(\zeta)'A(\xi) = O$ is given a special name; it is called the **polar** *of the point (ζ) with respect to the hypersurface $(\xi)'A(\xi) = O$.* This designation is also extended to the case in which (ζ) does lie on the hypersurface, but $(\zeta)'A \neq O$. The polar is then identical (by Theorem 1) to the tangent hyperplane. The point (ζ) is called the *pole* of the hyperplane $(\zeta)'A(\xi) = O$.

---

[1] Because $(\zeta)'A(\zeta) \neq O$, (ζ) does not belong to the hyperplane $(\zeta)'A(\xi) = O$. Thus, (ζ) and (η) are distinct and determine a line.

[2] (η) was a point of the hyperplane $(\zeta)'A(\xi) = O$. Hence,

$$(\eta)'A(\zeta) = (\zeta)'A(\eta) = O.$$

The resemblance between the terms 'pole' and 'polar,' and the term 'polarity' by which we denoted the correlations at the end of Chapter VII, is not accidental. For, if the hypersurface (4) is non-degenerate, then *every* point of P_n has a well-defined polar, and, in fact, the hyperplane coordinates $u_0, u_1, \ldots, u_n$ of the polar of a point (ζ) are given by

$$(u) = \varrho \cdot A \cdot (\zeta). \tag{13}$$

From the preceding chapter, we know that for $|A| \neq 0$ this equation represents a linear correlation, and moreover, since A is symmetric, it represents a *polarity*. The hypersurface (4) is then precisely the locus of all the points that lie in their own image hyperplane under the polarity (13) (cf. Chap. VII, p. 101 f.).

Thus, for a non-degenerate hypersurface of the second order, the relation 'pole $\rightleftarrows$ polar' satisfies, in particular, the defining properties of a linear correlation: The relation is one to one and leaves all cross ratios invariant. Further (by Chap. VII, Theorem 1), it associates a k-dimensional linear bundle with a k-dimensional linear space, and conversely. Since, as a polarity, it is also an involutory correlation in the sense of Chapter VII, it always takes a hyperplane into a hyperbundle whose kernel is precisely the pole of the hyperplane.

Dualization

The *dual* of a hypersurface of the second order is called a **hypersurface of the second class**. In accord with the duality principle, this term stands for the totality of all hyperplanes[3] whose coordinates $u_0, u_1, \ldots, u_n$ satisfy an equation of the form

$$\sum_{i,k=0}^{n} a_{ik}\, u_i\, u_k = 0 \tag{14}$$

or

$$(u)'\, A\, (u) = O, \tag{15}$$

[3] Despite the name hyper-'surface,' this dual concept is of course not that of a set of points, but of a set of hyperplanes.

where $A = (a_{ik})$ is a symmetric matrix. The dual to a tangent is a pencil of hyperplanes, called the *tangent pencil,* which have either just one hyperplane in common with the hypersurface of the second class or else belong to it entirely. Thus, if (u) is a fixed hyperplane of the hypersurface of the second class, all the tangent pencils through (u) fill out either the entire space (i.e., *every* pencil through (u) is a tangent pencil) or one single hyperbundle. In the first case, (u) is called a *double hyperplane*; in the second case, the kernel of the hyperbundle, which is of course a single point, is called a *tangent point.* Theorems 1 and 2 then hold true under this dual interpretation. In like manner, the concepts degenerate and non-degenerate hypersurface, pole, and polar can be dualized, together with the theorems concerning them.

The totality of the tangent hyperplanes to a non-degenerate hypersurface of the second order constitute a hypersurface of the second class. For it follows from $(\xi)'A(\xi) = O$ $(|A| \neq 0)$ that

$$(\xi)' A A^{-1} A (\xi) = O. \tag{16}$$

If we now let $u_0, u_1, \ldots, u_n$ denote the hyperplane coordinates of the tangent hyperplane to the hypersurface $(\xi)'A(\xi) = O$ at (ξ), we can, by Theorem 1, set $(u) = A(\xi)$, or $(u)' = (\xi)'A$. If we substitute this in (16), it follows that the equation

$$(u)' A^{-1} (u) = O \tag{17}$$

is necessarily satisfied by the coordinates of each one of the tangent hyperplanes. Conversely, if (17) is satisfied for some hyperplane (u), there exists, since $|A| \neq 0$, one and only one coordinate matrix (ξ) satisfying the relation $(u) = A(\xi)$. This (ξ) then also satisfies (16), i.e., represents a point of the hypersurface $(\xi)'A(\xi) = O$ of which (u) is a tangent hyperplane.

The dual argument shows that the equation $(\xi)'A(\xi) = O$ yields the totality of the tangent *points* of the hypersurface of the second class given by (17). We thus have the following theorem.

Theorem 5. *If $|A| \neq 0$, then the hypersurface of the second order given by $(\xi)'A(\xi) = O$ and the hypersurface of the second class defined by $(u)'A^{-1}(u) = O$ are related as follows: The second is exactly the totality of the tangent hyperplanes of the first, and the first is exactly the totality of tangent points of the second.*

Thus far we have not bothered to inquire if every equation of the form (3), i.e., $\sum_{i,k=0}^{n} a_{ik}\,\xi_i\,\xi_k = 0$, actually represents a set of points—in other words, if there always exist points that satisfy a given equation of the second degree. In *complex* P_n, to be sure, this will always be the case. For here, as we have seen, there is on *every* line at least one point whose coordinates satisfy the equation. In *real* P_n there will always be such points if $|a_{ik}| = 0$. For in this case there exist even special points of the hypersurface, namely the double points.

On the other hand, if equation (3), thought of as over the reals, has determinant $|a_{ik}| \neq 0$, there do not always necessarily exist points of real P_n that satisfy it. For example, the equation

$$\xi_0^2 + \xi_1^2 + \cdots + \xi_n^2 = 0$$

has $\xi_0 = \xi_1 = \cdots = \xi_n = 0$ as its only real solution, and this does not represent any point of P_n. However, it is neither customary nor desirable to bar this case from consideration. In the first place, every such hypersurface, while empty in the reals, has a non-empty complex part. In the second place, moreover, many theorems on hypersurfaces of the second order retain a perfectly sound meaning in this case, even when considered solely in the reals. For example, every point of real P_n still has a real polar,[4] and the correspondence 'pole $\rightleftarrows$ polar' is, in this case also, a correlation of real P_n.

If a real equation (3) has no real point as solution, it is said to represent an *imaginary* hypersurface of the second order. The dual expression is imaginary hypersurface of the second class.

Exercises

1. What is the meaning of the concepts hypersurface of the second order, polar, and polarity in P_1?

2. Show that a hypersurface of the second order in P_n which lies entirely in a hyperplane, consists solely of double points. From this it also follows, among other things, that on a non-imaginary non-degenerate hypersurface of the second order in P_n it is always possible to find $n+1$ linearly independent points.

3. Let there be given in P_n a hypersurface of the second order of rank $r > 1$ and a point P *not* on the hypersurface. Show that the tangents (= lines) to the hypersurface of the second order from P constitute a hypersurface of the second order of rank $r - 1$.

---

[4] The definition of the polar of the point (ζ) with respect to the hypersurface (4) has been given in a purely algebraic form as $(\zeta)'A(\xi) = O$.

CHAPTER IX

PROJECTIVE CLASSIFICATION OF HYPERSURFACES OF THE SECOND ORDER

Statement of the Problem

If a linear projectivity[1] of P_n is applied to a hypersurface of the second order, the totality of image points of all the points of the hypersurface itself constitutes a hypersurface of the second order. To see this, let us think of the hypersurface of the second order as given by the equation

$$(\xi)' A (\xi) = O, \qquad A' = A, \tag{1}$$

and the linear projectivity, as in Chapter VI, by

$$(\xi) = \varrho T (\xi^*). \tag{2}$$

If we eliminate (ξ) between (1) and (2), we obtain

$$(\xi^*)' T' A T (\xi^*) = O \tag{3}$$

as the equation for the image points. And this is indeed the equation of a hypersurface of the second order.

If a hypersurface F_1 can be transformed into a hypersurface F_2 by a linear collineation, then there also exists a linear projectivity, namely the inverse[2] of the given one, which transforms F_2 into F_1; hence we may make the following definition: Two hypersurfaces of the second order which can be transformed into each other by suitable linear collineations will be said to be **projectively equivalent**. Since this definition tells us

[1] A projectivity of the kind considered in Chapter VI is meant, mapping P_n onto itself.

[2] Here, and in what follows, the terms 'inverse' and 'product' are used in the general group-theoretical sense, as stipulated in § 19 of *Modern Algebra*.

nothing about the imaginary hypersurfaces of real P_n, we shall agree, furthermore, to call *any* two imaginary hypersurfaces of real P_n projectively equivalent.[3]

Every hypersurface of the second order is projectively equivalent to itself, since, for example, it is transformed into itself by the identity mapping of P_n. The most important property of the concept 'projective equivalence' is its *transitivity,* i.e.:

If F_1, F_2, F_3 are three hypersurfaces of the second order and if F_1 is projectively equivalent to F_2, and F_2 to F_3, then F_1 is projectively equivalent to F_3.

For the real imaginary hypersurfaces, this is true by definition. That it is true for the others follows immediately from the fact that linear collineations constitute a group. For if $\varkappa$ is a linear projectivity that takes F_1 into F_2 and $\varkappa'$ a second linear projectivity that takes F_2 into F_3, then the product $\varkappa' \cdot \varkappa$ (i.e., the collineation that results from first applying $\varkappa$ and then applying $\varkappa'$) takes F_1 into F_3.

A fundamental problem is to find necessary and sufficient conditions that two hypersurfaces of the second order be projectively equivalent. The solution of this problem will be the subject of the present chapter.

The problem as stated may be given another and somewhat broader formulation. This formulation is arrived at as follows. The transitivity of the relation of 'projeetive equivalence' has as a consequence a partition of the class of all hypersurfaces. More specifically, we define a *class of projectively equivalent hypersurfaces* to be a (non-empty) set of hypersurfaces of the second order having the following two properties:

1. Every two hypersurfaces of the class are projectively equivalent.
2. All hypersurfaces projectively equivalent to any one of the hypersurfaces of the class also belong to the class.

Property 2. can also be expressed as follows:

2*. A hypersurface of the class is never projectively equivalent to a hypersurface that does not belong to the class.

An *example* of such a class is the *set of all hypersurfaces of the second order projectively equivalent to a fixed hypersurface of the second order.* Properties 1. and 2. hold for this example merely by virtue of the transitivity of projective equivalence.

[3] A real non-empty hypersurface, of course, cannot be transformed into an imaginary surface by a real collineation. Cf. also p. 127, footnote 9.

This example immediately shows that *every hypersurface of the second order belongs to at least one class.* However, it is also true, conversely, that *every hypersurface of the second order appears in only one class.* Or to express it another way: *Two different classes always have an empty intersection.* For if $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are two classes with a common element F, then by property 1., every hypersurface of $\mathfrak{c}_2$ is projectively equivalent to F, and hence, since F also belongs to $\mathfrak{c}_1$, every element of $\mathfrak{c}_2$, by property 2., is contained in $\mathfrak{c}_1$. But the same result obtains if we reverse the roles of $\mathfrak{c}_1$ and $\mathfrak{c}_2$, so that $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are necessarily identical.

We see, then, that the totality of all hypersurfaces is so partitioned among our classes that each hypersurface occurs in one and only one class. Such a partitioning into classes is called *disjoint.* The problem posed above may now be incorporated into the following somewhat more ambitious one: To determine the classes of projectively equivalent hypersurfaces of the second order, and to discover simple criteria for deciding to which class any given hypersurface belongs. By determining the classes we mean giving a *complete system of representatives,* i.e., a system of hypersurfaces of the second order containing one and only one representative from each class. It will become apparent that this determination of classes is the crucial part of our task, from which the rest follows easily.

The connection of our problem with transformations of coordinates, though not necessary for our later work, is in itself both interesting and important. The conjecture that such a connection exists is suggested by the fact that the same algebraic relation exists between equations (1) and (3), representing projectively equivalent hypersurfaces of the second order, as between equations (4) and (5) of Chap. VIII (p. 107f.), having to do with a transformation of coordinates. To be sure, we do not yet know at this point whether two equations that represent projectively equivalent hypersurfaces of the second order in the same coordinate system must necessarily always be related in the same way as (1) and (3); for we do not yet know all the forms that are possible for an equation of a hypersurface of the second order in a fixed coordinate system. (See footnote 3 (p. 108) of Chapter VIII and Theorem 5 of Chapter X below.) Nevertheless, we are already in a position to state the following result:

Two equations $(\xi)'A(\xi) = 0$, $(\eta)'B(\eta) = 0$ referred to the same projective coordinate system represent projectively equivalent hypersurfaces of the second order if and only if they can also be looked upon as the equations of a single hypersurface referred to two (in general, different) coordinate systems.

Proof: Let F_1, F_2 designate the hypersurfaces represented, respectively, by the equations $(\xi)'A(\xi) = O$, $(\eta)'B(\eta) = O$ in the coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$. Assume they are projectively equivalent, and let $\varkappa$ be a fixed linear collineation that takes F_1 into F_2. Let the image points of Q_i, E under $\varkappa$ be Q_i^*, E^*. From Chap. V, Theorem 6, an image point Q^* in the coordinate system $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$ always has the same coordinates as its original Q in $(Q_0, Q_1, \ldots, Q_n \mid E)$. It follows at once from this that F_2, the image hypersurface of F_1, will be given by every equation in $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$ that represents F_1 in $(Q_0, Q_1, \ldots, Q_n \mid E)$ and thus, in particular, by $(\xi)'A(\xi) = O$. Thus, each of our two equations proves to be an equation of the *one* hypersurface F_2; in other words, the 'only if' part of our assertion is proved.

Conversely, if we think of $(\xi)'A(\xi) = O$ and $(\eta)'B(\eta) = O$ as given equations for F_2 in the coordinate systems $(Q_0^*, Q_1^*, \ldots, Q_n^* \mid E^*)$ and $(Q_0, Q_1, \ldots, Q_n \mid E)$ respectively, then the collineation $\varkappa$ must be determined as that linear projectivity which takes the points Q_i, E into Q_i^*, E^*, respectively. The *inverse* collineation, applied to F_2, yields an image hypersurface F_1; we show, as above, that this has the equation $(\xi)'A(\xi) = O$ in the coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$. Thus, the 'if' of our theorem is also proved.

Normal Forms. Complete System of Invariants

As a first step in solving our problem we must think of what simplification can be effected on the equation of a hypersurface of the second order by passing from form (1) to form (3). Since this auxiliary investigation does not pertain directly to hypersurfaces of the second order, but is a purely algebraic assertion about the behaviour of the expression on the left-hand side of the equation for such a hypersurface, let us not speak of hypersurfaces of the second order for a while but only of expressions of the type

$$\sum_{i,k=0}^{n} a_{ik}\, \xi_i\, \xi_k, \tag{4}$$

where $A = (a_{ik})$ is a fixed symmetric matrix $\neq O$ and $\xi_0, \xi_1, \ldots, \xi_n$ are variables. We may take the a_{ik} to be elements of an arbitrary field. Such an expression is called a *quadratic form* in the variables ξ_i.

In analogy with the way (3) was obtained from (1), we make a linear substitution in (4) of the form

$$\xi_i = \sum_{k=0}^{n} t_{ik}\, \xi_k^*, \qquad i = 0, 1, \cdots, n, \tag{5}$$

where the $\xi_0^*, \xi_1^*, \ldots, \xi_n^*$ are new variables and the t_{ik} are again elements of the ground field. We shall always assume that the matrix $T = (t_{ik})$ is non-singular and we accordingly call (5), as usual, a *non-singular* linear substitution.

The successive application of two non-singular substitutions can, of course, be replaced by a *single* such substitution.[4]

It can happen that all the square terms in (4) vanish, i.e., that $a_{ii} = 0$ for all $i = 0, 1, \ldots, n$. By assumption, however, not all a_{ik} with $i \neq k$ can also be 0. Let, say, $a_{01} \neq 0$. Then the non-singular substitution

$$\begin{aligned} \xi_0 &= \xi_0^* + \xi_1^*, \\ \xi_1 &= \xi_0^* - \xi_1^*, \\ \xi_i &= \xi_i^* \text{ for } i = 2, 3, \cdots, n \end{aligned} \tag{6}$$

transforms (4) into a new equation in which the terms $2\, a_{01}\, \xi_0^{*2}$ and $-2a_{01}\, \xi_1^{*2}$ occur, so that the new form actually does contain square terms.

We shall now prove the theorem most important for our purposes:

THEOREM 1. *A quadratic form* (4) *can always be transformed by an appropriate non-singular linear substitution* (5) *into the form*

$$c_0 \xi_0^{*2} + c_1 \xi_1^{*2} + \cdots + c_n \xi_n^{*2} \tag{7}$$

containing square terms only.

The proof is by induction on the number of variables. The theorem is true for one variable ξ_0. We assume it proved for n variables and show that it holds for $n + 1$ variables. By what was said above, we can certainly assume there is at least one square term having a coefficient different from zero. Let $a_{00} \neq 0$, say. Then we can write (4) in the form

$$a_{00} \left(\xi_0 + \frac{a_{01}}{a_{00}}\, \xi_1 + \cdots + \frac{a_{0n}}{a_{00}}\, \xi_n \right)^2 + \sum_{i,k=1}^{n} b_{ik}\, \xi_i\, \xi_k \tag{8}$$

---

[4] See for example *Modern Algebra,* § 19, equation (9).

where, provided all the b_{ik} do not vanish, $\sum_{i,k=1}^{n} b_{ik}\, \xi_i\, \xi_k$ is a certain quadratic form in *only* n variables, $\xi_1, \xi_2, \ldots, \xi_n$. The non-singular substitution

$$\xi_0 = \eta_0 - \frac{a_{01}}{a_{00}}\,\eta_1 - \frac{a_{02}}{a_{00}}\,\eta_2 - \cdots - \frac{a_{0n}}{a_{00}}\,\eta_n,$$
$$\xi_i = \eta_i \text{ for } i = 1, 2, \cdots, n$$

takes (8) into the form

$$a_{00}\,\eta_0^2 + \sum_{i,k=1}^{n} b_{ik}\,\eta_i\,\eta_k. \tag{9}$$

If all the coefficients of the second part, i.e., of

$$\sum_{i,k=1}^{n} b_{ik}\,\eta_i\,\eta_k,$$

vanish, then we are through. If not, then, according to the induction hypothesis, this part can be transformed by a suitable non-singular substitution, say by

$$\eta_i = \sum_{k=1}^{n} t_{ik}\,\xi_k^*, \qquad i = 1, 2, \cdots, n;\ |t_{ik}| \neq 0, \tag{10}$$

into the form

$$c_1\,\xi_1^{*2} + c_2\,\xi_2^{*2} + \cdots + c_n\,\xi_n^{*2}.$$

If the equation

$$\eta_0 = \xi_0^* \tag{11}$$

is adjoined to the equations in (10), it is seen that (10) and (11) together amount to a linear substitution with the matrix

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & t_{11} & t_{12} & \cdots & t_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix},$$

and that its determinant, because of the assumption in (10) that $|t_{ik}| \neq 0$, does not vanish. On the other hand, the substitution defined by (10) and (11) takes (9) into the form

$$a_{00}\,\xi_0^{*2} + c_1\,\xi_1^{*2} + \cdots + c_n\,\xi_n^{*2}.$$

Thus, albeit in a somewhat different notation, we have our desired result.

A form of the type

$$c_0 \xi_0^2 + c_1 \xi_1^2 + \cdots + c_n \xi_n^2 \tag{12}$$

can be still further simplified if we assume, as we shall from now on, that the coefficients of the forms and of the substitutions represent real or complex numbers. However, since the result is simpler in the field of complex numbers than in the field of real numbers, let us first consider the complex case. That is to say, the c_1 in (12) and the coefficients of the linear substitution to be used may be arbitrary complex numbers. Then we apply to (12) the substitution

$$\begin{aligned} \xi_i &= \frac{1}{\sqrt{c_i}}\, \xi_i^* \quad \text{for } c_1 \neq 0, \\ \xi_k &= \xi_k^* \qquad\quad \text{for all } k \text{ for which } c_k = 0. \end{aligned} \tag{13}$$

The matrix of this substitution has diagonal form, contains only non-zero elements along the main diagonal, and is thus non-singular. Upon application of (13), (12) takes on the form

$$d_0 \xi_0^{*2} + d_1 \xi_1^{*2} + \cdots + d_n \xi_n^{*2}, \tag{14}$$

where every d_i is either 0 or 1. Finally, by a further non-singular linear substitution, we can arrange that all the non-zero terms in (14) come first and all the zero terms last,[5] so that $d_0 = d_1 = \ldots = d_r = 1$, say, whereas $d_{r+1} = \ldots = d_n = 0$. We thus have Theorem 2.

THEOREM 2. *A complex quadratic form* (4) *can always be reduced to the form*

$$\xi_0^{*2} + \xi_1^{*2} + \cdots + \xi_r^{*2}$$

by a suitable complex non-singular substitution (5).

If the original form (12) is real and if only real coefficients are allowed in the desired substitution, then not quite as much can be achieved. Since $\sqrt{c_i}$ in (13) need not be real, we must replace (13) by

[5] To achieve this, it is merely necessary to rearrange the variables; thus, a non-singular substitution of the following form is indicated:

$$\xi_0 = \xi_{\nu_0}^*,\ \xi_1 = \xi_{\nu_1}^*,\ \ldots,\ \xi_n = \xi_{\nu_n}^*,$$

where all the numbers $0, 1, \ldots, n$ appear among the $\nu_0, \nu_1, \cdots, \nu_n$.

$$\xi_i = \frac{1}{\sqrt{c_i}} \xi_i^* \text{ for } c_i \neq 0, \tag{15}$$
$$\xi_k = \xi_k^* \qquad \text{for all } k \text{ for which } c_k = 0.$$

This substitution transforms (12) into a form (14) in which the d_i can take on only the values 0, 1, or -1. If we arrange the order of the terms so that all the terms with coefficient $+1$ come first, then those with -1, and lastly, those with coefficient 0, we have Theorem 3.

THEOREM 3. *A real quadratic form* (4) *can always be reduced to the form*

$$\xi_0^{*2} + \xi_1^{*2} + \cdots + \xi_k^{*2} - \xi_{k+1}^{*2} - \xi_{k+2}^{*2} - \cdots - \xi_r^{*2}$$

by a real non-singular substitution (5).

We now wish to consider how the results just proved for quadratic forms translate into results concerning our hypersurfaces of the second order. Since performing a non-singular linear substitution on the left-hand side of equation (1) amounts to nothing other than applying a linear projectivity to the hypersurface of the second order represented by (1), Theorem 2 translates at once into the following:

THEOREM 4. *Given a hypersurface of the second order in complex P_n, it is possible to find a projectively equivalent hypersurface of the second order having an equation*[6] *of the form*

$$\xi_0^2 + \xi_1^2 + \cdots + \xi_r^2 = 0. \tag{16}$$

Theorem 3 may be reinterpreted similarly. Here, however, it should be observed that if $k + h = r - 1$, then the hypersurfaces $\xi_0^2 + \xi_1^2 + \cdots + \xi_k^2 - \xi_{k+1}^2 - \cdots - \xi_r^2 = 0$ and $\xi_0^2 + \xi_1^2 + \cdots + \xi_h^2 - \xi_{h+1}^2 - \cdots - \xi_r^2 = 0$ are projectively equivalent. For, the second equation may be transformed into the first by multiplying through by -1 and then interchanging coordinates (which is a non-singular linear substitution). Thus, we need only consider the equations

$$\xi_0^2 + \xi_1^2 + \cdots + \xi_k^2 - \xi_{k+1}^2 - \cdots - \xi_r^2 = 0$$

in which $k + 1 \geqq r - k$. Thus we have Theorem 5.

[6] We have written ξ in place of ξ^*, since it is immaterial what symbols are used for the variables in the equation.

THEOREM 5. *In real P_n, given any hypersurface of the second order, it is possible to find a projectively equivalent hypersurface of the second order having an equation of the form*

$$\xi_0^2 + \xi_1^2 + \cdots + \xi_k^2 - \xi_{k+1}^2 - \cdots - \xi_r^2 = 0, \tag{17}$$

where $k + 1 \geqq r - k$.

This theorem is also true, of course, for the imaginary hypersurfaces of real P_n. For, by assumption, they are all projectively equivalent to the hypersurface $\xi_0^2 + \xi_1^2 + \cdots + \xi_n^2 = 0$.

We can now proceed directly to the solution of our problem. Suppose a hypersurface of the second order $(\xi)'A(\xi) = O$ is given which can be carried into the image hypersurface $(\xi^*)'B(\xi^*) = O$ by the linear collineation $(\xi) = \varrho \cdot T(\xi^*)$. Also, let g be a line and g^* the image line under the same collineation. Inasmuch as this collineation is a one-to-one mapping, the number of points of intersection of g with $(\xi)'A(\xi) = O$ must be exactly the same as those of g^* with $(\xi^*)'B(\xi^*) = O$; or, if g lies entirely in the first hypersurface, g^* must lie entirely in the second. And, in particular, if g is tangent to $(\xi)'A(\xi) = O$, then, and only then, will g^* be tangent to $(\xi^*)'B(\xi^*) = O$.

Further, a point P on $(\xi)'A(\xi) = O$ is or is not a double point according as all, or not all lines through P are tangent to the hypersurface. Since this latter property is invariant, the property appertaining to P of being a double point or of failing to be a double point must also be invariant under a linear collineation.

The totality of the double points of $(\xi)'A(\xi) = O$ will therefore be mapped precisely into the totality of all double points of $(\xi^*)'B(\xi) = O$ by the linear collineation $(\xi) = \varrho \cdot T(\xi^*)$. Both totalities are linear spaces which, by Chap. V, Theorem 1, have equal dimensionality. By Chap. VIII, Theorem 2, we thus have the following further result:

THEOREM 6. *If $(\xi)'A(\xi) = O$ and $(\xi^*)'B(\xi^*) = O$ are two projectively equivalent hypersurfaces of the second order, then the matrices A and B necessarily have the same rank.*[7]

The rank of a hypersurface of the second order[8] is thus an invariant with respect to linear collineations.

[7] This could also have been concluded from the fact that $(\xi)' A(\xi) = O$ and $(\xi^*)' B(\xi^*) = O$, referred to two suitable coordinate systems, can represent the same hypersurface. (cf. p. 119).

[8] In Chap. VIII we called the rank of matrix A the rank of the hypersurface $(\xi)'A(\xi) = O$.

At this point we are very nearly through as regards *complex* P_n. For here, by Theorem 4, every hypersurface of the second order is projectively equivalent to at least one of the $n + 1$ hypersurfaces (16). In other words, the classes represented by the hypersurfaces (16) already account for *all* the hypersurfaces of the second order. If we can just show that these $n + 1$ classes are disjoint, we shall have succeeded in determining all the classes, as we set out to do at the beginning of this chapter.

Now, by Theorem 6, two hypersurfaces will certainly belong to different classes if they have different rank. The hypersurfaces (16), however, all differ from each other in rank. To find the rank of any particular hypersurface of the form (16), we need only take a look at the matrix of equation (16); this has the following appearance:

$$\begin{pmatrix} 1 & & & & & & 0 \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & 0 \end{pmatrix} \quad (\text{the first } r+1 \text{ diagonal entries } 1:\ r+1 \text{ times}).$$

The rank of (16) is therefore $r + 1$, so that the truth of our assertion is self-evident.

It is also clear by now that the converse of Theorem 6 is likewise true. For the fact that, on the one hand, the hypersurfaces (16) represent all the classes and, on the other hand, all have different rank means that two different classes necessarily consist of hypersurfaces of different rank. Expressed somewhat differently: If two hypersurfaces of the second order have equal rank, they belong to the same class.

We can summarize our results as follows.

THEOREM 7. *In complex P_n there are in all $n + 1$ classes of projectively equivalent hypersurfaces of the second order. The equations* (16) *constitute a complete system of representatives of these classes. Each class is characterized by the rank of its hypersurfaces.*

All the questions posed at the outset are now answered for complex P_n. Let us now turn to the real case. In this case we know first of all, from Theorem 5, that all the classes are represented by the hypersurfaces (17). The question still to be answered is: How often does each class occur among the hypersurfaces (17)? Now, it will turn out that here, too, each class is represented only once. To show this, we will show that two equations of type (17) can only belong to the same class if they agree both in rank $r + 1$ and in the number $k + 1$ of positive squares. From this it will follow immediately that two different hypersurfaces (17) always belong to different clases, because they can not simultaneously have equal indices r and k.

Since we already know that two projectively equivalent hypersurfaces (17) must always have the same rank (Theorem 6), it only remains to prove the same for the index k. This is the burden of the following theorem, which characterizes this index as an invariant (and, in fact, as an invariant in a double sense).

THEOREM 8. *In real P_n, the following holds for a hypersurface represented in a projective coordinate system by* (17): *k is the greatest integer for which there exists a k-dimensional linear space of P_n having no point in common with the hypersurface* (17). *Also, $s = n - k - 1$ is the greatest integer for which there exists an s-dimensional linear space lying entirely in the hypersurface* (17).[9]

Proof: The $n - k$ equations

$$\xi_{k+1} = 0, \qquad \xi_{k+2} = 0, \cdots, \xi_n = 0 \tag{18}$$

define a k-dimensional linear space L having no (real) point in common with hypersurface (17). For from (17) and (18) it follows that

$$\xi_0^2 + \xi_1^2 + \cdots + \xi_k^2 = 0,$$

which is satisfied for real ξ_i only if we also have

$$\xi_0 = \xi_1 = \cdots = \xi_k = 0.$$

But no point exists in P_n all of whose coordinates vanish.

On the other hand, an $(n - k - 1)$-dimensional space L' lying in (17) can easily be given. Of the $k + 1$ coordinates in (17) which appear with a plus sign, set the first $r - k$, say, equal to $\xi_{k+1}, \xi_{k+2}, \cdots, \xi_r$, respectively, and the rest equal to zero. The resulting system of equations

[9] Accordingly, there is only one imaginary hypersurface among the hypersurfaces (17). Therefore (by Theorem 3), any two imaginary hypersurfaces can be transformed into each other by a non-singular linear substitution. This gives an algebraic meaning to our convention of regarding two imaginary hypersurfaces as projectively equivalent.

$$(19) \qquad \begin{aligned} \xi_0 = \xi_{k+1},\ \xi_1 = \xi_{k+2},\ \cdots,\ \xi_{r-k-1} = \xi_r, \\ \xi_{r-k} = 0,\ \xi_{r-k+1} = 0,\ \cdots,\ \xi_k = 0, \end{aligned}$$

of rank $k + 1$, does indeed define an $(n - k - 1)$-dimensional linear space all of whose points satisfy (17).

By Chap. II, Theorem 5, every linear space of dimension $> k$ must have points in common with L', and hence also with (17). On the other hand, every linear space whose dimension is $> n - k - 1$ must have points in common with L and therefore cannot belong entirely to the hypersurface. This completes the proof of Theorem 8.

With this, we have now determined the classes in the real case also. And we have proved, at the same time, that the rank and the second invariant k characterize the classes, since all hypersurfaces which have same rank and the same k always form exactly one class. We sum up the results for the real case in Theorem 9.

THEOREM 9. *In real P_n there exist as many classes of projectively equivalent hypersurfaces of the second order as equations of the form* (17). *More specifically, these equations constitute a complete system of representatives. Two hypersurfaces belong to the same class if and only if they are identical both in rank and in the invariant k characterized by Theorem* 8.

The representatives (16) and (17) of the classes of projectively equivalent hypersurfaces are often called *projective normal forms* of the hypersurfaces of the second order. By virtue of their relation to transformations of coordinates, as discussed on p. 120, the following result holds for normal forms: For every hypersurface of the second order there can be found a coordinate system with reference to which the hypersurface is given by an equation in normal form.

Let us here introduce a new term, which we shall make use of later. By *'a complete system of invariants for a set of mathematical constructs with respect to a group of mappings'* is meant a system of invariants whose equality for two of the constructs in question is sufficient[1] to insure that they can be transformed into each other by a mapping of the group. Thus, in our case, we have found a *complete system of invariants for hypersurfaces of the second order with respect to the linear projective group*; in complex P_n this is the single invariant, rank; in real P_n the system consists of rank together with the second invariant, k.

[1] That it is necessary is part of the concept 'invariant.'

In conclusion, it should be remarked that behind the invariance of the index k that appears in the *real* normal forms there lies a somewhat more general statement. Specifically, if we return from hypersurfaces of the second order to quadratic forms and raise the question, in connection with Theorem 3, whether two of the 'normal forms' that occur in Theorem 3 can be transformed into each other by a real non-singular linear substitution, the result we obtain differs from our previous result as follows: Two quadratic forms $\xi_0^2+\xi_1^2+\cdots+\xi_k^2-\xi_{k+1}^2-\cdots-\xi_r^2$ and $\xi_0^2+\xi_1^2+\cdots+\xi_{r-k-1}^2-\xi_{r-k}^2-\cdots-\xi_r^2$ (the one containing as many positive terms as the other contains negative) can *not* be transformed into each other, even though these forms, when set equal to 0, represent projectively equivalent hypersurfaces of the second order.[2] This somewhat more general result is called the Jacobi-Sylvester *Law of Inertia.* Since we shall not need it in the sequel, its proof will be left to the Exercises (cf. Exercise 2).

Related Questions

Just as hypersurfaces of the second order may be classified by the use of the projective group, so may various other geometrical structures. To illustrate, let us carry out this classification for two simple examples.

As our first example, let us consider ordered *linear quadruples,* i.e., systems of four *distinct* points on a line taken in a definite order. In exact analogy to the foregoing, we are to find the conditions under which two linear quadruples are projectively equivalent, i.e., can be transformed into each other by means of a linear projectivity.

It follows at once from the invariance of the cross ratio under linear projectivities that two linear quadruples are projectively equivalent only if they have the same cross ratio. But it turns out that this is also sufficient.

Let S_1, S_2, S_3, S_4 and $S_1^*, S_2^*, S_3^*, S_4^*$ be two linear quadruples with $\mathcal{R}(S_1 S_2 S_3 S_4) = \mathcal{R}(S_1^* S_2^* S_3^* S_4^*)$. If we can find a linear projectivity that takes S_1 into S_1^*, S_2 into S_2^*, and S_3 into S_3^*, then S_4 will automatically be taken into S_4^* because of the equality of the cross ratios. For if the image point of S_4 be denoted at first by T, then $\mathcal{R}(S_1^* S_2^* S_3^* T) = \mathcal{R}(S_1^* S_2^* S_3^* S_4^*)$, from which it follows that $T = S_4^*$.

---

[2] This is due to the fact that the necessary change of signs of all of the terms of its equation, which is possible for the hypersurface, cannot be effected by a linear substitution.

But Theorem 5 of Chapter V readily enables us to construct a linear projectivity that maps S_1, S_2, S_3 into S_1^*, S_2^*, S_3^*, respectively. In order to apply this theorem, let us set up a coordinate system in which S_1, S_2 are two fundamental points and S_3 is the point of intersection of the line joining S_1 and S_2 and the hyperplane determined by the unit point and the remaining $n-1$ fundamental points. Correspondingly, let S_1^*, S_2^* be two fundamental points of a second coordinate system and S_3^* the point of intersection of the line joining S_1^* and S_2^* and the hyperplane determined by the remaining fundamental points and the unit point. As soon as we have found two such coordinate systems, we are through. For the uniquely determined linear projectivity which maps the first coordinate system onto the second (Theorem 5, Chap. V) takes the points S_1, S_2, S_3 into S_1^*, S_2^*, S_3^*.

It only remains to show the existence of two such coordinate systems. This is done as follows. Let $\mathfrak{x}_0$ be a coordinate vector of S_1 and $\mathfrak{x}_1$ a coordinate vector of S_2. We so normalize $\mathfrak{x}_0$ and $\mathfrak{x}_1$ that $\mathfrak{y} = \mathfrak{x}_0 + \mathfrak{x}_1$ will be a coordinate vector of S_3.[3] If we now adjoin $n-1$ other vectors $\mathfrak{x}_2, \mathfrak{x}_3, \ldots, \mathfrak{x}_n$ to $\mathfrak{x}_0$, $\mathfrak{x}_1$ in such a way that all the $\mathfrak{x}_i$ are linearly independent, then the coordinate system with the $\mathfrak{x}_i$ as fundamental points and $\mathfrak{x}_0 + \mathfrak{x}_1 + \cdots + \mathfrak{x}_n$ as its unit point has the desired property. For $\mathfrak{y}$ can also be represented as a linear combination of the unit point and the points $\mathfrak{x}_2, \mathfrak{x}_3, \ldots, \mathfrak{x}_n$:

$$\mathfrak{y} = (\mathfrak{x}_0 + \mathfrak{x}_1 + \cdots + \mathfrak{x}_n) - \mathfrak{x}_2 - \mathfrak{x}_3 - \cdots - \mathfrak{x}_n = \mathfrak{x}_0 + \mathfrak{x}_1,$$

that is, $\mathfrak{y}$ also lies in the hyperplane determined by the unit point and the points $\mathfrak{x}_2, \mathfrak{x}_3, \ldots, \mathfrak{x}_n$. The second coordinate system, for S_1^*, S_2^*, S_3^*, can be found in exactly the same way.

Thus we see that *the linear quadruples fall into precisely as many projectively equivalent classes as there are different values of the cross ratio. The cross ratio is an invariant characterizing the classes.*

As our second example, let us classify in the same way the *linear projectivities* themselves. To clarify what is to be meant by projective equivalence of linear projectivities, we must state what we mean by the application of one linear projectivity τ to another linear projectivity σ.[4]

[3] This is certainly possible, since $\mathfrak{x}_0$, $\mathfrak{x}_1$, $\mathfrak{y}$ are three *distinct* points of a line.

[4] This is not to be confused with multiplication of linear projectivities.

Now, by this we shall mean the following: To pass from the correspondence $P \to \sigma(P)$, representing the mapping σ, to the correspondence $\tau(P) \to \tau[\sigma(P)]$. In other words, τ is always applied simultaneously to a point P and to its image $\sigma(P)$.

Another preassigned linear projectivity σ^* being given, if τ can be so chosen that the correspondence $\tau(P) \to \tau[\sigma(P)]$ is identical to the correspondence $P \to \sigma^*(P)$ defined by σ^* (i.e., the pair of points $\tau(P)$ and $\tau[\sigma(P)]$ always represent an original point and its image point under σ^*), then we shall say that σ and σ^* are projectively equivalent. In terms of a formula:

$$\sigma^*[\tau(P)] = \tau[\sigma(P)].$$

As this must be true for every P, it follows that

$$\sigma^*\tau = \tau\sigma, \qquad \sigma^* = \tau\sigma\tau^{-1}.$$

The question, When are two linear projectivities σ and σ^* projectively equivalent? then amounts to asking, When does there exist a linear projectivity τ such that $\sigma^* = \tau\sigma\tau^{-1}$?

Now let us formulate the problem analytically. To this end, let us think of the systems of equations of σ, σ^*, τ as expressed in terms of a fixed projective coordinate system of P_n. Let the matrices[5] of these systems of equations be A, A^*, B respectively. It then follows that the matrix of the system of equations of τ^{-1} is B^{-1} and that of $\tau\sigma\tau^{-1}$ is BAB^{-1}. But σ^* was to be equal to $\tau\sigma\tau^{-1}$. Thus, the projective equivalence of σ and σ^* means the following for the matrices A, A^*: When does there exist a non-singular matrix B such that BAB^{-1} is equal to A^*, up to a numerical factor ϱ?

But this question is essentially answered by Theorem 2 of *Modern Algebra*, § 26. Accordingly, *the necessary and sufficient condition for our problem is that there exist a number $\varrho \neq 0$ for which the characteristic matrices of A and ϱA^* have the same elementary divisors.*

For this criterion to be usable in actual practice, we would need to know how to find the elementary divisors of $\varrho A^* - uE$, for arbitrary ϱ, from those of $A^* - uE$. But the answer to this question has already been given in Chap. VI, Exercise 3.

[5] Cf. p. 88. It should not be forgotten that the matrix of a linear projectivity is unique only up to a factor $\neq 0$.

Exercises

1. Define projective equivalence of arbitrary ordered k-tuples of points of P_n just as was done for linear quadruples. Then show that:

k-tuples of linearly independent points are always projectively equivalent.

Two k-tuples of linearly dependent points, both of which, however, contain $k-1$ linearly independent points, are projectively equivalent if and only if corresponding subsets are always either linearly dependent or linearly independent together.

This condition, however, is no longer sufficient for k-tuples with at most $k-2$ linearly independent points. But in this case it is possible, as a generalization of the concept of cross ratios, to find invariants that characterize the classes. To see how this method applies to an example, consider those 5-tuples that lie in a plane and are such that three points of each 5-tuple are linearly independent. If P_1, P_2, P_3, P_4, P_5 is such a 5-tuple, then the ratios of the coordinates of point P_5 in the coordinate system $(P_1, P_2, P_3 \mid P_4)$ are invariants under linear projectivities and characterize the classes of these 5-tuples.

In similar fashion, discuss the other cases for 5-tuples. Generalize the method to arbitrary k-tuples.

2. The Jacobi-Sylvester law of inertia, mentioned on p. 129, states that *a quadratic form in the normal form given in Theorem 3, i.e.,*

$$\xi_0^{*2} + \xi_1^{*2} + \cdots + \xi_k^{*2} - \xi_{k+1}^{*2} - \cdots - \xi_r^{*2}, \qquad 0 \leqq k \leqq r \leqq n,$$

can never be transformed into a second form of this type, say $\xi_0^2 + \xi_1^2 + \cdots + \xi_h^2 - \xi_{h+1}^2 - \cdots - \xi_r^2$, *by a real non-singular linear substitution*

$$(*) \qquad \xi_i^* = \sum_{\nu=0}^{n} a_{i\nu}\,\xi_\nu, \qquad i = 0, 1, \cdots, n,$$

if $k \neq h$.

For proof, let us assume the opposite. Thus, the first form becomes equal to the second upon substitution of $(*)$. This yields the following identity in the ξ_i.

$$(\ast\ast) \qquad \begin{aligned} &\left(\sum_{\nu=0}^{n} a_{0\nu}\,\xi_\nu\right)^2 + \left(\sum_{\nu=0}^{n} a_{1\nu}\,\xi_\nu\right)^2 + \cdots + \left(\sum_{\nu=0}^{n} a_{k\nu}\,\xi_\nu\right)^2 + \xi_{h+1}^2 + \cdots + \xi_r^2 \\ &= \xi_0^2 + \xi_1^2 + \cdots + \xi_h^2 + \left(\sum_{\nu=0}^{n} a_{k+1,\nu}\,\xi_\nu\right)^2 + \cdots + \left(\sum_{\nu=0}^{n} a_{r\nu}\,\xi_\nu\right)^2. \end{aligned}$$

If now, for certain ξ-values, the left-hand side of $(\ast\ast)$ is equal to 0, then it necessarily follows from the form of this equation that all the ξ_i must $= 0$ for $i = 0, 1, \ldots, n$. But this is in contradiction to Theorem 5 of *Modern Algebra*, § 6, which says that for $k < h$ there assuredly are $n+1$ real numbers $\xi_0, \xi_1, \cdots, \xi_n$, *not* all zero, which satisfy the homogeneous equations

$$\sum_{\nu=0}^{n} a_{0\nu}\,\xi_\nu = 0, \qquad \xi_{h+1} = 0,$$

$$\sum_{\nu=0}^{n} a_{1\nu}\,\xi_\nu = 0, \qquad \xi_{h+2} = 0,$$

$$\cdots\cdots\cdots \qquad \cdots\cdots$$

$$\sum_{\nu=0}^{n} a_{k\nu}\,\xi_\nu = 0, \qquad \xi_r = 0.$$

3. We have given a complete answer, both in the real case (Ex. 2) and the complex case, to the question of when two quadratic forms can be transformed into each other by a non-singular linear transformation. For certain geometrical applications (cf. Chap. X, Ex. 4), the same question is of interest for ordered *pairs* of quadratic forms. We can write a quadratic form in matrix notation as $(\xi)'A(\xi)$. This is transformed into $(\xi)=T(\eta)$ by the non-singular linear substitution $(\eta)'T'AT(\eta)$. Stated in terms of matrices alone, our question is:

When does there exist a non-singular matrix T such that, given the ordered pairs of symmetric matrices A, B, and A^*, B^*, $A^*=T'AT$ and $B^*=T'BT$ hold simultaneously?

In the *complex* case, with the added assumptions $|B|\neq 0$ and $|B^*|\neq 0$, then the answer is simple: *If and only if the polynomial matrices $A-uB$ and A^*-uB^* have the same elementary divisors.*

That the condition is necessary follows at once, by *Modern Algebra,* § 25, Theorem 4 from the equation

$$(T'AT)-u(T'BT)=T'(A-uB)\,T.$$

To show that the condition is also sufficient, first note that the general theorem follows from the special case in which $B=B^*=E$. This is an easy consequence of Theorem 2 of this chapter (if we make use of $|B|\neq 0$, $|B^*|\neq 0$). If we now assume the equality of the elementary divisors of $A-uE$ and A^*-uE, it can be shown that the equations $T'AT=A^*$, $T'T=E$ are solvable for T if and only if the equations

$$(*) \qquad XA=AX, \quad X'X=W'W$$

are solvable for X, where W is a matrix for which $WAW^{-1}=A^*$ (whose existence follows from *Modern Algebra,* § 26, Theorem 2). *Hint:* Set $X=TW$.

Finally, in order to solve (*), first note that from $WAW^{-1}=A^*$ and the symmetry of A and A^* we have $(W'W)A=A(W'W)$. From this it may readily be deduced that if there exists a polynomial $f(u)$ such that $[f(W'W)]^2=W'W$, then $X=f(W'W)$ is a solution of (*).

In Exercise 4 it will be shown how a polynomial $f(u)$ can be found such that for any arbitrary non-singular complex matrix Z, $[f(Z)]^2=Z$.

4. In Exercise 3 we used the theorem: *For every non-singular complex matrix Z there always exists a polynomial $f(u)$ with complex coefficients that satisfies the equation* $[f(Z)]^2 = Z$.

If $m(u)$ is the minimal polynomial of Z, the theorem will be proved if we can find an $f(u)$ for which $[f(u)]^2 - u$ is divisible by $m(u)$.

First, consider the case in which $m(u)$ is a power of a linear factor. As the minimal polynomial of a non-singular matrix, $m(u)$ cannot have u as a factor; hence, $m(u) = (u-\alpha)^k$ with $\alpha \neq 0$. Now, let us write the desired polynomial $f(u)$ in the form

$$f(u) = c_0 + c_1(u-\alpha) + \cdots + c_{k-1}(u-\alpha)^{k-1}.$$

If we then arrange $[f(u)]^2 - u = [f(u)]^2 - (u-\alpha) - \alpha$ according to powers of $u-\alpha$ and set the constant term and all the coefficients of $(u-\alpha), (u-\alpha)^2, \ldots, (u-\alpha)^{k-1}$ equal to 0, we obtain equations from which $c_0, c_1, \ldots, c_{k-1}$ can be computed in succession.

In the general case, let $m(u) = (u-\alpha_1)^{k_1}(u-\alpha_2)^{k_2}\cdots(u-\alpha_r)^{k_r}$, with $\alpha_i \neq 0$ for every i. If we now define $m_i(u) = \dfrac{m(u)}{(u-\alpha_i)^{k_i}}$, there exist r polynomials $g_i(u)$ for which

$$g_1(u)m_1(u) + g_2(u)m_2(u) + \cdots + g_r(u)m_r(u) = 1.$$

If, finally, we let $h_i(u) = g_i(u)m_i(u)$, then

a) $h_i(u)h_k(u)$ is divisible by $m(u)$ for $i \neq k$;
b) $[h_i(u)]^2 - h_i(u)$ is divisible by $m(u)$.

Lastly, if we determine, as in the first case, polynomials $f_i(u)$ for which $[f_i(u)]^2 - u$ is always divisible by $(u-\alpha_i)^{k_i}$, then

$$f(u) = \sum_{i=1}^{r} f_i(u)\,h_i(u)$$

is the desired polynomial, for which $[f(u)]^2 - u$ is divisible by $m(u)$. ***Hint:*** Apply a) and b) to the identity

$$\left[\sum_{i=1}^{r} f_i h_i\right]^2 - u = \left\{\left[\sum_{i=1}^{r} f_i h_i\right]^2 - \sum_{i=1}^{r} f_i^2 h_i\right\} + \sum_{i=1}^{r}(f_i^2 - u)h_i + u\left(\sum_{i=1}^{r} h_i - 1\right).$$

5. The theorem set forth in Exercise 3 is valid for the case $B = B^* = E$ even if we restrict ourselves to real forms and real substitutions. The general theorem, however, no longer holds. Give counter-examples.

CHAPTER X

PROJECTIVE PROPERTIES OF HYPERSURFACES OF THE SECOND ORDER

Every property proved for *one* hypersurface of the second order will hold true for all the hypersurfaces of its class provided the property is invariant with respect to linear collineation. Properties and theorems that are invariant with respect to all the transformations of the linear projective group are themselves called '*projective.*' In order to discover projective properties of hypersurfaces of the second order, it will suffice to study one representative of each class considered. The advantage in doing this is obvious: one can try to select a representative so that its equation is especially tractable and convenient for investigating a particular question. Of course, such a choice of a suitable representative can also be interpreted as a transformation of coordinates, as is evident from the argument on p. 120.

For many projective theorems, the normal forms afford such suitable representatives. As a first example, consider the *hypersurfaces of the second order of rank* 1. Both in complex and in real P_n they constitute exactly one class, with the normal form $\xi_0^2 = 0$. This equation, being completely equivalent to $\xi_0 = 0$, represents this latter hyperplane. Every point of this hyperplane is a double point of $\xi_0^2 = 0$. These statements about the hypersurface $\xi_0^2 = 0$ are invariant with respect to linear collineations and therefore hold for all hypersurfaces of the class. Thus we have the general result:

A hypersurface of rank 1 *consists precisely of all the points of a hyperplane; every point is a double point.*

Let us now consider *hypersurfaces of rank* 2. Here, too, we can draw conclusions about their nature from the normal forms. The hypersurfaces of rank 2 form a single class in *complex* P_n and two classes in *real* P_n.

Let us consider the real case first. Here the two classes are represented by the normal forms $\xi_0{}^2 + \xi_1{}^2 = 0$, $\xi_0{}^2 - \xi_1{}^2 = 0$. The first is satisfied only by real points for which $\xi_0 = \xi_1 = 0$. Thus it consists precisely of all the points of the linear space of dimension $n-2$ given by the equations $\xi_0 = 0$, $\xi_1 = 0$. Again, every point is a double point. As above, analogous statements hold for every hypersurface of this class.

The normal form $\xi_0{}^2 - \xi_1{}^2 = 0$ of the other class in real P_n can be taken at the same time as a representative of the *single* complex class of rank 2. Thus, every projective statement that is true for the hypersurface $\xi_0{}^2 - \xi_1{}^2 = 0$ is true both for all the hypersurfaces of the real class it represents and for *all* complex hypersurfaces of the second order with rank 2. Now, the equation $\xi_0{}^2 - \xi_1{}^2 = 0$ can also be written in the form $(\xi_0 + \xi_1)(\xi_0 - \xi_1) = 0$. Since this equation is satisfied if and only if *at least* one of the equations $\xi_0 + \xi_1 = 0$ and $\xi_0 - \xi_1 = 0$ is true, the hypersurface $\xi_0{}^2 - \xi_1{}^2 = 0$ consists precisely of all the points of the *two distinct hyperplanes* $\xi_0 + \xi_1 = 0$ and $\xi_0 - \xi_1 = 0$. The double points are all the points for which $\xi_0 = \xi_1 = 0$, i.e., the points of intersection of the two hyperplanes.

Summing this up, we have:

In complex P_n, a hypersurface of the second order of rank 2 consists of a pair of hyperplanes, the intersection of which constitute the totality of its double points. In real P_n, there exists the additional case in which the hypersurface consists solely of an $(n-2)$-dimensional linear space of double points.[1]

As a final example of the degenerate hypersurfaces of the second order of P_n, let us consider those of rank n. They have exactly one double point. Their normal forms are

$$(1) \qquad \xi_0^2 + \xi_1^2 + \cdots + \xi_k^2 - \xi_{k+1}^2 - \cdots - \xi_{n-1}^2 = 0.$$

If $k = n-1$ in this equation, then in real P_n the hypersurface consists of the double point only. In the other cases, i.e., those of the remaining real classes and that of the single complex class, we consider the intersection of the hypersurface (1) with the coordinate hyperplane $\xi_n = 0$.

[1] In the latter case, the real equation of the hypersurface, considered as an equation in complex P_n, represents two 'conjugate complex' hyperplanes that intersect precisely in the real double points. In the case of $\xi_0^2 + \xi_1^2 = 0$, the conjugate complex hyperplanes are given by $\xi_0 + i\xi_1 = 0$ and $\xi_0 - i\xi_1 = 0$.

The same correspondence exists between the points of this hyperplane and the homogeneous n-tuples formed from the first n coordinates $\xi_0, \xi_1, \ldots, \xi_{n-1}$ of these points as exists between the points of P_{n-1} and the homogeneous n-tuples. This coordinate hyperplane may itself be regarded, accordingly, as P_{n-1}. The intersection, which, as regards the first n coordinates of its points, is again given by equation (1), therefore represents nothing other than a hypersurface of the second order in P_{n-1}, and as such, moreover, is non-degenerate.

If we now join the double point, which lies outside the hyperplane $\xi_n = 0$ (indeed, its coordinates are $\xi_0 = \xi_1 = \cdots = \xi_{n-1} = 0,\ \xi_n = 1$), with any point of the intersection just considered, the entire line lies on the hypersurface of P_n. For if $\xi_0, \xi_1, \ldots, \xi_{n-1}, \xi_n = 0$ are the coordinates of a point of this intersection, the coordinates $\eta_0, \eta_1, \ldots, \eta_n$ of a point of the connecting line in question can be written as

$$(2)\quad \eta_0 = \lambda\xi_0 + \mu\cdot 0,\ \eta_1 = \lambda\xi_1 + \mu\cdot 0, \cdots, \eta_{n-1} = \lambda\xi_{n-1} + \mu\cdot 0,$$
$$\eta_n = \lambda\cdot 0 + \mu\cdot 1.$$

But these η_i always satisfy (1), since $\xi_0, \xi_1, \ldots, \xi_{n-1}$ do so. Conversely, if the numbers $\eta_0, \eta_1, \ldots, \eta_n$ are the coordinates of any point of the hypersurface (1) other than the double point, then $\eta_0, \eta_1, \ldots, \eta_{n-1}, 0$ represents a point of intersection of the hypersurface (1) with the hyperplane $\xi_n = 0$ which is such that the line joining the double point with this point of intersection passes through the given point.

Thus we see that if we draw all the lines from the double point to the points of intersection of the hypersurface (1) with the hyperplane $\xi_n = 0$, then these lines constitute the whole hypersurface. For $n \geqq 4$, such a hypersurface is called a (*quadratic*) *hypercone*; in the case of P_3, it is called simply a (*quadratic*) *cone*; in P_2, it is a pair of lines; and in P_1, a double point.

Our discussion has yielded, as an incidental result, that the non-degenerate hypersurfaces of the second order of P_{n-1} can be obtained as the intersections of the hypersurface defined in P_n by (1) and a hyperplane of P_n (namely, $\xi_n = 0$). In particular, this is true for the hypersurfaces of P_2, i.e., the curves of the second order, which are for this reason called *conic sections*. The term is also applied to the degenerate curves of the second order, since they can be generated in the same way: they are formed when the cutting plane passes through the vertex, i.e., the double point of the cone. In the sequel, we shall usually use the shorter term 'conic section' rather than 'curve of the second order.'

Projective Generation of Conic Sections

The preceding discussion takes care of all the hypersurfaces of the second order in P_1 and the degenerate ones in P_2. We now turn our attention to the remaining case, of the non-degenerate conic sections in P_2. Again, all our statements will be projective in nature.

Let us first find out whether all the points of a line can belong to a non-degenerate conic section in P_2. Let us assume that the line g is part of a conic section. If the conic section consists of g alone, then all the points of g are double points,[2] and the conic section is accordingly degenerate. However, if there exists a point P of the conic section that is not on g, then there also exists at least one tangent h of the conic section which passes through P. Let the point of intersection of g and h be Q. Since at least two tangents pass through Q (namely g and h), no uniquely determined tangent hyperplane[3] exists at Q. Consequently (see Chap. VIII, Theorem 1), Q is a double point, and thus, again, the conic section is degenerate.

We see, then, that an entire line can never belong to a non-degenerate conic section of P_2. Every tangent therefore has just one point in common with the conic section.

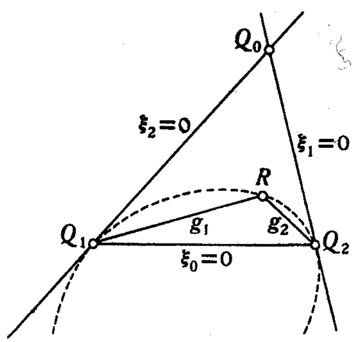

Fig. 10

Let us now consider two different points Q_1, Q_2 of a non-degenerate conic section. The tangents at Q_1 and Q_2 are uniquely determined. Let these tangents intersect in a point Q_0 (Fig. 10). We take Q_0, Q_1, Q_2 as the fundamental points of a projective coordinate system. In this coordinate system, the equation of the conic section takes on a characteristic form.

Let us write the equation, to begin with, in the form

$$(3) \qquad \sum_{i,k=0}^{2} a_{ik}\,\xi_i\,\xi_k = 0, \qquad a_{ik} = a_{ki}.$$

Since the point Q_1 lies on the conic section, (3) must be satisfied by $\xi_1 = 1$,

[2] For then *every* line of the plane distinct from g cuts the line g in just *one* point and is thus a tangent.

[3] A tangent hyperplane in P_2 is a line, and thus a tangent line.

$\xi_0 = \xi_2 = 0$. This shows that $a_{11} = 0$. Likewise, from the fact that Q_2 lies on the conic section, it follows that $a_{22} = 0$. Hence, the equation of the tangent at Q_1 is, by (3), $a_{10}\xi_0 + a_{12}\xi_2 = 0$. On the other hand, we know that this tangent is just the side Q_0Q_1 of the fundamental triangle, or $\xi_2 = 0$. It follows that $a_{10} = 0$, $a_{12} \neq 0$. Similarly, the tangent at Q_2 is given, on the one hand, by $a_{20}\xi_0 + a_{21}\xi_1 = 0$ and, on the other hand, by $\xi_1 = 0$, whence it follows that $a_{20} = 0$. Equation (3) is thus reduced to the form

$$a_{00}\,\xi_0^2 + 2\,a_{12}\,\xi_1\,\xi_2 = 0. \tag{4}$$

Now in (4), a_{00} must certainly be different from zero, since the point Q_0 does not lie on the conic section.[4] Hence, we may divide (4) by a_{00}. But then, merely by altering the unit point, we can absorb[2] the coefficient $-2a_{12}/a_{00}$ into the ξ_2 (or into the ξ_1), so that our equation finally takes on the form

$$\xi_0^2 - \xi_1\,\xi_2 = 0. \tag{5}$$

The fundamental points of our coordinate system are still Q_0, Q_1, Q_2.

We now consider the pencils of lines with Q_1 and Q_2 as carriers. Let us call them b_1 and b_2, respectively. If we join any point $R \neq Q_1$ of the conic section (5) to Q_1, we obtain a line g_1 of the pencil b_1 (Fig. 10). *If we let R run through all the points $\neq Q_1$ of the conic section, g_1 will run through every line of the pencil b_1 except Q_1Q_0 and will do so just once.* For, every line g_1 of b_1 distinct from Q_1Q_0 has *exactly one* other point R in addition to Q_1 in common with the conic section, since g_1 is not a tangent. This discussion also applies, of course, to b_2.

Thus, if we let the lines g_1 of b_1 and g_2 of b_2 correspond to each other whenever g_1 and g_2 pass through the same point R of the conic section,

---

[4] For the tangent Q_0Q_2 has only the one point Q_2 in common with the conic section.

[5] This may be accomplished, say, by the coordinate transformation

$$\xi_0^* = \xi_0, \qquad \xi_1^* = \xi_1, \qquad \xi_2^* = -\frac{2\,a_{12}}{a_{00}}\,\xi_2 .$$

This transformation leaves the fundamental points fixed, while the point $\xi_0 = 1$, $\xi_1 = 1$, $\xi_2 = -\dfrac{a_{00}}{2\,a_{12}}$ becomes the unit in the ξ^*-coordinates.

distinct from Q_1 and Q_2, a *one-to-one* correspondence is defined which makes each line of b_1 distinct from Q_1Q_0 and Q_1Q_2 correspond to a line of b_2 distinct from Q_2Q_0 and Q_2Q_1, and conversely. It is an obvious step to extend this correspondence to a one-to-one relation between the *whole* of b_1 and the *whole* of b_2 by stipulating that the line Q_1Q_0 of b_1 shall have Q_2Q_1 as its image in b_2 and the line Q_1Q_2 of b_1 shall have Q_2Q_0 as its image in b_2.[6] We now assert the following:

This correspondence is a linear projectivity between b_1 and b_2.

In order to prove this, we first observe that every line of b_1 must have an equation of the form

$$\lambda \xi_0 + \mu \xi_2 = 0. \tag{6}$$

For, a linear equation $\lambda \xi_0 + \nu \xi_1 + \mu \xi_2 = 0$ will be satisfied by the coordinates of the point Q_1 (i.e., by $\xi_0 = \xi_2 = 0$, $\xi_1 = 1$) if and only if $\nu = 0$. Similarly,[7]

$$\lambda' \xi_1 + \mu' \xi_0 = 0 \tag{7}$$

always represents a line of b_2, and every such line has an equation of this form. We now show that we again obtain our old correspondence between b_1 and b_2 if we let each line of (6) correspond with that line of (7) for which $\lambda' = \lambda$, $\mu' = \mu$.

For $\lambda = 1$, $\mu = 0$, we obtain in (6) the line Q_1Q_2 of b_1 and for $\lambda' = 1$, $\mu' = 0$ we obtain in (7) the line Q_2Q_0 of b_2. But for $\lambda = \lambda' = 0$, $\mu = \mu' = 1$, Q_1Q_0 and Q_2Q_1 correspond to each other. This is in accord with our stipulation above. However, if $\lambda = \lambda'$ and $\mu = \mu'$ are both different from zero, then Q_2 does *not* lie on (6) nor Q_1 on (7), so that the point of intersection R of (6) and (7) must be different from Q_1 and Q_2. It remains to prove that this point of intersection lies on the conic section (5). The coordinates η_0, η_1, η_2 of the point of intersection R, by virtue

[6] These agreements and the foregoing can be comprehended in one statement, as follows: g_1 of b_1 and g_2 of b_2 shall correspond to each other whenever the pair of lines g_1, g_2 have *precisely* the points Q_1, Q_2 and the point of intersection of g_1, g_2 in common with the conic section.

[7] The reason why we employ a notation that makes the coefficients of ξ_1 and ξ_0 in (7) correspond to the coefficients of ξ_0 and ξ_2, respectively in (6) will soon become apparent.

of (6) and (7), clearly satisfy the two equations

$$(8)\qquad \begin{aligned} \lambda\eta_0+\mu\eta_2 &= 0, \\ \lambda\eta_1+\mu\eta_0 &= 0. \end{aligned}$$

Since λ and μ are both different from zero, it follows that the determinant of this system of homogeneous linear equations in λ, μ vanishes:

$$(9)\qquad \begin{vmatrix} \eta_0 & \eta_2 \\ \eta_1 & \eta_0 \end{vmatrix} = 0.$$

This means that the η_i satisfy equation (5), as was to be proved.

It is now easy to see that our correspondence is a linear projectivity. The coordinate vectors (hyperplane coordinates) of the three lines $\xi_0=0$, $\xi_1=0$, $\xi_2=0$ are

$$\begin{aligned} \xi_0 &= 0: & \mathfrak{u}_0 &= \{1, 0, 0\}, \\ \xi_1 &= 0: & \mathfrak{u}_1 &= \{0, 1, 0\}, \\ \xi_2 &= 0: & \mathfrak{u}_2 &= \{0, 0, 1\}. \end{aligned}$$

Our correspondence, as defined between equations (6) and (7), now means simply this: To the lines of b_1 with coordinate vector $\lambda\mathfrak{u}_0+\mu\mathfrak{u}_2$ there corresponds the line $\lambda\mathfrak{u}_1+\mu\mathfrak{u}_0$ of b_2. We need only recall Theorem 9 of Chapter V to see at once that we are indeed dealing with a linear projectivity.[8]

The following converse of what we have proved holds:

Let there be given any two pencils of lines b_1 and b_2, whose carriers we denote by Q_1, Q_2. We assume $Q_1 \neq Q_2$. In addition, let a linear projectivity be given between b_1 and b_2. Each pair of corresponding lines intersect in a point; *the locus of all these points of intersection is a conic section.*

---

[8] The second part of the theorem quoted, Theorem 9 of Chapter V, can be restated as follows: If $\mathfrak{u}_1$, $\mathfrak{u}_2$ are any two distinct hyperplanes of a pencil b_1 and $\mathfrak{v}_1$, $\mathfrak{v}_2$ two hyperplanes of a pencil b_2, then the correspondence $\lambda\mathfrak{u}_1+\mu\mathfrak{u}_2 \rightleftarrows \lambda\mathfrak{v}_1+\mu\mathfrak{v}_2$ always yields a linear projectivity between b_1 and b_2, namely, that which takes the three hyperplanes $\mathfrak{u}_1$, $\mathfrak{u}_2$, $\mathfrak{u}_1+\mathfrak{u}_2$ into $\mathfrak{v}_1$, $\mathfrak{v}_2$, $\mathfrak{v}_1+\mathfrak{v}_2$, respectively.

In order to discover in the course of the proof when the resulting conic section is degenerate and when it is not, we distinguish between the case in which the line g through Q_1, Q_2 (which belongs to *both* pencils) corresponds, under the given projectivity to g itself and the case in which it does not. Let us first consider the case in which g does *not* correspond to itself;[9] i.e., g, thought of as a line of b_1, has as its image in b_2 a line $h_2 \neq g$ (Fig. 11) and, similarly, thought of as a line of b_2, has an image $h_1 \neq g$ in b_1. Let Q_0 denote the point of intersection of h_1 and h_2.

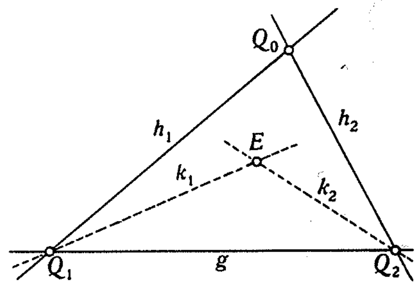

Fig. 11

Let us, in addition, select a line k_1 of b_1, distinct from g and h_1. Let its image in b_2 be k_2. Let E be the point of intersection of k_1 and k_2. We then introduce the coordinate system $(Q_0, Q_1, Q_2 \mid E)$. The lines of the pencil b_1 may now again be represented in the form (6) and those of b_2 in the form (7). And, just as before, we consider, besides the given projectivity, the correspondence that arises by letting each line of (6) correspond to that line of (7) for which $\lambda' = \lambda$, $\mu' = \mu$. As we already know, this correspondence is also a linear projectivity and takes the lines h_1, g of the pencil b_1 into the lines g, h_2 of b_2. Furthermore, for $\lambda = 1$, $\mu = -1$, the line (6) is identical with k_1, because the unit point E $(\xi_0 = \xi_1 = \xi_2 = 1)$ lies on this line. Likewise, the line (7) coincides with k_2 for $\lambda' = 1$, $\mu' = -1$. We thus have the result that the prescribed linear projectivity and that defined between (6) and (7) by the relation $\lambda' = \lambda$, $\mu' = \mu$ coincide in three pairs of corresponding lines, namely in (h_1, g) (g, h_2), (k_1, k_2). Therefore, by Chap. V, Theorem 9, our two projectivities are identical.

But we have proved above that the points of intersection of corresponding lines under the correspondence between (6) and (7) defined by $\lambda' = \lambda$, $\mu' = \mu$, fill out the conic section (5). The desired locus in our present case is thus a non-degenerate conic section.[1]

---

[9] This was automatically the case in the preceding discussion.

[1] Obviously also a non-imaginary conic section.

It remains to consider the case in which the line g corresponds to itself under the given linear projectivity between b_1 and b_2. Then let h_1, h_2 and k_1, k_2 (Fig. 12) be two further pairs of corresponding lines. Let the points of intersection of h_1, h_2 and of k_1, k_2 be respectively S, T. Denote the line through S, T by j. Let us now consider the new correspondence which makes the line g_1 of b_1 and the line g_2 of b_2 correspond if they meet j in the same point U (Fig. 12). This correspondence, again, is a linear projectivity (and even a perspectivity), it coincides with the given linear projectivity in three pairs of corresponding lines, namely in

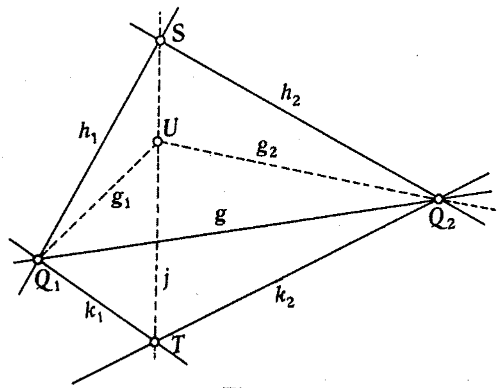

Fig. 12

(h_1, h_2), (g, g), (k_1, k_2), and it must accordingly be identical with the given linear projectivity.

This time, the locus of the points of intersection of corresponding pairs of lines therefore consists of the two lines g and j. Thus, it is a degenerate conic section of rank 2. Summing up, we can state:

THEOREM 1. *If two pencils of lines with distinct kernels are linearly projective to each other, then the locus of the points of intersection of corresponding lines is a conic section: a conic section of rank 2 if the common line of the two pencils corresponds to itself; otherwise, of rank 3. Every non-imaginary conic section of rank 2 or 3 can in fact be generated in this fashion, and, in the case of a non-degenerate conic section, any two of its points can be chosen as the carriers of the generating pencils.*[2]

[2] This was the very first result to be proved. In the degenerate case, where, as in Fig. 12 say, the conic section consists of the lines g and j, the carriers of the pencils must be distinct from the intersection of g and j, and must either both lie on g or both on j.

The significance of this theorem lies in the fact that it affords a means of constructing a conic section, given five of its points. More specifically, let Q_1, Q_2, Q_3, Q_4, Q_5 be given in the projective plane. We set ourselves the problem of finding a conic section passing through all five points Q_i. If among these five points there are three that lie on a line, then every conic section that contains all the Q_i must also contain this line and is therefore degenerate. In any such case the problem of finding all the conic sections that pass through the Q_i is trivial.

We may therefore assume at the outset that no three of the five points Q_i are collinear. Then there exists no degenerate conic section containing the five points.[3] In order to find the (necessarily non-degenerate) conic section of the desired kind, let us consider, say, the pencils of lines

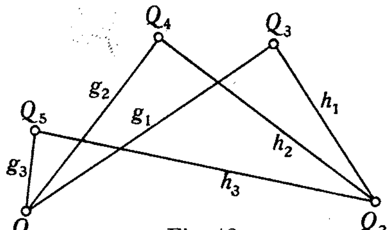

Fig. 13

with carriers Q_1, Q_2, which we may denote by b_1, b_2. The three lines Q_1Q_3, Q_1Q_4, Q_1Q_5 (Fig. 13), which we may call g_1, g_2, g_3, belong to b_1, whereas the lines Q_2Q_3, Q_2Q_4, Q_2Q_5, which we will call h_1, h_2, h_3, belong to b_2. By virtue of our assumption concerning the Q_i, the lines g_1, g_2, g_3 are distinct; likewise, $h_i \neq h_k$ for $i \neq k$. Hence, there exists one and only one linear projectivity between b_1 and b_2 that takes the g_i into the h_i for $i = 1, 2, 3$. Since every conic section that passes through all the Q_i is necessarily generated ('generated' in the sense of Theorem 1) by this linear projectivity between b_1 and b_2, there exists one and only one conic section of this kind.

Furthermore, a simple construction can be found for our linear projectivity that enables us to find the image line in b_2 of any line of b_1, and conversely. To this end, let the lines g_3 and h_1, say,[4] i.e., Q_1Q_5 and Q_2Q_3,

---

[3] For, any such conic section would have to consist either of one line or of a pair of lines. If five points, however, are divided among at most two lines, there must be three that are collinear.

[4] Of course, we can here choose any pair g_i, h_k, with $i \neq k$.

intersect at S (Fig. 14). Moreover, let us draw the line through Q_3Q_4, which we shall call j, and the line Q_4Q_5, which we may call k. None of

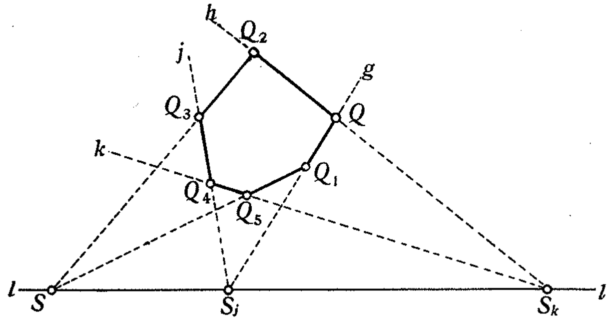

Fig. 14

the points S, Q_1, Q_2 lies on k or on j.[5] Therefore if l is any line through S, intersecting j in S_j and k in S_k, we can always make the line g, which joins Q_1 with S_j, and the line h, which joins Q_2 with S_k, correspond to each other. It is easy to see (by repeated application of Theorem 1 of Chap. IV) that this correspondence represents a linear projectivity between the pencils b_1 and b_2. If we now let l coincide successively with SQ_3, SQ_4, SQ_5, it results that our correspondence takes g_1 into h_1, g_2 into h_2, and g_3 into h_3 and is thus precisely the required linear projectivity.

This construction of the projectivity in question permits us to find as many additional points of the conic section as we wish. We need merely let any two corresponding lines g and h (Fig. 14) intersect, and the point of intersection Q will be a point of the conic section.

Our construction admits of still another important interpretation, as follows. If we think of the line l in Fig. 14 as fixed and, as before, denote the point of intersection of g and h by Q, then the hexagon $Q_1\,Q\,Q_2\,Q_3\,Q_4\,Q_5$, all of whose vertices lie on the conic section, has the following property: The points of intersection S, S_j, S_k of the three pairs of opposite sides all lie on l. Since, given any conic section, we could begin our construction with any five points Q_1, Q_2, Q_3, Q_4, Q_5 and then obtain any other point Q as our sixth point by a suitable choice of l, we have proved the **Theorem of Pascal**, which may be stated as follows:

THEOREM 2. *The points of intersection of the three pairs of opposite sides of a hexagon inscribed in a non-degenerate conic section always lie on a line.*

[5] Suppose, for example, that S were to lie on k or j. Then either Q_1, Q_4, Q_5 or Q_2, Q_3, Q_4 must lie on a line. And so forth.

Note that this theorem remains valid if Q coincides with Q_1 or Q_2. The 'side' QQ_1, or QQ_2, of the hexagon in this case means (according to the above construction) the tangent to the conic section at Q_1 and Q_2, respectively. This remark gives the clue to a method of constructing the tangent at a given point (namely, Q_1) of a conic section, a method that could also have been obtained, to be sure, directly from the linear projectivity between b_1 and b_2.

We made the assumption that no three of the five points Q_i should lie on a line. The reader can easily convince himself, however, that our argument remains valid, word for word, if instead of this assumption we make the somewhat weaker one that the point Q_1 is not collinear with any two of the points Q_3, Q_4, Q_5 and that, similarly, Q_2 and every two of the points Q_3, Q_4, Q_5 are always linearly idependent. To be sure, the conic section obtained by the above construction can then be degenerate. Of particular interest is the now admissible degenerate case in which Q_4, Q_1, and Q_2 are collinear. Then the points Q_3, Q_5, Q must necessarily also be collinear, and we obtain Pascal's Theorem for a pair of lines, the very case of Pascal's Theorem that we encountered in Chap. IV, Exercise 2.

The theorem *dual* to that of Pascal is called the **Theorem of Brianchon**. Taking Chap. VIII, Theorem 5 into account, it goes as follows:

THEOREM 3. *The lines joining the three pairs of opposite vertices of a hexagon circumscribed about a non-degenerate conic section all intersect in a point.*

The Families of Lines on Non-degenerate Surfaces of the Second Order in P_3

We now turn to P_3. Here, too, the discussion at the beginning of this section covers all the degenerate cases (rank 1, 2, $3 = n$).[6] As regards non-degenerate surfaces of the second order of rank 4, we shall derive, in what follows, some descriptive results pertaining to the families of straight lines which are to be found on such surfaces.

In real P_3, according to Chap. IX, Theorem 8, there is only one class of non-degenerate surfaces of the second order on which straight lines are

[6] We refrain from enumerating these cases separately at this point. However, it is recommended that the student do this and clarify for himself the geometrical picture in each case, and that he do the same for P_1 and P_2. In the next chapter, we shall give a comprehensive table of all types of conic sections and surfaces of the second order (cf. pp. 176 and 177).

to be found. The normal form of this class is the surface

$$\xi_0^2 + \xi_1^2 - \xi_2^2 - \xi_3^2 = 0. \tag{10}$$

In complex P_3, the non-degenerate surfaces constitute but a single class, of which we can also take (10) as representative.

Equation (10) does not have the most convenient form, however, for the investigation we are about to make. We accordingly subject this equation to the linear collineation

$$\begin{aligned} \xi_0 &= \eta_0 + \eta_3, \\ \xi_2 &= \eta_0 - \eta_3, \\ \xi_3 &= \eta_1 + \eta_2, \\ \xi_1 &= \eta_1 - \eta_2. \end{aligned} \tag{11}$$

This transforms (10) into the projectively equivalent hypersurface

$$\eta_0 \cdot \eta_3 - \eta_1 \cdot \eta_2 = 0. \tag{12}$$

We now examine this surface. All our assertions will be projective in nature, i.e., invariant under linear collineations, and thus they hold both for every non-degenerate surface of complex P_3 and for every real surface of the real class represented by (10). Moreover, we can also find a coordinate system for each of these surfaces in which it is even represented by equation (12) (cf. p. 119).

Using two fixed numbers[7] λ and μ that do not both vanish and the variables $\eta_0, \eta_1, \eta_2, \eta_3$, we form the system of equations

$$\begin{aligned} \lambda\eta_0 + \mu\eta_1 &= 0, \\ \lambda\eta_2 + \mu\eta_3 &= 0. \end{aligned} \tag{13}$$

The rank of this system of equations is 2, and it therefore represents a straight line in P_3. It can be seen immediately that *for any arbitrary choice of λ and μ the line* (13) *always lies on the surface* (12). For since λ, μ are not both zero, we must have, for the determinant formed from the coordinates $\eta_0, \eta_1, \eta_2, \eta_3$ of an arbitrary point of the line,

$$\begin{vmatrix} \eta_0 & \eta_1 \\ \eta_2 & \eta_3 \end{vmatrix} = 0, \tag{14}$$

i.e., (12) is satisfied.

---

[7] Real or complex numbers, according as we are dealing with real or complex P_3.

Since λ and μ may be arbitrary, provided only they are not both zero, we have thus found a whole 'family' (13) of straight lines that lie on the surface (12). *Through each point of the surface* (12) *there passes one and only one line of the family* (13). For, given η_0, η_1, η_2, η_3 satisfying (12) and not all zero, the equations (13), with the given η_i substituted in them and with the λ, μ considered as the unknowns, determines λ, μ up to a constant of proportionality.[8]

Since only one line of (13) thus passes through each point of (12), it immediately follows that *two distinct lines of the family* (13) *can never have a point in common, i.e., they are always skew to each other.*

If, instead of (13), we take the equations

$$\begin{aligned} \lambda\eta_0 + \mu\eta_2 &= 0, \\ \lambda\eta_1 + \mu\eta_3 &= 0, \end{aligned} \tag{15}$$

we can draw, word for word, exactly the same conclusions. Thus, we see that (15) again yields a family of lines lying entirely on (12) and having the same properties as (13).

But what relation does there exist between the lines of the families (13) and (15)? Let g be any line of (13) and h any line of (15). We claim that *g and h have exactly one point in common.* For if we denote by λ, μ the coefficients of g in (13) and by λ', μ' the coefficients of h in (15), then the matrix of the intersection of g and h, i.e., the matrix of all the four equations together, is

$$\begin{pmatrix} \lambda & \mu & 0 & 0 \\ 0 & 0 & \lambda & \mu \\ \lambda' & 0 & \mu' & 0 \\ 0 & \lambda' & 0 & \mu' \end{pmatrix}. \tag{16}$$

The determinant of (16) is readily found to be 0. On the other hand, since only one number at most of each of the number pairs λ, μ and λ', μ' can be zero, there always exist third-order sub-determinants of (16) which are $\neq 0$. For example, if $\lambda \neq 0$ and $\lambda' \neq 0$, such a sub-determinant may

[8] Since the rank of the matrix $\begin{pmatrix} \eta_0 & \eta_1 \\ \eta_2 & \eta_3 \end{pmatrix}$ is equal to 1, the solution vectors $\{\lambda, \mu\}$ of (13) constitute a one-dimensional linear vector space.

be obtained by striking out the third row and the last column. In every case, the rank of (16) is thus 3, so that g and h have one and only one point in common.

The lines of the families (13) *and* (15) *constitute all of the lines that lie on the surface* (12); *there are no others.* For let k be any line belonging entirely to (12), let P be a point of k, and e the tangent plane[9] at P. Through P there passes one line g of the family (13) and one line h of the family (15). All three of the lines, k, g, h must certainly lie entirely in e.[1] If we can show that e has *only* the points of the lines g and h in common with the surface (12), then we shall have proved that k must be identical with one of these two lines.

Now, every line g' of the family (13) that is *distinct from* g is skew to g, and accordingly cannot lie in e. Consequently, g' has only one point in common with e, and this must necessarily lie on h, because the intersection of g' with h is not empty. However, since the lines of the family (13) contain *all* the points of the surface, no point of the surface (12) other than points of g and h can belong to e, as was to be proved.

Once again, we summarize our results, as follows:

THEOREM 4. *Every surface of the second order in P_3 that can be represented*[2] *in a suitable coordinate system by equation* (12) *has the following properties: The lines that lie entirely in the surface fall into two families, of such a nature that through every point of the surface there passes one and only one line of each of the families. Every two lines of the same family are skew, whereas two lines from different families always intersect each other in a single point. The two lines, one from each family, that pass through a fixed point of the surface determine the tangent plane at that point and, for its part, this plane has no other points in common with the surface than these same two lines.*

Thus a tangent plane can never belong in its entirety to such a surface. Consequently, neither can any plane whatsoever of P_3, for it would then be a tangent plane. We thus have the general result: *A non-degenerate surface of the second order in P_3 can never contain an entire plane.*[3]

---

[9] It is uniquely determined, since we are dealing with a non-degenerate surface.

[1] They are indeed tangents at the point P. Cf. Chap. IX, p. 110.

[2] As we have already pointed out, in *complex* P_3 these surfaces are *all* the non-degenerate surfaces, and in real P_3 they are the surfaces belonging to the class that has the normal form (10).

[3] For surfaces of real P_3 this is a consequence of Theorem 8 of the preceding chapter.

Let us try to gain still further insight into the relation between the two families of lines. Consider a fixed line g of family (13). Every plane e that contains g will be intersected by all the other lines of the same family and must thus contain points of the surface (12) that do not belong to g. Let Q be one such point. The line h of the family (15) that passes through Q also intersects g and must therefore lie entirely in e. Hence e is the tangent plane at the point of intersection of g and h. Of course, h is uniquely determined by e and, conversely, e by h. Thus we see the following:

We obtain a one-to-one correspondence between the pencil of all planes through g *and the family of lines* (15) *if we put into correspondence with every plane through* g *that line of the family* (15) *that lies entirely in the plane.*

Aside from g, let there now be given two other lines g_1, g_2 of the family (13), distinct both from each other and from g. We obtain a one-to-one correspondence between the points of g_1 and the points of g_2 if we let a point of g_1 and a point of g_2 correspond to each other whenever they are the points of intersection of g_1 and g_2 with one and the same line of the family (15). Under this correspondence two associated points always lie in the same plane through g. The correspondence is thus a perspectivity and, as such, a linear projectivity between g_1 and g_2 (cf. Chap. IV, Theorem 1 and p. 59). This is customarily expressed as follows:

The lines of family (15) *intersect the lines of family* (13) *in a perspective set of points.*

Of course, the family (13) does the same with respect to the family (15).

Now let us consider also the two pencils of planes b_1, b_2 having the two lines g_1, g_2 as kernels. A one-to-one correspondence is also established between b_1 and b_2 if we always let two planes correspond to each other whenever they pass through one and the same line of the second family (15). Two corresponding planes intersect g in the same point, and therefore *the correspondence between* b_1 *and* b_2 *is again, by Chap.* IV, *Theorem* 1, *a linear projectivity.*

What we have proved suggests the following method of generating surfaces of the second order. If we start with two arbitrary skew lines g_1, g_2 and prescribe a linear projectivity between g_1 and g_2 and then connect by a straight line the points that correspond under this projectivity, we obtain a family of straight lines. We maintain that *this family of straight lines spans a non-degenerate surface of the second order.*

To prove this, consider three different pairs of points that correspond to each other under the given linear projectivity, say: Q_1, Q_1^*; Q_2, Q_2^*; Q_3, Q_3^* (Fig. 15). Denote the line through Q_i, Q_i^* by h_i. Then every two of the three lines h_1, h_2, h_3 are skew.[4] Thus, in particular, the four points Q_1, Q_1^*, Q_3, Q_3^* do not lie in a plane. Furthermore, a point E on h_2 distinct from Q_2 and Q_2^*, lies in none of the four planes determined by any three of the points Q_1, Q_1^*, Q_3, Q_3^*.[5]

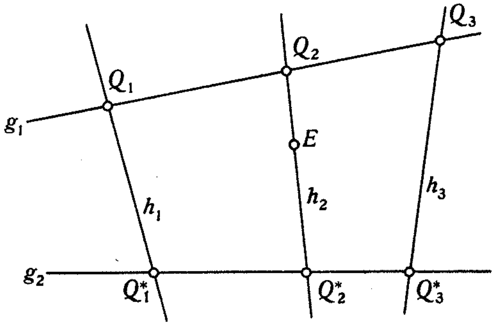

Fig. 15

We can therefore choose the points Q_1, Q_1^*, Q_3, Q_3^* as the fundamental points and E as the unit point of a projective coordinate system. For the sake of definiteness, let us set

$$Q_0' = Q_1,\ Q_1' = Q_1^*,\ Q_2' = Q_3,\ Q_3' = Q_3^*.$$

Let us now consider, in the coordinate system $(Q_0', Q_1', Q_2', Q_3' \mid E)$, the surface of the second order defined by the equation (12). On this surface lie all the points for which $\eta_0 = \eta_1 = 0$, i.e., the points of the line through Q_2', Q_3', namely h_3. Likewise, on this surface lie all the points with $\eta_2 = \eta_3 = 0$, i.e., h_1; in addition, all the points with $\eta_1 = \eta_3 = 0$, i.e., g_1; and, finally, g_2 as well. Furthermore, E lies on the surface, so that the line h_2 has the three points Q_2, E, Q_2^* in common with the surface and consequently must lie entirely in the surface.

All the lines of Fig. 15 thus lie on the surface (12). Now the lines h_1, h_2, h_3 (since each is skew to the other two) must belong to one of the two families of lines on (12) and g_1, g_2 must belong to the other. In addition to this, we know that the family of lines to which the h_i belong intersect g_1 and g_2 in perspective point sets. This perspectivity, how-

[4] For if h_1, h_2, say, lay in a plane, then g_1 and g_2 would necessarily belong to that same plane, contrary to the assumption that g_1, g_2 are skew.

[5] For example, the plane determined by Q_1, Q_1^*, Q_3 can have but one point in common with h_2. But this point is Q_2, and hence cannot be E.

ever, carries the points Q_1, Q_2, Q_3 into Q_1^*, Q_2^*, Q_3^* and accordingly, as a *linear* projectivity, must be identical with the *given* projectivity (Chap. V, Theorem 5). Thus, our assertion is proved.

The following theorem can be directly reduced to this last result.

Let there be given two pencils of planes, b_1, b_2, with skew carriers g_1, g_2. Let there also be given a linear projectivity between b_1 and b_2. *Then the family consisting of all the lines of intersection of two corresponding planes spans a non-degenerate surface of the second order.*[6]

Indeed, it is easy to see that these lines of intersection give rise to a linear projectivity between g_1 and g_2. For if the plane e of b_1 corresponds under the given projectivity to the plane e^* of b_2, then the line of intersection of e and e^* must pass through the point of intersection Q of e^* and g_1 as well as the point of intersection Q^* of e and g_2. Now, the mappings $e \rightleftarrows Q^*$ and $e^* \rightleftarrows Q$ are perspectivities. And since $e \rightleftarrows e^*$ is a linear projectivity, it follows immediately from Chap. IV, Theorem 1 that the mapping $Q \rightleftarrows Q^*$ is likewise a linear projectivity.

We should like to make here the following additional observation: Let three lines g_1, g_2, g_3 be given, each skew to the others. All the lines that intersect g_1, g_2, g_3 simultaneously can be obtained in the following way. Consider a plane e through g_3; it intersects g_1 in the point Q, say, and g_2 in the point Q^*. Then the line QQ^* also intersects g_3 (because it lies entirely in e). The correspondence $Q \rightleftarrows Q^*$ is a perspectivity between g_1 and g_2. This shows the following:

If g_1, g_2, g_3 are three lines, each skew to the others, then the family of all lines that simultaneously intersect g_1, g_2, g_3 spans a non-degenerate surface of the second order.

The lines g_1, g_2, g_3 themselves belong to the second family of lines of this surface; every line of this family can be obtained by taking the common intersector of three arbitrary lines of the first family.

The Determinacy of the Equation of a Hypersurface of the Second Order

In conclusion, let us clear up the question of when the equation of a hypersurface of the second order of P_n is '*essentially*' unique, i.e., determined up to a common factor of all the coefficients. It should be

[6] This theorem is the dual of the preceding one.

observed at the outset that the property of determining its equation in this sense is a projective property of the hypersurface of the second order and is therefore applicable to *all* the hypersurfaces of one class (of projectively equivalent hypersurfaces) whenever it holds true for any single example of the class. For consider a hypersurface F_1 which is represented in a fixed projective coordinate system by the two essentially[7] different equations $(\xi)'A(\xi) = O$ $(\xi)'B(\xi) = O$. Let the linear collineation $(\xi) = \varrho T(\xi^*)$ take F_1 into F_2. Then F_2 is represented, in the same coordinate system, both by the equation $(\xi^*)' T' A T(\xi^*) = O$ and by the equation $(\xi^*)' T' B T(\xi^*) = O$. These two equations are again, however, essentially different.[8]

To settle our question, therefore, we need only consider one representative for each class. Let us now investigate the equation

$$(17) \quad \xi_0^2 + \xi_1^2 + \cdots + \xi_k^2 - \xi_{k+1}^2 - \cdots - \xi_r^2 = 0, \qquad k+1 \geqq r-k.$$

For $k = r \geqq 1$ this equation will represent, in *real* P_n, a hypersurface which is also given by, for example, every equation $\xi_0^2 + \xi_1^2 + \cdots + c\,\xi_k^2 = 0$ with $c > 0$. Thus, the hypersurface in question can in these cases be represented in real P_n by essentially different equations. We claim, however, that these are the only cases of this kind, i.e., that the following theorem holds:

THEOREM 5. *In complex P_n the equation $\sum_{i,k=0}^{n} a_{ik}\,\xi_i\,\xi_k = 0$ of a hypersurface of the second order is always determined up to a common nonvanishing factor of all the coefficients. This is the case in real P_n only for the hypersurfaces of those classes that can be represented by an equation* (17) *in which either $r = 0$ or $k < r$.*

In order to prove this, we have to compare a fixed equation of the form (17) with the equation

$$(18) \qquad \sum_{i,k=0}^{n} a_{ik}\,\xi_i\,\xi_k = 0$$

under the assumption that (17) and (18) represent the same hypersurface of the second order in a fixed projective coordinate system. It then

---

[7] This means, of course, that the matrices A and B do not differ simply by a numerical factor.

[8] For it would follow from an equation of the form $T'AT = \varrho T'BT$ that $A = \varrho B$.

follows, first of all, that *in* (18) *whenever at least one of the subscripts i, k is greater than r, we must always have* $a_{ik} = 0$. For assume that a_{jh} were $\neq 0$ and that $h > r$. From the fact that $h > r$ it follows that the fundamental point with the coordinates

$$\xi_h = 1, \xi_0 = \xi_1 = \cdots = \xi_{h-1} = \xi_{h+1} = \cdots = \xi_n = 0$$

is a double point of (17) (cf. Chap. VIII, Theorem 2). On the other hand, the double points of (18), which must be precisely the same[9] as those of (17), are given by the equations

$$\sum_{k=0}^{n} a_{ik}\, \xi_k = 0, \qquad i = 0, 1, \cdots, n.$$

But if we had $a_{jh} \neq 0$, then the $(j+1)$-st of these equations would not be satisfied for

$$\xi_h = 1, \ \xi_0 = \xi_1 = \cdots = \xi_{h-1} = \xi_{h+1} = \cdots = \xi_n = 0.$$

Thus, we see that equation (18) must necessarily be of the form

$$\sum_{i,k=0}^{r} a_{ik}\, \xi_i\, \xi_k = 0. \tag{19}$$

Furthermore, we must have $a_{ii} \neq 0$ in (19) for $i = 0, 1, \ldots, r$. For if we had $a_{hh} = 0$ $(0 \leqq h \leqq r)$, then

$$\xi_h = 1, \ \xi_0 = \cdots = \xi_{h-1} = \xi_{h+1} = \cdots = \xi_n = 0$$

would be a solution of (19) without being a solution of (17).

For $r = 0$, (17) has the form $\xi_0{}^2 = 0$, and we are through. In the contrary case, we have in real P_n only the equations with $k < r$ to consider. We can always make this assumption in complex P_n as well, for in this latter space there exists, in every class for which $r > 0$, a hypersurface of the form (17) with $k < r$.

---

[9] For, the definition of a double point as a point which is such that each line through it is a tangent, is independent of the form of the equation.

Thus, from now on we may assume $0 < k < r$ in (17). Let us now choose two fixed indices j, h such that $0 \leqq j \leqq k$; $k+1 \leqq h \leqq r$. We claim that for two such indices we have

$$a_{jh} = 0, \tag{20}$$

$$a_{jj} = -a_{hh}. \tag{21}$$

For, the two points[1] with the coordinates $\xi_j = 1$, $\xi_h = \pm 1$, $\xi_i = 0$ for $i \neq j, h$ satisfy equation (17). They must therefore also satisfy (19), which is possible only if the equations

$$a_{jj} + 2\,a_{jh} + a_{hh} = 0,$$
$$a_{jj} - 2\,a_{jh} + a_{hh} = 0$$

hold. By subtracting these equations, we first obtain (20) and then, immediately, (21).

If we hold the index h in (21) fixed while letting j run from 0 to k, we see that all a_{jj} for $j = 0, 1, \ldots, k$ are equal to one another; denote their common value by, say, $c \neq 0$. Then all the a_{hh} for $h = k+1, k+2, \ldots, r$ are equal to $-c$. Thus,

$$a_{00} = a_{11} = \cdots = a_{kk} = c, \tag{22}$$

$$a_{k+1,k+1} = a_{k+2,k+2} = \cdots = a_{rr} = -c. \tag{23}$$

It only remains to show that $a_{jh} = 0$ also when j, h are both $\leqq k$ or both $> k$. To see this, say, in the first case, consider the point having the coordinates $\xi_j = 1$, $\xi_h = 1$, $\xi_r = +\sqrt{2}$, but $\xi_i = 0$ for all $i \neq j, h, r$. These coordinates satisfy (17) and hence must also satisfy (19). Upon substituting in (19) it follows at once, taking into consideration what we have already proved above, that $a_{jh} = 0$. We have thus shown that (19) has the form

$$c\,(\xi_0^2 + \xi_1^2 + \cdots + \xi_k^2 - \xi_{k+1}^2 - \cdots - \xi_r^2) = 0,$$

as was to be proved.

---

[1] Since the points used here have *real* coordinates, our conclusion holds both for complex P_n and for real P_n. The same is true further on.

Exercises

1. In the coordinate system $(Q_0, Q_1, \ldots, Q_n \mid E)$ of P_n the linear space of dimension k determined by the $k+1$ points $Q_0, Q_1, \ldots, Q_k$ can be looked upon as a projective P_k whose points are given by the coordinate $(k+1)$-tuple $\xi_0, \xi_1, \ldots, \xi_k$. Now let there be given a non-degenerate hypersurface of the second order in this P_k (that is, a hypersurface of rank $k+1$); call it H. If we now construct in P_n the $(n-k)$-dimensional spanning space through a point of H and the $n-k$ fixed points $Q_{k+1}, Q_{k+2}, \ldots, Q_n$, then the totality of the points of all such spanning spaces constitute a hypersurface of the second order of rank $k+1$ in P_n. Every hypersurface of the second order in P_n of rank $k+1$ can be generated in this way.

2. Show that in both real and complex P_n if a linear space of dimension p lies on a hypersurface of the second order, the hypersurface is at most of rank $2(n-p)$.

3. Let F be a hypersurface of the second order in P_n. Let $Q_0, Q_1, \ldots, Q_k$ be $k+1$ points not on F with the property that the polar of each point with respect to F contains the other k points. Show that if $k+1$ is less than the rank of F, an additional point Q_{k+1} can always be found, lying in the intersection of the polars of all the Q_i, which lies neither on F nor in the spanning space of the points $Q_0, Q_1, \ldots, Q_k$. The polar of Q_{k+1} passes through $Q_0, Q_1, \ldots, Q_k$.

This fact makes possible, for example, the construction of a so-called *polar-simplex* with respect to a given non-degenerate F, i.e., $n+1$ linearly independent points of such a nature that the hyperplane through any n of them is the polar of the remaining point. The first point may be chosen as any arbitrary point not on F; the others are then restricted by the rule given above.

What is the equation of a non-degenerate hypersurface of the second order in a coordinate system whose fundamental points form a polar simplex?

4. Let a linear collineation of P_n be associated with an ordered pair F_1, F_2 of non-degenerate hypersurfaces of the second order in the following way. From a point P we pass to its polar h with respect to F_2 and then determine the pole P^* of h with respect to F_1. The mapping $P \rightarrow P^*$ is then the linear projectivity in question. In *complex* P_n the following is true: If F_1, F_2 and F_1^*, F_2^* are two pairs of non-degenerate hypersurfaces of the second order, then there exists a linear projectivity which simultaneously takes F_1 into F_1^* and F_2 into F_2^* if and only if the two collineations associated with the pairs are projectively equivalent.

Hint: Use Exercise 3 of Chapter IX and Theorem 5 of the present chapter.

The theorem does not hold in general in real P_n. Cf. Exercise 5 of Chapter IX.

CHAPTER XI

THE AFFINE CLASSIFICATION OF HYPERSURFACES OF THE SECOND ORDER

In Chapter IX we answered the question of when two hypersurfaces of the second order could be transformed into each other by a linear projectivity of P_n. In the present chapter we wish to pose a question of the same kind, in which a restriction is placed on the allowable projectivities. Specifically, we shall now admit only a subgroup of the linear projective group, namely the affine group (Chap. VI). Our question will give rise, just as in Chap. IX, to a partition of the hypersurfaces of the second order into classes, but a partition that is finer than the one we then obtained and from which we shall gain new knowledge about our hypersurfaces.

Let us recall at the outset that by an affine transformation in P_n we mean a linear projectivity of P_n onto itself that maps the improper hyperplane into itself[1] (Chap. VI). Since an affine transformation is accordingly nothing other than a special linear projectivity, it always takes a hypersurface of the second order into a hypersurface of the second order. If the hypersurface F_1 is mapped into the hypersurface F_2 by a certain affine transformation, the *inverse* affine transformation maps F_2 into F_1.

We say that two hypersurfaces of the second order that can be mapped into each other by a suitable affine transformation are **affinely equivalent**. Obviously, every hypersurface of the second order is affinely equivalent to itself. Furthermore, the concept 'affinely equivalent' is transitive, i.e., if F_1 is affinely equivalent to F_2 and F_2 to F_3, then F_1 is also affinely equivalent to F_3. This is an immediate consequence of the group property of affine transformations.

[1] Again we here expressly emphasize (cf. footnote 4 on p. 97) that in this and the following chapter we always mean by an affine transformation a one-to-one mapping of P_n onto itself. The mapping induced in affine R_n by such a mapping is thus always a *non-singular* affine transformation in the sense of *Modern Algebra*, p. 180.

As in Chap. IX, we now obtain a partition into classes. We call a non-empty set C of hypersurfaces of the second order a *class of affinely equivalent hypersurfaces* if C possesses the following two properties:

a) Every two hypersurfaces of C are affinely equivalent.

b) A hypersurface in C is never affinely equivalent to a hypersurface not in C.

We may also interpret b) as: Every hypersurface of the second order that is affinely equivalent to a hypersurface in C, itself belongs to C.

It is easy to show, just as in Chapter IX, that this partition into classes is disjoint, i.e., that every hypersurface of the second order belongs to one and only one class. We again set ourselves the problem of determining all of the classes.

Determination of the Classes

Let

(1)
$$\sum_{i,k=0}^{n} a_{ik}\,\xi_i\,\xi_k = 0$$

be a hypersurface of the second order in P_n. Let the coordinate system be chosen in such a way that the equation of the improper hyperplane is $\xi_0 = 0$. Then all the improper points of the hypersurface (1) satisfy the equation

(2)
$$\sum_{i,k=1}^{n} a_{ik}\,\xi_i\,\xi_k = 0.$$

This is a homogeneous quadratic equation in the n variables $\xi_1, \xi_2, \ldots, \xi_n$ only. Since the improper hyperplane is nothing but a projective P_{n-1} consisting[2] of the homogeneous n-tuples $\xi_1, \xi_2, \ldots, \xi_n$, the equation (2), again, in general represents a hypersurface of the second order in this P_{n-1}. By 'in general,' we mean, aside from the case in which all the coefficients of (2) vanish. In order not to have to make special mention of this latter case each time, we shall speak of the 'hypersurface' (2) in this case also, meaning thereby the entire improper plane.[3]

---

[2] As long as we are speaking only of the improper points of P_n, we can simply drop the first coordinate ξ_0, since it is zero in our coordinate system for all improper points.

[3] For example, if we take $n = 3$, then (1) represents a surface in P_3 and (2), in general, a conic section in the improper plane of P_3. If in particular, however, (1) is a pair of planes one of which is the improper plane itself, then we have the situation in which all the coefficients of (2) vanish.

Hypersurface (2), considered as a hypersurface of a space P_{n-1}, has its own rank, namely, the rank of the matrix

$$(3) \qquad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix};$$

the significance of the rank s of this matrix in relation to (2), is one that we already know (Chap. VIII): namely, that $(n-1)-s$ is the dimension of the set of double points of (2).[4] Such a double point of (2) need not necessarily, however, be a double point of (1). For a double point of (2), regarded as a hypersurface of the improper space P_{n-1}, is just any point of (2) at which every improper[5] line of P_n is tangent to (2). Needless to say, however, not every proper line of P_n through such a point need be tangent to (1).[6] Nevertheless, every improper line that is tangent to (2) is also tangent to (1), and we thus have a geometrical interpretation for the rank s of (3) which pertains *solely to* (1), namely:

$n-1-s$ is the dimension of the linear space of all the *improper* points Q of the hypersurface (1) having the property that every *improper* line through Q is tangent to (1).

Let there now be given an affine transformation of P_n, and let us apply this transformation to (1). Since an affine transformation, being a projectivity, takes tangents into tangents and, being affine, also takes improper elements (points, lines, etc.) into improper elements, it follows from the meaning of s just given, that s is invariant under affine transformations.

We could also deduce the invariance of s, independently of the meaning it has for (1), as follows: An affine transformation of P_n induces in the improper hyperplane a linear projectivity of this hyperplane onto itself. Hence, if we have two affinely equivalent hypersurfaces of the second order in P_n, their improper parts, considered as hypersurfaces of the improper space P_{n-1}, are projectively equivalent. These improper parts, as hypersurfaces of P_{n-1}, must accordingly have the same rank.

---

[4] In the case that all the coefficients of (2) vanish, we consider every point of the improper plane as a double point of (2).

[5] By an improper linear space we mean (Chap. I) a linear space of P_n consisting solely of improper points.

[6] For example, if (1) is a pair of planes in P_3 *one* of which is the improper plane itself, then the totality of double points of (2) is the entire improper plane, whereas the totality of double points of (1) is only the line of intersection of the two planes.

Besides this invariant s, we shall also have to take into consideration, of course, the rank r of (1) itself. Moreover, in real P_n, we shall also have to bring into the picture the invariant k of Chap. IX, Theorem 8, and the corresponding quantity h for (2). Thus, k is the maximal dimension of those linear spaces of P_n that have no point in common with (1), whereas h is the maximal dimension of the *improper* linear spaces that have an empty intersection with (1) (or with (2)). The notation just given for all four invariants will be adhered to in the sequel.

We shall show that in complex P_n, r and s already constitute a complete system of invariants, i.e., r and s determine uniquely the classes of affinely equivalent hypersurfaces of the second order. In real P_n, r and s do not suffice, and we obtain a complete system of invariants only by taking all four of the quantities r, s, k, h. We shall carry through the proof of these facts by again transforming the equations of the hypersurfaces of the second order into certain normal forms.

In making the transformation into normal form, we of course make use, in accordance with our present problem, of affine transformations only. As we know from Chap. VI, in a coordinate system in which $\xi_0 = 0$ is the improper hyperplane, every affine transformation can be represented by a system of equations of the form

$$\eta_0 = \xi_0, \tag{4}$$
$$\eta_i = \sum_{k=0}^{n} t_{ik}\, \xi_k, \qquad i = 1, 2, \cdots, n.$$

We always have $\xi_0 = \eta_0 = 0$ for an *improper* point of P_n; hence, in order to observe the effect of our affine transformation on the improper points, we need only know how to compute the last n coordinates $\eta_1, \eta_2, \ldots, \eta_n$ of an improper image point, given the last n coordinates $\xi_1, \xi_2, \ldots, \xi_n$ of its improper pre-image. By (4), this is done by means of the equations

$$\eta_i = \sum_{k=1}^{n} t_{ik}\, \xi_k, \qquad i = 1, 2, \cdots, n. \tag{5}$$

Therefore we may regard equations (5) as the equations of the linear projectivity induced in the improper hyperplane by the affine transformation (4).

We need now only apply substitution (5) to equations (2) in order to see what effect the affine transformation (4) has on the improper portion of hypersurface (1). Equation (2) is then transformed into the equation of the image (itself improper) of the improper portion of (1).

We know from Chapter IX that we can always find a non-singular substitution of the type (5) that transforms equation (2) into the form[7]

$$\eta_1^2 + \eta_2^2 + \cdots + \eta_j^2 - \eta_{j+1}^2 - \cdots - \eta_s^2 = 0, \tag{6}$$

where s is the rank of matrix (3).[8] In the complex case, we can always make $j = s$ in (6), whereas in the real case j can take any value from 1 to s,[9] and it is even possible for only negative squares to appear in (6).

Now assume the ξ_i in (5) so chosen that, as a result of applying (5), (2) goes over into (6). Then we adjoin to the equations (5) the one further equation

$$\eta_0 = \xi_0. \tag{7}$$

Taken together, (5) and (7) define an affine transformation in P_n that we should like to apply to (1). We know the effect of the linear substitution represented by (5) and (7) upon the terms in (1) that do not contain ξ_0 (in other words, upon the terms comprising (2)). As for the remaining terms, which do contain ξ_0, the application to them of this substitution gives rise only to terms in which η_0 appears. We thus see that the affine transformation represented by (5) and (7) takes (1) into a hypersurface whose equation is

$$\eta_1^2 + \eta_2^2 + \cdots + \eta_j^2 - \eta_{j+1}^2 - \cdots - \eta_s^2 = b\,\eta_0^2 + 2\sum_{\nu=1}^{n} b_\nu\, \eta_0\, \eta_\nu. \tag{8}$$

---

[7] In order to actually carry out the substitution, we must, of course, first solve (5) for $\xi_1, \xi_2, \ldots, \xi_n$ and then substitute.

[8] If $s = 0$, (2) is already of this form. Accordingly we can, if we wish, assume at the outset that $s > 0$. For, the conclusion that is to be drawn from what has been said (i.e., the possibility of transforming (1) into (8)) is trivial for $s = 0$.

[9] In Chapter IX, multiplication of (6) if necessary by -1 and the making of a further affine transformation enabled us to bring about that the number of positive squares should not be less than the number of negative ones. It serves our purpose here, however, to make no use of this possibility for the time being.

Here, b, b_ν are constants which hold no special interest for us.

In this equation we can also make the terms $b_\nu \eta_0 \eta_\nu$ vanish for $\nu = 1, 2, \ldots, s$. For we can write (8) as

$$(\eta_1 - b_1 \eta_0)^2 + (\eta_2 - b_2 \eta_0)^2 + \cdots + (\eta_j - b_j \eta_0)^2$$
$$-(\eta_{j+1} + b_{j+1}\eta_0)^2 - \cdots - (\eta_s + b_s \eta_0)^2 = c \cdot \eta_0^2 + 2 \sum_{\nu=s+1}^{n} b_\nu \eta_0 \eta_\nu,$$

where

$$c = b + b_1^2 + b_2^2 + \cdots + b_j^2 - b_{j+1}^2 - \cdots - b_s^2 .$$

This shows at once that the affine transformation

$$\begin{aligned} &\zeta_\nu = \eta_\nu && \text{for } \nu = 0 \text{ and } \nu = s+1, s+2, \cdots, n, \\ (9) \quad &\zeta_\nu = \eta_\nu - b_\nu \eta_0 && \text{"} \quad \nu = 1, 2, \cdots, j, \\ &\zeta_\nu = \eta_\nu + b_\nu \eta_0 && \text{"} \quad \nu = j+1, \cdots, s \end{aligned}$$

takes the hypersurface (8) into

$$(10) \quad \zeta_1^2 + \zeta_2^2 + \cdots + \zeta_j^2 - \zeta_{j+1}^2 - \cdots - \zeta_s^2 = c \cdot \zeta_0^2 + 2 \sum_{\nu=s+1}^{n} b_\nu \zeta_0 \zeta_\nu .$$

Further treatment of this equation calls for the distinguishing of various cases. As our first case, let us consider the possibility that $c = b_{s+1} = b_{s+2} = \ldots = b_n = 0$ and thus that all the terms except the first s squares vanish. In complex P_n, where $j = s$, this is already the final normal form. In the case of the reals we are likewise through if $j \geqq s - j$. Otherwise, just as in the proof of Theorem 5 of Chapter IX, we can multiply (10) through by -1, make a further affine transformation of P_n (a permutation of the variables, with ζ_0 held fixed), and thus see to it that the condition $j \geqq s - j$ is always satisfied.

As a second possibility, let us assume that not all the b_ν in (10) vanish and that c is arbitrary. For simplicity, we may assume that $b_{s+1} \neq 0$, as this can be achieved by a suitable relabelling of the variables. Moreover, in the reals, we may also assume that $j \geqq s - j$, which can always be achieved in the way just described. We now define an affine transformation in P_n by

$$(11)\qquad \begin{aligned} \zeta'_\nu &= \zeta_\nu \text{ for } \nu = 0, 1, \cdots, s \text{ and } \nu = s+2, \cdots, n, \\ \zeta'_{s+1} &= \frac{c}{2}\zeta_0 + b_{s+1}\zeta_{s+1} + b_{s+2}\zeta_{s+2} + \cdots + b_n\zeta_n, \end{aligned}$$

which reduces[1] equation (10) to

$$(12)\qquad \zeta_1'^2 + \cdots + \zeta_j'^2 - \zeta_{j+1}'^2 - \cdots - \zeta_s'^2 = 2\,\zeta_0'\,\zeta_{s+1}'.$$

This is sufficient for our purposes.

There remains to be considered the third, and last, case, in which all the $b_\nu = 0$ in (10) (for $\nu = s+1, \ldots, n$), but $c \neq 0$. In the reals, it is now more advantageous *not* to stipulate that $j \geqq s - j$ (for this would necessitate a new breakdown into cases, according to the sign of c). Instead of this, we now avail ourselves in the real case of the possibility of multiplying equation (10) by -1 to insure that c is always positive, so that we can assume this to be true in (10) at the outset. In the complex case, however, we shall assume, as before, that $j = s$, and we leave c unaltered. Then the equations

$$\begin{aligned} \zeta_0' &= \sqrt{c}\cdot\zeta_0, \\ \zeta_\nu' &= \zeta_\nu \text{ for } \nu \neq 0 \end{aligned}$$

define, in both the real and the complex cases, an affine transformation[2] which takes (10) into

$$(13)\qquad \zeta_1' + \zeta_2'^2 + \cdots + \zeta_j'^2 - \zeta_{j+1}'^2 - \cdots - \zeta_s'^2 = \zeta_0'^2,$$

and we have thus reduced these hypersurfaces as well to a normal form.

In order to get a clear picture of the situation, let us summarize what we have proved thus far.

---

[1] For example, if (11) be substituted in (12), we obtain (10). The substitution (11) is non-singular, whence the necessity of the assumption $b_{s+1} \neq 0$.

[2] We may take either of the two values of the root.

In complex P_n every hypersurface of the second order is affinely equivalent to at least one of the hypersurfaces[3] *represented by the following forms of equation:*

$$\xi_1^2+\xi_2^2+\cdots+\xi_s^2=0, \tag{14}$$

$$\xi_1^2+\xi_2^2+\cdots+\xi_s^2=2\,\xi_0\,\xi_{s+1}, \tag{15}$$

$$\xi_1^2+\xi_2^2+\cdots+\xi_s^2=\xi_0^2. \tag{16}$$

In real P_n, however, every hypersurface of the second order is affinely equivalent to at least one of the following hypersurfaces:[4]

$$\xi_1^2+\xi_2^2+\cdots+\xi_j^2-\xi_{j+1}^2-\cdots-\xi_s^2=0, \qquad j\geqq s-j, \tag{17}$$

$$\xi_1^2+\xi_2^2+\cdots+\xi_j^2-\xi_{j+1}^2-\cdots-\xi_s^2=2\,\xi_0\,\xi_{s+1}, \qquad j\geqq s-j, \tag{18}$$

$$\xi_1^2+\xi_2^2+\cdots+\xi_j^2-\xi_{j+1}^2-\cdots-\xi_s^2=\xi_0^2, \qquad 0\leqq j\leqq s. \tag{19}$$

In equations (14) and (17), any of the numbers 1, 2, ..., n is an admissible value for s. In (15), (16), and (18), (19) it can of course happen that all the terms on the left-hand side vanish; this happens when the rank s of (3) is zero. On the other hand, however, equations of the form (15) and (18) cannot occur if $s=n$.

Thus far, we have only shown that equations of the forms (14) through (16) and (17) through (19), in the real and complex cases respectively, contain representatives of all the classes of affinely equivalent hypersurfaces of the second order. We do not yet know, however, how often each class occurs among them. What we now wish to show is that each class actually occurs only once, i.e., that any two different equations of the forms (14) through (16) and (17) through (19), respectively, always belong to different classes.

We first prove this for complex P_n, that is, for equations (14) through (16). To this end, we use the invariants r and s. The invariant s, i.e., the rank of the matrix (3), is, for each of the equations (14) through (16), just the number of squares on the left-hand side of the equation. As for the invariant r, i.e., the rank of the full matrix of the equation, we may compute it directly, obtaining:

for an equation of the form (14), $r=s$;
for an equation of the form (15), $r=s+2$;
for an equation of the form (16), $r=s+1$.

[3] We again employ ξ for the variables.

[4] $j=0$ (in (19)) means that the left-hand side of (19) consists solely of negative terms.

From this it follows at once that two distinct equations from among (14) through (16) can never coincide in both invariants r and s, and thus can never belong to the same class of affinely equivalent hypersurfaces of the second order. This shows at the same time that in complex P_n, r and s together constitute a complete system of invariants for hypersurfaces of the second order with respect to the affine group of P_n. For to any two possible[5] fixed pair of values of r, s there belongs just *one* equation of type (14) through (16) and hence just *one* class of affinely equivalent hypersurfaces.

To show the same for equations (17) through (19) with respect to real P_n, we first observe that for equation (17) we must have $r = s$; for (18), $r = s + 2$; and for (19), $r = s + 1$. All that this implies is that two equations of type (17) through (19) can coincide in the two invariants r and s only if the equations are either both of type (17), both of type (18), or both of type (19).

Now, in order to show that two *different* equations, both of form (17), must belong to different classes, we must have recourse to the invariants k or h mentioned earlier (p. 160). Let us take, say, k. From Theorem 8 of Chap. IX,[6] it follows that $k = j - 1$. Thus, two *different* equations (17) can not have the same k and consequently can not be affinely equivalent.

For the sake of variety, in the case of equations of the kind (18), let us take the invariant h. Since we obtain for the improper part of a hypersurface given by (18) exactly the same equation as in the case of (17) (namely, the left-hand side of (18) set equal to zero), it follows that for (18) $h = j - 1$ if $s \neq 0$.[7] Thus, two *different* equations of the form (18) as well, cannot represent affinely equivalent hypersurfaces.

It remains to prove the same for equations of the form (19). Here it does not suffice to deal with just one of the invariants k, h; we must consider both. We can again determine them by Theorem 8 of Chapter IX, but we must now keep in mind that in that theorem the number of positive

[5] As we have just seen, we necessarily have $0 \leqq r - s \leqq 2$. Moreover, in the case $s = 0$, the only possible values of r are 1 or 2, and for $s = n$ the only possible values are $r = n$ or $r = n + 1$.

[6] For k is equal to the number of positive squares in the normal form, diminished by one.

[7] We need only consider $s \neq 0$. For, the case $s = 0$ in equations of type (18) is *characterized* by the fact that every improper point belongs to the hypersurface. Consequently, $0 = 2\xi_0\xi_1$ cannot be affinely equivalent to any of the remaining equations (18).

squares in the normal form was assumed to be at least as great as the number of negative ones.[8] In order to take this into account, let us think of the squares that appear in the equations under consideration as always being so transposed that the number of positive squares is not less than the number of negative ones. For the cases in which $s \neq 0$,[9] it is easily computed that

$$(20)\qquad \begin{aligned} k &= s - j \qquad \text{for } j \leqq \frac{s}{2}, \\ k &= j - 1 \qquad \text{" } j > \frac{s}{2}; \end{aligned}$$

$$(21)\qquad \begin{aligned} h &= s - j - 1 \qquad \text{" } j < \frac{s}{2}, \\ h &= j - 1 \qquad \text{" } j \geqq \frac{s}{2}. \end{aligned}$$

From this it follows that

$$(22)\qquad \begin{aligned} k - h &= 1 \qquad \text{for } j \leqq \frac{s}{2}, \\ k - h &= 0 \qquad \text{" } j > \frac{s}{2}. \end{aligned}$$

Now, consider two equations (19) with two different j, say $j = j_1$ in one case and $j = j_2$ in the other ($j_1 \neq j_2$). Let us assume that both hypersurfaces agree in their invariant k. Then it follows from (20) that the numbers j_1, j_2 cannot both be $\leqq s/2$ nor both $> s/2$. Thus, let $j_1 \leqq s/2$, $j_2 > s/2$, say. Then (22) shows that the invariant h is different for the two hypersurfaces. Hence the hypersurfaces cannot be affinely equivalent.

Thus, we have now determined the classes in the reals as well. In addition, we see that in every case the class is uniquely determined by specifying the values of the four invariants r, s, k, and h.

Gathering our results together once more, we state them as follows:

In complex P_n there exist precisely as many classes of affinely equivalent hypersurfaces of the second order as there are types of equation (14) *through* (16). *Exactly one representative of each class is to be found*

[8] This was actually made use of in the proof of this theorem. How?

[9] Again, we see in advance that the case $s = 0$ cannot be affinely equivalent to any of the other equations (19).

among these equations. The invariants r and s constitute a complete system of invariants for hypersurfaces of the second order with respect to the affine group of complex P_n.

In real P_n there are as many classes of affinely equivalent hypersurfaces of the second order as there are types of equation (17) *through* (19). *Again, each class is represented exactly once among these equations. The invariants r, s, k, h constitute a complete system of invariants.*

Affine Geometry

Those theorems and properties that remain invariant under the mappings of the linear projective group were designated, in Chapter X, as *projective.* We proceed in exactly the same way with respect to the affine group, referring to statements or properties that are invariant under affine transformations as *affine.* By *affine geometry* we mean, correspondingly, the totality of affine theorems.

Anything that is invariant under all linear projectivities is *a fortiori* also invariant under affine transformations. That is to say, every concept and theorem of projective geometry is also a concept or theorem of affine geometry. But the converse is by no means true. For example, the concept 'parallel' is an affine concept, but not a projective one. For, linear spaces that are parallel do not lose that property if we apply an affine transformation to them; however, they may very well go over into non-parallel linear spaces under a suitable linear projectivity which is not an affine transformation. Similarly, the above invariants s and h of a hypersurface of the second order are affine invariants but not projective invariants.

The study of the cross ratio of four points of a line one of which is an improper point leads to another important affine invariant. If Q_1, Q_2, Q_3, say, are three *proper* points of a line g and Q_u denotes the improper point of g, then we define the *distance ratio* $\mathcal{D}(Q_1\, Q_2\, Q_3)$ of these three points to be the value of the expression

$$\mathcal{D}(Q_1\, Q_2\, Q_3) = \mathcal{R}(Q_1\, Q_2\, Q_3\, Q_u). \tag{23}$$

If we apply an affine transformation which takes the elements g, Q_1, Q_2, Q_3, Q_u into g^*, Q_1^*, Q_2^*, Q_3^*, Q_u^*, respectively, then our fourth point Q_u^* is again an improper point, and it follows that

$$\begin{aligned}\mathcal{D}(Q_1\,Q_2\,Q_3) &= \mathcal{R}(Q_1\,Q_2\,Q_3\,Q_u)\\ &= \mathcal{R}(Q_1^*\,Q_2^*\,Q_3^*\,Q_u^*)\\ &= \mathcal{D}(Q_1^*\,Q_2^*\,Q_3^*).\end{aligned}$$

The distance ratio is thus an affine invariant[1] *of three points on a line.*

To clarify the behaviour of this affine invariant, let us compare *its role in affine geometry* with that of the cross ratio in projective geometry. According to Chapter IX, the classes of projectively equivalent linear quadruples are characterized by the cross ratio itself. Thus, in projective geometry, linear quadruples fall into an infinite number of classes—to be specific, as many classes as there are possible values of the cross ratio. A quite different situation obtains for linear *triples* (i.e., ordered systems of three distinct points on a line). Such a triple can be mapped by a suitable linear projectivity of P_n into any other linear triple.[2] Thus, linear triples constitute only a single class in projective geometry. If we pass over to affine geometry, however, the linear triples now also fall into an infinite number of classes of affinely equivalent triples.

If we first consider the triples that contain improper points, we easily see that those consisting of three improper points constitute one single class, themselves, whereas those containing just one improper point constitute three classes, characterized by the position of the improper point in the triple. This can easily be proved by a refinement of the method of Chap. IX, pp. 129-130. For example, if S_1, S_2, S_3 and S_1^*, S_2^*, S_3^* are two triples to be mapped into each other, of which S_2 and S_2^* are improper points and the rest proper, the coordinate system used on p. 130 must be so chosen that, with the exception of S_1 and S_1^*, all the fundamental points are improper. Then the linear projectivity that takes the one coordinate system into the other automatically is an affine transformation.

Now, for linear triples whose points are all proper, the distance ratio plays the same role as the cross ratio does for quadruples. To be specific, we again have the result that two ordered linear triples of proper points can be carried into each other by an affine transformation of P_n if and only if they have the same distance ratio. *In short: The distance ratio is in this case the invariant that characterizes the classes.* That this condition is necessary is an immediate consequence of the affine invariance of the

[1] It is not, of course, a projective invariant, for an arbitrary linear projectivity need not necessarily take the improper point Q_u of g into the improper point of g^*.

[2] Cf. Chap. IX, p. 130.

distance ratio. That it is also sufficient may be shown as follows: Let S_1, S_2, S_3 and S_1^*, S_2^*, S_3^* be two linear triples of proper points for which $\mathcal{D}(S_1 S_2 S_3) = \mathcal{D}(S_1^* S_2^* S_3^*)$. Further, let S_4 and S_4^* be the respective improper points of the lines containing the triples. Then we also have $\mathcal{R}(S_1 S_2 S_3 S_4) = \mathcal{R}(S_1^* S_2^* S_3^* S_4^*)$. Now let the triple S_2, S_3, S_4 be mapped by an affine transformation, as in the preceding paragraph, into S_2^*, S_3^*, S_4^*. In view of the equality of the cross ratios, S_1 also goes over into S_1^*, which proves our result.

To justify the *name* 'distance ratio,' let us derive an intuitive interpretation that this invariant has in the real case. To this end, we introduce the definition of euclidean length in the proper part of real P_n, i.e., in affine R_n. If, then, Q_1, Q_2, Q_3 is a linear triple of proper points of real P_n, we shall prove that

$$|\mathcal{D}(Q_1 Q_2 Q_3)| = \overline{Q_1 Q_3} : \overline{Q_2 Q_3}, \tag{24}$$

where $\overline{Q_1Q_3}$, $\overline{Q_2Q_3}$ denote euclidean lengths. This formula is in itself sufficient justification for the name. However, in this formula, the sign of the distance ratio is still without meaning. We can take care of this too by considering, in place of the distances, the affine vectors $\overrightarrow{Q_1Q_3}$ and $\overrightarrow{Q_2Q_3}$. For affine vectors, then, we will have the following relation:

$$\overrightarrow{Q_1 Q_3} = \mathcal{D}(Q_1 Q_2 Q_3) \cdot \overrightarrow{Q_2 Q_3}. \tag{25}$$

In euclidean R_n, equation (24) is a direct consequence of (25). We need therefore only prove the latter. To compute the components of the vectors in question we must introduce *non-homogeneous* coordinates for the points Q_1, Q_2, Q_3. We do this by normalizing the first of the $n+1$ homogeneous coordinates to 1. Thus we have, say,

$$Q_1 = [1, x_1, x_2, \cdots, x_n], \qquad Q_2 = [1, y_1, y_2, \cdots, y_n]. \tag{26}$$

The x_i, y_i are then the affine coordinates of these points. Since we are assuming that $Q_1 \neq Q_2$, the affine coordinates of Q_3 can be taken to be of the form $x_i + \lambda(y_i - x_i)$; thus,

$$Q_3 = [1, (1-\lambda)x_1 + \lambda y_1, (1-\lambda)x_2 + \lambda y_2, \cdots, (1-\lambda)x_n + \lambda y_n]. \tag{27}$$

In order to find the improper point Q_u of the line g determined by Q_1, Q_2, we require a linear combination of the coordinate vectors of Q_1, Q_2 whose first component vanishes. This may be had merely by forming the difference of the two coordinate vectors (26). Thus, we can write

(28) $$Q_u = [0, y_1 - x_1, y_2 - x_2, \cdots, y_n - x_n].$$

We then obtain (cf. Chap. IV, p. 53)

(29) $$\mathcal{D}(Q_1 Q_2 Q_3) = \mathcal{R}(Q_1 Q_2 Q_3 Q_u) = \frac{\lambda}{\lambda - 1}.$$

On the other hand, computing the affine vectors $\overrightarrow{Q_1Q_3}$, $\overrightarrow{Q_2Q_3}$ from the last n coordinates of the points (26), (27), we find

(30) $$\begin{aligned} \overrightarrow{Q_1 Q_3} &= \lambda \{y_1 - x_1, y_2 - x_2, \cdots, y_n - x_n\}, \\ \overrightarrow{Q_2 Q_3} &= (\lambda - 1)\{y_1 - x_1, y_2 - x_2, \cdots, y_n - x_n\}. \end{aligned}$$

In conjunction with (29), this immediately yields the validity of equation (25).

As a particular case of (25), if $\mathcal{D}(Q_1 Q_2 Q_3) = -1$,[3] then $\overrightarrow{Q_1Q_3} = \overrightarrow{Q_3Q_2}$; i.e., Q_3 *is the midpoint of the segment* Q_1Q_2. We shall shortly have occasion to make considerable use of this fact.

It is possible to give meaning to equations (24) and (25) in the complex case, by extending to the complex case the definition of euclidean length and the concept of an affine vector. This is more than we need, however. We shall make use only of the meaning of the distance ratio for the special case $\mathcal{D}(Q_1 Q_2 Q_3) = -1$. This meaning, however, can be extended at once to the complex case if we make the following definition: If (26) are any two proper points of complex P_n (x_i, y_i being arbitrary complex numbers), then by *the midpoint of the segment* Q_1Q_2 we shall mean that point Q_3 which is given by (27) for $\lambda = \frac{1}{2}$. *Then in the complex space* P_n *also the equation* $\mathcal{D}(Q_1 Q_2 Q_3) = -1$ *is completely equivalent with the statement that* Q_3 *is the midpoint of the segment* Q_1Q_2.

By virtue of the meaning of the distance ratio in the reals, we see that a large number of theorems of elementary geometry are affine theorems. All the theorems in which only the concepts 'parallel' and 'distance ratio' occur and concepts derivable from them, such as 'midpoint,' 'parallelogram,' etc. are certainly theorems of affine geometry. Examples of such theorems are: 'The diagonals of a parallelogram bisect each other,' and 'The lines joining the midpoints of the successive sides of a quadrilateral form a parallelogram.' It is easy to find further examples of this kind.

---

[3] I.e., Q_1, Q_2 and Q_3, Q_u separate each other harmonically.

We shall now use the meaning obtained for the distance ratio to give a more intuitive meaning to some obvious invariants of hypersurfaces of the second order. In doing so, we restrict ourselves to non-degenerate hypersurfaces. Let us recall the polarity that is associated with every such hypersurface; to every hyperplane of P_n it assigns a point, the pole. In particular, this is also true for the improper hyperplane. Now, the pole of the improper hyperplane can itself be either an improper or a proper point. In the first case, the improper hyperplane is a tangent hyperplane to the given hypersurface of the second order; in the second case, it is not. It is obvious that the property of having or not having the improper hyperplane as a tangent hyperplane is an affine invariant of a hypersurface of the second order.

The (non-degenerate) hypersurfaces that have the improper hyperplane as a tangent hyperplane are called *paraboloids*. Which of the affine normal forms (14) through (19) represent paraboloids? The normal forms (14) and (17) represent only degenerate hypersurfaces. In the case of the non-degenerate hypersurfaces of the types (16) and (19) (i.e., those for which $s = n$), the pole of the improper hyperplane is a proper point, namely the point $\xi_0 = 1$, $\xi_i = 0$ for $i = 1, 2, \dots, n$. There remain only the non-degenerate normal forms of types (15) and (18), for which $s = n - 1$. All of these latter are paraboloids. For, the improper point $\xi_n = 1$, $\xi_0 = \xi_1 = \cdots = \xi_{n-1} = 0$ has the hyperplane $\xi_0 = 0$ as its polar, i.e., the improper hyperplane is a tangent hyperplane at exactly that point.

As we have just ascertained, for the hypersurfaces that belong to the classes represented by the normal forms (16) and (19) the pole of the improper hyperplane is proper. The pole in these cases is not, of course, a point of the hypersurface. Let F be a fixed hypersurface of this kind. Let the pole of the improper hyperplane be denoted by M. If we draw a line g through M, then by Chap. VIII, Theorem 3, this line will be tangent to F if and only if g and F have an improper point in common. If g is not a tangent, either g will not cut F at all (which is possible only in the real space P_n) or it will have two different proper points in common with F. Consider the latter case, and let P, Q be the proper points of intersection of g with F. Then, by Chap. VIII, Theorem 4, the points P, Q are harmonically separated by M and the improper point of g. Consequently, $\mathcal{D}(PQM) = -1$, i.e., M is the midpoint of the segment PQ. Observe that this holds for any line g through M that cuts F in two proper points. A (proper) point having this property is called *the center of the hypersurface of the second order*. Thus we have shown that *the non-degenerate hypersurfaces of the classes represented by* (16) *and* (19) *have at least one center.*

Conversely, it is clear that *these hypersurfaces can not have more than one center,* aside from the case of the imaginary hypersurfaces. For, by virtue of the property of a center, such a point must necessarily be the pole of the improper hyperplane and, as such, is determined uniquely. Likewise, it follows that *the paraboloids have no center.* The hypersurfaces of the classes represented by the non-degenerate normal forms (16) and (19) are accordingly referred to as *central hypersurfaces.*[4]

Certain other concepts are closely bound up with that of a center. A straight line through the center of a central hypersurface is called a *diameter.* A hyperplane containing the center is called a *diametral hyperplane.* The pole of a diametral hyperplane lies on the polar of the center and is accordingly an improper point. *A diametral hyperplane and a diameter* are said to be *conjugate* if the diameter passes through the pole of the diametral hyperplane.

Let h be a diametral hyperplane of a central hypersurface and g any line parallel to the conjugate diameter. From Theorems 3 and 4 of Chap. VIII, we have immediately the following further result:

If g is tangent to the hypersurface, then the point of tangency lies in h. If, however, g intersects the hypersurface in two points P, Q, then the midpoint of the segment PQ lies in h.

This statement remains valid for a paraboloid as well, if we take h to mean a hyperplane through the point of tangency (= pole) of the improper hyperplane and g to mean any line through the pole of h.

In the two-dimensional space P_2 every 'diametral hyperplane' of a central conic is itself a diameter. Two diameters g and h of a central conic in P_2 are conjugate if the improper point of each is the pole of the other. The improper points of g and h, together with the midpoint M, constitute a *polar triangle* of the conic section, i.e., a triangle in which each vertex is the pole of the opposite side.

The n-dimensional counterpart of a polar triangle is the *polar simplex* with respect to a hypersurface of the second order in P_n. By this we mean $n + 1$ linearly independent points (these are the vertices of the simplex) having the property that each is the pole of the hyperplane passing through the other n points. If n diameters of a central hypersurface of P_n have the property that their improper points, together with the center, form a polar simplex, we say that they constitute an *n-hedral of conjugate diameters.* Similarly, the sides (= hyperplanes) of a polar

[4] Here the name is restricted to non-degenerate hypersurfaces. It is sometimes used for degenerate hypersurfaces as well.

simplex, n of whose vertices are improper points, is called an *n-tuple of conjugate diametral hyperplanes.*

The existence of an n-hedral of conjugate diameters with respect to a non-degenerate central hypersurface reduces to the existence of a polar simplex with one vertex at the center of the hypersurface. For instance, in the case of the non-degenerate normal forms (16) and (19), the fundamental points of the coordinate system constitute such a polar simplex. Aside from this, the procedure described in Exercise 3, Chap. X affords a means of constructing all the n-hedrals of conjugate diameters, provided we begin with the center as the first vertex.

As a final example of concepts closely related to the concept of center, we mention the asymptotes. *Asymptotes* are those diameters that are at the same time tangent to the given hypersurface. They are obtained (according to Chap. VIII, Theorem 3) by connecting the center with the improper points of the hypersurface.

If the equation of the central hypersurface is given in the normal form (16) or in the normal form (19), the asymptotes can be found readily. For the center is then given by $\xi_0 = 1$, $\xi_1 = \xi_2 = \ldots = \xi_n = 0$, whereas every *improper* point of the hypersurface must satisfy the equation

$$\xi_1^2 + \xi_2^2 + \cdots + \xi_j^2 - \xi_{j+1}^2 - \cdots - \xi_n^2 = 0. \tag{31}$$

Now, if $\xi_0 = 0, \xi_1, \xi_2, \ldots, \xi_n$ are the coordinates of a definite improper point of the hypersurface, a point on the asymptote through this point has the coordinates $\mu, \lambda\xi_1, \lambda\xi_2, \ldots, \lambda\xi_n$, where μ, λ are two numbers of the field. From this we see the following:

A point lies on an asymptote if and only if it satisfies equation (31). *Thus, equation* (31) *represents the locus of all asymptotes.* It is itself a hypersurface of the second order, namely, a hypercone. It is called the *asymptotic cone.*[5]

Whereas the principle of duality was fully applicable to the developments of the two preceding chapters, this is no longer the case as regards the arguments of the present chapter. The validity of the duality principle is confined to projective geometry. *The concepts and theorems of affine geometry are not, in general, dualizable.* For, the transition from projective to affine geometry involved the distinguishing of the improper

[5] If $j = n$ or $j = 0$, i.e., if the same sign appears before all the squares in (31), then the asymptotic cone is 'imaginary,' and its vertex alone (i.e., the center of the hypersurface) is real.

elements, in which points and hyperplanes are no longer dealt with in symmetrical fashion (cf. Chap. III). It should be noted, in this connection, that the principle of duality is for the same reason not applicable to the developments of the next chapter, in which we shall study a geometry that subsumes affine geometry.

In conclusion, we append two tables containing, respectively, all of the affine normal forms of the conic sections in the real space P_2 and all of the affine normal forms of the surfaces of the second order in the real space P_3 (pp. 176, 177). The following should be observed about these tables.

The arrangement is such that all the equations of type (17) come first, then those of type (18), and lastly those of type (19).

The numbers r, s, k, h denote the invariants we discussed above (p. 160). We have put $h = -1$ in those cases for which $s = 0$. This is the value which would be given by, for example, formulas (21).

Cases 7, 14, and 15 of the second table represent cones with an improper vertex (= double point). Such cones are called *cylinders.*[6] In the case of cylinders, the lines through the vertex that lie on the surface are all parallel. The names 'parabolic,' 'elliptic,' and 'hyperbolic' cylinder are justified by the fact that the intersections with the plane $\xi_3 =$ a constant (or $x_3 =$ a constant) are parabolas, ellipses, and hyperbolas, respectively.

In the cases 3 and 7 of the first table, it is customary to say that the conic section represents a pair of conjugate complex lines with real point of intersection. Similarly, in cases 3 and 11 of the second table, we speak of two conjugate complex planes with real line of intersection; in case 5, of an imaginary cone with real vertex; and in case 13, of an imaginary cylinder.

As for the rest, the names and non-homogeneous equation forms will already be familiar to the reader from elementary analytic geometry.

Exercises

1. Show that in the real space P_n there are $n^2 + 3n + 1$ different classes of affinely equivalent hypersurfaces of the second order. How many are there in the complex space P_n?

2. Show that ordered k-tuples (with fixed $k \leqq n + 1$) of linearly independent *proper* points of P_n form a single class with respect to the affine group of P_n.

---

[6] More generally, in P_n a hypercone with an improper vertex is called a *hypercylinder.*

This is no longer true for k-tuples consisting of linearly dependent points. However, it is possible, by the use of Exercise 1 of Chap. IX, to derive for such k-tuples invariants which characterize the classes in a way similar to that in which the distance ratio was derived from the cross ratio. Let $S_1, S_2, \ldots, S_k$ and $T_1, T_2, \ldots, T_k$, say, be two such ordered k-tuples. Denote by U_{ik} and V_{ik} the improper points of the lines S_iS_k and T_iT_k, respectively. Then show that the necessary and sufficient condition for the existence of an affine transformation of P_n taking S_i into T_i for $i = 1, 2, \ldots, k$ is the existence of a linear projectivity that not only takes each S_i into T_i but also each U_{ik} into the corresponding V_{ik}. Thus, a complete system of projective invariants for the totality of all the points S_i, U_{ik} is at the same time a complete system of affine invariants for the k-tuples of the S_i.

3. If the k in the equation

$$\xi_1^2 + \xi_2^2 + \cdots + \xi_j^2 - \xi_{j+1}^2 - \cdots - \xi_n^2 = k \cdot \xi_0^2$$

is allowed to vary over the entire (real or complex) field, we obtain a family of hypersurfaces of the second order of P_n all with the same center and the same asymptotic cone. Show that if a diameter and a diametral hyperplane are conjugate with respect to a single one of these hypersurfaces, they are conjugate with respect to all of the others. One easily proved consequence of this is that the chords cut out on an arbitrary line by a non-degenerate central hypersurface and its asymptotic cone have a common midpoint.

4. A non-degenerate hypersurface of the second order divides all the proper points of P_n that do not belong to the hypersurface, as follows, into two classes. We write the normal form of the equation of the hypersurface in non-homogeneous coordinates—in the case of a central hypersurface, for example, we write

$$x_1^2 + x_2^2 + \cdots + x_j^2 - x_{j+1}^2 - \cdots - x_n^2 = 1$$

—and then put into one class all the points $(x_1, x_2, \ldots, x_n)$ for which

$$(*) \qquad x_1^2 + x_2^2 + \cdots + x_j^2 - x_{j+1}^2 - \cdots - x_n^2 < 1$$

and define the other class, similarly, by means of

$$(⁑) \qquad x_1^2 + x_2^2 + \cdots + x_j^2 - x_{j+1}^2 - \cdots - x_n^2 > 1.$$

If we call a set of points of the affine space R_n *convex* whenever, given any two points of the set, the entire line joining them belongs to the set, we can state the following result:

If $j < n$, neither of the two classes (*) and (⁑) is convex; if $j = n$, (*) is convex but (⁑) is not.

The same holds true for paraboloids, if the classes are defined by the equations

$$(*) \qquad x_1^2 + x_2^2 + \cdots + x_j^2 - x_{j+1}^2 - \cdots - x_{n-1}^2 < 2x_n,$$

$$(⁑) \qquad x_1^2 + x_2^2 + \cdots + x_j^2 - x_{j+1}^2 - \cdots - x_{n-1}^2 > 2x_n.$$

The Affine Normal Forms of the Conic Sections in Real P_2

Ref. No.	The Homogeneous Equation	The Non-Homogeneous Equation	r	s	k	h	Name or Description of the Conic Section	Description of the Improper Portion
1	$\xi_1^2 = 0$	$x_1^2 = 0$	1	1	0	0	A proper line (of double points)	A point (double point)
2	$\xi_1^2 - \xi_2^2 = 0$	$x_1^2 - x_2^2 = 0$	2	2	0	0	A pair of lines with proper intersection	A pair of points
3	$\xi_1^2 + \xi_2^2 = 0$	$x_1^2 + x_2^2 = 0$	2	2	1	1	A proper point (double point)	Imaginary
4	$0 = 2\xi_0\xi_1$	$0 = 2x_1$	2	0	0	-1	A proper line and the improper line	The improper line
5	$\xi_1^2 = 2\xi_0\xi_2$	$x_1^2 = 2x_2$	3	1	1	0	A parabola	A point (= tangent point of the parabola with the improper line)
6	$0 = \xi_0^2$		1	0	0	-1	The improper line (all points double points)	The improper line
7	$-\xi_1^2 = \xi_0^2$	$-x_1^2 = 1$	2	1	1	0	An improper point (double point)	A point
8	$\xi_1^2 = \xi_0^2$	$x_1^2 = 1$	2	1	0	0	Two parallel lines	A point (double point)
9	$-\xi_1^2 - \xi_2^2 = \xi_0^2$	$-x_1^2 - x_2^2 = 1$	3	2	2	1	Imaginary	Imaginary
10	$\xi_1^2 - \xi_2^2 = \xi_0^2$	$x_1^2 - x_2^2 = 1$	3	2	1	0	Hyperbola	A pair of points
11	$\xi_1^2 + \xi_2^2 = \xi_0^2$	$x_1^2 + x_2^2 = 1$	3	2	1	1	Ellipse	Imaginary

The Affine Normal Forms of the Surfaces of the Second Order in Real P_3

Ref. No.	The Homogeneous Equation	The Non-Homogeneous Equation	r	s	k	h	Name or Description of Surface	Description of the Improper Portion
1	$\xi_1^2 = 0$	$x_1^2 = 0$	1	1	0	0	A proper plane (of double points)	A line (every point a double point)
2	$\xi_1^2 - \xi_2^2 = 0$	$x_1^2 - x_2^2 = 0$	2	2	0	0	A pair of planes with proper intersection	A pair of lines
3	$\xi_1^2 + \xi_2^2 = 0$	$x_1^2 + x_2^2 = 0$	2	2	1	1	A proper line (of double points)	A point (double point)
4	$\xi_1^2 + \xi_2^2 - \xi_3^2 = 0$	$x_1^2 + x_2^2 - x_3^2 = 0$	3	3	1	1	A cone	A non-degenerate conic section
5	$\xi_1^2 + \xi_2^2 + \xi_3^2 = 0$	$x_1^2 + x_2^2 + x_3^2 = 0$	3	3	2	2	A proper point (double point)	Imaginary
6	$0 = 2\,\xi_0\,\xi_1$	$0 = 2\,x_1$	2	0	0	—1	A proper plane and the improper plane	The improper plane
7	$\xi_1^2 = 2\,\xi_0\,\xi_2$	$x_1^2 = 2\,x_2$	3	1	1	0	A parabolic cylinder	A line
8	$\xi_1^2 - \xi_2^2 = 2\,\xi_0\,\xi_3$	$x_1^2 - x_2^2 = 2\,x_3$	4	2	1	0	A hyperbolic paraboloid	A pair of lines
9	$\xi_1^2 + \xi_2^2 = 2\,\xi_0\,\xi_3$	$x_1^2 + x_2^2 = 2\,x_3$	4	2	2	1	An elliptic paraboloid	A point
10	$0 = \xi_0^2$		1	0	0	—1	The improper plane (every point a double point)	The improper plane
11	$-\xi_1^2 = \xi_0^2$	$-x_1^2 = 1$	2	1	1	0	An improper line (of double points)	A line
12	$\xi_1^2 = \xi_0^2$	$x_1^2 = 1$	2	1	0	0	Two parallel planes	A line (every point a double point)
13	$-\xi_1^2 - \xi_2^2 = \xi_0^2$	$-x_1^2 - x_2^2 = 1$	3	2	2	1	An improper point (double point)	A point
14	$\xi_1^2 - \xi_2^2 = \xi_0^2$	$x_1^2 - x_2^2 = 1$	3	2	1	0	A hyperbolic cylinder	A pair of lines
15	$\xi_1^2 + \xi_2^2 = \xi_0^2$	$x_1^2 + x_2^2 = 1$	3	2	1	1	An elliptic cylinder	A point
16	$-\xi_1^2 - \xi_2^2 - \xi_3^2 = \xi_0^2$	$-x_1^2 - x_2^2 - x_3^2 = 1$	4	3	3	2	Imaginary	Imaginary
17	$\xi_1^2 - \xi_2^2 - \xi_3^2 = \xi_0^2$	$x_1^2 - x_2^2 - x_3^2 = 1$	4	3	2	1	Hyperboloid of two sheets	A non-degenerate conic section
18	$\xi_1^2 + \xi_2^2 - \xi_3^2 = \xi_0^2$	$x_1^2 + x_2^2 - x_3^2 = 1$	4	3	1	1	Hyperboloid of one sheet	A non-degenerate conic section
19	$\xi_1^2 + \xi_2^2 + \xi_3^2 = \xi_0^2$	$x_1^2 + x_2^2 + x_3^2 = 1$	4	3	2	2	An ellipsoid	Imaginary

CHAPTER XII

THE METRIC CLASSIFICATION OF HYPERSURFACES OF THE SECOND ORDER

In the preceding chapters we have seen, in the case of two fundamental examples, how the concept of a group of transformations can be used to classify and arrange geometrical configurations and theorems. This was done by basing our investigation on a definite group of transformations (the projective group, the affine group). With respect to this group a specified class of geometrical figures (hypersurfaces of the second order, linear quadruples of points, linear triples of points) were partitioned into classes of equivalent figures, i.e., figures that can be mapped onto each other by the transformations of the group. A geometrical property of these figures was considered as belonging to the *geometry defined by the group* whenever the property was invariant under all the transformations of the group. The problem of studying a particular kind of geometrical figure in a certain geometry consequently amounts to finding the invariants of the figures of this kind with respect to the group defining the geometry. These principles, which have lent clarity and perspective to the manifold theorems and problems of geometry, were developed by Felix Klein in his famous **Erlanger Programm**.

In this chapter we shall make one final application of these fundamental principles. The group that we shall now take as our basic group is essentially the group of motions in euclidean R_n. The geometry belonging to this group is called **euclidean geometry**.

The Group of Motions as a Subgroup of the Projective Group

Our discussion will at first be confined to the real space P_n. For the sake of clarity, let us fix upon a definite coordinate system to be used in all that follows, and in fact let us take the *natural* coordinate system mentioned on p. 32, in which each point $[\xi_0, \xi_1, \ldots, \xi_n]$ of P_n has precisely the coordinates $\xi_0, \xi_1, \ldots, \xi_n$.

The proper points of P_n, i.e., those in which $\xi_0 \neq 0$, constitute the affine space R_n (Chap. II). Just as in § 7 of *Modern Algebra,* we introduce into this affine part of P_n the euclidean definition of length. The distance between two proper points $S = [\xi_0, \xi_1, \cdots, \xi_n]$ and $Q = [\eta_0, \eta_1, \cdots, \eta_n]$ ($\xi_0 \neq 0$, $\eta_0 \neq 0$), with the non-homogeneous coordinates $x_i = \xi_i/\xi_0$, $y_i = \eta_i/\eta_0$ ($i = 1, 2, \ldots, n$), is given by

$$\overline{SQ} = {}_{+}\sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} = \frac{1}{|\xi_0\, \eta_0|} \cdot {}_{+}\sqrt{\sum_{i=1}^{n}(\eta_i\, \xi_0 - \xi_i\, \eta_0)^2}.$$

Distance is not defined between two improper points or between an improper and a proper point.

Let us now return to the rigid motions of the euclidean space R_n (*Modern Algebra,* § 12). We know, from *Modern Algebra* § 13, that every such motion is also an affine transformation and, indeed, a non-singular affine transformation. In consequence, whatever was said about non-singular affine transformations in Chapter VI of the present volume holds true *a fortiori* for motions. Thus, every motion of the euclidean space R_n—that is to say, of the proper part of the real space P_n—is induced by one and only one linear projectivity of P_n. Such a projectivity obviously maps the improper hyperplane onto itself, i.e., it belongs to the affine group of P_n. A linear projectivity of the real space P_n that induces a motion in the proper part of P_n is itself called *a motion of* P_n.

Since such a motion in the real space P_n is characterized by the fact that it leaves euclidean distances invariant in the proper part of P_n, it follows at once that the product of two such motions and the inverse of a motion is itself a motion. The rigid motions thus constitute a subgroup of the linear projective group of P_n and, indeed, even a subgroup of the affine group of P_n.

We should now like to ascertain what the system of equations of a rigid motion looks like when expressed in the above fixed projective coordinate system. A rigid motion, as a special kind of affine transformation, can certainly always be represented by a system of equations of the form

$$(1) \qquad \begin{aligned} \eta_0 &= \xi_0, \\ \eta_1 &= t_{10}\, \xi_0 + t_{11}\, \xi_1 + t_{12}\, \xi_2 + \cdots + t_{1n}\, \xi_n, \\ \eta_2 &= t_{20}\, \xi_0 + t_{21}\, \xi_1 + t_{22}\, \xi_2 + \cdots + t_{2n}\, \xi_n, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \eta_n &= t_{n0}\, \xi_0 + t_{n1}\, \xi_1 + t_{n2}\, \xi_2 + \cdots + t_{nn}\, \xi_n \end{aligned}$$

(cf. Chap. VI). However, since not every affine transformation is a rigid motion, certain additional conditions must be satisfied by the coefficients of (1). We assert the following:

The system of equations (1) *represents a rigid motion if and only if the square matrix*

$$T = \begin{pmatrix} t_{11}, & t_{12}, & \cdots, & t_{1n}, \\ t_{21}, & t_{22}, & \cdots, & t_{2n}, \\ \cdot & \cdot & \cdot & \cdot \\ t_{n1}, & t_{n2}, & \cdots, & t_{nn} \end{pmatrix} \tag{2}$$

is orthogonal.

Proof: The mapping induced in the proper part of P_n by (1) is, in affine coordinates (cf. Chap. VI, p. 97),

$$y_i = t_{i0} + \sum_{k=1}^{n} t_{ik} x_k, \qquad i = 1, 2, \cdots, n. \tag{3}$$

Since (1) was given in the natural coordinate system of P_n, (3) is expressed in the natural coordinates of the euclidean space R_n (cf. footnote 3 on p. 97), that is, in a *cartesian* coordinate system. In such a coordinate system, however, (3) can represent a rigid motion if and only if the matrix (2) is orthogonal.

We now proceed to a study of hypersurfaces of the second order in the geometry associated with the group of motions of P_n just discussed. Again, the chief problem is the classification of the hypersurfaces with respect to this group, a classification usually referred to as *euclidean-metric* or, simply, *metric*.

Metric Classification of Hypersurfaces of the Second Order

If two hypersurfaces of the second order in real P_n can be taken into each other by a rigid motion, we call them **congruent**. We shall again tackle the problem of determining the classes of congruent hypersurfaces by trying to find a representative, i.e., a normal form, for each class. A complete system of invariants for hypersurfaces of the second order with respect to the group of motions of projective P_n will then be obtained automatically.

The procedure for obtaining these normal forms bears a strong resemblance to that of the preceding chapter. There our first step was to change the hypersurface

$$\sum_{i,k=0}^{n} a_{ik}\,\eta_i\,\eta_k = 0 \tag{4}$$

by a suitably chosen affine transformation, into an equation of the form (8) on p. 161. Since we no longer have arbitrary affine transformations at our disposal, but only rigid motions, we can no longer achieve quite as much as before. Nevertheless, the part of the equation containing no η_0 can still be reduced, as we shall soon see, to pure squares. However, we will now have as coefficients of the square terms not just ± 1, as in the preceding chapter, but arbitrary real numbers.

We can carry out a rigid motion on (4) by replacing the η_i by the right-hand sides of (1).[1] In doing so we are again concerned first of all with the behaviour of the improper part of the hypersurface, i.e., with that part of equation (4) that is free of η_0 and that part of the result of our substitution that is free of ξ_0. Thus, we first pay attention to what result the substitution obtained from (1) (cf. equation (5) on p. 160), namely,

$$\eta_i = \sum_{k=1}^{n} t_{ik}\,\xi_k, \qquad i = 1, 2, \cdots, n, \tag{5}$$

will have when applied to the expression

$$\sum_{i,k=1}^{n} a_{ik}\,\eta_i\,\eta_k . \tag{6}$$

The simplification that may be achieved is easier to see if we make use of matrix notation. Then (6) may be written as

$$(\eta)'\,A\,(\eta), \tag{7}$$

where (η) now, of course, stands for that n-by-n (square) matrix whose first column consists of the elements $\eta_1, \eta_2, \ldots, \eta_n$ (no η_0 this time) and whose remaining entries are zeros, while A stands for the n-by-n matrix

[1] More precisely, this means we apply to (4) the inverse of the motion (1). For we are going over from the η_i to the ξ_i (cf. Chap. IX).

$$(8) \qquad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Moreover, using the matrix (2), we can write the substitution (5) in the form

$$(9) \qquad (\eta) = T(\xi).$$

Application of (9) now takes (7) into

$$(10) \qquad (\xi)' T' A T(\xi).$$

In (10) are contained all the terms free of ξ_0 which result from application of the substitution (1) to (4). But it was our purpose to choose the motion (1) in such a way that in the expression resulting from the substitution the part free of ξ_0 should consist of the square terms only. This means, as regards (10), that $T'AT$ must have diagonal form. Thus our problem amounts to this: Given the real symmetric matrix A, to find an orthogonal matrix T such that $T'AT$ be of diagonal form. But that this is always possible has been shown in *Modern Algebra*, § 24, Theorem 13.

Thus, we have shown the following:

Every hypersurface of the second order of real P_n is congruent to at least one of the hypersurfaces of the form[2]

$$(11) \qquad \sum_{i=1}^{s} a_i \xi_i^2 = b\,\xi_0^2 + 2 \sum_{i=1}^{n} b_i \xi_i \xi_0, \qquad a_i \neq 0.$$

The rest of our discussion as well directly parallels that of Chap. XI. We first write (11) in a slightly different form, as follows:

$$(12) \qquad \sum_{i=1}^{s} a_i \left(\xi_i - \frac{b_i}{a_i}\,\xi_0\right)^2 = c\,\xi_0^2 + 2 \sum_{i=s+1}^{n} b_i \xi_i \xi_0,$$

where $c = b + \sum_{i=1}^{s} \frac{b_i^2}{a_i}$. We next carry out the substitution

---

[2] Cf. equation (8) on p. 161.

$$
(13) \qquad \begin{aligned} \xi_i' &= \xi_i && \text{for } i = 0 \text{ and } i = s+1, s+2, \cdots, n, \\ \xi_i' &= -\frac{b_i}{a_i}\xi_0 + \xi_i && \;\;\text{"}\;\; i = 1, 2, \cdots, s \end{aligned}
$$

This merely amounts to applying another rigid motion[3] to (12). For simplicity, we drop the primes from our variables after making the substitution. We thus see that every hypersurface of the second order is always congruent to one of the form

$$
(14) \qquad \sum_{i=1}^{s} a_i \xi_i^2 = c\xi_0^2 + 2\sum_{i=s+1}^{n} b_i \xi_i \xi_0 .
$$

Now, just as in Chapter XI, we distinguish three different cases.

Case One: All the coefficients of the right-hand side of (14) vanish. Then we may assume to begin with that the number of positive terms on the left-hand side of (14) is at least as great as the number of negative terms; this can be brought about, if need be, by multiplying the equation through by -1. By further multiplication by a non-zero constant, we make the largest $a_i > 0$ equal to 1. We then permute[4] the variables so that the a_i are arranged in descending order of magnitude. In the further subcase in which the number of positive terms is equal to the number of negative terms, we can also insure that the *first* of the sums $a_1 + a_s$, $a_2 + a_{s-1}$, ... that is not zero be positive.[5] Taking all this into account, the final form that we achieve may be written as follows

$$
(15) \left\{ \begin{array}{ll} a_1\xi_1^2 + a_2\xi_2^2 + \cdots + a_s\xi_s^2 = 0, & 1 \leqq s \leqq n, \\ 1 = a_1 \geqq a_2 \geqq \cdots \geqq a_k > 0 > a_{k+1} \geqq \cdots \geqq a_s, & k \geqq s - k; \\ \multicolumn{2}{l}{\textit{where, moreover, if } s = 2k, \textit{ the first non-vanishing sum } a_1 + a_s,} \\ \multicolumn{2}{l}{a_2 + a_{s-1}, \ldots \textit{ is positive.}} \end{array} \right.
$$

[3] More precisely, a translation of the euclidean space R_n, as is easily seen if the equations be written in non-homogeneous coordinates, as in (3).

[4] A permutation of variables amounts to applying a rigid motion.

[5] For if the number of positive terms is equal to the number of negative terms, the conditions already attained are not changed by multiplication of the equation by $1/a_s$ and reversal of the numbering of the variables $\xi_1, \xi_2, \ldots, \xi_s$; that is, they hold for the new equation that results from the carrying out of these two operations. This means, however, that the above additional condition can always be satisfied if it is not satisfied to begin with.

Case Two: In (14), all the $b_i = 0$, but $c \neq 0$. In this case, we divide the equation through by c. If we simply write a_i instead of a_i/c and make a suitable permutation of the variables, we arrive at the following form of the equation:[6]

$$(16) \qquad \sum_{i=1}^{s} a_i \xi_i^2 = \xi_0^2, \qquad \begin{array}{r} a_1 \geqq a_2 \geqq \cdots \geqq a_k > 0 > a_{k+1} \geqq \cdots \geqq a_s, \\ 0 \leqq s \leqq n, \quad 0 \leqq k \leqq s. \end{array}$$

Case Three: The b_i are *not all* equal to zero, and c is arbitrary. In this case, we first observe that by multiplying the equation through by a suitable non-zero constant we can always insure that

$$(17) \qquad \sum_{i=s+1}^{n} b_i^2 = 1.$$

We then apply the substitution

$$(18) \qquad \begin{aligned} \xi_i' &= \xi_i \quad \text{for } i = 0, 1, 2, \cdots, s, \\ \xi_{s+1}' &= \frac{c}{2}\,\xi_0 + \sum_{i=s+1}^{n} b_i \xi_i , \end{aligned}$$

which takes (14) into the form (where we again simply write ξ_i in place of ξ_i'):

$$(19) \qquad \sum_{i=1}^{s} a_i \xi_i^2 = 2\,\xi_0\,\xi_{s+1}.$$

However, the question now arises: Can this substitution be interpreted as a rigid motion? This will be the case provided we can extend equations (18) by the adjunction of $n - (s+1)$ other equations in the variables $\xi_{s+2}', \ldots, \xi_n'$ to a system of equations for which the matrix corresponding to (2) is orthogonal. But this means that we are to make the $(s+1)$-by-$(s+1)$ matrix

$$(20) \qquad \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & b_{s+1} & b_{s+2} & \cdots & b_n \end{pmatrix}$$

[6] $s = 0$ shall, of course, mean that the equation is of the form $0 = \xi_0^2$.

into an orthogonal matrix by adding to it $n - s - 1$ suitable further rows. But since, by virtue of (17), the row vectors of (20) already constitute a normalized orthogonal system, the supplementation is always possible, by § 11 of *Modern Algebra*.

Further, it is always possible to change the sign on one side only of (19). For this may be done by applying the motion (of reflection) given by

$$\text{(21)} \qquad \begin{aligned} \xi'_{s+1} &= -\xi_{s+1}, \\ \xi'_i &= \xi_i \text{ for } i \neq s+1. \end{aligned}$$

Hence we can treat the left-hand side of (19) in the same way as in case one, so that we finally arrive at the following form for the equation in this third and last case:

$$\text{(22)} \left\{ \begin{array}{l} \displaystyle\sum_{i=1}^{s} a_i \xi_i^2 = 2\,\xi_0\,\xi_{s+1}, \\ a_1 \geqq a_2 \geqq \cdots \geqq a_k > 0 > a_{k+1} \geqq \cdots \geqq a_s, \end{array} \quad k \geqq s-k,\ 0 \leqq s \leqq n; \right.$$
$$\textit{in addition, if } s = 2k, \textit{ the first non-vanishing sum } a_1 + a_s,\ a_2 + a_{s-1}, \ldots \textit{ is positive.}$$

What we have shown thus far can be summed up as follows:

Every hypersurface of the second order of real P_n is congruent to at least one of the hypersurfaces represented by equations (15), (16), *and* (22).

We now wish to see whether or not any two of the expressions we have arrived at can be congruent. Let us begin by considering some cases that display a special kind of behaviour. Such equations are those of the form (15) in which $k = s$ and those of the form (16) in which all the coefficients on the left-hand side are negative.

If $k = s$ in (15), so that all the coefficients are positive, then the hypersurface represented by (15) (remember that we are concerned with the real space) consists simply of the linear space

$$\xi_1 = 0,\ \xi_2 = 0,\ \cdots,\ \xi_s = 0.$$

This is also represented, however, by the equation

$$\text{(23)} \qquad \xi_1^2 + \xi_2^2 + \cdots + \xi_s^2 = 0.$$

Thus, for a given s, all equations of the type (15) *for which $k = s$ represent, not merely congruent, but identical hypersurfaces of the second* *order.*[7] We accordingly regard (23) as the normal form of all these hypersurfaces. It is also evident that the hypersurface (23) cannot be congruent to any of those of the type (15) with $k < s$, nor to any of the hypersurfaces of the types (16) or (22), for it is not even affinely equivalent to any of them.

Similarly, for a fixed s, all hypersurfaces of the type (16) *in which the coefficients of the left-hand side are all negative, are identical with*

$$-\xi_1^2 - \xi_2^2 - \cdots - \xi_s^2 = \xi_0^2 \tag{24}$$

and are not congruent to any other hypersurface (15), (16), *or* (22).

The special cases just considered are precisely those for which, in accordance with Chap. X, Theorem 5, the equation of the hypersurface is not uniquely determined. For all other equations of the types (15), (16), and (22), however, according to the same theorem, the equation *is* uniquely determined. In particular, there are no two different ones among them that represent the *same* hypersurface. Let us now assume that two different ones among these equations represent *congruent* hypersurfaces. These hypersurfaces, then, can be carried into each other by a rigid motion; and since their equations are essentially[8] uniquely determined, we can make the following statement:

If among the remaining equations (15), (16), and (22) there exist two different ones representing congruent hypersurfaces of the second order, then there must also exist a substitution (1) with orthogonal matrix (2) that takes one of these equations into a multiple of the other.

But we shall prove that, on the contrary, *a substitution of the form* (1) *with orthogonal matrix* (2) *can never transform an equation of the form* (15), (16), *or* (22) *into a multiple of another equation of this kind.*[9]

---

[7] In the real space P_n, be it noted! (15) with $k = s$ and (23) do *not* in general coincide in their complex parts. If this be taken into consideration, the situation takes on a completely different aspect. See footnote 9 below.

[8] I.e., up to a non-zero constant of proportionality.

[9] This purely algebraic statement and the proof that follows holds true even for those cases of (15) and (16) already treated. Thus although, say, equation (15) with $k = s$ represents the same hypersurface in the reals as does (23), it is not possible to transform an arbitrary equation (15) into (23) by a real substitution (1) with orthogonal matrix (2). This fact is in itself of interest and may be interpreted as follows: If we alter the concept of congruence for the special cases of (15) and (16) already treated to the extent that two of these hypersurfaces are to be considered congruent only if there exists a rigid motion that brings into coincidence not only their real parts but at the same time their complex parts as well, then we also have that no two of the cases of (15) with $k = s$ and no two of the cases of (16) with $k = 0$ can be congruent.

We will thus have proved that among the remaining equations of the form (15), (16), and (22), there are no two that are congruent.

Suppose, then, that there is given an equation of the kind in question, say

$$a_1 \eta_1^2 + a_2 \eta_2^2 + \cdots + a_s \eta_s^2 = P, \tag{25}$$

where P stands for either 0 or η_0^2 or $2\,\eta_0\,\eta_{s+1}$. For simplicity, we have again written this initial equation in terms of the variables η_i, in order to be able to use the substitution (1) without having to alter the notation.

We now assume that we have found a substitution (1) with orthogonal matrix (2) that transforms (25) into a multiple of an equation of the form (15), (16), or (22). We write the result of the substitution in the form

$$c \cdot (b_1 \xi_1^2 + b_2 \xi_2^2 + \cdots + b_s \xi_s^2) = c \cdot P'. \tag{26}$$

Since the performing of substitution (1) always amounts to the application of an affine transformation, it is clear from Chap. XI that $P' = 0,\ = \xi_0^2$, or $= 2\,\xi_0\,\xi_{s+1}$ according as $P = 0,\ = \eta_0^2$, or $= 2\,\eta_0\,\eta_{s+1}$, respectively. As regards the b_i, we can assume that they satisfy the subsidiary conditions associated with whatever form of equation (15), (16), or (22) is under consideration. And clearly, the left-hand sides of equations (25) and (26) have exactly the same number of terms, this number being, in fact, even an affine invariant.

Again let us consider first of all only the terms free of η_0, i.e., the left-hand side of equation (25), and compare them, as before, to the terms free of ξ_0 in equation (26), which is formed from (25) as a result of the substitution. We recall that in the general case the expression (7) represents the part of the initial equation that is free of η_0 and (10) the part of the equation that, after the substitution, is free of ξ_0. In our present case, we must, in (7), set A equal to the matrix

$$A = \begin{pmatrix} a_1 & & & & & & 0 \\ & a_2 & & & & & \\ & & \ddots & & & & \\ & & & a_s & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & 0 \end{pmatrix} \tag{27}$$

while, by (26), for the matrix $T'AT$ of (10) we have

$$
(28) \qquad T'AT = c \cdot \begin{pmatrix} b_1 & & & & & & 0 \\ & b_2 & & & & & \\ & & \ddots & & & & \\ & & & b_s & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & 0 \end{pmatrix}.
$$

Since T is assumed to be orthogonal, i.e., $T' = T^{-1}$, we have in explicit form, after multiplying (28) on the left by T,

$$
(29) \begin{pmatrix} a_1 t_{11} & a_1 t_{12} & \cdots & a_1 t_{1n} \\ a_2 t_{21} & a_2 t_{22} & \cdots & a_2 t_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_s t_{s1} & a_s t_{s2} & \cdots & a_s t_{sn} \\ 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 0 \end{pmatrix} = c \cdot \begin{pmatrix} b_1 t_{11} & b_2 t_{12} & \cdots & b_s t_{1s} & 0 & \cdots & 0 \\ b_1 t_{21} & b_2 t_{22} & \cdots & b_s t_{2s} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_1 t_{n1} & b_2 t_{n2} & \cdots & b_s t_{ns} & 0 & \cdots & 0 \end{pmatrix}.
$$

Since the a_i and the b_i are all different from zero, it follows at once from a comparison of the two sides of (29) that

$$
(30) \qquad \begin{aligned} t_{i,s+1} &= t_{i,s+2} = \cdots = t_{i,n} = 0 \quad \text{for } i = 1, 2, \cdots, s, \\ t_{s+1,i} &= t_{s+2,i} = \cdots = t_{n,i} = 0 \quad \text{"} \quad i = 1, 2, \cdots, s. \end{aligned}
$$

This means the following:

I. *The matrix* (2) *of our substitution must necessarily have the form*

$$
(31) \qquad T = \left(\begin{array}{ccc:ccc} t_{11} & \cdots & t_{1s} & & & \\ \vdots & & \vdots & & 0 & \\ t_{s1} & \cdots & t_{ss} & & & \\ \hdashline & & & t_{s+1,s+1} & \cdots & t_{s+1,n} \\ & 0 & & \vdots & & \vdots \\ & & & t_{n,s+1} & \cdots & t_{n,n} \end{array}\right).
$$

The next consequence we derive from our assumption is the following:

II. *The coefficients* $t_{10}, t_{20}, \ldots, t_{s0}$ *in* (1) *must also all vanish.*

To see this, consider all the terms $\xi_0 \xi_i$ with $1 \leqq i \leqq s$ that result from (25) when we make the substitution. By virtue of I, terms of this type can certainly come only from the left-hand side of (25), whatever form P may have. On the other hand, all of these terms must cancel out, because the final equation is of the form (26). In other words, for every i between 1 and s we have

$$(32) \qquad 2 \sum_{\nu=1}^{s} a_{\nu}\, t_{\nu 0}\, t_{\nu i} = 0, \qquad i = 1, 2, \cdots, s.$$

Since the s-by-s matrix

$$(33) \qquad \begin{pmatrix} t_{11} & \cdots & t_{1s} \\ \vdots & & \vdots \\ t_{s1} & \cdots & t_{ss} \end{pmatrix}$$

is non-singular ((31) being non-singular), our result follows from (32) (by *Modern Algebra,* § 9, Theorem 4).

In the case in which (25) is an equation of type (22), observe that we have the following further result:

III. *If* $P = 2\,\eta_0\,\eta_{s+1}$ *in the right-hand side of* (25) (*whence, in* (26), $P' = 2\,\xi_0\,\xi_{s+1}$), *then the coefficients* $t_{s+1,0}$ *and* $t_{s+1,i}$ *with* $i > s+1$ *in* (1) *also vanish.*

For otherwise, the substitution would produce on the right-hand side of (25) terms containing $\xi_0{}^2$ and $\xi_0 \xi_i$, $i > s + 1$, which, by virtue of I and II, would have nothing to cancel out against. In that case, however, the result of the substitution could not have form (26).

Now let us return to equation (28). Since $T' = T^{-1}$, passing from the matrix A to $T'AT$ represents a transformation in the sense of the matrix calculus. But the roots of the characteristic polynomial are invariant under such a transformation. On the other hand, both A and $T'AT$ are diagonal matrices and thus contain in their principal diagonals precisely the roots of the characteristic polynomial. Hence the totality of terms in the principal diagonals of these two matrices must be the same.

It should be noted, further, that the a_i and the b_i, by virtue of the stipulations regarding the forms of equations (15), (16), (22), are arranged in descending order of magnitude. Consequently, the numbers

$c \cdot b_1, c \cdot b_2, \ldots, c \cdot b_s$ are also arranged in descending or ascending order of magnitude according as c is positive or negative. Thus, if $c > 0$, $c \cdot b_1$ must equal the greatest of the a_i, i.e., a_1; $c \cdot b_2$ the greatest of the elements $a_2, a_3, \ldots, a_s$, i.e., a_2; and so forth. On the other hand, if $c < 0$, then $c \cdot b_1 = a_s$, $c \cdot b_2 = a_{s-1}$, etc. Thus we have IV:

IV: *If* $c > 0$, *then in* (25) *and* (26) *we must have* $a_i = c \cdot b_i$ *for* $i = 1, 2, \ldots, s$, *whereas if* $c < 0$, *then* $a_i = c \cdot b_{s+1-i}$ *for* $i = 1, 2, \ldots, s$.

We can, further, prove the following:

V: *The constant* c *of* (26) *is necessarily* ± 1, *and whenever, in the case of* (25) (*and* (26)), *we are dealing with an equation of the form* (16), *it is necessarily even equal to* $+1$.

In proving V, we first consider the case $P = \eta_0^2$, $P' = \xi_0^2$. By II, no terms containing $\xi_0{}^2$ arise from applying the substitution to the left-hand side of (25), and the right-hand side simply goes into $\xi_0{}^2$. Hence, $c = +1$.

Secondly, consider the case $P = 2\,\eta_0\,\eta_{s+1}$. It then follows from I and III that $t_{s+1,\,s+1} = \pm 1$ because of the orthogonality of matrix T. Consequently, the right-hand side of (25) goes directly over into $\pm 2\,\xi_0\,\xi_{s+1}$. By I, there is no contribution to this term from the left-hand side. Hence, c is ± 1.

In the last case $P = P' = 0$, by virtue of the subsidiary conditions of (15), we must certainly have $a_1 = b_1 = 1$. Hence, if $c > 0$, it follows immediately from IV that $c = +1$. If, however, $c < 0$, we need to recall that the number of positive a_i is not smaller than the number of negative ones, and that this is likewise true for the b_i. But on the other hand, since by IV $a_i = c \cdot b_{s+1-i}$, the number of the positive a_i must be equal to the number of negative b_i and the number of negative a_i equal to the number of positive b_i. From this it follows that both for the a_i and for the b_i precisely half the terms are > 0 and half < 0. Thus, $c < 0$ is possible only if, in the notation of (15), $s = 2k$. But then, according to (15) the first non-vanishing of the sums $a_1 + a_s, a_2 + a_{s-1}, \ldots$ is positive. The same must hold for the sequence made up of the b_i: $b_1 + b_s, b_2 + b_{s-1}, \ldots$. Since by IV, however, $a_i + a_{s+1-i} = c \cdot (b_i + b_{s+1-i})$ and therefore $b_i + b_{s+1-i}$ is always negative whenever $a_i + a_{s+1-i}$ is positive, all the sums $a_i + a_{s+1-i}$ must be zero. In particular, $a_1 + a_s = 0$, so that $a_s = -1$. But by IV $c \cdot b_1 = a_s$, and thus $c = -1$, which is what we wished to show.

For the case $c = -1$, we then have VI.

VI. *The constant c can* $= -1$ *only if equation* (25) *is of the type* (15) *or* (22), *the number of positive terms and of negative terms on the left-hand side are equal and, in addition,* $a_i = -a_{s+1-i}$ *for* $i = 1, 2, \ldots, s$.

That $c = -1$ only under the conditions stated above, has just been shown for the case $P = P' = 0$. For the case $P = 2\,\eta_0\,\eta_{s+1}$, the desired result may be obtained by a precisely similar argument.

From IV and VI it now follows immediately that in the case $c = -1$, we must have

$$(34) \qquad a_1 = b_1,\ a_2 = b_2,\ \cdots,\ a_s = b_s.$$

That is to say, (26) is then nothing but (25) itself multiplied by -1. But for the case $c = +1$, (25) and (26) are, by IV, identical. We thus have the final result:

VII. *A substitution* (1) *with orthogonal matrix* (2) *which, when applied to* (25) *gives an end result of the form* (26), *either leaves the equation* (25) *unaltered or multiplies it by* -1.

In no case, then, can the end result be a multiple of any of the equations (15), (16), or (22) different from (25). Hence, our proof is complete.

The question of euclidean-metric normal forms of the hypersurfaces of the second order is now settled. We have seen that *a complete system of representatives of the classes of congruent hypersurfaces is given by*

equations (23), (24);
equations of the type (15) *with* $k \neq s$;
equations of the type (16) *with* $k > 0$;
all equations of the type (22).

Once we know the normal forms, we can immediately give a *complete system of invariants* with respect to the group of rigid motions. *The four invariants of Chap.* XI *and the coefficients* a_i *that occur in our present normal forms constitute such a system.* To within a common normalizing factor the a_i are equal to the characteristic roots of the matrix (8). The process of reduction to normal form given above contains a procedure for determining the normalizing factor.

We add a few remarks about the non-degenerate central hypersurfaces. Their metric normal forms are given by

$$a_1 \xi_1^2 + a_2 \xi_2^2 + \cdots + a_n \xi_n^2 = \xi_0^2. \tag{35}$$

It is clear that the fundamental simplex of our coordinate system is a polar simplex (Chap. XI) with respect to every hypersurface (35). The fundamental point $Q_0 = [1, 0, 0, \ldots, 0]$ is the center of (35). The lines joining Q_0 to the n remaining fundamental points $Q_1, Q_2, \ldots, Q_n$ thus form an n-tuple of conjugate diameters.

On the other hand, the n lines Q_0Q_i, $i = 1, 2, \ldots, n$, are such that, in terms of euclidean measurement of angle, every two are perpendicular to each other. Such an n-tuple of conjugate diameters is called *a system of principal axes*. Since the fact of the existence of a system of principal axes is invariant under rigid motions, such a system of axes certainly exists for all hypersurfaces of the second order that are congruent to a hypersurface of the form (35). But that means for all central hypersurfaces.

The transformation of a central hypersurface into the normal form (35) by means of a rigid motion is called, in accordance with this interpretation, a **principal axis transformation**, or a *transformation to principal axes*. This term is then extended to non-central hypersurfaces as well by taking it to mean, more generally, the transformation of a hypersurface of the second order into euclidean-metric normal form. Finally, this expression is also used for the algebraic essence of this transformation, i.e., for the statement given in Theorem 13, § 24, of *Modern Algebra.*

Moreover, as in Chap. IX, p. 119, the reduction of a hypersurface of the second order to normal form can also be thought of as a *transformation of coordinates.* This transformation of coordinates then represents the transition from the natural coordinate system to a new coordinate system whse fundamental points are the n improper points of the lines of a system of principal axes and the center of the hypersurface. For the non-homogeneous coordinates of the proper points, this transformation of coordinates is given by the system of equations (3); it thus amounts to passing from one cartesian coordinate system to another cartesian coordinate system.

If we determine the points of intersection of the lines Q_0Q_i with the hypersurface (35), we obtain the following values for their coordinates:

$$\xi_0 = 1, \quad \xi_i = \frac{1}{\sqrt{a_i}}, \quad \xi_k = 0 \textit{ for } k \neq 0 \textit{ and } k \neq i. \tag{36}$$

The two values of the square root give the two points of intersection. The root is a real number, however, only if $a_i > 0$. In this case, the euclidean distance between the two points of intersection is $\frac{2}{|\sqrt{a_i}|}$. This number is accordingly called the *length of the i-th principal axis*. This name is retained even in the case $a_i < 0$. Thus, we have obtained a geometrical interpretation for the invariants a_i for the case of a central hypersurface.

The Absolute

From now on, we suppose the dimension n of P_n to be greater than one. In euclidean geometry those non-degenerate central hypersurfaces are of special importance whose a_i are all positive and equal to each other. The normal form of such a hypersurface is thus as follows:

$$(37) \qquad a\,(\eta_1^2 + \eta_2^2 + \cdots + \eta_n^2) = \eta_0^2 \qquad a > 0.$$

The hypersurfaces of the class represented by (37) are called **hyperspheres** (*spheres* in P_3, *circles* in P_2). They are characterized by the fact that the distance from the center to a point on the surface is the same for every point of the surface. This property can easily be read off from the form of equation (37) (say, by writing it in non-homogeneous form). The distance from the center to a point on a given hypersphere is called the *radius* of the hypersphere.

The general equations of hyperspheres are likewise easy to characterize. For, by definition, the equation of every hypersphere can be obtained by applying to the normal form (37) a suitable substitution (1) with orthogonal matrix (2). Under such a substitution, however, that part of (37) that is free of η_0 cannot change in form. For the matrix A which appears in the expression (7) is, in the case of (37), a multiple of the unit matrix. And the unit matrix does not change under transformation. *The equation of a hypersphere must accordingly be of the form*[1]

$$(38) \qquad a\,(\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2) = \sum_{i=0}^{n} b_i\, \xi_0\, \xi_i .$$

[1] This form of the equation may be derived with equal ease from the above characterization.

Conversely, every equation of the form (38) represents either a hypersphere or a hypersurface of the type[2]

$$\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2 = 0. \tag{39}$$

This can be seen at once by reduction to the normal form.

In the reals, no hypersurface of the form (38) has a point in common with the improper hyperplane. This is no longer the case, however, if we consider equation (38) over the complex field, that is, if we extend the hypersurface into the complex domain by the addition of all the points of complex P_n that satisfy this equation. The intersection of the improper hyperplane with (38) is then represented by the equations

$$\xi_0 = 0, \qquad \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2 = 0. \tag{40}$$

This defines a hypersurface of the second order in the improper hyperplane whose dimension is lower by one than that of the hyperplane. In particular, in P_2 (40) represents a pair of (imaginary) improper points, and in P_3 an (imaginary) circle in the improper plane.

Now what is of significance here is that the set of points represented by (40) is independent of the coefficients that appear in (38). This means that *all* hypersurfaces of the type (38), and in particular all real hyperspheres, intersect the improper hyperplane in the set of points given by (40). This set of points is therefore common to all hyperspheres.

Conversely, a hypersurface of the second order that intersects the improper hyperplane in the set of complex points (40) can always be represented by an equation of the type (38). If we set ξ_0 equal to zero in a hypersurface $\sum_{i,k=0}^{n} a_{ik}\xi_i\xi_k = 0$, the resulting equation can represent the same set of points as (40) in complex improper P_{n-1} only if it differs from $\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2 = 0$ solely by a constant non-zero factor (Chap. X, Theorem 5).

Let us now return to the rigid motions of real P_n. Every such motion and, more generally, every affine transformation of real P_n, can be extended into the complex domain. This simply means that we consider the real system of equations (1) not just in real P_n but in complex P_n; and in the complex domain as well, the equations are referred to the natural coordinate system of p. 32 f. Such an affine transformation,

[2] In the reals, this equation represents a single point; in the complex field, a hypersphere with a real vertex.

an affine transformation in the complex space P_n that is specialized to the extent that the coefficients of its system of equations are all real is called a ***real** affine transformation of the complex space P_n.* Among all the affine transformations of the complex space P_n, the real affine transformations are characterized by the fact that they always map real points into real points.

Among these real affine transformations of the complex space P_n are to be found, in particular, the real rigid motions, i.e., those mappings of complex P_n that are represented in our coordinate system by a system of equations (1) with real coefficients and *orthogonal* matrix (2). Such a motion always takes the set of points (40) into itself. For a real hypersphere certainly goes into a real hypersphere,[3] and hence the improper part of a real hypersphere goes into the improper part of a real hypersphere, that is, (40) goes into itself. Because of this invariance under the group of rigid motions, the set of points (40) is called the **absolute** of euclidean geometry.

The question immediately arises whether the rigid motions, as a subclass of the class of real affine transformations, are characterized by the invariance of the absolute or whether there exist other real affine transformations that map the absolute into itself. To decide this question, consider a real affine transformation (1) that takes every point of (40) into a point of (40). Since the mapping induced in the improper hyperplane by (1) (cf. Chap. XI) is given by the system of equations

$$(41)\qquad \begin{aligned} \eta_1 &= t_{11}\,\xi_1 + t_{12}\,\xi_2 + \cdots + t_{1n}\,\xi_n, \\ \eta_2 &= t_{21}\,\xi_1 + t_{22}\,\xi_2 + \cdots + t_{2n}\,\xi_n, \\ &\cdots\cdots\cdots\cdots\cdots \\ \eta_n &= t_{n1}\,\xi_1 + t_{n2}\,\xi_2 + \cdots + t_{nn}\,\xi_n \end{aligned} \qquad (\eta_0 = \xi_0 = 0!)$$

our question amounts to this: What substitutions of the type (41) take the equation $\eta_1^2 + \eta_2^2 + \ldots + \eta_n^2 = 0$ into a multiple[4] of itself? Expressed in matrix notation, our question takes the following form: Find a substitution $(\eta) = T(\xi)$ which, when applied to the equation $(\eta)' \cdot (\eta) = O$, yields a result that differs from the original equation only by a constant factor. But the result of the substitution may be written $(\xi)' \cdot T'T \cdot (\xi) = O$, so that our condition means that $T'T$ is to be a multiple of the unit matrix. But the matrix T is the matrix of (41), i.e., the

[3] By virtue of the definition of hyperspheres as all the hypersurfaces of the second order that are congruent to a hypersurface of the type (37).

[4] The equation $\eta_1^2 + \eta_2^2 + \ldots + \eta_n^2 = 0$ of the absolute, as the equation of a hypersurface of the second order in the *improper* complex space P_{n-1}, is determined up to a non-zero constant of proportionality.

matrix (2). Thus we see the following: *The absolute is invariant under the mapping* (1) *if and only if matrix* (2) *satisfies the equation*

$$T' \cdot T = c \cdot E, \tag{42}$$

where c is a constant and E is the unit matrix.

It follows immediately from (42) that $|T'| \cdot |T| = c$, so that $c = |T|^2$. Consequently, since T was assumed real, c must be positive. Hence the square root of c is a real number, and we can write (42) in the form

$$\left(\frac{T}{\sqrt{c}}\right)' \cdot \left(\frac{T}{\sqrt{c}}\right) = E. \tag{43}$$

But this implies that $\dfrac{T}{\sqrt{c}}$ is an orthogonal matrix. Thus we have the following result:

A real affine transformation (1) *leaves the absolute invariant if and only if* (2) *is a multiple of an orthogonal matrix.*

The real affine transformations that map the absolute into itself are called *similarity transformations.* Our result shows that there is a close connection between the similarity transformations and the group of motions of real P_n. In order to stress this similarity we shall show that certain fundamental invariants of euclidean geometry, i.e., invariants of the group of rigid motions, are also invariant under similarity transformations. First of all:

A similarity transformation leaves all angles invariant. For if the matrix (2) is a multiple of an orthogonal matrix, say $T = c \cdot T^*$ (with $c \neq 0$ and T^* orthogonal), we decompose (1) into the product of two mappings by first applying the mapping

$$\xi_0' = \xi_0, \qquad \xi_i' = c \cdot \xi_i \quad \text{for } i = 1, 2, \cdots, n \tag{44}$$

and then

$$\eta_0 = \xi_0', \qquad \eta_i = t_{i0}\, \xi_0' + \sum_{k=1}^{n} \left(\frac{t_{ik}}{c}\right) \xi_k', \quad i = 1, 2, \cdots, n. \tag{45}$$

We see immediately that all vectors of the proper part of P_n go over into their multiples by c under the mapping (44), and thus the angle between two such vectors, as defined in § 7 of *Modern Algebra,* remains unchanged. But the mapping (45) is a rigid motion (its matrix is T^*) and thus in turn leaves the angle invariant.

This discussion also shows that under our similarity transformations all distances are multiplied by the constant c. For the mapping (44) has this property, while (45) leaves distances invariant. It follows that *equality* of distances is invariant under similarity transformations. Even more: *All relations of the form* $\overline{QR} = \lambda \cdot \overline{ST}$ *between two distances are unaltered by a similarity transformation.*[5]

Moreover, the characterization of similarity transformations by the invariance of the absolute permits us to think of the invariants of these mappings as projective relations to the absolute. We should like to make clear what the main points are of this reinterpretation, taking angle as our example.

Let us begin with two non-vanishing vectors $\mathfrak{x} = \{x_1, x_2, \ldots, x_n\}$ and $\mathfrak{y} = \{y_1, y_2, \ldots, y_n\}$ of the affine space R_n. Each of the two vectors determines a family of lines parallel to it, all having the same improper point in common. Let U_x and U_y denote the respective improper points of the families of lines parallel to $\mathfrak{x}$ and $\mathfrak{y}$. The *homogeneous* coordinates of these points are (cf., for example, Chap. I)

$$(46)\qquad U_x = [0, x_1, x_2, \cdots, x_n], \qquad U_y = [0, y_1, y_2, \cdots, y_n].$$

If $\mathfrak{x}$ and $\mathfrak{y}$ are not parallel, which we shall assume to be the case, then $U_x \neq U_y$. The line determined by these two points belongs entirely to the improper space P_{n-1} and therefore intersects the absolute in two (distinct or coincident) points. Let the points of intersection be V_1, V_2. We shall now express the angle between the vectors $\mathfrak{x}$ and $\mathfrak{y}$ in terms of the cross ratio $\mathcal{R}(U_x\, U_y\, V_1\, V_2)$.

Since the angle between $\mathfrak{x}$ and $\mathfrak{y}$ as well as the points U_x and U_y are independent of the length of the vectors $\mathfrak{x}$ and $\mathfrak{y}$, we shall assume, to simplify our computation, that both vectors are of unit length. That is, we assume that

$$(47)\qquad \sum_{i=1}^{n} x_i^2 = 1, \qquad \sum_{i=1}^{n} y_i^2 = 1.$$

[5] For an arbitrary affine transformation, this is in general true only if the four points Q, R, S, T are collinear.

Under this assumption, the angle α formed by the two vectors is given by

$$\cos\alpha = \sum_{i=1}^{n} x_i y_i. \tag{48}$$

Now, in order to determine the cross ratio mentioned above, we must express the points V_i as linear combinations of the points U_x, U_y. Since the U_x, U_y, as real points,[6] are distinct from the necessarily non-real V_1, V_2, we can write the coordinates of the points V_i in the form

$$\lambda x_i + y_i. \tag{49}$$

Since, moreover, the V_i lie on the absolute, they must satisfy its equation, i.e., we must have

$$\sum_{i=1}^{n} (\lambda x_i + y_i)^2 = 0,$$

or

$$\lambda^2 \left(\sum_{i=1}^{n} x_i^2\right) + 2\lambda \left(\sum_{i=1}^{n} x_i y_i\right) + \sum_{i=1}^{n} y_i^2 = 0.$$

In view of (47) and (48), it follows that

$$\lambda^2 + 2\lambda \cdot \cos\alpha + 1 = 0. \tag{50}$$

From (50) we obtain the two possible values for λ in (49)

$$\begin{aligned} V_1:\ \lambda_1 &= -\cos\alpha - i\sin\alpha,\\ V_2:\ \lambda_2 &= -\cos\alpha + i\sin\alpha. \end{aligned} \tag{51}$$

We now obtain immediately as the value of the required cross ratio

$$\begin{aligned} \mathcal{R}(U_x U_y V_1 V_2) &= \frac{\lambda_2}{\lambda_1} = \frac{\cos\alpha - i\sin\alpha}{\cos\alpha + i\sin\alpha}\\ &= [\cos(-\alpha) + i\sin(-\alpha)] \cdot [\cos\alpha + i\sin\alpha]^{-1}\\ &= \cos(-2\alpha) + i \cdot \sin(-2\alpha). \end{aligned} \tag{52}$$

---

[6] The assumption that $\mathfrak{x}$, $\mathfrak{y}$ are vectors of the affine space R_n implies that the numbers x_i, y_i are real.

The simple relation thus disclosed between the angle α and $\mathcal{R}(U_x U_y V_1 V_2)$ can be written in the form

(53) $$\mathcal{R}(U_x U_y V_1 V_2) = e^{-2i\alpha}$$

if we borrow the identity

$$e^{i\omega} = \cos\omega + i\sin\omega$$

from analysis.[7]

Formula (52) (or (53)) is due to Laguerre. It reveals a surprising connection between projective and euclidean-metric geometry. This connection can be developed further. For example, it is easy to see that *perpendicularity* of two vectors $\mathfrak{x}$ and $\mathfrak{y}$ means that the cross ratio of the points U_x, U_y, V_1, V_2 has the value -1 and thus that the four points form a harmonic set. Moreover, perpendicularity of a hyperplane and a line means that the improper parts of each represent pole and polar with respect to the absolute. And so forth.

The possibility of such a reinterpretation of euclidean-metric concepts by means of the Laguerre formula paves the way, too, for new *'non-euclidean'* geometries which, from this point of view, are entitled to positions of equal standing alongside euclidean geometry.

In fact, the characterization of the similarity transformations by the invariance of the absolute immediately suggests the following generalization: We take, instead of the absolute, any other (degenerate or nondegenerate) hypersurface of the second order and consider all projectivities that map this hypersurface onto itself. These projectivities constitute a subgroup of the projective group. Then to each such subgroup there belongs a particular geometry. In this way, we obtain a whole host of new geometries, the study of which is readily accessible precisely to the extent that their essential invariants can be derived from projective relations to the hypersurface considered as fixed. The importance of the Laguerre formula lies in the fact that it shows how metric concepts can be introduced to describe such geometries, whose properties, to be sure, are often widely divergent from those of euclidean geometry.[8]

[7] If V_1 and V_2 be interchanged (in P_n, with $n > 2$, they are in fact indistinguishable without further information), the cross ratio goes into its reciprocal, i.e., the minus sign disappears in the exponent in (53).

[8] This train of thought, which we cannot pursue here in any detail, was initiated by Cayley and Klein and has proved to be an extraordinarily fruitful approach. In particular, it afforded a new approach to the non-euclidean geometries that were already known, these appearing as special cases of the above. In this connection, the reader is referred to Felix Klein's admirable book, *Vorlesungen über nicht-euklidische Geometrie* (Berlin, 1928; New York, 1960).

Exercises

1. Let η_1, η_2 be two real[9] planes in complex P_3 and let g denote their line of intersection. Argue that through g there pass just two (conjugate complex) planes φ_1, φ_2 whose improper lines are tangents to the absolute of three-dimensional euclidean geometry. Then show that

$$\mathcal{R}(\varphi_1\,\varphi_2\,\eta_1\,\eta_2) = e^{2i\alpha},$$

where α is one of the two angles formed by the planes. If φ_1 and φ_2 are interchanged in the formula, α is replaced by its supplement.

Also prove the n-dimensional generalization of this. The angle formed by two hyperplanes is to be understood to mean the angle between two vectors that are perpendicular to the respective hyperplanes.

2. Consider a fixed real non-degenerate hypersurface of the second order F in the complex space P_n. Every real line g that is not tangent to F cuts F in two points V_1, V_2, which are either both real or conjugates complex of each other. If Q_1, Q_2 are two other *real* points of g, then $\mathcal{R}(V_1\,V_2\,Q_1\,Q_2)$ is real in the first case and is a complex number of absolute value 1 in the second case, as is easy to see. We therefore set $\mathcal{R}(V_1\,V_2\,Q_1\,Q_2) = r\cdot(\cos\varphi + i\cdot\sin\varphi)$ with $r > 0$ and, since $e^{i\varphi} = \cos\varphi + i\cdot\sin\varphi$, define: $\ln\mathcal{R}(V_1\,V_2\,Q_1\,Q_2) = \ln r + i\cdot\varphi$ (where $\ln r$ denotes the natural logarithm of r).

If we start with two real points Q_1, Q_2, the points V_1, V_2 are determined except for their order (as the points of intersection of the line through Q_1, Q_2 with F). Now, with the use of a constant c, we construct the function $E(Q_1Q_2) = c\cdot\ln\mathcal{R}(V_1\,V_2\,Q_1\,Q_2)$. Now, $E(Q_1\,Q_2)$ is called the *non-euclidean distance* between the points Q_1, Q_2 relative to the non-euclidean '*absolute*' F. Show that

a) $E(Q_1\,Q_2)$ is unique only up to sign and an added integral multiple of $2\pi i\cdot c$.
b) If $Q_1 = Q_2$ $(Q_\iota \neq V_1, V_2)$, then $E(Q_1\,Q_2) = 0$.
c) If Q_1, Q_2, Q_3 are three real points of a line, then $E(Q_1\,Q_2) + E(Q_2\,Q_3) = E(Q_1\,Q_3)$ if, in each of these expressions, the same order is retained for the points of intersection V_1, V_2 of g with F.

A *non-euclidean angular measure* relative to F as absolute may be introduced in an exactly dual way (cf. Exercise 1). The angle $A(\eta_1\,\eta_2)$ between two hyperplanes η_1, η_2 is similarly defined by the formula $A(\eta_1\,\eta_2) = c\cdot\ln\mathcal{R}(\varphi_1\,\varphi_2\,\eta_1\,\eta_2)$, where φ_1, φ_2 are the two tangent hyperplanes to F belonging to the pencil of planes determined by η_1, η_2. Statements dual to a), b), and c) above hold for this angular measure.

We obtain two important special cases of this non-euclidean measure if we take for F the hypersurfaces $x_1^2 + x_2^2 + \cdots + x_n^2 = \pm 1$ (in affine coordinates). We speak of *elliptic* or *hyperbolic* measure, according as the minus or the plus sign occurs on the right. In the elliptic case F is imaginary; if one then takes c in the definition of

[9] By this is meant, of course, a plane that can be represented by an equation with real coefficients with respect to a real coordinate system—say, the natural coordinate system—of the complex space P_3. In what follows, the concepts of a 'real' hypersurface of the second order and a 'real' linear space are to be understood analogously.

$E(Q_1 Q_2)$ to be a pure imaginary number, then all the distances $E(Q_1 Q_2)$ are real (for real Q_i). In the hyperbolic case c is taken as a real quantity; then $E(Q_1 Q_2)$ is real for all pairs of points in the 'interior' of F (interior: $x_1^2 + x_2^2 + \cdots + x_n^2 < 1$). The constant used in the definition of the 'angle' $A(\eta_1 \eta_2)$ is taken as a pure imaginary in both cases, with analogous consequences.

3. The non-euclidean distances and angles defined in Exercise 2 are obviously invariant under all linear projectivities of P_n that map F into itself (the so-called 'non-euclidean motions' relative to F). But also, conversely, if $E(Q_1 Q_2) = E(Q_1' Q_2')$, then there always exists a linear projectivity of P_n that takes F into itself and also maps Q_1 into Q_1' and Q_2 into Q_2'. A corresponding dual statement holds true for the angle.

INDEX